

A Hybrid Principal Component Analysis–Entropy Approach to Balanced Weighting in Single Usability Metric for E-Learning Evaluation



Muhammad Hamka^{1,2*}, Purwanto Purwanto³, Adi Wibowo⁴

¹ Doctoral Program of Information System, Diponegoro University, Semarang 50275, Indonesia

² Department of Informatics Engineering, Universitas Muhammadiyah Puwokerto, Purwokerto 53182, Indonesia

³ Department of Chemical Engineering, Diponegoro University, Semarang 50275, Indonesia

⁴ Department of Informatics, Diponegoro University, Semarang 50275, Indonesia

Corresponding Author Email: muhammadhamka@students.undip.ac.id

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310324>

ABSTRACT

Received: 4 December 2025

Revised: 10 February 2026

Accepted: 23 March 2026

Available online: 31 March 2026

Keywords:

e-learning usability, e-learning platforms, Single Usability Metric, hybrid Principal Component Analysis–Entropy weighting, task-level usability evaluation, Z-score normalization

Usability evaluation in e-learning often encounters a discrepancy between subjective perception and actual performance, leading to biased or incomplete assessments. The System Usability Scale (SUS) measures perceived usability but fails to account for task-level performance, whereas traditional Single Usability Metric (SUM) formulations depend on assumption-based weighting methods that may skew the significance of indicators, leading to potentially misleading conclusions about the usability of e-learning systems. This paper introduces a hybrid Principal Component Analysis–Entropy (PCA)–Entropy methodology to improve SUM by concurrently analyzing variance structure and information dispersion. Usability testing was performed on nine representative learning management systems (LMS) tasks with 73 student participants and 6 professional evaluators. Objective data (completion rate, error rate, time on task) and subjective satisfaction (SUS) were normalized utilizing task-level Z-scores. PCA was utilized to uncover latent correlations among usability factors, while entropy weighting enhanced their relative significance and reduced dominance bias. The findings indicate that effectiveness is the primary usability dimension (weight = 0.4294), followed by efficiency and satisfaction (both = 0.2423). Task-level study reveals Resource (16.29%), Collaboration (41.83%), and Assessment (40.04%) as significant usability impediments, while Interaction (67.55%) and Log In (66.88%) demonstrate consistent performance. These findings indicate that the combination of structural variance and informational diversity yields a more stable, interpretable, and diagnostically accurate usability metric, especially in heterogeneous conditions, although sensitivity to normalization assumptions and dimensional scope persists.

1. INTRODUCTION

The rapid development of digital learning technologies has had a significant impact on higher education globally [1]. E-learning platforms and learning management systems (LMS) are increasingly used to create flexible and scalable learning environments, enabling students to access educational materials and interact with teachers and classmates online [1, 2].

The efficacy of these systems relies not only on their technological capabilities but also on the ease of human interaction, which includes factors such as user interface design, accessibility features, and the quality of communication tools available for interaction between students and instructors. In this context, usability is essential in assessing learners' abilities to effectively complete learning tasks in digital environments [1, 3]. Usability is typically defined as the degree to which users can accomplish their objectives effectively, efficiently, and satisfactorily when engaging with a system [4]. Previous studies emphasize that

usability significantly influences user experience, system adoption, and the sustainability of e-learning platforms [4, 5]. Poor usability in LMS can result in navigation challenges, ineffective interaction, and diminished learning efficacy [3, 5].

The domain of human–computer interaction (HCI) provides multiple approaches for evaluating the usability of digital systems [5, 6]. These methodologies encompass heuristic evaluation, usability testing, user experience surveys, and task-oriented performance assessments [3–5]. These techniques aim to assess the effectiveness of user interactions with digital systems and the simplicity of job completion inside an interface [5, 7]. In e-learning systems, usability assessment is essential for identifying interaction problems that may affect learners' engagement and performance [1, 2, 4], such as confusing navigation, unclear instructions, or slow response times, which can hinder the overall learning experience.

Among the available usability evaluation approaches, questionnaire-based methods have become the most prevalent due to their simplicity and effectiveness [3]. The System

Usability Scale (SUS) is a standardized questionnaire designed to evaluate users' subjective perceptions regarding the usability of a system [8]. This instrument is among the most commonly utilized tools. SUS has been extensively employed in the assessment of various digital systems, including educational technologies and LMS, due to its ability to swiftly and accurately gauge perceived utility [9, 10]. Nonetheless, subjective usability assessments, like the SUS, primarily center on users' perceptions and may not truly reflect the efficiency with which users accomplish tasks within a system [7].

Although subjective usability assessments offer important insights into user views, they may not accurately represent real task performance during system interaction [7, 8]. Objective performance measurements are crucial in usability testing to comprehensively comprehend user interactions with a system during actual job execution [7, 11].

These indicators often encompass job completion rate, time spent on task, error rate, and user satisfaction [4, 12]. Nonetheless, assessing these indications in isolation may yield only restricted understanding of the system's overall usability [5]. Incorporating multiple usability metrics could therefore result in a more comprehensive understanding of system usability and user engagement [2, 4].

To deal with this issue, researchers have proposed composite usability metrics that integrate multiple usability indicators into a unified evaluation framework [13, 14]. The Single Usability Metric (SUM) is a widely recognized approach that incorporates essential usability metrics—such as completion rate, task duration, error rate, and user satisfaction—into a standardized usability score [13, 15].

By standardizing several usability metrics into a comparable scale, SUM enables the combination of several usability metrics into a single measure, providing examiners to acquire a comprehensive, interpretable evaluation encompassing the overall efficacy of system usability [13, 14].

Despite its advantages, researchers have identified certain methodological issues related to the weighting of usability metrics within the SUM framework [14]. Traditional SUM solutions often rely on equal weighting or assumption-based aggregation of usability metrics, which can distort the relative importance of individual indicators [16].

Prior studies indicate that calculating completion rates using the conventional SUM formulation may induce bias, resulting in inflated usability scores and potentially misleading evaluation outcomes [17]. Consequently, systems may seem functional based on the cumulative score, yet, nevertheless, harbor fundamental usability issues [14, 17].

Considering these limitations, there is an increasing need for refined weighing methodologies to enhance the reliability of composite usability assessment [18]. Consequently, data-driven weighing methodologies have been investigated to more accurately reflect the relative significance of usability characteristics [19, 20]. Principal component analysis and entropy-based weighting have been widely employed to determine objective weights for evaluation indicators [19, 21].

This work proposes a hybrid Principal Component Analysis (PCA)–Entropy weighting method to enhance the computation of the SUM for assessing e-learning systems [21]. The suggested method integrates PCA with entropy weighting to enhance the balanced representation of usability indicators and augment the reliability of usability evaluation outcomes [16, 19].

2. LITERATURE REVIEW

2.1 Usability evaluation in e-learning system

Usability has emerged as an essential aspect for evaluating the sustainability and effectiveness of e-learning platforms in higher education [22, 23]. As LMS increasingly facilitate instructional activities, the quality of user interaction considerably impacts learning experiences and system acceptance. Previous research consistently demonstrates that usability influences user engagement, learning effectiveness, and prolonged system usage in digital learning environments [1, 4, 5].

In usability engineering, system usability is often assessed through three fundamental dimensions: effectiveness, efficiency, and satisfaction [9, 24]. These dimensions pertain to users' ability to perform tasks accurately, accomplish objectives with minimal effort or time, and perceive the system as satisfactory during interaction [25-27]. Empirical research on LMS usability often utilizes these parameters to evaluate student interaction with educational platforms during actual learning activities [24, 28, 29].

Nonetheless, despite the broad acknowledgment of usability as a critical determinant of success in e-learning systems, prior research indicates significant methodological variation in the evaluation of usability [3, 5, 10]. Numerous research studies predominantly utilize perception-based instruments, while some adopt task-based performance indicators, such as completion rates and time on task, to assess usability in e-learning environments [7, 9, 11]. This methodological variability has prompted researchers to investigate more holistic usability evaluation frameworks that integrate subjective and objective metrics [2, 7].

2.2 Limitations of perception-based usability evaluation (System Usability Scale)

Among the various usability evaluation instruments, the SUS stands out as one of the most commonly utilized tools for assessing perceived usability [10]. The SUS comprises a standardized questionnaire with ten Likert-scale items aimed at assessing users' perceptions of system usability [8, 30]. SUS has been widely used to evaluate various digital systems, including LMS and adaptive e-learning platforms, due to its simplicity and reliability [31, 32].

Numerous studies have utilized perception-based assessment to examine usability and user approval in educational technology contexts [6, 9]. For instance, Simon et al. [10] indicated that SUS is among the most commonly utilized tools for evaluating student experiences on LMS platforms. Alghabban and Hendley [9] established that perceived usability substantially affects the acceptance of adaptive e-learning systems. In addition to standardized usability questionnaires, perception-based evaluation has been conducted through survey-based analytical models. Al-Adwan et al. [28] investigated the efficacy of higher education e-learning systems by structural equation modeling, focusing on users' perceptions, behavioral intentions, and satisfaction dimensions.

Despite their widespread adoption, perception-based usability evaluation methods highlight several methodological shortcomings [3, 22, 30]. The SUS predominantly measures users' subjective perceptions rather than their actual performance during job execution [7]. As a result, critical usability metrics such as task completion rate, interaction

mistakes, and time on task are rarely explicitly assessed [11, 12, 33]. Secondly, subjective usability assessments may be affected by previous experience, familiarity with analogous systems, as well as system-related interaction constraints, or anticipations concerning system performance [5, 34]. Consequently, perceived usability ratings may not consistently align with real usability performance [7, 35].

Empirical evidence corroborates this contradiction. Altin Gumussoy et al. [7] noted that subjective evaluations of usability frequently differ from objective measures of task performance. The findings indicate that dependence exclusively on perception-based measurements may yield an inadequate depiction of system usability, since they exclude critical elements like user efficiency and mistake rates, which are more precisely assessed by objective usability metrics [7, 11, 35].

2.3 Emergence of composite usability metrics: The Single Usability Metric framework

To deal with the disadvantages of single-method usability evaluation, experts have suggested composite usability metrics that combine various usability indicators into a cohesive evaluation framework [13, 14]. The SUM, introduced by Sauro and Kindlund [13, 36], stands out as one of the most acknowledged methodologies in the field.

SUM integrates various task-oriented usability metrics—like completion rate, error rate, time on task, and user satisfaction—into a uniform usability score using Z-score transformation [18]. By aggregating multiple usability indicators into a single metric, SUM allows for the evaluation of usability performance across systems, tasks, or user groups using a comparable measurement scale [36].

Previous studies underscore the advantages of composite usability metrics in providing a more comprehensive representation of the quality of user interactions [15, 36]. Specifically, Albert and Tullis [15] highlight that integrated usability metrics allow for the capture of various dimensions of user experience that individual usability indicators alone cannot reveal.

Nonetheless, in spite of its theoretical benefits, SUM presents methodological difficulties concerning the consolidation of usability indicators [14]. Specifically, establishing suitable weights for various usability dimensions continues to be a significant challenge in calculating the SUM score [14, 17], as different dimensions may have varying levels of importance depending on the context of use and user needs [16].

2.4 Weighting challenges in Single Usability Metric

Conventional implementations of SUM often depend on uniform weighting or assumption-driven aggregation of usability metrics. These methodologies implicitly presume that all usability dimensions contribute uniformly to total usability performance. Empirical research indicates that the significance of various usability indicators may differ based on the interaction setting [14, 16].

Numerous research studies have underscored potential biases linked to conventional SUM formulations [14, 17]. In addition, Pearson [17] illustrated that completion-rate bias can arise when completion measurements predominate the overall usability score, potentially obscuring usability issues associated with interaction errors or ineffective task

performance. Similarly, Van Waardhuizen et al. [14] observed that the aggregation of usability data without suitable weighting mechanisms may skew the interpretation of usability performance, leading to potentially misleading conclusions about user experience and satisfaction.

The weighting approach used to aggregate usability factors significantly influences the reliability and interpretability of SUM, as demonstrated by the results. As a result, researchers have progressively investigated data-driven methodologies for establishing indicator weights, such as machine learning techniques and statistical analyses, to ensure a more accurate representation of usability performance [19, 21].

2.5 Data-driven weighting approaches: Principal Component Analysis and entropy

Recently, methods that utilize data-driven weighting, derived from multi-criteria decision-making studies, have seen a growing application in assessing the significance of indicators within intricate evaluation frameworks [21]. Principal Component Analysis (PCA) stands out as a prominent statistical method employed to uncover latent structures within multidimensional datasets [37, 38].

PCA identifies the weights of indicators through an examination of the variance structure and correlation patterns present among variables, effectively minimizing redundancy among correlated indicators [19]. Through the extraction of principal components that account for the greatest amount of variance in the dataset, PCA offers an impartial foundation for assessing the significance of evaluation criteria.

A commonly employed method is entropy weighting, which assesses the significance of indicators by examining the extent of information variability across variables [19, 20]. Indicators that exhibit higher variability provide more information and, consequently are assigned greater weights during the evaluation process. Entropy weighting has found extensive application across diverse decision-making scenarios, such as urban planning, healthcare evaluation, and industrial assessment [20].

While both PCA and entropy weighting offer objective methods for assessing the significance of indicators, each approach comes with its own set of limitations. PCA highlights the structure of variance, yet it might neglect the informational diversity present among indicators [19]. On the other hand, entropy weighting emphasizes the distribution of information while failing to explicitly consider the relationships between evaluation indicators, which can lead to an incomplete understanding of how these indicators interact and affect the overall assessment [19, 21].

2.6 Theoretical integration and research gap

The complementary characteristics of PCA and entropy weighting have inspired the creation of hybrid weighting methodologies that integrate the strengths of both techniques. In multi-criteria decision-making contexts, hybrid PCA–entropy models have proven the ability to yield more equitable and objective weighting outcomes by concurrently accounting for variance structure and information dispersion among assessment indicators. Pliego-Martínez et al. [21] suggested an integrated PCA–entropy weighting method to enhance the reliability of composite indicator assessment.

Despite these methodological advancements, usability evaluation frameworks still face constraints in their use of

hybrid PCA–entropy weighting. Present studies on usability assessment in e-learning systems predominantly depend on either subjective perception-based tools like SUS or composite usability indices such as SUM [5, 10]. SUS accurately measures perceived usability, but it fails to represent task-level

interaction performance during real system utilization [7]. In contrast, SUM incorporates various task-oriented usability metrics but generally depends on predetermined or assumption-driven weighting systems, which may inject bias into the overall usability score [13, 17].

Table 1. Comparison of related studies and current work

Study	Focus	Approach	Key Limitation	Significance to This Research
Tania et al. [23]	E-learning adoption	Perception-based review	No operational usability metrics	Motivates usability-driven evaluation
Al-Adwan et al. [28]	HE e-learning success	SEM (subjective evaluation)	Lacks task-level performance data	Complements with objective usability metrics
Ferreira et al. [24, 29]	Usability dimensions	Controlled usability experiments	No composite usability index	Supports SUM-based integration
Simon et al. [10]	LMS usability & UX	Scoping review	No weighting framework	Motivates structured usability modelling
Talib et al. [5]	LMS UX	Literature review	No quantitative aggregation	Reinforces need for data-driven weighting
Novák et al. [6]	UX evaluation methods	Systematic review	No unified usability scoring model	Supports unified SUM scoring
Pearson [17]	SUM bias	Bias analysis	Single-metric dominance in SUM	Motivates improved SUM weighting
Nasr and Zahabi [32]	SUMA framework	SUM extension	Core weighting strategy unchanged	Shows extensibility of SUM
Pliego-Martínez et al. [21]	Indicator weighting	PCA–entropy hybrid weighting	Not applied to usability evaluation	Provides hybrid weighting basis
Torres-Molina and Seyam [41]	LMS usability prediction	Machine learning using interaction logs	Requires extensive interaction datasets and demands extensive interaction datasets, training and implementing complex models presents challenges, and particularly regarding the limited interpretability of usability metrics	Highlights complementary role of usability metric frameworks
This study	E-learning usability evaluation	Hybrid PCA–entropy weighting integrated with SUM	Addresses the identified limitations	Improves the computation of the Single Usability Metric (SUM) through hybrid PCA–entropy weighting

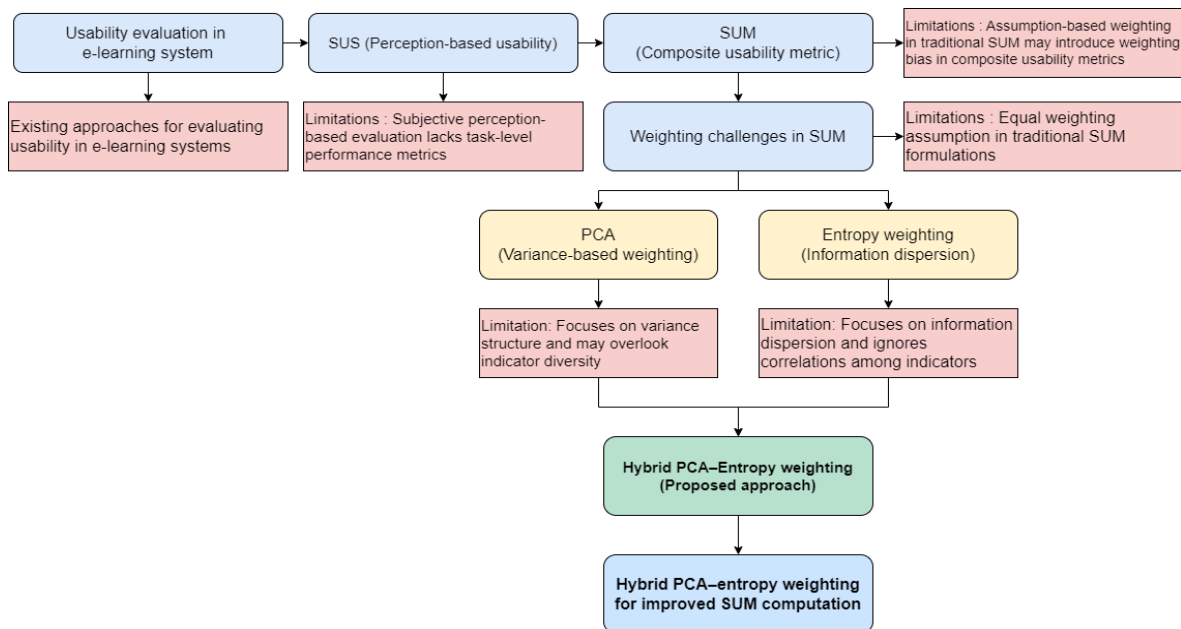


Figure 1. Hybrid Principal Component Analysis–Entropy weighting framework for improved Single Usability Metric (SUM) computation

Recent progress in usability analytics has explored machine learning and learning analytics techniques to analyze interaction logs, gaze data, and behavioral patterns during user

engagement [39, 40]. These methodologies provide extensive analysis of user behavior and assist in the automatic identification of usability problems [41]. Nevertheless,

numerous machine learning-based usability evaluation methods predominantly emphasize behavioral pattern detection or predictive modeling instead of generating interpretable and standardized usability metrics. Moreover, these methodologies frequently necessitate comprehensive data preprocessing, hand labeling, and context-specific training datasets, hence constraining their generalizability and incorporation into established usability evaluation frameworks.

Thus, a systematic and interpretable data-driven weighting methodology is required to consolidate various usability indications into a singular usability index. This study proposes a hybrid PCA–Entropy weighting approach to improve the computation of the SUM for evaluating usability in e-learning systems. To clarify the identified research gap, Table 1 summarizes representative studies and their key limitations, while Figure 1 presents the research framework illustrating how the proposed hybrid PCA–Entropy weighting approach addresses these limitations within the SUM framework.

3. METHODS

This study proposes a structured usability evaluation framework for higher education e-learning platforms by

refining the SUM through a hybrid PCA–Entropy weighting approach. The methodology combines objective task-based usability metrics with subjective user perceptions to overcome the shortcomings of traditional SUM implementations that depend on static or heuristic weighting systems. The research procedure consisted of four main phases: common task and error model design, participant selection, data collection, and usability data analysis. The overall research framework is illustrated in Figure 2, while the computational workflow of the proposed PCA–Entropy-enhanced SUM is presented in Figure 3.

3.1 Common task and error model design

Common task scenarios have been developed for demonstrating the fundamental features of e-learning platforms in higher education. The identification of tasks was guided by previous empirical research on LMS usability and interaction patterns, together with international usability standards pertinent to task-based usability assessment. Prior research consistently delineates numerous fundamental tasks of LMS, encompassing authentication, course navigation, content access, communication and collaboration, assessment submission, feedback retrieval, search capabilities, and user control features [40–43].

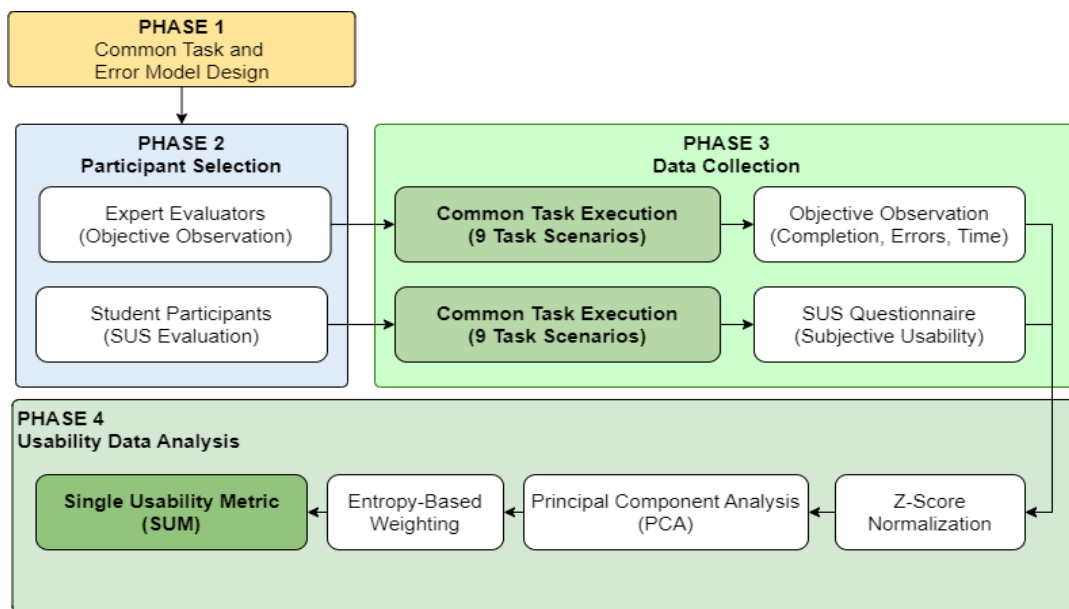


Figure 2. Proposed research framework for task-level usability evaluation and Single Usability Metric (SUM)

Drawing from this synthesis of the literature, nine representative task scenarios were developed to illustrate the typical academic workflows that students engage in while interacting with an LMS environment. The tasks were structured to effectively capture essential functional interactions while ensuring experimental feasibility throughout the usability testing process [42].

To guarantee the representativeness and clarity of the chosen tasks, the scenarios underwent a thorough review and validation process conducted by six e-learning administrators who possess professional experience in managing LMS platforms within higher education [44, 45]. The validation process emphasized the importance of task relevance, the clarity of instructions, and the alignment with actual user interactions [46]. Consensus among experts was reached via a thorough review process, enhancing the content validity of the

chosen tasks.

Error measurement was integrated as a fundamental element of task effectiveness within the SUM framework [33, 41]. Rather than examining individual error occurrences, potential errors were predetermined and classified into functional categories to improve the clarity of the analysis. The identified categories included navigation errors, feedback-related errors, problems with system responsiveness, inconsistencies in interface design, and failures related to search functionalities [4, 7, 44, 46, 47]. The task scenarios and associated error categories were derived from prior usability studies in e-learning systems [45–50]. Table 2 illustrates the correlation between task scenarios and the categorized task-specific errors that were observed empirically during the execution of tasks.

Table 2. Common e-learning task scenarios and functional error categories

No.	Category	Description	Usability Aspect	Observed Error Categories
1	Log In	Users log in to the LMS using student credentials	Accessibility	Feedback visibility, input validation, authentication flow
2	Navigation	Users access course areas and discussion forums	Navigation & Information Architecture	Navigation structure, mobile responsiveness, content visibility
3	Collaboration	Users post or reply to discussion forums	Interactivity	Action visibility, content submission flow, system feedback
4	Communication	Users communicate via chat features	Interactivity	Interface clarity, responsiveness, message handling
5	Interaction	Users access learning materials and quizzes	Efficiency	Content discoverability, system responsiveness
6	Assessment	Users submit course assignments	Effectiveness	Submission workflow, confirmation feedback, deadline visibility
7	Feedback	Users review instructor feedback	Feedback Mechanism	Notification visibility, content accessibility
8	Interface	Users modify profile and account settings	User Control & Flexibility	Menu discoverability, input validation, confirmation feedback
9	Resource	Users search for courses or learning activities	Search Functionality	Search accuracy, filtering, result presentation

3.2 Participant

To get a full picture of usability from both objective and subjective points of view, the participants were split into two groups. The initial group comprised six e-learning administrators hailing from various higher education institutions, chosen deliberately due to their extensive professional experience in the management and evaluation of e-learning platforms. The sample size aligns with recognized usability inspection standards, indicating that a group of five to six expert evaluators is adequate for uncovering most usability issues in task-oriented assessments [51-53].

A total of 72 undergraduates were recruited from LMS-supported courses at the same institutions as the expert evaluators for the second group. Participants were selected based on their genuine engagement with the e-learning platforms, ensuring that they represented authentic end users with direct interaction experience. This course-based recruitment method is consistent with previous research on the evaluation of perceived usability in LMS, wherein students evaluate systems integrated into their academic experiences [31].

Participants in the study were tasked with assessing subjective usability through the SUS. A total of twelve respondents were gathered from each institution, and this sample size is deemed sufficient, as prior studies have indicated that SUS yields reliable outcomes even with comparatively small samples [31, 53].

The deliberate distinction between expert evaluators and student participants was crafted to differentiate between objective performance measurement and subjective perception assessment. Specialized evaluators meticulously documented task-oriented usability metrics, such as completion rates, error frequencies, and time spent on tasks, maintaining methodological rigor throughout the observation process. In contrast, the student participants acted as genuine end users and were tasked with delivering subjective usability assessments via the SUS. This design minimizes the likelihood of bias and cognitive load that could occur when participants are tasked with performing activities while also assessing their own performance.

3.3 Data collection

The period of data collection was February to May 2025.

The evaluated e-learning platforms were employed by all participants to complete the same nine predefined task scenarios, thereby guaranteeing a consistent interaction context among the groups.

Expert evaluators employed a standardized protocol to capture objective usability data through structured observation. The task execution was screen-recorded to guarantee accuracy and facilitate verification, and the time spent on the task was measured from the commencement to the completion using a digital timer. Errors were classified into functional categories using a predefined error coding manual that was derived from prior usability studies. These categories included navigation, feedback, system responsiveness, interface consistency, and search-related issues. The coding scheme was consistently applied by all evaluators as a result of their training.

A five-point Likert scale was employed to collect subjective usability data from student participants using the SUS, which was administered promptly following the completion of the task scenarios. The resulting metrics represent complementary dimensions of the same user experience and support their integration within the SUM framework, although data were collected from two distinct groups that interacted with the same system, tasks, and conditions.

3.4 Usability metrics and standardization

This study's usability evaluation was separated into three main dimensions: effectiveness, efficiency, and satisfaction [24, 29]. The effectiveness was evaluated using task completion rate (CR) and error rate, which signify accomplishing objectives and interaction quality, respectively [54]. A multiplicative formulation was adopted instead of a simple arithmetic average, whereby the incidence of errors proportionally diminishes the contribution of successful job completion. This method averts exaggerated effectiveness ratings when elevated completion rates coincide with significant mistakes. Effectiveness is calculated as specified in Eq. (1), indicating the degree of task success attained with minimal interaction errors, in accordance with recognized usability evaluation criteria [4, 12, 54].

$$Eff = (CR \times (1 - error_{rate})) \times 100\% \quad (1)$$

Efficiency was measured by the ratio of task completion rate to average task completion time, following the guidelines of ISO/IEC 25062:2006 [19]. To allow comparison across tasks, efficiency values were standardized to a 0–100% scale. The efficiency formula is presented in Eq. (2).

$$Effcy = \frac{\frac{CR}{T}}{Effcy_{max}} \times 100\% \quad (2)$$

The SUS was employed to assess satisfaction. Individual SUS scores were computed using the standard scoring procedure, where adjusted item scores were summed and multiplied by a constant factor [30]. The computation of the SUS score is detailed in Eq. (3), and the mean SUS score was calculated to reflect the overall perceived usability of the system.

$$SUS = 2.5 \times \left[\sum_{n=1}^5 (U_{2n-1} - 1) + (5 - U_{2n}) \right] \quad (3)$$

Prior to their integration into the SUM, all usability metrics underwent standardization through Z-score transformation [18]. The computation of a Z-score entails subtracting the sample mean from the observed value and subsequently dividing the result by the standard deviation, as specified in Eq. (4). The standardized values enabled direct comparison and aggregation across multiple usability dimensions. Usability metrics were standardized via Z-score normalization at the task level (refer to Eq. (4)). The standardized values were then integrated at the usability dimension level using PCA and entropy-based weighting to calculate the SUM.

$$Z = \frac{x - \mu}{\sigma} \quad (4)$$

3.5 Principal Component Analysis-Entropy weighting and Single Usability Metric

This study employs the integration of PCA and entropy weighting as a complementary analytical framework to address the limitations associated with single-method weighting approaches [19, 21]. PCA is utilized to identify the underlying correlation structure among usability dimensions and to reduce redundancy by projecting the data into orthogonal components [37, 38]. The resulting loading matrix is interpreted as a structural representation of each dimension’s contribution to the latent variance space [38]. After extracting the loading matrix, these values are normalized and transformed into proportional distributions to enable their application within the entropy weighting scheme. Entropy is subsequently used to quantify the informational contribution of each dimension based on its dispersion characteristics [19, 21]. In this framework, PCA defines the structural relationships among variables, while entropy refines their relative importance, ensuring that the final weights reflect both variance structure and information diversity [19]. This hybridization is not merely a procedural extension, but a structured integration that improves the balance and interpretability of usability weighting compared to the independent use of PCA or entropy. Based on this framework,

the computational process begins with PCA to extract the variance structure of the usability data.

PCA was applied to aggregated Z-scores of effectiveness, efficiency, and satisfaction to analyze the variance structure and minimize redundancy among usability dimensions. The PCA procedure commenced with the calculation of the covariance matrix [55], as indicated in Eq. (5), followed by the decomposition of eigenvalues and eigenvectors using the characteristic polynomial formulation presented in Eq. (6). The standardized data were subsequently projected into the principal component space, as specified in Eq. (7).

Principal components were retained based on cumulative explained variance, as determined by Eq. (8) [37]. The first two principal components (PC1 and PC2) were selected to construct the decision matrix, as they captured the dominant variance structure while preserving the interpretability of the original usability dimensions. The loading values obtained from PC1 and PC2 were then used as input for entropy-based weighting [21]. The loadings, which indicate the structured contribution of each usability dimension, provide the foundation for the ensuing entropy weighting process [38].

$$Cov(X_i, Y_j) = \frac{\sum_{k=1}^n (X_{ik} - \bar{X}_i)(Y_{jk} - \bar{Y}_j)}{n-1} \quad (5)$$

$$(C - \lambda I)v = 0 \quad (6)$$

$$Y = X \times V \quad (7)$$

$$R^2 = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (8)$$

The entropy weighting method was employed to assess information dispersion and to quantify the relative significance of each usability feature based on the normalized loading-derived decision matrix [19]. The process included the normalization of the decision matrix and the calculation of entropy values, as specified in Eq. (9). The entropy for each criterion was calculated using Eq. (10), after which criterion weights were determined using Eq. (11) [19].

$$e_{ij} = -\frac{1}{\ln(n)} \sum_{k=1}^n \frac{n_{ik}}{n} \ln\left(\frac{n_{ik}}{n}\right) \quad (9)$$

$$H_j = \frac{1}{m} \sum_{i=1}^m e_{ij} \quad (10)$$

$$W_j = \frac{1 - H_j}{m - \sum_{k=1}^m H_k} \quad (11)$$

The refined SUM was calculated by aggregating weighted Z-scores of the usability metrics. The SUM score for each task was computed using Eq. (12), facilitating the identification of usability priorities at the task level. A global SUM value indicating overall system usability was derived by aggregating task-level SUM scores, as outlined in Eq. (13). The computational workflow of the proposed hybrid weighting method is depicted in Figure 3 for enhanced clarity.

$$SUM_k = (Z_{sat(k)} \times W_{sat}) + (Z_{Eff(k)} \times W_{Eff}) + (Z_{Effcy(k)} \times W_{Effcy}) \quad (12)$$

$$SUM = \frac{1}{n} \sum_{k=1}^n SUM_{Z,k} \quad (13)$$

Figure 3 illustrates the computational workflow of the proposed PCA–entropy-enhanced SUM, highlighting the stepwise conversion of usability data into hybrid weights. The approach begins with raw usability measurements, including completion rate, error rate, duration on task, and satisfaction scores, which are adjusted using Z-score normalization to enable comparability across dimensions. PCA is employed to

elucidate the underlying variance structure and to discern the primary components that encapsulate the predominant correlations among usability factors. The derived loading matrix is regarded as a structural representation of variable contributions and is further normalized to form a decision matrix.

The entropy weighting approach is employed to assess the informative contribution of each usability parameter according to its dispersion characteristics. This sequential integration facilitates the inclusion of variance-based relationships and information diversity in the weighting process, yielding a more balanced and interpretable usability evaluation. The calculated hybrid weights are employed to consolidate standardized usability measures into the final SUM, resulting in a coherent and interpretable evaluation of system usability.

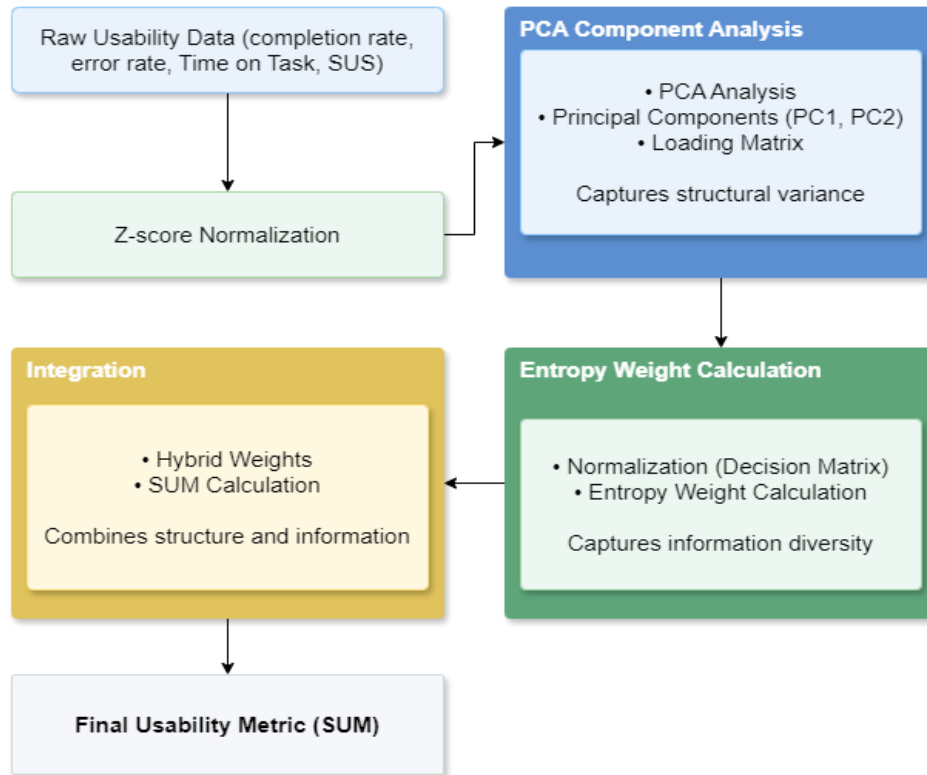


Figure 3. Computational workflow of the proposed Principal Component Analysis–Entropy-enhanced SUM, illustrating the integration of Principal Component Analysis (PCA)-based structural analysis and entropy-based weighting

4. RESULTS AND DISCUSSION

4.1 Effectiveness

The effectiveness was assessed using Eq. (1), which combines *CR* and *error_{rate}*, thereby facilitating the evaluation of task performance based on both successful task completion and interaction precision. This integrated formulation illustrates that effectiveness involves both achieving objectives and maintaining quality interaction, in line with established usability evaluation frameworks [7, 54]. Incorporating error behavior, the effectiveness metric offers a more holistic view of user success than metrics based only on completion [17, 33].

This study analyzes the correlation between task completion and interaction errors in influencing effectiveness, as outlined in Table 3. The use of both measures facilitates a more

equitable evaluation of usability performance, as users may accomplish tasks yet still encounter interaction challenges that diminish the overall quality of their experience.

Table 3 shows that core tasks like logging in, accessing the interface, and navigating were completed successfully by all users, indicating they achieved their intended goals. However, the effectiveness scores varied because of differences in error rates, suggesting that high completion rates don't necessarily mean the engagement was of high quality [7, 35].

Previous research suggests that completion alone can lead to inflated usability assessments if interaction problems are ignored [7, 17, 35]. Conversely, other studies indicate that high completion rates can be a good indicator of usability, especially in controlled environments or for simple tasks [12, 33]. The specific tasks used in this study help explain the differing views on this issue.

In contrast to simplified experimental environments, e-learning tasks generally encompass multi-step interactions,

navigation choices, and the interpretation of system feedback [4, 24, 29]. Consequently, users can complete tasks successfully even when interaction errors occur, which reduces the quality of the interaction and ultimately decreases its effectiveness.

Although Table 3 provides a numerical comparison of effectiveness indicators, a clearer understanding of the relationship between completion rate, error rate, and overall effectiveness can be gained through visual analysis. Figure 2 shows the performance effectiveness across different tasks, highlighting how changes in the error rate affect overall effectiveness, even when completion rates are consistently high.

In addition, Figure 4 demonstrates a consistent inverse relationship between error rate and efficacy across all the tasks that were evaluated. Tasks with elevated error rates showed diminished effectiveness, irrespective of their completion rates. This trend underscores the significance of interaction quality in usability assessment and corroborates prior results

that minimizing user errors is crucial for enhancing overall usability performance [24, 33]. In contrast to controlled usability studies where task completion is closely linked to efficacy, the findings of this study indicate that similar correlations may not apply in more intricate interaction situations, such as e-learning systems [24, 29].

Table 3. Completion rate, error rate, and effectiveness scores for each evaluated task

Task	CR	error _{rate}	Effectiveness
LogIn	100%	0%	100%
Interaction	83.33%	26.67%	61.11%
Communication	66.67%	25%	50%
Interface	100%	20%	80%
Navigation	100%	11.11%	88.89%
Feedback	83.33%	30.56%	57.87%
Assessment	83.33%	26.67%	61.11%
Collaboration	100%	33.33%	66.67%
Resource	33.33%	6.25%	31.25%

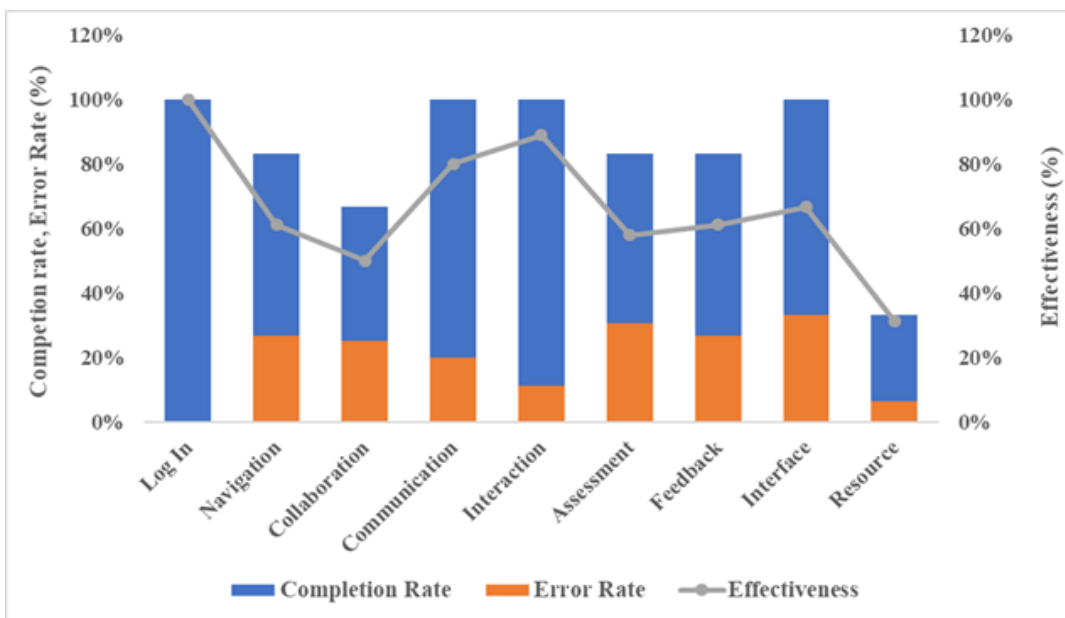


Figure 4. Effectiveness performance of each task

The login task earned the greatest effectiveness score, indicating a coherent workflow and intuitive interaction design. This discovery corroborates previous studies suggesting that well-organized and familiar activities enhance usability performance [15, 24]. Nevertheless, other research indicates that authentication procedures may create usability obstacles due to supplementary security measures, such as CAPTCHA, which might heighten interaction complexity [34, 55]. The disparity noted in this study may suggest that the assessed platforms effectively reconcile security demands with usability, hence reducing superfluous interaction friction. The interface task demonstrated a disparity between task completion and effectiveness, with participants successfully finishing the task despite experiencing interaction difficulties. This finding is consistent with existing research indicating that ambiguous interface designs and insufficient feedback can increase cognitive load, thereby leading to interaction errors [48]. Conversely, in simpler interface designs, such inconsistencies are typically less pronounced, suggesting that interface complexity substantially affects efficacy outcomes. Substantial usability limitations were observed in the resource

task, which was marked by both low completion rates and reduced effectiveness [24, 29]. This result aligns with other studies emphasizing the importance of search visibility and navigational clarity in promoting effective user interaction [42, 49].

Nevertheless, certain research indicates that search-related tasks can attain satisfactory usability provided alternate navigation routes are accessible [33, 42]. The lack of accessible search options and alternate navigation methods may account for the markedly reduced effectiveness found. A comparable tendency was observed in the collaborative task. Despite users successfully completing the job, the elevated error rate resulted in diminished effectiveness. This discovery corroborates previous research indicating that intricate interaction processes elevate the probability of user errors [50]. Conversely to systems designed for task completion, where errors hinder the accomplishment of specific tasks, users engaged in collaborative environments might tolerate interaction inefficiencies due to the social and exploratory nature of the undertaking [26, 28]. This explains the observed phenomenon of sustained completion rates despite a decline in

interaction quality, which ultimately diminishes overall effectiveness. The findings reveal a consistent divergence between task completion and interaction quality. High completion rates can coexist with substantial interaction errors, leading to moderate or low efficacy, especially when users encounter difficulties in comprehending the content or effectively navigating the system. This observation suggests that the effectiveness of e-learning systems is contingent not only on task completion but also on error reduction, interface transparency, and the quality of feedback provided. Consequently, integrating completion rate and error rate into a unified effectiveness metric provides a more meaningful and reliable evaluation of usability performance, thereby establishing a solid foundation for further investigation into efficiency, satisfaction, and the consolidated SUM.

4.2 Efficiency

Efficiency was evaluated through the application of Eq. (2), which incorporated both completion rate and time on task (ToT), thereby accounting for both task success and temporal performance. This formulation is predicated on the notion that efficiency signifies the judicious utilization of time and cognitive resources during task execution, a fundamental aspect of usability assessment [27, 33]. Consequently, by considering both the degree of task completion and the duration required, the efficiency metric offers a more comprehensive perspective on operational effectiveness compared to methodologies that exclusively focus on temporal aspects [33].

Table 4. Completion rate, time on task, and efficiency scores for each evaluated task

Task	CR	ToT (Minutes)	Effcy
Log In	100%	0.41	46.62%
Navigation	83.33%	0.21	76.16%
Collaboration	66.67%	0.88	14.56%
Communication	100%	0.56	34.41%
Interaction	100%	0.23	85.19%
Assessment	83.33%	0.67	23.76%
Feedback	83.33%	0.16	100%
Interface	100%	0.47	40.83%
Resource	33.33%	0.22	29.49%

Table 4 presents the completion rates, average time spent on each task, and efficiency metrics for the evaluated tasks. The findings indicate that numerous tasks, such as login procedures, interface configuration, and communication activities, achieved a 100% completion rate yet exhibited notably low efficiency scores. This pattern indicates that successful task completion does not inherently reflect efficient performance when execution time is extended [24, 29], may also indicate the presence of completion-rate bias in composite usability metrics [17].

This finding aligns with previous research suggesting that efficiency should be assessed as a composite of task success and time-based performance, rather than as separate criteria [27]. Nonetheless, some studies indicate that completion rate may function as a surrogate for usability in structured or low-complexity settings, where execution time is generally more consistent [12, 33].

The intricate nature of interaction dynamics within e-learning systems accounts for the divergence observed between these viewpoints. Unlike the structured settings of

controlled environments, real-world applications often present navigational challenges, the interpretation of system feedback, and multi-step procedures, thereby extending execution time [24, 29] even when tasks are successfully accomplished. Consequently, this suggests that efficiency, in these contexts, is substantially influenced by interaction design, rather than being solely determined by the outcomes of task completion.

Despite achieving complete success, the login task exhibited relatively low efficiency. This result is consistent with prior research that has shown that authentication processes can extend the duration of a task by introducing additional interaction steps and system feedback requirements [51]. Nevertheless, login duties are generally more efficient in systems with simplified authentication mechanisms, indicating that security-related interaction complexity significantly contributes to the reduction of time-based performance [33, 34].

In contrast, tasks featuring more straightforward interaction paths and clearer procedures, like those involving feedback and interaction, achieved higher efficiency scores. Prior research has shown that streamlined navigation, intuitive interface design, and minimal interaction steps substantially reduce execution time and improve efficiency. This observation is consistent with the findings [16, 24]. Nevertheless, duties that involve procedural complexity or ambiguous system feedback are less likely to generate such efficiency gains, particularly in tasks that require multiple steps or where users receive unclear instructions on how to proceed [24, 29].

Despite moderate completion rates, the collaboration and assessment assignments exhibited significantly low efficiency scores. This finding further corroborates prior research that complex, multi-stage interaction processes increase task duration and decrease efficacy [23, 24]. Nevertheless, users in this study were able to complete these tasks despite the prolonged execution time, in contrast to highly structured systems where inefficiencies may preclude task completion [12, 33]. This implies that users may adjust to inefficient interaction processes, but at the expense of reduced operational performance and increased effort [26, 28].

The interface task exhibited moderate efficiency, which can be attributed to the challenges associated with navigation accessibility, insufficient system feedback, and limited responsiveness, particularly on mobile devices. This finding aligns with studies that highlight the importance of responsive interface design and cross-platform optimization in reducing interaction time [23, 34]. Nevertheless, the performance limitations are typically minimized in systems with well-optimized interfaces, suggesting that the quality of the interface design is a critical determinant of efficiency outcomes [24, 29].

A consistent misalignment between task completion and time-based performance is revealed by the efficiency analysis. Extended execution times often correlate with high completion rates, thereby diminishing overall efficiency. This finding suggests that the effectiveness of e-learning systems is contingent not only on task success but also on the clarity of feedback, the ease of interactions, and the optimization of workflows. Consequently, incorporating completion rate and time on task into a single efficiency metric provides a more comprehensive evaluation of usability performance, thereby establishing a consistent relationship among the aggregated SUM, efficiency, and effectiveness.

4.3 Satisfaction

The satisfaction dimension was evaluated using the SUS, as defined in Eq. (3), which reflects users' perceived usability during interaction [8, 30]. To maintain methodological consistency within the SUM framework, SUS scores were normalized by Z-score normalization in conjunction with effectiveness and efficiency measurements (Eq. (4)) [18], thereby enabling direct comparisons across a range of usability attributes.

The data in Table 5 indicate that user satisfaction does not invariably align with objective usability measures. This finding contradicts earlier studies that posited a universal relationship among enhanced task performance, execution efficiency, and elevated user satisfaction [7, 24, 29]. The observed variations across different tasks suggest that satisfaction is not exclusively dictated by performance efficiency or accuracy; rather, it is influenced by users' subjective assessments of interaction quality and perceived utility [7, 35].

Table 5. Z-scores of usability metrics for each evaluated task

Task	Z-Sat	Z-Eff	Z-Effcy
Log In	-1.121	1.719	-0.124
Navigation	-0.556	-0.266	0.926
Collaboration	1.890	-0.833	-1.264
Communication	0.746	0.698	-0.558
Interaction	-1.409	1.152	1.246
Assessment	0.660	-0.431	-0.937
Feedback	-0.687	-0.266	1.773
Interface	0.626	0.018	-0.330
Resource	-0.149	-1.791	-0.733

Table 5 presents the standardized Z-scores for satisfaction, effectiveness, and efficiency, calculated for each task. This allows for a comparison of the relationships between personal opinions and measurable performance results. The results reveal a consistent divergence between subjective satisfaction and objective usability assessments.

Several tasks exhibit contrasting performance characteristics, where heightened enjoyment does not correspond with strong effectiveness or efficiency outcomes. The collaboration task, for instance, achieved the highest satisfaction score ($Z = 1.890$), yet it simultaneously presented below-average effectiveness ($Z = -0.833$) and efficiency ($Z = -1.264$). Conversely, tasks such as login and interaction, while demonstrating significant efficacy and efficiency, are associated with low satisfaction ratings.

These differing patterns suggest that how users feel about something doesn't always directly relate to how well they perform a task [7, 35]. This implies that what users think is important is influenced by factors beyond just how efficiently and effectively they complete the task. In contrast, controlled usability tests usually show a stronger connection between performance measures and user satisfaction. This is likely because the tasks are structured and focused on achieving specific results [12, 33]. The difference observed in this study can be attributed to the complexities of real-world e-learning environments, where users engage in multi-step processes, interpret system inputs, and adapt to system constraints.

Figure 5 further depicts the distribution of satisfaction scores pertaining to the collaborative task, highlighting that the majority of users provided positive feedback, notwithstanding the presence of identifiable inefficiencies. The collaboration

task achieved a quality level of 97.06%, characterized by a minimal fault area, thereby indicating a substantial degree of perceived usability.

This divergence suggests that users may perceive collaborative components as both valuable and engaging, even when interaction procedures exhibit inefficiency [26, 28]. Previous investigations into e-learning interaction have indicated that collaborative environments often emphasize social interaction, communication, and the perceived value of learning, which can contribute to user satisfaction regardless of performance efficiency [26, 28].

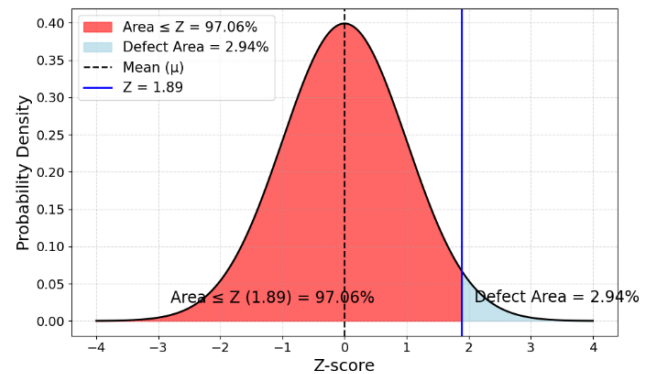


Figure 5. Distribution of the satisfaction Z-score for the collaboration task and its corresponding defect area

Therefore, this finding highlights a key difference between how usable something seems and how well it actually works. Users are willing to accept some inefficiencies if the technology helps them achieve important interaction goals. This inconsistency can be explained by the social aspects of collaborative work. In these situations, users often prioritize interaction and communication over efficiency and straightforward procedures [26].

Subsequent examination of the login task uncovers a contrasting trend, wherein elevated efficacy ($Z = 1.719$) fails to correspond with high pleasure ($Z = -1.121$). This suggests that merely completing tasks successfully is inadequate for fostering a favorable user perception. The findings suggest that additional interaction difficulties, such as complex authentication processes and unclear feedback systems, negatively affect user satisfaction, even though they help with completing tasks. This observation is consistent with prior research indicating that security-related interactions can increase cognitive load and reduce perceived usability [51, 55].

A similar misalignment is present in the interaction task, characterized by high effectiveness and efficiency alongside minimal enjoyment. This pattern suggests that usability performance alone doesn't fully capture the quality of an experience [7, 35]. Things like a consistent interface, how quickly the system responds, and how well it works on mobile devices can greatly affect how users perceive the system. This is true even when tasks are completed successfully. This supports earlier research that shows satisfaction comes from the overall interaction, not just from looking at individual performance metrics [7, 35].

Figure 6 further illustrates the occurrence of outlier patterns across various activities, particularly in login and interaction contexts. In these situations, strong objective performance doesn't match positive user evaluations. These patterns reinforce the argument that satisfaction is influenced by

experiential and contextual factors beyond measurable usability performance.

The resource task shows a notable difference, with low effectiveness ($Z = -1.791$), low efficiency ($Z = -0.733$), and a similar level of satisfaction ($Z = -0.149$). This indicates that users might offset inadequate system performance by exerting additional effort, such as engaging in recurrent navigation or extending task execution time. Previous research attributes this behavior to deficient search capabilities, ambiguous navigation frameworks, and poor feedback systems, compelling users to adjust rather than discontinue the activity [29, 42, 49].

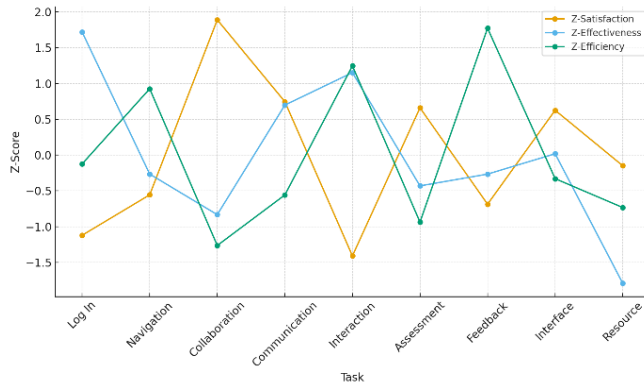


Figure 6. Outlier patterns in the Z-score correlations among satisfaction, effectiveness, and efficiency over tasks

The findings, in essence, indicate a consistent divergence between satisfaction levels and objective usability measures. This suggests that perceived usability is shaped by contextual, cognitive, and experiential elements that transcend mere task execution. Consequently, satisfaction should not be viewed as a straightforward indicator of efficiency or effectiveness; instead, it reflects the perceived value of the interaction and the overall user experience.

The SUM framework provides a more comprehensive and balanced evaluation of usability by integrating subjective satisfaction with objective performance metrics [13-15]. In e-learning, this integration is especially important. The quality of interactions depends not only on how well they work and how accurate they are, but also on how engaged the learner is, how useful they find the content, and the specific learning situation.

4.4 Principal Component Analysis—Entropy weighting

The three main aspects of usability—effectiveness, efficiency, and satisfaction—show different but related patterns of behavior when tasks are used to test usability. Prior studies have shown that composite usability measures can exhibit bias or redundancy when weighting algorithms are uniform or based in the SUM [36]. This study employs a hybrid PCA—Entropy weighting method to address this limitation, integrating variance-based structural analysis with information-theoretic dispersion to provide a more equitable and data-driven weighting framework.

The preliminary phase involves the analysis of the correlations between standardized usability measures. To clarify the linear relationship structure between satisfaction, efficacy, and efficiency, a covariance matrix was generated using Z-score normalized values (see Table 6).

The covariance matrix presented in Table 6 indicates

specific relational patterns among usability dimensions. The significant negative correlations between Z-satisfaction and Z-efficiency ($r = -0.763$), and between Z-satisfaction and Z-effectiveness ($r = -0.476$), indicate a systematic misalignment between users' subjective perceptions and objective task performance outcomes. This finding shows that satisfaction judgments are affected by more than just efficiency or task success. They are also affected by experiential factors, such as how familiar the user is with the interface and how easy it is to interact with it. This supports previous usability research [7, 35].

Table 6. Covariance among usability metrics

Matrix	Z-Sat	Z-Eff	Z-Effcy
Z-Sat	1	-0.476	-0.763
Z-Eff	-0.476	1	0.320
Z-Effcy	-0.763	0.320	1

The moderate positive correlation between Z-effectiveness and Z-efficiency ($r = 0.320$) indicates that greater task success is typically linked to enhanced efficiency, thereby supporting the conceptual alignment between these two objective performance metrics [24, 33]. The identified covariance patterns support the use of PCA for extracting underlying usability structures, thereby minimizing redundant metric contributions.

Eigenvalue decomposition was conducted on the covariance structure to identify the principal components that encapsulate the most significant variance in the usability dataset. The eigenvalue vector and the corresponding explained variance ratios are displayed in Table 7.

Table 7. Eigenvalue vector

λ_1	λ_2	λ_3
2.063	0.721	0.216
0.688	0.240	0.072

The eigenvalue distribution proves that the first two principal components (PC1 and PC2) account for approximately 92.8% of the total variance, significantly surpassing the widely recognized 50% threshold for usability analysis [17]. This suggests that the most significant information contained in the original usability metrics can be accurately represented by only two components, thus confirming the dimensional reduction process.

PC1 accounts for the largest proportion of variance, representing a dominant latent dimension in usability perception, whereas PC2 shares complementary information not captured by PC1 alone. This finding supports the retention of both components in future entropy-based weighting to maintain analytical completeness. A scree plot of eigenvalues was constructed to visually assess the variance distribution across principal components and to support the decision about component retention, as illustrated in Figure 7.

Figure 7 depicts a considerable decrease in eigenvalues subsequent to PC2, followed by an almost horizontal trajectory from PC3 onwards. This pattern indicates that PC3 and following components provide negligible new information and can be omitted without substantial information loss. The scree plot thus supports the retention of PC1 and PC2 for entropy weighting, ensuring analytical simplicity while preserving representational sufficiency [21, 27].

An examination of the eigenvector loadings for each usability metric was conducted to interpret the semantic meaning of the retained components. Figure 8 illustrates the roles of satisfaction, effectiveness, and efficiency in relation to PC1 and PC2.

The eigenvector loading pattern indicates that satisfaction and efficiency have a stronger association with PC1, while effectiveness shows a significant loading on PC2. This distinction suggests that PC1 mainly represents users' perceptions of interaction fluency and system responsiveness, which corresponds with the satisfaction–efficiency construct. Conversely, PC2 reflects the accuracy of tasks and the reliability of their completion, aligning with the dimension of effectiveness.

While PC2 may not be the primary carrier of variance, its significant contribution underscores that effectiveness is a separate and essential dimension of usability that cannot be fully captured by perceptual metrics alone. This finding offers a data-driven rationale for the persistent prominence of effectiveness in composite usability evaluations [13, 36].

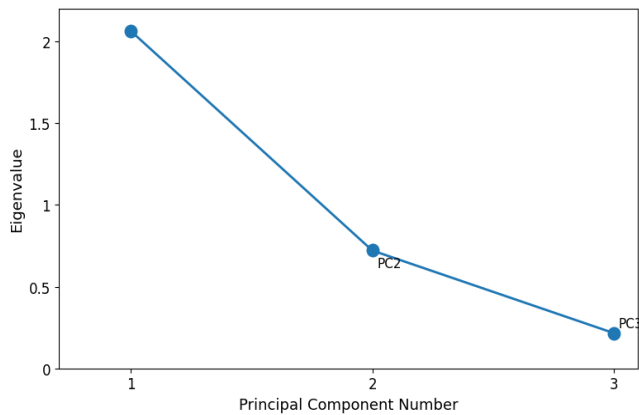


Figure 7. Scree plot of eigenvalues

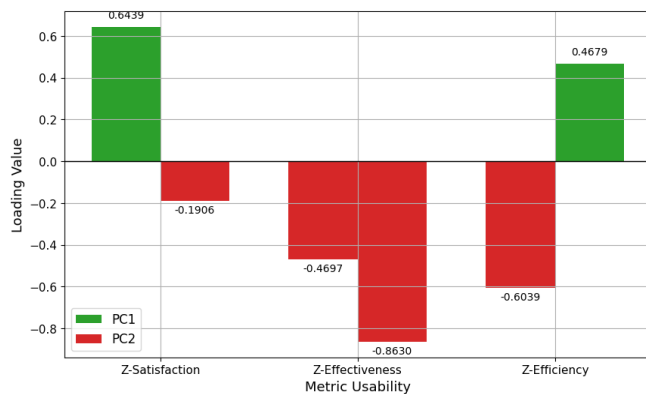


Figure 8. Eigenvectors loading of each metrics

Subsequent to analyzing the PCA, the component loadings that were retained were normalized to create a decision matrix appropriate for entropy calculation. Table 8 displays the normalized matrix.

Normalization ensures that usability metrics contribute proportionally to the entropy calculation by removing scale effects and minimizing dominance caused by differences in magnitude [21]. This step is critical for converting PCA outputs into probabilistic distributions, allowing entropy to effectively measure information dispersion across components.

Table 8. Normalized decision matrix

Usability Metric	PC1	PC2
Satisfaction	0.3748	0.1252
Effectiveness	0.2734	0.5671
Efficiency	0.3518	0.3077

Entropy values have been determined using the normalized decision matrix to measure the level of uncertainty related to each principal component. Table 9 presents the resulting entropy values.

Table 9. Entropy values obtained for each principal component

Principal Component	Entropy e_{ij}
PC1	0.664
PC2	0.620

The entropy results reveal that PC2 has a lower entropy score of 0.620 when compared to PC1's 0.664, indicating enhanced informational diversity and discriminatory capability. Lower entropy indicates that PC2 more effectively distinguishes between usability metrics, underscoring its significance despite contributing less variance in PCA. This finding underscores the complementary relationship between PCA and entropy: PCA focuses on variance dominance, whereas entropy reflects informational heterogeneity [19, 21].

The final stage integrates PCA-derived structure with entropy-derived information diversity to derive the hybrid weights for each usability metric. Tables 10 and 11 present the resulting component weights and aggregated usability weights.

Table 10. Entropy weights of PC1 and PC2

Principal Component	Entropy H_j	W_j
PC1	0.664	0.469
PC2	0.620	0.531
Total		1

Table 11. Hybrid Principal Component Analysis (PCA)–Entropy weights for each usability metric

Usability Metric	W_i
Satisfaction	0.2423
Effectiveness	0.4294
Efficiency	0.2423

The final hybrid weighting outcomes (Table 11) reveal that efficacy received the highest weight (0.4294), followed by satisfaction and efficiency, each with a weight of 0.2423. This distribution demonstrates a theoretically coherent prioritizing of usability variables, emphasizing task accomplishment while maintaining fair representation of experience factors. PCA-only weighting predominantly reflects variance structure, while entropy-only weighting may excessively highlight dispersion, potentially resulting in disproportionate contributions from specific dimensions. The hybrid PCA–Entropy method synthesizes different viewpoints, yielding a more equitable and structurally informed weighting system that alleviates the shortcomings of each separate technique.

Both theoretical and empirical considerations substantiate the predominance of effectiveness, as task completion

accuracy is the fundamental objective of usability in e-learning environments, and efficiency and satisfaction serve as additive experiential factors [15, 36]. The hybrid technique offers a more stable and interpretable weighting framework, hence improving the construct validity of the resultant SUM scores. Table 12 illustrates a comparative analysis of the robustness and interpretability of this weighting scheme, alongside PCA-only and entropy-only techniques.

Table 12. Comparison of Principal Component Analysis (PCA), entropy, and hybrid weights for usability dimensions

Usability Metric	PCA Weight	Entropy Weight	Hybrid (PCA-Entropy) Weight
Satisfaction	0.381	0.595	0.2423
Effectiveness	0.301	0.092	0.4294
Efficiency	0.317	0.313	0.2423

The entropy-only method, as illustrated in Table 12, allocates an excessively high weight to pleasure while significantly undervaluing efficacy, revealing its susceptibility to data variability [19, 21]. PCA-only weighting yields a more equitable distribution according to variance structure; nonetheless, it fails to comprehensively reflect the variations in information across several usability criteria [37].

The proposed hybrid PCA–Entropy method integrates both perspectives to establish a more equitable and rational approach to weighting [21]. The effectiveness of the hybrid paradigm is substantiated by theoretical and empirical evidence [16]. The main goal of usability in e-learning environments is the accuracy of task completion, whereas efficiency and enjoyment serve as supplementary experiential variables after successful task execution [16]. The hybrid weighting technique adeptly encapsulates this hierarchical relationship, demonstrating both performance-based significance and a data-driven framework. This comparison demonstrates that the suggested hybrid weighting methodology significantly mitigates the possibility of single-method bias, hence enhancing the stability and reliability of usability evaluation results [19, 21].

4.5 Final Single Usability Metric score

Following the establishment of hybrid PCA–Entropy weights (Section 4.4), the final SUM was obtained by integrating effectiveness, efficiency, and satisfaction into a single usability indicator. The SUM score aggregates standardized task-level Z-scores (Table 5) with the corresponding usability metric weights (Table 11), offering a comprehensive and comparable evaluation of overall usability performance across e-learning tasks [13, 18]. The composite method complies with recognized SUM formulas and ISO-based assessment of usability principles [13-15]. This integration allows for a more thorough interpretation of usability by incorporating the observed discrepancies between objective performance and subjective perception, as identified in Sections 4.1–4.3.

The SUM values presented in Table 13 demonstrate considerable variability in the overall usability performance across the tasks, reflecting the combined influence of satisfaction, efficiency, and efficacy. Tasks like Interaction (67.55%) and Log In (66.88%) attained the highest SUM scores, suggesting a relatively equitable distribution among the usability dimensions. Consequently, these results imply that both objective performance and favorable user perception

are concurrently shaped by coherent interaction pathways, clearly defined task structures, and effective system feedback mechanisms [15, 24, 29].

Table 13. Single Usability Metric (SUM) for each task

Task	SUM Z-Score	SUM (%)
Log In	0.4367	66.88%
Navigation	-0.0247	49.02%
Collaboration	-0.2062	41.83%
Communication	0.3455	63.52%
Interaction	0.4553	67.55%
Assessment	-0.2524	40.04%
Feedback	0.1489	55.92%
Interface	0.0793	53.16%
Resource	-0.9825	16.29%

In contrast, the tasks of Resource (16.29%), Collaboration (41.83%), and Assessment (40.04%) reflect substantially lower SUM scores. The observed low values represent compounded usability deficiencies due to misalignment among usability dimensions, particularly in situations in which high completion rates are accompanied by elevated error rates, extended task duration, or reduced user comfort [14, 17]. Sections 4.1–4.3 demonstrate that these tasks reveal limited discoverability, procedural complexity, and insufficient system feedback, which collectively decrease overall usability performance.

The SUM scores were benchmarked in accordance with the established usability evaluation conventions for standardized composite metrics to facilitate meaningful interpretation [13-15]. An SUM value of approximately 50% is interpreted as marginal usability in this study, suggesting that the system is functional but has inefficiencies or interaction issues that could potentially harm the user experience over time [14, 15]. This threshold-based interpretation provides a more pragmatic framework for determining usability priorities than a mere numerical metric comparison.

Feedback (55.92%) and interface (53.16%) represent tasks that, while conditionally acceptable, require tailored improvements. The findings of this investigation indicate that the observed moderate usability performance might mask interaction inefficiencies that are exacerbated with extended system utilization [24, 29]. Furthermore, the Resource task's SUM score, which is notably diminished, highlights significant usability concerns, specifically regarding navigation clarity and search feedback systems, which substantially degrade the overall user experience [29, 42, 49].

A qualitative evaluation of robustness was undertaken through an examination of the stability of usability metric superiority across various weighting systems. The comparative analysis of PCA-only, entropy-only, and hybrid PCA-Entropy methodologies demonstrates that efficacy consistently emerges as the most significant usability dimension [19, 21]. In contrast to single-method approaches, the hybrid model yields a more balanced distribution by diminishing the influence of variance magnitude (PCA-only) and sensitivity to dispersion (entropy-only) [19, 42]. Consequently, this suggests that the hybrid weighting technique offers a more robust and dependable representation of usability structure.

The observed stability suggests that the derived SUM rankings are not substantially influenced by slight alterations in weighting assumptions, thereby validating the resilience of the proposed approach. The hybrid PCA–Entropy framework

provides a reliable and easily understandable basis for usability evaluation, negating the necessity for additional statistical resampling or complex validation techniques [16, 21].

These results imply that the integration of objective performance metrics with subjective satisfaction through a hybrid weighting scheme enables a more comprehensive and diagnostically valuable usability assessment [13, 15, 16]. This method enhances the interpretability of SUM outcomes and enables more accurate diagnosis of task-level usability problems in e-learning systems [14, 15].

5. CONCLUSIONS

This study advances the theoretical understanding of usability metric integration by demonstrating that usability dimensions contribute unevenly when considering both variance structure and information dispersion. The consistent importance of effectiveness suggests that task completion accuracy constitutes the core of usability, while efficiency and satisfaction are secondary attributes.

The proposed hybrid PCA–Entropy method proves most effective in situations characterized by partial correlation and heterogeneous variability in usability metrics. In such instances, PCA identifies underlying structural correlations, while entropy corrects for dispersion bias. The utility of the hybrid approach diminishes in homogeneous datasets or when significant multicollinearity is evident.

The task-level SUM offers pragmatic diagnostic information for system enhancement. The resource task (16.29%) reveals significant shortcomings in search capabilities and navigation clarity, indicating a necessity for focused redesign. Conversely, high-performing assignments demonstrate the significance of optimized workflows and constructive feedback in attaining balanced usability performance.

Z-score normalization presupposes comparability and near-normality among usability indicators, potentially introducing bias in restricted or skewed data distributions. The examination is confined to three usability dimensions—effectiveness, efficiency, and satisfaction—omitting other notions like learnability and accessibility. Despite the framework's structural extensibility, its efficacy in higher-dimensional usability models remains unverified and may impact weight stability. Moreover, the methodology relies on linear correlations, which may constrain its capacity to identify non-linear interactions among usability criteria, potentially leading to incomplete assessments of user experience and usability outcomes.

Future research should explore alternative normalization methods that are resilient to non-normal data, expand the framework to include more usability factors, and evaluate the model across various e-learning platforms. It is advisable to incorporate non-linear modeling techniques to more effectively represent intricate usability dynamics.

ACKNOWLEDGMENT

The authors gratefully acknowledge Universitas Muhammadiyah Purwokerto for its financial support of this research. Appreciation is also expressed to Unit Pengelola E-learning of Universitas Muhammadiyah Brebes, Universitas

Surakarta, Universitas 'Aisyiyah Surakarta, Lembaga Pengembangan Akademik of Universitas Muhammadiyah Purwokerto, Lembaga Pengembangan dan Penjaminan Mutu Pendidikan of Universitas Diponegoro, and Biro Sistem Informasi of Institut Teknologi dan Bisnis Muhammadiyah Purbalingga for their support in providing observational data and respondents for e-learning satisfaction survey in this study.

REFERENCES

- [1] Altalbe, A. (2021). Antecedents of actual usage of e-learning system in high education during COVID-19 pandemic: Moderation effect of instructor support. *IEEE Access*, 9: 93119-93136. <https://doi.org/10.1109/ACCESS.2021.3087344>
- [2] Banowosari, L.Y., Utama, K.A.B. (2018). Evaluation of user engagement in e-learning standardization and conformity assessment using subjective and objective measurement. In 2018 Third International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, pp. 1-6. <https://doi.org/10.1109/IAC.2018.8780479>
- [3] Vlachogianni, P., Tselios, N. (2022). Perceived usability evaluation of educational technology using the system usability scale (SUS): A systematic review. *Journal of Research on Technology in Education*, 54(3): 392-409. <https://doi.org/10.1080/15391523.2020.1867938>
- [4] Manik, L.P. (2024). Exploring usage-based and usability metrics for user experience for sustainable e-learning systems. *E3S Web of Conferences*, 501: 02003. <https://doi.org/10.1051/e3sconf/202450102003>
- [5] Talib, E.A.H., Santosa, P.I., Wibirama, S. (2023). Evaluation of learning management systems based on usability and user experience: A systematic literature review. In 2023 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia, pp. 691-696. <https://doi.org/10.1109/ISITIA59021.2023.10221015>
- [6] Novák, J.Š., Masner, J., Benda, P., Šimek, P., Merunka, V. (2024). Eye tracking, usability, and user experience: A systematic review. *International Journal of Human-Computer Interaction*, 40(17): 4484-4500. <https://doi.org/10.1080/10447318.2023.2221600>
- [7] Altin Gumussoy, C., Pekpazar, A., Esengun, M., Bayraktaroglu, A.E., Ince, G. (2022). Usability evaluation of TV interfaces: Subjective evaluation vs. objective evaluation. *International Journal of Human-Computer Interaction*, 38(7): 661-679. <https://doi.org/10.1080/10447318.2021.1960093>
- [8] Drew, M.R., Falcone, B., Baccus, W.L. (2018). What does the system usability scale (SUS) measure? Validation using think aloud verbalization and behavioral metrics. In *International Conference of Design, User Experience, and Usability*, pp. 356-366. https://doi.org/10.1007/978-3-319-91797-9_25
- [9] Alghabban, W.G., Hendley, R. (2022). Perceived level of usability as an evaluation metric in adaptive e-learning: A case study with dyslexic children. *SN Computer Science*, 3(3): 238. <https://doi.org/10.1007/s42979-022-01138-5>
- [10] Simon, P., Jiang, J., Fryer, L.K. (2024). Measurement of higher education students' and teachers' experiences in

- learning management systems: A scoping review. *Assessment & Evaluation in Higher Education*, 49(4): 441-452. <https://doi.org/10.1080/02602938.2023.2266154>
- [11] Walldén, S., Mäkinen, E., Raisamo, R. (2016). A review on objective measurement of usage in technology acceptance studies. *Universal Access in the Information Society*, 15(4): 713-726. <https://doi.org/10.1007/s10209-015-0443-y>
- [12] Kortum, P., Hebl, M., Oswald, F.L. (2014). Applying usability measures to assess textbooks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1): 1346-1350. <https://doi.org/10.1177/1541931214581281>
- [13] Sauro, J., Kindlund, E. (2005). Using a single usability metric (SUM) to compare the usability of competing products. In *Proceedings of the Human Computer Interaction International Conference (HCII)*, pp. 1-9.
- [14] Van Waardhuizen, M., McLean-Oliver, J., Perry, N., Munko, J. (2019). Explorations on single usability metrics. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, Scotland UK, pp. 1-8. <https://doi.org/10.1145/3290607.3299062>
- [15] Albert, W., Tullis, T.S. (2023). Combined and comparative metrics. *Interactive Technologies, Measuring the User Experience (Third Edition)*, pp. 217-241. <https://doi.org/10.1016/B978-0-12-818080-8.00009-1>
- [16] Alabbas, A., Alomar, K. (2025). A weighted composite metric for evaluating user experience in educational chatbots: Balancing usability, engagement, and effectiveness. *Future Internet*, 17(2): 64. <https://doi.org/10.3390/fi17020064>
- [17] Pearson, C.J. (2023). A completion rate conundrum: Reducing bias in the single usability metric. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 67(1): 1407-1411. <https://doi.org/10.1177/21695067231194328>
- [18] Sauro, J., Kindlund, E. (2005). Making sense of usability metrics: Usability and six sigma. In *Proceedings of the 14th Annual Conference of the Usability Professionals Association*, pp. 1-10.
- [19] Wu, R.M., Zhang, Z., Yan, W., Fan, J., et al. (2022). A comparative analysis of the principal component analysis and entropy weight methods to establish the indexing measurement. *PLOS One*, 17(1): e0262261. <https://doi.org/10.1371/journal.pone.0262261>
- [20] Puška, A., Lukić, M., Božanić, D., Nedeljković, M., Hezam, I.M. (2023). Selection of an insurance company in agriculture through hybrid multi-criteria decision-making. *Entropy*, 25(6): 959. <https://doi.org/10.3390/e25060959>
- [21] Pliego-Martínez, O., Martínez-Rebollar, A., Estrada-Esquivel, H., de la Cruz-Nicolás, E. (2024). An integrated attribute-weighting method based on PCA and entropy: Case of study marginalized areas in a city. *Applied Sciences*, 14(5): 2016. <https://doi.org/10.3390/app14052016>
- [22] Harper, S.B., Dorton, S.L. (2021). A pilot study on extending the sus survey: Early results. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1): 447-451. <https://doi.org/10.1177/1071181321651162>
- [23] Tania, K.D., Abdullah, N.S., Ahmad, N., Sahmin, S. (2022). Continued usage of e-learning: A systematic literature review.
- [24] Ferreira, J.M., Acuña, S.T., Dieste, O., Vegas, S., Santos, A., Rodríguez, F., Juristo, N. (2020). Impact of usability mechanisms: An experiment on efficiency, effectiveness and user satisfaction. *Information and Software Technology*, 117: 106195. <https://doi.org/10.1016/j.infsof.2019.106195>
- [25] Kortum, P., Johnson, M. (2013). The relationship between levels of user experience with a product and perceived system usability. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1): 197-201. <https://doi.org/10.1177/1541931213571044>
- [26] Neha, Kim, E. (2023). Designing effective discussion forum in MOOCs: Insights from learner perspectives. *Frontiers in Education*, 8: 1223409. <https://doi.org/10.3389/educ.2023.1223409>
- [27] Wani, A.A., Abeer, F. (2025). Application of machine learning techniques for warfarin dosage prediction: A case study on the MIMIC-III dataset. *PeerJ Computer Science*, 11: e2612. <https://doi.org/10.7717/peerj-cs.2612>
- [28] Al-Adwan, A.S., Albelbisi, N.A., Hujran, O., Al-Rahmi, W.M., Alkhalifah, A. (2021). Developing a holistic success model for sustainable e-learning: A structural equation modeling approach. *Sustainability*, 13(16): 9453. <https://doi.org/10.3390/su13169453>
- [29] Ferreira, J.M., Rodríguez, F.D., Santos, A., Dieste, O., Acuña, S.T., Juristo, N. (2022). Impact of usability mechanisms: A family of experiments on efficiency, effectiveness and user satisfaction. *IEEE Transactions on Software Engineering*, 49(1): 251-267. <https://doi.org/10.1109/TSE.2022.3149586>
- [30] Lewis, J.R. (2018). The system usability scale: Past, present, and future. *International Journal of Human-Computer Interaction*, 34(7): 577-590. <https://doi.org/10.1080/10447318.2018.1455307>
- [31] Orfanou, K., Tselios, N., Katsanos, C. (2015). Perceived usability evaluation of learning management systems: Empirical evaluation of the system usability scale. *The International Review of Research in Open and Distributed Learning*, 16(2): 227-246. <https://doi.org/10.19173/irrodl.v16i2.1955>
- [32] Nasr, V., Zahabi, M. (2024). Development of a single usability metric that accounts for accessibility (SUMA). In *Human Factors in Design, Engineering, and Computing*, 159: 889-896. <https://doi.org/10.54941/ahfe1005655>
- [33] Tullis, T., Albert, B. (2013). Performance metrics. *Measuring the User Experience*, pp. 71-107. <https://doi.org/10.1016/B978-0-12-818080-8.00004-2>
- [34] Reddy, A., Cheng, Y. (2024). User perceptions of CAPTCHAs: University vs. internet users. In *IFIP Annual Conference on Data and Applications Security and Privacy*, pp. 290-297. https://doi.org/10.1007/978-3-031-65172-4_18
- [35] Georgsson, M., Staggers, N. (2016). Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. *Journal of the American Medical Informatics Association*, 23(1): 5-11. <https://doi.org/10.1093/jamia/ocv099>
- [36] Sauro, J., Kindlund, E. (2005). A method to standardize

- usability metrics into a single score. In Proceedings of the SIGCHI conference on Human factors in computing systems, Portland, Oregon, USA, pp. 401-409. <https://doi.org/10.1145/1054972.1055028>
- [37] Konishi, T. (2025). Means and issues for adjusting principal component analysis results. *Algorithms*, 18(3): 129. <https://doi.org/10.3390/a18030129>
- [38] Chipman, H.A., Gu, H. (2005). Interpretable dimension reduction. *Journal of Applied Statistics*, 32(9): 969-987. <https://doi.org/10.1080/02664760500168648>
- [39] Tretow-Fish, T.A.B., Khalid, M.S. (2023). Methods for evaluating learning analytics and learning analytics dashboards in adaptive learning platforms: A systematic review. *Electronic Journal of e-Learning*, 21(5): 430-449. <https://doi.org/10.34190/ejel.21.5.3088>
- [40] Tao, L., Cukurova, M., Song, Y. (2025). Learning analytics in immersive virtual learning environments: A systematic literature review. *Smart Learning Environments*, 12(1): 43. <https://doi.org/10.1186/s40561-025-00381-6>
- [41] Torres-Molina, R., Seyam, M. (2024). A hybrid approach for usability evaluation of learning management systems using machine learning algorithms. In 2024 IEEE Frontiers in Education Conference (FIE), Washington, DC, USA, pp. 1-9. <https://doi.org/10.1109/FIE61694.2024.10893160>
- [42] Wang, J., Antonenko, P., Celepkolu, M., Jimenez, Y., Fieldman, E., Fieldman, A. (2019). Exploring relationships between eye tracking and traditional usability testing data. *International Journal of Human-Computer Interaction*, 35(6): 483-494. <https://doi.org/10.1080/10447318.2018.1464776>
- [43] Simon, N., Carbonera, B.J., Custodio, B. (2016). A comparative study on the usability of educational platforms used by instructors in the university of the Philippines. In *Advances in Human Factors, Business Management, Training and Education: Proceedings of the AHFE 2016 International Conference on Human Factors, Business Management and Society*, July 27-31, 2016, Walt Disney World®, Florida, USA, pp. 187-194. https://doi.org/10.1007/978-3-319-42070-7_18
- [44] Daramola, O., Oladipupo, O., Afolabi, I., Olopade, A. (2017). Heuristic evaluation of an institutional e-learning system: A Nigerian case. *International Journal of Emerging Technologies in Learning*, 12(3): 26-42. <https://doi.org/10.3991/ijet.v12i03.6083>
- [45] Alzghaibi, H. (2023). Usability of health IT for health and medical students: A systematic review. *Informatics in Medicine Unlocked*, 38: 101200. <https://doi.org/https://doi.org/10.1016/j.imu.2023.101200>
- [46] Doubleday, E.G., O'Loughlin, V.D., Doubleday, A.F. (2011). The virtual anatomy laboratory: Usability testing to improve an online learning resource for anatomy education. *Anatomical Sciences Education*, 4(6): 318-326. <https://doi.org/10.1002/ase.252>
- [47] Lai, L.L., Lin, S.Y. (2017). An analysis for difficult tasks in e-learning course design. In *International Conference on HCI in Business, Government, and Organizations*, pp. 171-180. https://doi.org/10.1007/978-3-319-58481-2_14
- [48] Haryanto, T., Sholihah, M.A., Yuniarto, D., Sopandi, A., Kaffah, F.M., Subiyakto, A. (2023). Evaluating the effectiveness and efficiency of a website using cognitive walkthrough method. In 2023 11th International Conference on Cyber and IT Service Management (CITSM), Makassar, Indonesia, pp. 1-6. <https://doi.org/10.1109/CITSM60085.2023.10455690>
- [49] Almajali, D., Al-Okaily, M., Barakat, S., Al-Zegaier, H., Dahalin, Z.M. (2022). Students' perceptions of the sustainability of distance learning systems in the post-COVID-19: A qualitative perspective. *Sustainability*, 14(12): 7353. <https://doi.org/10.3390/su14127353>
- [50] Almukhaylid, M., Suleman, H. (2020). Socially-motivated discussion forum models for learning management systems. In *Conference of the South African Institute of Computer Scientists and Information Technologists 2020*, Cape Town, South Africa, pp. 1-11. <https://doi.org/10.1145/3410886.3410902>
- [51] Kew, S.N., Tasir, Z. (2021). Analysing students' cognitive engagement in e-learning discussion forums through content analysis. *Knowledge Management & E-Learning*, 13(1): 39-57. <https://doi.org/10.34105/j.kmel.2021.13.003>
- [52] Anthony Jr, B., Kamaludin, A., Romli, A., Raffei, A.F.M., Phon, D.N.A.E., Abdullah, A., Ming, G.L. (2022). Blended learning adoption and implementation in higher education: A theoretical and systematic review. *Technology, Knowledge and Learning*, 27(2): 531-578. <https://doi.org/10.1007/s10758-020-09477-z>
- [53] Albert, W., Tullis, T. (2010). *Measuring the User Experience*. San Francisco: Morgan Kaufmann, pp. 195-216. <https://doi.org/10.1016/B978-0-12-818080-8.00006-6>
- [54] Furman, S.M., Stanton, B.C., Theofanos, M.F., Libert, J.M., Grantham, J.D. (2017). *Contactless Fingerprint Devices Usability Test*. US Department of Commerce, National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.IR.8171>
- [55] Alqurni, J. (2023). Assessing the usability of E-learning software among university students: A study on student satisfaction and performance. *International Journal of Information Technology and Web Engineering*, 18(1): 1-26. <https://doi.org/10.4018/IJITWE.329198>

NOMENCLATURE

<i>Eff</i>	Effectiveness
<i>error</i>	Error Rate
<i>Effcy</i>	Efficiency
<i>CR</i>	Completion Rate
\bar{T}	Average Time on Task
<i>SUS</i>	SUS Score
<i>U</i>	Score attributed to i^{th} SUS item question
<i>Z</i>	Z-Score
<i>x</i>	observed value for each task within the corresponding usability metric
<i>Cov</i>	Covariance between usability metrics X in Eq. (5) is one of the two metrics involved in the covariance (e.g., effectiveness), where X in Eq. (7) refer to normalised data matrix, with dimensions $n \times m$
<i>X</i>	Y in Eq. (5) is one of the two metrics involved in the covariance (e.g., efficiency), where Y in Eq. (7) refer to the data matrix projected into the principal component space, with dimensions $n \times k$
<i>Y</i>	

\bar{X}	Average values for variable X_i
\bar{Y}	Average values for variable Y_i
C	Covariance matrix of data measured as $m \times m$
v	Eigenvector of covariance matrix C
I	Identity matrix of the same size as C
V	Diagonal matrix with a value of 1 on the main diagonal
R	Eigenvector matrix
e	Cumulative explained variance
ln	Entropy of the normalized data matrix
n	Natural logarithm of n number
H	number of usability components
W	Entropy of each Principal Component
m	Weight of each Principal Component
SUM	Number of principal components retained. In this work is 2
	SUM score

Greek symbols

λ	Eigenvalue of the i^{th} principal component
μ	Mean value for each task within every

	usability aspect
σ	Standard deviation for each usability component

Subscripts

$rate$	The percentage of errors encountered during task completion
max	The maximum efficiency ratio is employed as the normalization threshold for scaling efficiency values to the 0–100% range
n	Denotes the SUS questionnaire's five odd–even item pairs iteration index
i	The task index for each usability metric
j	Index of primary component retained
k	k in Eq. (9) and (11) are probability values range from 1 to n for each usability metric, where k in Eq. (12) and (13) refer to index corresponding to each task contributing to the composite SUM calculation
z	The task-level SUM score computed from each task's z-score