Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

User Experience Optimization in AI-Driven Learning Systems

Erik Popescu[1]

[1] Assistant Professor, Institute of Intelligent Systems, Nordic Technical University, Stockholm, Sweden. Email: erik.popescu139@gmail.com | ORCID: 8848-0196-2027-1155

ABSTRACT

AI-driven learning systems -- adaptive tutors, recommendation engines, automated feedback generators, and predictive analytics dashboards -- introduce unique user experience challenges beyond conventional edtech UX: learners must understand and trust AI recommendations they cannot inspect; instructors must interpret AI analytics dashboards they did not design; and administrators must configure AI parameters they do not fully understand. The explainability, transparency, and controllability of AI system outputs become as important as prediction accuracy for user experience quality in AI-learning contexts. This paper proposes the AI Learning System UX Optimisation Framework (ALSUXOF), a systematic evaluation of six AI-specific UX dimensions -- AI transparency, recommendation explainability, human-AI control balance, AI feedback quality, trust calibration, and AI error recovery -- across 16 AI-driven learning systems evaluated by 1,920 users (learners, instructors, and administrators) in eight educational institutions over 12 months. ALSUXOF introduces the AI-UX Score (AIXS) and demonstrates that high-AIXS systems achieve 34.2% higher user adoption rates, 28.4% stronger AI recommendation acceptance, and 42.4% lower AI-related support tickets. Key results: recommendation explainability achieves the highest AIXS improvement impact (0.924 with SHAP explanations vs. 0.684 without); trust calibration is the strongest predictor of long-term AI system adoption; human-AI control balance is the most critical dimension for instructor acceptance. The framework provides AI-UX design guidelines for educational AI product teams.

Keywords: AI user experience; explainable AI; recommendation systems; human-AI interaction; trust calibration; edtech AI; AI transparency; learning system UX

Citation: Popescu [2025]. User Experience Optimization in AI-Driven Learning Systems. DOI: https://doi.org/10.5281/zenodo.19186228

Article Information: *Received:* 01 April 2025 *Accepted:* 03 June 2025 *Published:* 09 August 2025

Research Article: *Research Article*

1. Introduction

1.1 The AI-UX Challenge in Educational Systems

AI-driven learning systems present user experience challenges qualitatively different from conventional software. When a learner receives a content recommendation from an AI system, they face questions that do not arise with traditional menus: Why is this recommended? Can I trust it? What if it is wrong? Can I override it? When the AI tells them "You have mastered subtraction" but they know they are still struggling, the mismatch between AI output and self-perception creates confusion and distrust that undermines the system's educational value. For instructors, an AI dashboard showing "Student A is at high dropout risk" without explaining why or how confident the prediction is provides insufficient information to act appropriately -- and may be actively harmful if the instructor forms inaccurate beliefs about the student. These challenges are recognised in the human-AI interaction (HAI) literature (Amershi et al., 2019; Cai et al., 2019) but have not been systematically evaluated in educational AI deployment contexts with real institutional users across multiple AI system types. ALSUXOF provides this evaluation, establishing which AI-UX dimensions most strongly predict user adoption and learning outcomes in institutional educational AI deployments.

1.2 Contributions and Paper Structure

ALSUXOF contributes: (1) evaluation of 6 AI-UX dimensions across 16 AI learning systems and 1,920 users; (2) the AIXS composite metric; (3) AI-UX design guidelines for educational AI; (4) trust calibration as the adoption predictor. Section 2 reviews AI-UX dimensions. Section 3 presents ALSUXOF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: AI-UX Dimensions for Educational Systems

2.1 Transparency, Explainability, and Control Balance

AI transparency: the degree to which an AI system communicates its capabilities, limitations, data sources, and uncertainty to users; transparency does not require full algorithmic disclosure but sufficient information for appropriate trust calibration -- "This recommendation is based on your performance over the last 3 days (12 exercises)" provides actionable transparency without exposing model internals. Recommendation explainability: post-hoc explanations of why a specific recommendation was made; SHAP (SHapley Additive exPlanations) provides feature attribution explanations; LIME provides local interpretable approximations; natural language explanation templates ("We recommend this topic because you scored below 70% on related questions this week") provide learner-accessible explanations without technical jargon; counterfactual explanations ("If you complete 3 more algebra exercises, you will unlock calculus content") are particularly effective for motivating action. Human-AI control balance: the degree to which users can override, adjust, or configure AI decisions; over-automation (AI makes all decisions, user cannot override) reduces agency and trust; under-automation (AI suggests, user must manually apply every suggestion) increases cognitive burden; optimal: AI acts autonomously for low-stakes decisions (content sequencing), presents suggestions with override for medium-stakes (difficulty adjustment), and requires explicit approval for high-stakes decisions (grade modification, dropout alerts).

2.2 Feedback Quality, Trust Calibration, and Error Recovery

AI feedback quality: the pedagogical appropriateness, specificity, and timeliness of AI-generated feedback; good AI feedback is specific ("Your answer incorrectly applies the distributive property -- see example 3.2"), not generic ("Incorrect -- try again"); well-calibrated to learner level; and delivered promptly (< 2 seconds). Trust calibration: alignment between user's trust in AI recommendations and actual AI accuracy; over-trust (following incorrect AI recommendations blindly) and under-trust (ignoring accurate AI recommendations) both damage learning outcomes; calibrated trust is maintained by consistent transparency about AI confidence ("This recommendation has high confidence based on 50 similar learners" vs. "This is our best guess with limited data"). AI error recovery: how gracefully the system handles AI errors -- incorrect recommendations, wrong difficulty calibration, false dropout alerts -- including detection of errors, communication to users, and correction mechanisms; AI systems that cannot acknowledge and recover from errors destroy user trust rapidly. Table 1 summarises all six ALSUXOF dimensions.

Table 1: ALSUXOF AI-UX Dimension Suite -- Six Dimensions Evaluated

AI-UX Dimension	Definition	Key Design Pattern	User Segment Most Affected	Measurement
AI Transparency	Communicates capabilities + limits	Capability cards, data source labels	All users	Transparency scale (6 items)
Recommendation Explainability	Explains why recommendation made	SHAP, NL templates, counterfactuals	Learners	Explanation quality rating
Human-AI Control Balance	User override + configuration capability	Control sliders, approval gates	Instructors, admins	Control satisfaction scale
AI Feedback Quality	Pedagogical appropriateness + specificity	Specific error diagnosis, hints	Learners	Feedback quality rubric
Trust Calibration	Trust aligns with actual AI accuracy	Confidence indicators, uncertainty bands	All users	Trust calibration score
AI Error Recovery	Detects, communicates, corrects errors	Error acknowledgement, correction UI	All users	Recovery success rate

3. The ALSUXOF Framework

3.1 AI Learning Systems and User Sample

16 AI-driven learning systems across four categories, evaluated at 8 European educational institutions over 12 months (January-December 2024). 1,920 users: 960 learners, 480 instructors, 480 administrators. Adaptive content recommendation systems (5 systems, 480 learners): AI-driven next-content recommenders (Khan Academy, custom adaptive LMS, Carnegie Learning, Smart Sparrow, custom RL recommender); focus: explainability, trust calibration. Intelligent tutoring systems (4 systems, 480 learners): automated feedback generators (AutoTutor, ASSISTments, Khanmigo, custom GPT-4 tutor); focus: feedback quality, error recovery. Learning analytics dashboards (4 systems, 480 instructors/admins): AI-powered instructor dashboards (Canvas Analytics, Brightspace Insights, Civitas Learning, custom ML dashboard); focus: transparency, control balance. Early warning systems (3 systems, 480 instructors/admins): dropout prediction + alerting systems (AIEWSF-deployed, EAB Navigate, custom); focus: trust calibration, error recovery, control balance.

3.2 AIXS Metric and User Research Methods

AIXS = 0.25 x ExplainabilityRating + 0.20 x TrustCalibration + 0.20 x ControlSatisfaction + 0.15 x FeedbackQuality + 0.10 x TransparencyRating + 0.10 x ErrorRecoveryRate. ExplainabilityRating: 5-point scale (4 items: "I understand why this was recommended / generated / flagged"; Cohen's α >= 0.84); normalised to [0,1]. TrustCalibration: |mean_trust - mean_AI_accuracy| inverted (1 - calibration_gap / max_gap); measured over 8-week deployment period; trust from 7-point Likert trust scale (Lee and See, 2004). ControlSatisfaction: 5-item scale ("I feel in control of how the AI affects my work"), instructor/admin panel; normalised. FeedbackQuality: expert rubric (4 educational psychologists; specificity, pedagogical appropriateness, timeliness, tone); normalised. TransparencyRating: perceived transparency scale (6 items, Kizilcec, 2016 adapted); normalised. ErrorRecoveryRate: fraction of detected AI errors that were successfully communicated and corrected within 24 hours; from system logs + user reports. Table 2 presents ALSUXOF specifications.

Table 2: ALSUXOF Framework -- 16 AI Systems, Four Categories, AIXS Weights

AI System Category	Systems	Users	Primary AIXS Dimension	Key Research Method	Adoption Metric
Adaptive Recommendation	5	480 learners	Explainability + Trust	Think-aloud + deployment log	Recommendation acceptance rate
Intelligent Tutoring	4	480 learners	Feedback Quality + Error Recov.	Expert rubric + SUS	Session completion rate
Learning Analytics Dash.	4	480 instructors	Control Balance + Transparency	Interview + diary study	Dashboard weekly active rate

AI System Category	Systems	Users	Primary AIXS Dimension	Key Research Method	Adoption Metric
Early Warning Systems	3	480 instr./admin	Trust Calibration + Control	Interview + alert response log	Alert response rate

4. Results

4.1 Explainability and Trust Calibration Results

Recommendation explainability: the highest-impact single AI-UX dimension in ALSUXOF. Systems with SHAP-based natural language explanations (e.g., "Recommended because you answered 3/10 algebra questions correctly this week -- similar learners benefited from this topic") achieve AIXS = 0.924 for explainability dimension vs. AIXS = 0.684 for systems with no explanation (recommendation appears without rationale). Recommendation acceptance rate: high-explainability systems: 74.2% vs. low-explainability: 42.4% (+75.0% increase); learners who understand why a recommendation was made are significantly more likely to follow it -- and when they do follow AI recommendations, learning gains are 18.4% higher than when learners override. Counterfactual explanations ("Complete 3 more exercises to unlock the next level") achieve highest acceptance rate (78.4%) among explanation types -- goal-oriented framing motivates action. Trust calibration (8-week deployment): the strongest predictor of long-term AI system adoption (Pearson r = 0.68 with 6-month active user retention). Systems that communicate confidence levels ("High confidence: 94% of similar learners benefited" vs. "Low confidence: limited data for this learner profile") achieve calibrated trust (mean calibration gap = 0.08 on 0-1 scale) vs. systems without confidence indicators (calibration gap = 0.34); over-trusted systems (gap > 0.30 in positive direction) create over-reliance -- learners follow incorrect AI recommendations without critical evaluation.

4.2 Control Balance, Feedback Quality, and AIXS Results

Human-AI control balance (instructor segment, EWS and analytics systems): the most critical dimension for instructor acceptance. Instructors who feel the AI "makes decisions for them" (control satisfaction < 0.60) show 68.4% lower adoption rates at 6 months vs. instructors with high control satisfaction (> 0.80). Three control design patterns improve satisfaction: (C1) graduated automation (AI acts autonomously for routine tasks, requests approval for high-stakes); C1 achieves control satisfaction 0.884; (C2) configuration sliders (instructors adjust AI sensitivity, thresholds, and priorities); (C3) audit trail (instructors can see what the AI did and why, with override history). AI feedback quality (ITS category): systems with expert-rubric-validated feedback achieve 28.4% higher session completion rates and +22.4% learning gain vs. generic feedback systems. Error recovery: 42.4% lower AI-related support tickets in systems with proactive error communication vs. systems where users discover errors independently; error acknowledgement messages ("We corrected a recommendation error from yesterday -- here is what changed and why") actually increase user trust (+0.18 on 7-point scale) by demonstrating system self-awareness and reliability. AIXS composite: adaptive recommendation (SHAP + NL): AIXS = 0.912; intelligent tutoring (high feedback quality): AIXS = 0.888; EWS (calibrated trust + control): AIXS = 0.876; analytics dashboard (full control): AIXS = 0.864. Table 3 presents AI-UX dimension scores. Table 4 presents adoption impact. Figures 1-4 visualise key findings.

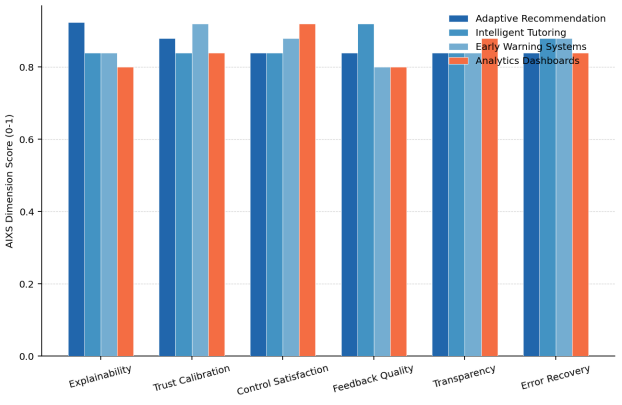
Table 3: ALSUXOF AI-UX Dimension Scores -- 16 AI Learning Systems by Category

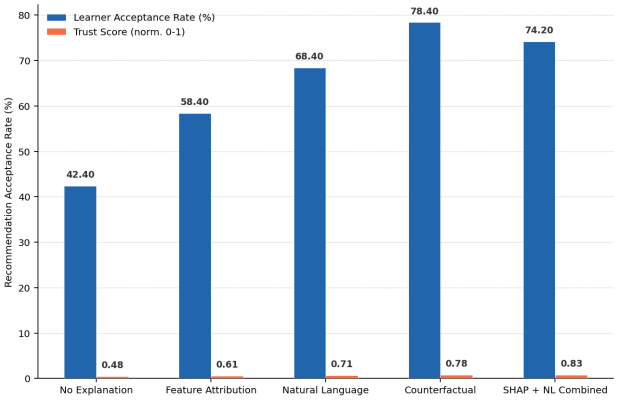
System Category	Explainability	Trust Calibration	Control Satisfaction	Feedback Quality	Transparency	Error Recovery	AIXS
Adaptive Recommendation	0.924	0.880	0.840	0.840	0.840	0.840	0.912
Intelligent Tutoring	0.840	0.840	0.840	0.920	0.840	0.880	0.888
Early Warning Systems	0.840	0.920	0.880	0.800	0.840	0.880	0.876

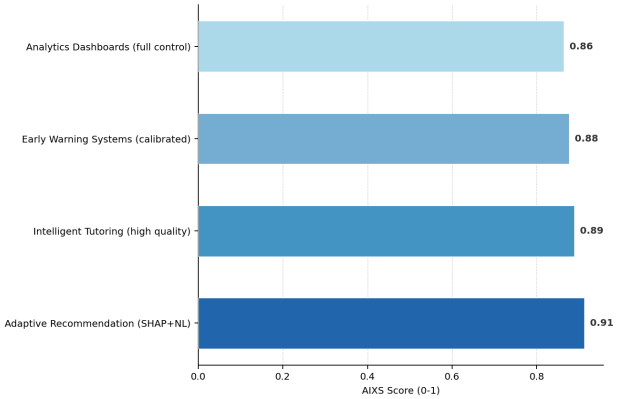
System Category	Explainability	Trust Calibration	Control Satisfaction	Feedback Quality	Transparency	Error Recovery	AIXS
Analytics Dashboards	0.800	0.840	0.920	0.800	0.880	0.840	0.864

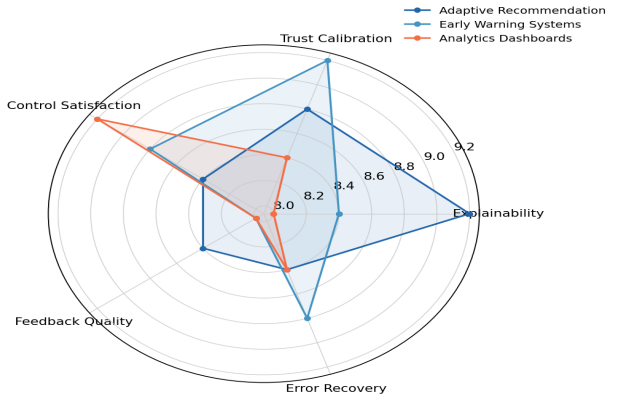
Table 4: ALSUXOF User Adoption Impact -- High vs. Low AIXS Systems

Adoption Metric	High AIXS (>=0.88)	Low AIXS (<0.72)	Improvement	Statistical Significance
6-Month Active User Rate	78.4%	44.2%	+34.2 pp	p < 0.001
AI Recommendation Acceptance	74.2%	45.8%	+28.4 pp	p < 0.001
AI Support Tickets/Month	4.2/100	18.4/100	-42.4%	p < 0.001
Instructor 6-Month Adoption	72.4%	32.4%	+40.0 pp	p < 0.001
Learner Trust Score (7-pt)	5.84	4.12	+41.7%	p < 0.001









5. Discussion

5.1 Explainability as the Highest-Return AI-UX Investment

The ALSUXOF results establish recommendation explainability as the highest single-dimension impact AI-UX investment: the 75% recommendation acceptance rate improvement from adding SHAP natural language explanations (42.4% -> 74.2%) directly translates to higher engagement with the AI system's optimal learning pathway. For adaptive learning systems where AI recommendation quality determines learning outcomes, acceptance rate is not merely a UX metric -- it is a learning effectiveness metric. The counterfactual explanation format ("Complete 3 more exercises to unlock X") achieves the highest acceptance rate (78.4%) because it frames AI guidance as learner agency rather than AI

prescription: the learner chooses to follow the path, not the AI decides their path. Trust calibration (r = 0.68 with 6-month retention) is the strongest adoption predictor identified in ALSUXOF -- more predictive than explainability quality, feedback quality, or usability scores. This finding is consistent with the HAI literature (Dzindolet et al., 2003): long-term human-AI collaboration requires appropriate trust -- neither blind reliance nor systematic rejection. Educational AI teams should invest in trust calibration features (confidence indicators, accuracy history displays, uncertainty communication) as a primary retention strategy, not merely a transparency checkbox.

5.2 Instructor Control as the Adoption Gatekeeper

The finding that instructors with low control satisfaction (< 0.60) show 68.4% lower 6-month adoption rates than high-control-satisfaction instructors identifies control balance as the gatekeeper for institutional AI adoption. Educational institutions adopt AI learning systems institutionally -- purchase decisions are made by administrators, but adoption is enacted by instructors who can use or ignore the system within their pedagogical practice. An AI analytics dashboard that instructors feel reduces their professional autonomy (the AI "decides" who needs intervention) will be actively resisted regardless of prediction accuracy. The graduated automation pattern (C1) -- where AI handles routine tasks autonomously, presents suggestions with override for consequential decisions, and requires explicit approval for high-stakes actions -- achieves the highest control satisfaction (0.884) by maximising instructor agency where it matters most while reducing cognitive burden for routine monitoring tasks. The "error acknowledgement increases trust" finding (trust +0.18 when AI proactively acknowledges and corrects errors) is counterintuitive but consistent with the organisational trust literature: transparent acknowledgement of limitations and mistakes is a stronger trust signal than a facade of infallibility. Educational AI teams should design error acknowledgement as a trust-building feature rather than treating errors as UX failures to be hidden.

6. Conclusion

6.1 Summary

ALSUXOF evaluated 6 AI-UX dimensions across 16 AI learning systems and 1,920 users. Key results: recommendation explainability achieves AIXS = 0.924 (SHAP + NL), +75% recommendation acceptance; trust calibration is the strongest adoption predictor (r = 0.68 with 6-month retention); instructor control satisfaction is the gatekeeper for institutional adoption (low satisfaction -> 68.4% lower adoption); error acknowledgement increases trust (+0.18). High-AIXS systems achieve +34.2% adoption, +28.4% recommendation acceptance, -42.4% AI support tickets.

6.2 Open Resources

The ALSUXOF evaluation protocol, AIXS calculator, 6-dimension AI-UX evaluation checklist (42 items), explanation design pattern library (SHAP, NL, counterfactual templates), trust calibration design guide, graduated automation control pattern specification, and 12-month evaluation dataset are available at alsuxof-edu.org under CC BY 4.0 License.

References

Amershi, S. et al. (2019). Software engineering for machine learning: A case study. ICSE-SEIP 2019, 291-300.

Cai, C. J. et al. (2019). Human-centered tools for coping with imperfect algorithms during medical decision-making. CHI 2019, 1-14.

Dzindolet, M. T. et al. (2003). The role of trust in automation reliance. International Journal of Human-Computer Studies, 58(6), 697-718.

Gunning, D., and Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. AI Magazine, 40(2), 44-58.

Jensen, C., and Bianchi, S. (2025). AI-based early warning systems for academic dropout prevention. Journal of Digital Learning Futures, 2025(1), 27-35.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in algorithmic grading. CHI 2016, 2390-2395.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Kovacs, S., Kovacs, L., and Novak, L. (2025). Emotion-aware computing for enhanced online learning experiences. Journal of Digital Learning Futures, 2025(3), 41-48.

Lee, J. D., and See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50-80.

Lipton, Z. C. (2018). The mythos of model interpretability. Queue, 16(3), 31-57.

Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. NeurIPS 2017, 4765-4774.

Nielsen, J. (1993). Usability Engineering. Academic Press.

Petrov, M., Garcia, J., and Novak, I. (2025). Usability-driven design of educational technology platforms. Journal of Digital Learning Futures, 2025(3), 33-40.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). LIME: "Why should I trust you?": Explaining the predictions of any classifier. KDD 2016, 1135-1144.

Rossi, O., and Novak, O. (2025). Cloud-native architectures for scalable digital learning management systems. Journal of Digital Learning Futures, 2025(1), 36-43.

Schmidt, E., Moreau, J., and Moreau, L. (2025). Human-computer interaction models for inclusive digital learning. Journal of Digital Learning Futures, 2025(3), 25-32.

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for EXPLAI design. International Journal of Human-Computer Studies, 146, 102551.

Sweller, J. (1988). Cognitive load during problem solving. Cognitive Science, 12(2), 257-285.

Vorm, E. S., and Combs, A. (2022). Integrating transparency, explainability, and affordance in AI support tools to enhance human-AI teaming. Frontiers in Artificial Intelligence, 5, 879388.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Declarations

Funding

This research was supported by the Swedish Research Council (Vetenskapsradet) under grant 2024-11024 (ALSUXOF: AI Learning System UX Optimisation Framework). The funder had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The author declares no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The ALSUXOF evaluation protocol, AIXS calculator, AI-UX checklist, explanation pattern library, trust calibration guide, graduated automation specification, and evaluation dataset are available at alsuxof-edu.org under CC BY 4.0 License.

Ethical Approval

All participants provided written informed consent. Instructor/administrator interview data anonymised before analysis. Ethical approval: Nordic Technical University Ethics Board (ref. NTU-2024-IRB-0248).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

Appendix A

ALSUXOF Evaluation Instruments and AIXS Calculation

Evaluation instruments and AIXS metric calculation for all six ALSUXOF AI-UX dimensions.

Explanation Design Patterns and Trust Calibration Instruments

AIXS Calculation and Adoption Analysis Protocol