TULSSAA Journal

Vol. 13 Issue No. 1, August 2026

DOI- https://doi.org/10.3126/tulssaa.v13i1.98608

TULSSAA, Nepal

Analyzing the System Usability Scale in Open-Source Information Systems: An Item-Level Psychometric Assessment of Post-Adoption Usability of KOHA ILS

Prof. Dr. Subarna Shakya

Rajendra Neupane

Email : rajpane45@gmail.com

Abstract

System Usability Scale (SUS) is still an industry standard tool for fast software usability testing. But its psychometric performance in the long-term post-adoption cycles, in developing countries institutional structures is less investigated. This study is based on the Usability Theory and Technology Acceptance Model (TAM), and narrows the scope to the SUS to conduct a separate psychometric diagnostics of the open source KOHA Integrated Library System (ILS). In 2025, a follow-up sample of 20 librarians with long-term experience with KOHA ILS was given the same 10 item SUS after an initial baseline assessment of the sample. To get more precise confidence intervals for the regression, 1000 resamples were used in bootstrapping, given the small sample size. The mean SUS score of the system was 61.6, which shows that the system has acceptable usability level. An item-level analysis showed that the respondents agreed on the positive rating of the system integration, but there were still some problems with regard to the system complexity and system learnability according to the rating of the negatively worded items. Based on exploratory regression analysis, there were positive correlations between user competence and interface customization and SUS scores. The results have to be considered with some caution due to the small sample size.

Key Words: TAM, SUS, KOHA, ILS, Learnability, Post- adoption

1. Introduction

A long-term challenge in information systems management is to quantify user experience (UX) over a prolonged period of time after the adoption. To overcome costs of licenses, public and academic institutions in developing countries heavily depend on open-

source software (OSS) such as the KOHA Integrated Library System (ILS). The knowledge of software usability is of fundamental importance to avoid the systemic operational failures. Objective task metrics yield technical baseline data; psychometric evaluation scales measure subjective cognitive friction, anxiety and ultimate user satisfaction.

Developed by John Brooke in 1996, the System Usability Scale (SUS) is a standard measure for speedily assessing usability. But modern UX literature often simplifies the final SUS score into a single, unchanging metric, while concealing the changes that occur to each of the individual psychometric dimensions when used for the first time. This study helps to separate the SUS to observe its diagnostic behaviour in the library environment in Nepal. This paper reveals particular patterns of system complexity that are masked by simpler composite scores, through an analysis of item-level score distributions.

2. Theoretical Background and Psychometric Foundations

The Technology Acceptance Model (TAM) is the primary paradigm used to guide this study, which states that users' acceptance and continued use of information systems is largely determined by their Perceived Ease of Use (PEOU). TAM is complemented by the System Usability Scale (SUS), which offers a standardized measure of perceived usability, to get an empirical assessment of people's perceptions of system usability after the system has been implemented.

The standard SUS has 10 items, 5 of which are negatively worded (Items 2, 4, 6, 8, and 10), and 5 positively worded (Items 1, 3, 5, 7, and 9). This alternating item format helps reduce acquiescence response bias and allows for careful consideration of each item rather than a uniform response.

The SUS could also be considered at the subscales level, following the two factor interpretation proposed by Lewis and Sauro (2009), which consist of:

Usability (Items 1, 2, 3, 5, 6, 7, 8, and 9): Provides users' perspective of the overall usability of the system, effectiveness, efficiency, ease of use and confidence in system's ability in performing library tasks.

Learnability (Items 4, 10): How easily users can learn to use the system, including the perceived complexity and technical support/training required.

This interpretation is intended to be used in conjunction with the overall SUS score to give more detailed information about the dimensions of usability, without destroying the overall usability measure of the SUS score.

3. Methodology

3.1. Evaluation Context and Sample

This study is a revisit for evaluating KOHA ILS post-adoption in 2025. The participants consisted of a baseline sample of librarians, and users who remained using the system over time. After a period of extended operational experience with KOHA, 20 of the original cohort completed the SUS evaluation form. The interpretation of the findings should be considered as a follow-up assessment, as presented in this paper is only a part of the repeated-measures analysis of the longitudinal scale. The number of participants in the follow-up group was reduced from the original group due to availability of participants during the study period. The limitations of sample size and generalizability of the follow-up sample do not detract from the strength of the statistical results, but do give some evidence to explore post-adoption usability issues of experienced users of KOHA. Qualitative baseline interviews were completed to add context to the organizations, in addition to the quantitative survey administration. The study is a post adoption diagnostic exploratory study because of bounded sample constraints (n = 20). There were 20 librarians in the baseline cohort who completed the follow-up assessment. While this small sample size decreases the power and generalizability of the study, it still reflects the individuals whose use of KOHA persisted throughout the study. Therefore, the results are considered as preliminary data and should not be interpreted as population estimates.

3.2. Instrument Scoring Mechanics

The standard 10-item SUS scale was administered on a 5-point Likert scale (1 = "Strongly Disagree" to 5 = "Strongly Agree"). Individual item responses were normalized to a 0–4 scale standard as follows:

- For odd-numbered (positive**)** items: $Xi = Score_i - 1$
- For even-numbered (negative) items: $Xi = 5 – Score_i$

SUS Score= $\sum_1^{10}(Xi)$*2.5

where:

- X_i = the normalized score for item *i* (after converting odd and even items)
- $\sum_1^{10}(Xi)$= the sum of the normalized scores for all 10 SUS items
- Multiplying by 2.5 converts the total from a 0–40 scale to a 0–100 scale.

4. Results: Item-Level SUS Diagnostics

4.1. Composite Score Re-Calculation

The converted mean item values yield a summed total of 24.64 across the 10 items. Applying the standard scaling multiplier (24.64 times 2.5) produces a corrected Composite Mean SUS Score of 61.6 (Range: 42.5–87.5). A score of 61.6 falls below the industry acceptability benchmark (68.0), classifying the system's current implementation state as marginal.

4.2. Positive vs. Negative Item Divergence

Deconstructing individual item means (0–4 scale) reveals a pronounced psychological split among users.

Table 1: Item-Level Diagnostic Breakdown (n = 20)

Item Configuration	Structural Item Statement (Abbreviated)	Converted Mean (0–4 Scale)
Item 1 (Positive)	Desire for frequent use of KOHA ILS	3.42
Item 2 (Negative)	Found the KOHA ILS unnecessarily complex	1.85
Item 3 (Positive)	Perceived ease of use of the platform	3.15
Item 4 (Negative)	Would need technical support to use the system	1.41
Item 5 (Positive)	System functions are well integrated	3.22
Item 6 (Negative)	Thought there was too much inconsistency	2.10
Item 7 (Positive)	Most people would learn to use KOHA quickly	2.88
Item 8 (Negative)	Found the system very cumbersome to use	2.35
Item 9 (Positive)	Felt very confident using the interface	3.04
Item 10 (Negative)	Needed to learn many things before getting started	1.22

ITEM-LEVEL CONVERTED VALUES (0-4 Scale Breakdown)

Positive Items (1, 3, 5, 7, 9) = 3.14
Negative Items (2, 4, 6, 8, 10) = 1.79

*Note: Lower scores on converted negative items indicate high operational friction.

The 5 positive items yielded an average score of 3.14 out of 4.00, demonstrating high appreciation for system integration (Item 5 = 3.22) and user willingness to engage with the system (Item 1 = 3.42).

In contrast, the 5 negative items averaged 1.79 out of 4.00. Lower scores on Items 4 and 10 suggest that respondents continued to perceive difficulties related to learning and independent system use. These findings indicate areas where additional training and interface refinement may improve the user experience

5. Statistical Predictors of Psychometric Usability

The factors influencing the overall usability were analyzed using a multiple linear regression model, in which the overall System Usability Scale (SUS) score of 0 to 100 was considered the dependent variable (DV) and the independent variables (DVs) were User Competence (X1), Interface Customization (X2), and Management Support (X3).

The standard regression assumptions were pretested before estimating the model. Scatterplots indicated possible linear relationships between the predictors and the dependent variable. Multicollinearity was assessed by the Variance Inflation Factors (VIFs) and were found not to exceed the usual values. There were no significant departures from homoscedasticity as revealed by residual plots. These diagnostics should be treated with some caution, due to the small sample size.

The model is specified as:

$$\text{SUS Score}=\beta 0+\beta 1X1+\beta 2X2+\beta 3X3+\varepsilon$$

where:

- SUS Score = Overall System Usability Scale score (0–100), the dependent variable.
- $\beta 0$ = Intercept (constant).
- X1 = User Competence.
- X2 = Interface Customization.
- X3 = Management Support.

- β1, β2, β3 = Regression coefficients representing the expected change in the SUS score associated with a one-unit increase in each predictor, holding the other variables constant.
- ε = Random error term.

Given the sample size (n = 20), bootstrapping with 1,000 resamples was conducted to generate stable parameters and 95% confidence intervals.

Table 2: Bootstrapped Regression Weights for SUS Outcomes

Predictor Parameter	**β Weight**	**Standard Error**	**95% Bootstrapped CI**	**Significance (p)**
Intercept (β0)	12.40	4.50	[3.40, 21.40]	< 0.05
User Competence (X1)	4.20	1.10	[2.10, 6.40]	< 0.01
Interface Customization (X2)	3.80	0.90	[1.90, 5.60]	< 0.01
Management Support (X3)	0.80	1.40	[-1.80, 3.50]	> 0.05

User Competence, Interface Customization, and Management Support were measured using composite scores derived from multiple Likert-scale questionnaire items included in the broader KOHA usability survey. Higher scores indicated greater perceived competence, higher levels of interface customization, and stronger organizational support.

To examine factors associated with post-adoption usability, an exploratory multiple linear regression model was developed using the composite SUS score as the dependent variable. The regression should be interpreted as an exploratory analysis because of the relatively small follow-up sample.

Prior to model estimation, assumptions of linear regression were examined. Scatterplots indicated approximate linear relationships between predictors and the dependent variable. Variance Inflation Factors (VIFs) were below commonly accepted thresholds, indicating no serious multicollinearity. Residual plots suggested no substantial heteroscedasticity or influential outliers. Given the limited sample size, these diagnostic procedures should be regarded as indicative rather than definitive.

Because the follow-up sample was relatively small (n = 20), bias-corrected bootstrapping with 1,000 resamples was employed to obtain more robust confidence intervals for regression coefficients. Bootstrapping improves estimation of sampling variability under small-sample conditions but does not eliminate limitations associated with statistical power. Consequently, the regression findings should be interpreted as exploratory.

Bootstrapped parameter estimates confirm that User Competence (beta = 4.20, $p < 0.01$) and Interface Customization (beta = 3.80, $p < 0.01$) serve as primary direct positive drivers of long-term SUS scores. Management Support (beta = 0.80, $p > 0.05$) did not demonstrate a statistically significant direct effect on psychometric usability outcomes.

6. Discussion

6.1. Resolving the Monolithic SUS Paradox.

A primary paradox of open-source information system post-adoption is that the item-level split brings about a great deal of frustration, yet there is a lot of value. The overall system has a composite score of 61.6, which is below the standard usability scores. Behind the high composite scores, though, item-level diagnostics reveal that high utility scores (TAM's Perceived Usefulness) are due to a significant level of complexity in the backend (Item 10) and dependency on technical (Item 4). Librarians are dedicated to using the software, but still have high cognitive load.

6.2. The Role of System Customization

The Interface Customization (beta = 3.80) stands out as having a significant predictive value, indicating a clear path for intervention. Open source platforms provide a lot of administrative flexibility, but by default a lot of metadata fields are unused and clutter the user's experience. Customization of layout and workflow of the interface directly affects Perceived Ease of Use (PEOU) and reduces complexity noted in Items 2, 4 and 10.

6.3. Management Support is an Indirect Enabler.

Management Support was not directly statistically significant (beta = 0.80, $p > 0.05$), due to the fact that this variable is not directly involved in the development of the new product. The analysis of the baseline qualitative interview data shows that institutional support is a control of basic infrastructure, including the quality of server operation, power quality and maintenance funds. Administrators don't use the day-to-day cataloging modules, but they are responsible for the structural support that helps ensure the success of softwarization and training interventions.

7. Conclusion and Recommendations

Granular, item-level psychometrics have to be used in conjunction with composite measures of usability in order to assess open source software across the lifespan. KOHA Integrated Library System (ILS) is a usable system, but the corrected SUS score value is 61.6 points which indicates that there are still usability issues that require attention particularly learnability, complexity of the user interface and efficient operation. The results indicate that while the system works, and is indeed widely accepted, user experience improvements must be targeted to specific usability problems and not just through general performance metrics.

Based on the findings the following recommendations are made to make KOHA usable in libraries in Nepal:

Break Down Composite Metrics: Be familiar with providing item level analysis of SUS along with overall SUS scores that can help to identify usability problems that may be masked by composite scores. The answers to each individual query item enable librarians and system administrators to highlight opportunities to enhance learnability, efficiency, error prevention and satisfaction.

Streamline Administrative Interfaces: delete unnecessary database fields from the administrative interface, to simplify cataloguing/circulation screens, and to set up workflows that are more relevant to the local library's needs. A user-friendly interface could help reduce errors in operation, simplify tasks and improve efficiency, and help to ease the user's cognitive load.

Advanced User Competency: Provide beyond basic training, and continuing professional development on advanced KOHA features, troubleshooting, customizing KOHA and generating reports. The building of in-house technical expertise will aid to increase the self-help level, user confidence and reduce the need for external technical support.

Ensure Sustainable Technical Infrastructure: Commit regular to the institution to server maintenance, software updates, data backup and security management. Strong commitment towards management and proper resource allocation are key factors in ensuring reliability, reducing downtime and supporting the sustainability of deployments of KOHA.

Combined, these will help libraries move beyond a successful system adoption to a successful and sustained system use. Through ongoing assessment of usability, as well as strategic improvements to the KOHA user interface, capacity development and institutional support, libraries can further improve user satisfaction, as well as operational efficiency and maximise the long-term value of KOHA Integrated Library System.

References

Bansode, S. Y., & Periera, S. (2018). Usability evaluation of open-source integrated library systems: A comparative study of Koha and NewGenLib. *Journal of Library & Information Technology*, *38*(3), 195–201.

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

Chaddah, A., & Gupta, M. (2022). Implementation and user perceptions of Koha ILS in developing economies: Empirical evidence from South Asia. *International Journal of Information Management*, *63*, Article 102114.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Eyiolorunshe, T., Hould, S., & Aboyade, M. (2023). Socio-technical challenges in the adoption of open-source software in university libraries. *Library Hi Tech*, *41*(2), 489–505.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM: Focus on bootstrapping parameters. *European Business Review*, *31*(1), 2–24.

Mumford, E. (2000). A socio-technical approach to systems design. *Requirements Engineering*, *5*(2), 125–133.

Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.

Shrestha, N., & Maharjan, P. (2024). Digital library automation trends in Nepal: Evaluating the trajectory of open-source platforms. *Nepalese Journal of Library and Information Science*, *8*(1), 45–61.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315.