China and WTO Review (CWR) Journal
2026, (12)
ISSN: 2384-4388
https://cwto.net
Research Article

Competency Dashboards for Tracking Progress in Hands-On Learning Programs

Emily Carter [1*], Daniel Brooks [1], Rachel Mitchell [1*]

[1] Graduate School of Education, Stanford University, Stanford, USA
* **Corresponding Author**: r.mitchell@institution.edu

ARTICLE INFO

ABSTRACT

This study examines how competency dashboards can support progress tracking and instructional decision-making in hands-on learning programs. The study is designed to involve approximately 800 students and 50 instructors from 40 hands-on learning courses over one semester, collecting around 18,000 task-completion records, 10,500 practice logs, 6,000 teacher feedback entries, 4,200 competency milestone records, 2,800 learner self-assessment entries, and 1,600 pre-test and post-test performance scores. Key variables include task completion rate, competency milestone achievement, feedback frequency, feedback adoption rate, practice intensity, error-reduction rate, self-assessment accuracy, learner self-management score, instructor intervention frequency, and final performance improvement. The study applies learning analytics, dashboard usability evaluation, latent growth modeling, cluster analysis, multilevel regression, and path analysis to examine whether dashboard use improves progress awareness, self-regulated learning, and instructional adjustment. Dashboard effectiveness is evaluated through learner progress visibility, milestone completion rate, learning gain, feedback-use efficiency, self-management improvement, and instructor decision accuracy. The innovation of this study lies in replacing single final grades with dynamic competency tracking, allowing instructors and learners to observe how specific practical skills develop over time and to adjust learning plans based on real-time progress evidence.

Keywords: Competency Dashboard, Hands-On Learning, Progress Tracking, Learning Analytics, Practical Competencies, Self-Regulated Learning, Instructional Decision-Making, Learning Gain.

INTRODUCTION

Hands-on learning is a central component of vocational education, engineering, health training, laboratory instruction, design, and creative practice. In these settings, competence develops through repeated performance, guided practice, feedback, reflection, and correction of errors rather than through a single final assessment. Conventional grades can summarize overall achievement, but they provide limited information about how individual skills develop over time, where performance begins to stagnate, or which specific competencies still require support. Competency-based assessment responds to this limitation by linking learning outcomes to observable performance and clearly defined skill standards [1,2]. Research on collaborative teaching and teaching-resource efficiency has also emphasized that effective practical education depends on the coordinated use of instructional resources, feedback, and learning support [3]. This is particularly important in hands-on courses, where students may progress at different rates and require different types of guidance. Learning analytics dashboards offer a possible way to make this progression more visible by integrating task records, practice data, feedback, self-assessment, and performance results within a single interface [4,5]. Rather than relying only on end-of-course scores, such systems can provide continuous evidence of what students have completed, how their performance is changing, and where additional intervention may be needed.

Student-facing dashboards commonly present indicators such as task completion, study time, assessment results, and progress toward course goals. Their value, however, depends on whether the information is understandable and useful for learning decisions. Previous research has shown that students generally value clear progress information and practical guidance, while comparative indicators such as peer ranking may be less helpful [6,7]. Other studies have noted that many dashboards remain largely descriptive and present data without explaining the causes of difficulty or suggesting appropriate next actions [8,9]. Recent reviews indicate a gradual shift toward learner-centered dashboard design, with greater attention to learning needs, instructional goals, user participation, and the alignment between displayed indicators and meaningful educational decisions [10,11]. This development is important because a dashboard should not function merely as a reporting interface. In practical learning environments, it should help learners understand what they have mastered, what remains incomplete, and how their next actions may influence later performance. This role is closely related to self-regulated learning. Effective learners are expected to set goals, monitor their progress, adjust strategies, respond to feedback, and reflect on outcomes. Learning analytics can support these processes when the information presented is sufficiently timely and interpretable. Evidence suggests that analytics can reveal relationships between course design and learner strategies [12], while explainable models can generate feedback and action-oriented recommendations rather than simple risk labels [13,14]. Feedback is also more effective when it is timely, clear, and connected to later action [15]. These findings suggest that the educational value of dashboards lies not in visualization alone but in their capacity to support a cycle of observation, interpretation, adjustment, and improvement. For hands-on learning, this cycle may include repeated practice, correction of technical mistakes, revision of procedures, and gradual achievement of competency milestones. Evidence regarding actual learning gains remains inconsistent. Reviews of learning analytics interventions have reported that fewer than half of evaluated systems produced clear improvements in learning outcomes [16]. Other studies found that real-time analytics support could strengthen self-regulated learning behavior without producing statistically significant gains in final achievement [17,18]. In contrast, analytics-based feedback has also been associated with more regular study behavior, higher task completion, and improved test performance [19]. These mixed findings indicate that dashboard effectiveness depends on more than system availability. The type of data displayed, the timing of feedback, the quality of interpretation, and the nature of the learning task can all influence whether students change their behavior. In practical courses, this issue becomes more important because skill development is cumulative and often nonlinear. A student may complete all required tasks while still repeatedly making the same procedural error, or may receive a satisfactory final score despite weak performance in one critical competency. Aggregate course indicators are therefore insufficient for diagnosing practical learning needs. Teacher-facing dashboards are intended to support more informed instructional decisions, but their usefulness depends on how instructors interpret and act on the data. Teacher adoption has been linked to perceived usefulness, ease of use, and institutional support [20]. Dashboards have also been used to help instructors examine teamwork, collaboration patterns, and group-learning processes [21]. More recent work suggests that teachers with stronger monitoring and diagnostic skills are better able to identify learning problems from dashboard information [22]. Related studies have emphasized the importance of teacher training, carefully selected indicators, transparent data presentation, and alignment between analytics and course objectives [23,24]. These findings show that dashboards do not automatically improve teaching quality simply by providing more information. Their value depends on whether the indicators correspond to meaningful pedagogical decisions and whether instructors can translate the displayed patterns into timely interventions. Several limitations remain in current dashboard research. A large proportion of studies still rely primarily on learning management system data, including logins, page views, clicks, time-on-platform, and submission records. These indicators are useful for describing online participation but do not directly measure practical competence. A student may access course materials frequently without demonstrating improvement in laboratory technique, equipment handling, clinical procedure, design execution, or other hands-on skills. Many studies are also limited to one course, one institution, a relatively small sample, or a short implementation period. Such designs provide limited evidence about differences across learners, instructors, and course types. Pre-test and post-test comparisons further reduce a complex learning process to two measurement points and cannot capture periods of rapid improvement, stagnation, regression, or delayed progress across repeated practice stages. Current evaluations also tend to examine isolated outcomes. Satisfaction, usability, dashboard frequency, or final achievement are often analyzed separately, while feedback adoption, error reduction, self-assessment accuracy, competency milestones, teacher intervention, and instructional decision quality are rarely considered within the same analytical framework. This separation makes it difficult to explain how dashboard use produces learning effects. A learner may view the dashboard regularly but ignore the feedback, while another may access it less often yet make substantial changes to practice behavior. Similarly, an instructor may identify a weak learner accurately but intervene too late to affect performance. The mechanism between dashboard use and learning improvement therefore requires closer examination. Another important limitation is that many dashboard systems implicitly treat learners as a relatively homogeneous group. In hands-on courses, however, competence may develop through markedly different trajectories. Some students improve steadily, some show rapid early gains followed by stagnation, some progress slowly but consistently, and others may continue to struggle despite repeated practice. Average scores can conceal these differences and lead to overly general instructional responses. A more useful competency dashboard should therefore support longitudinal tracking and distinguish different patterns of development rather than reporting only class-level averages. This requires analytical methods capable of modeling change over time, identifying heterogeneous learner groups, and accounting for the fact that students are nested within different courses and taught by different instructors.

The present study addresses these limitations by examining competency-dashboard use among approximately 800 students, 50 instructors, and 40 hands-on courses. The dataset integrates task completion records, repeated practice logs, teacher feedback, competency milestones, learner self-assessment, dashboard-use records, instructional interventions, and performance scores. Latent growth modeling is used to examine how competence changes across repeated stages of practice rather than reducing development to a single pre-test and post-test comparison. Cluster analysis is applied to identify distinct learner progress patterns, allowing differences in development trajectories to be examined directly. Multilevel regression is used to account for the hierarchical structure of the data, including students nested within courses and instructors, while path analysis examines the relationships among dashboard use, feedback adoption, self-management, teacher intervention, and learning gain. The study therefore seeks to determine not only whether competency dashboards are associated with better learning outcomes, but also how they support progress awareness, practical skill development, self-regulated learning, and instructor decision accuracy. Its significance lies in shifting dashboard evaluation from general activity monitoring toward continuous evidence of competency development. By connecting longitudinal performance records with learner behavior and instructional action, the study provides a framework for identifying individual learning needs, supporting targeted intervention, and improving the timing and accuracy of teaching decisions in hands-on education.

MATERIALS AND METHODS

Participants and Study Setting

The study was carried out over a 16-week semester in 40 hands-on courses offered by vocational colleges, applied universities, and adult education centers. The courses covered engineering practice, laboratory training, health skills, digital production, and creative workshops. About 800 students and 50 instructors took part in the study. Each course included 15–30 students and followed a task-based teaching plan with clear skill milestones. Students were included when they were formally enrolled, completed the baseline test, and attended at least 80% of the practical sessions. Students were excluded when they withdrew from the course, missed the pre-test or post-test, or had more than 30% missing activity records. Instructors were required to have at least one year of teaching experience and to use the same scoring guide throughout the semester. The dataset included about 18,000 task records, 10,500 practice logs, 6,000 teacher feedback entries, 4,200 skill milestone records, 2,800 self-assessment entries, and 1,600 paired test scores. Student age, prior experience, course type, class size, instructor experience, and baseline score were also recorded.

Experimental Design and Control Conditions

A course-level controlled design was used to reduce contact between the two study groups. The 40 courses were first matched by course type, class size, and mean baseline score. They were then assigned to either the dashboard group or the control group. The dashboard group included 20 courses and about 400 students. The control group included the other 20 courses with a similar number of students. Students in the dashboard group could view task completion, practice frequency, skill milestones, feedback status, error trends, and differences between self-ratings and teacher ratings. Instructors could view class progress, delayed milestones, repeated errors, low feedback use, and students who needed support. The dashboard was updated after each practical session. The control group followed the same teaching plan, task rules, scoring standards, and feedback schedule. However, students and instructors received progress information through standard course pages and regular score reports. Both groups used the same teaching materials and skill rubrics. Access to continuous dashboard information was therefore the main difference between the two groups. Baseline balance was checked using student background data, prior experience, pre-test scores, and initial self-management scores.

Measurement Procedures and Quality Control

Learning progress was measured using system records, teacher ratings, student reports, and practical tests. Task completion rate was defined as the share of assigned tasks completed within the required time. Practice intensity was measured by the number and duration of practice sessions each week. A skill milestone was recorded when a student met the required standard for a specific task. Feedback frequency referred to the number of teacher comments received. Feedback adoption rate was defined as the share of comments followed by a revision or corrected performance. Error-reduction rate was measured from changes in repeated task errors across sessions. Students completed a self-assessment after each major task. Self-assessment accuracy was calculated from the difference between student ratings and teacher ratings. Self-management was measured at the beginning, middle, and end of the semester. The scale covered goal setting, time planning, progress checking, and strategy adjustment. Instructor decision accuracy was assessed by comparing intervention decisions with later task results and milestone achievement. Three subject teachers reviewed the skill rubrics before data collection. A pilot study was conducted in two courses that were not included in the main sample. All instructors received the same scoring guide and rated sample tasks before the study began. At least 10% of the practical assessments were rated by two instructors. Duplicate records, invalid timestamps, extreme activity times, and missing values were checked each week. Survey scales with a Cronbach's alpha below 0.70 were not included in the final analysis.

Data Processing and Model Formulation

All records were linked using anonymous student, course, and instructor codes. Duplicate events were removed, and time records were converted into weekly values. Activity values above the 99th percentile were limited to reduce the effect of extreme observations. Missing background data were handled by multiple imputation when the missing rate was below 10%. Missing outcome scores were not imputed when no valid measurement was available. Continuous variables were standardized before cluster analysis and path analysis. Learning improvement was measured using normalized gain (1):

$$g_i = \frac{Post_i - Pre_i}{S_{max} - Pre_i} \tag{1}$$

where g_i is the learning gain of student *i*, Pre_i and $Post_i$ are the pre-test and post-test scores, and S_{max} is the highest possible score. Changes in skill performance were examined using a three-level growth model because repeated records were grouped within students, and students were grouped within courses (2):

$$Y_{tij} = \beta_0 + \beta_1 Time_t + \beta_2 Dashboard_j + \beta_3 (Time_t \times Dashboard_j) + \beta_4 X_{ij} + u_j + v_{ij} + \varepsilon_{tij} \tag{2}$$

where Y_{tij} is the skill score at time *t* for student *i* in course *j*. $Dashboard_j$ shows the study group, and X_{ij} includes baseline score, prior experience, attendance, class size, and instructor experience. The terms u_j, v_{ij}, and ε_{tij} represent course-level variation, student-level variation, and random error. The interaction term β_3 shows whether the rate of skill growth differed between the dashboard and control groups.

Statistical Analysis and Model Checks

Descriptive statistics were used to summarize student features, course conditions, dashboard use, and learning outcomes. Baseline differences between the two groups were tested using independent-sample tests and standardized mean differences. Latent growth models were used to examine changes in milestone achievement, self-management, and skill scores during the semester. Cluster analysis was used to identify common progress patterns based on task completion, practice intensity, error reduction, feedback adoption, and milestone achievement. The number of clusters was selected using the silhouette score, Bayesian information criterion, and the practical meaning of each group. Multilevel regression was used to test dashboard effects while accounting for differences among students, courses, and instructors. Path analysis was used to examine the links among progress visibility, feedback adoption, self-management, teacher intervention, and learning gain. Model fit was checked using the comparative fit index, Tucker–Lewis index, root mean square error of approximation, and standardized root mean square residual. Sensitivity tests excluded students with low attendance, courses with incomplete records, and extreme dashboard-use values. Effect sizes and 95% confidence intervals were reported together with ppp-values. Statistical significance was set at 0.05. The analyses were completed in R and Mplus. Participation was voluntary, and all personal information was removed before analysis. The study followed institutional rules for research involving students.

RESULTS AND DISCUSSION

Competency Growth and Milestone Achievement

After data screening, 786 students were included in the analysis, with 394 in the dashboard group and 392 in the control group. The two groups showed no significant differences in baseline practical scores, prior learning experience, attendance, or initial self-management scores (all "$p>0.05$"). Competency scores increased in both groups during the semester, but the dashboard group achieved a larger improvement. Its mean score rose from 52.6 at baseline to 82.4 at the final assessment, while the control group increased from 52.9 to 75.8. The interaction between group and time was significant ("$\beta=0.27,95$"). The final milestone completion rate was also higher in the dashboard group than in the control group (84.7% vs. 71.9%, "$p<0.001$"). As shown in **Figure 1**, the dashboard presented task progress, performance differences, and current learning status in a clear form. In this study, these indicators were linked to practical skill milestones and repeated errors. Li et al. (2026) reported that both descriptive and prescriptive dashboards improved learning, although prescriptive feedback had a stronger effect on learning strategies. The current results show that descriptive progress information can also support practical performance when it is closely linked to course tasks [25,26]. Compared with dashboards based mainly on login records or test scores, the competency dashboard showed which skills had been completed, which errors remained, and which tasks required further practice.

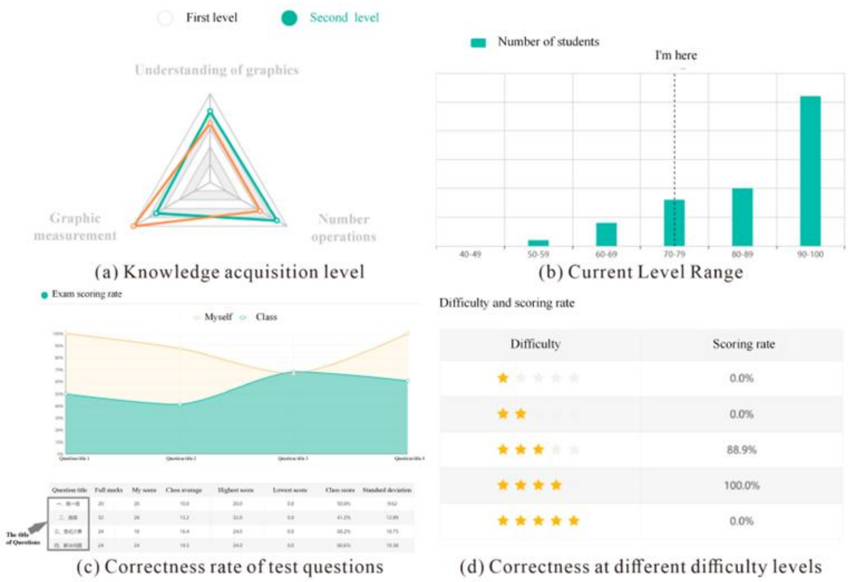

Figure 1. Dashboard Showing Student Performance, Progress, and Current Learning Status

Feedback Adoption and Practice Regulation

Dashboard use was also related to changes in feedback use and practice behavior. The feedback adoption rate was 76.2% in the dashboard group and 55.4% in the control group ("$p<0.001$"). Repeated task errors decreased by 48.1% among dashboard users, compared with 30.7% in the control group. Students in the dashboard group completed more practice tasks before the deadline and had fewer long gaps between sessions. Their self-management score increased by 0.61 standard deviations, while the control group increased by 0.22 standard deviations. Progress visibility was positively related to feedback adoption ("$\beta=0.41,p<0.001$"), and feedback adoption was positively related to learning gain ("$\beta=0.36,p<0.001$"). The indirect effect was 0.15. This result shows that progress information was useful mainly when students used teacher comments to correct their work [27,28]. Rüdian et al. (2025) found that learning analytics did not always improve learning outcomes. Better results were more common when students received clear guidance instead of raw activity data. A similar pattern was found here. Students could connect each weak skill with a teacher comment, a required correction, and a later milestone. The dashboard therefore supported later practice rather than only recording earlier activity.

Learner Patterns and Instructor Intervention

Cluster analysis identified four learning patterns: consistent mastery, delayed recovery, irregular practice, and persistent difficulty. Consistent mastery accounted for 38.9% of the sample and was linked to regular practice, high feedback adoption, and steady milestone completion. Delayed recovery included 24.8% of students who made limited progress in the early weeks but improved after extra practice and teacher support. Irregular practice accounted for 21.6%, while persistent difficulty accounted for 14.7%. The share of students in the persistent-difficulty group was lower in the dashboard group than in the control group (9.8% vs. 19.6%). Teacher identification of students with continuing skill gaps also improved. The sensitivity of intervention decisions increased from 0.71 under standard score reports to 0.86 with dashboard support. The median detection time fell from four practical sessions to two sessions. **Figure 2** shows the teacher-facing view used to identify at-risk learners and review individual activity before intervention. Qi et al. (2026) found that TEADASH helped teachers review student progress and use learning data in teaching decisions. Their study covered five courses at two Swedish universities. The current study used a larger sample and linked dashboard use to measured intervention accuracy [29,30]. The findings also show that low scores alone were not enough for teacher decisions. Better decisions were made when teachers considered milestone delay, repeated errors, feedback use, and practice regularity together.

Figure 2. Teacher Dashboard for Identifying at-Risk Students and Reviewing Individual Learning Records

Model Results and Practical Implications

The final path model showed acceptable fit ("CFI=0.955,TLI=0.946,RMSEA=0.039andSRMR=0.034S"). Progress visibility was positively related to feedback adoption (" β =0.41,p<0.001") and self-management (" β =0.30,p<0.001"). Feedback adoption had the strongest direct link with learning gain (" β =0.36,p<0.001"), followed by self-management ("β=0.24,p=0.003"). Teacher intervention was also related to milestone completion ("β=0.23,p=0.005"). In contrast, dashboard visit frequency had no significant direct effect on final performance after practice behavior and feedback use were included in the model ("β=0.08,p=0.14"). This result shows that frequent dashboard access alone did not improve learning. The main benefit came from using the information to revise work and adjust practice. The mean usability score was 81.3 for students and 78.9 for instructors. Progress and milestone views received the highest ratings, while crowded comparison displays received lower scores. These findings support the use of simple dashboard views that show the current skill gap and the next required task. The study covered only one semester, and the courses used different practical rubrics. Future research should examine long-term skill retention, skill transfer, and dashboard use across several semesters.

CONCLUSION

The results indicate that competency dashboards can improve progress tracking, feedback use, self-management, and teaching decisions in hands-on learning. Students in the dashboard group achieved higher skill scores, completed more milestones, used more teacher feedback, and reduced repeated errors faster than those in the control group. However, viewing progress data alone did not lead to better performance. Learning gains were mainly linked to task revision, added practice, and better time planning after dashboard use. Teachers also identified skill gaps earlier when milestone delays, error records, and feedback use were presented together. The main contribution of this study is the use of continuous skill records instead of a single final grade. By combining task results, practice logs, teacher feedback, self-ratings, and test scores, the dashboard provided a clearer record of skill growth over time. This method may be useful in vocational education, laboratory courses, health training, engineering practice, and other task-based programs. The study was limited to one semester, and the courses used different skill rubrics. Differences in teacher experience and student use may also have affected the results. Further research should test long-term skill retention, skill transfer, and dashboard use across institutions and longer study periods.

REFERENCES

Qi, C., & Qiao, X. (2026, May). From Traditional Machine Learning to Large Language Models: The Evolution of Infrastructure from an Engineering Perspective. In 2026 6th International Symposium on Computer Technology and Information Science (ISCTIS) (pp. 503-506). IEEE.

Ober, T., Williams, K., & Courey, K. (2026). Considerations for Advancing Research on Competency-Based Assessment in K-12 Education. Competency-Based Education Research Journal, 3(6).

Lin, T., Chen, A., & Spivack, Z. (2026). Research on Collaborative Piano Teaching Models and Teaching Resource Efficiency Optimization in Public Art Education Systems. Available at SSRN 6284818.

Ghader Azad, H. (2025). Enhancing Student Engagement through Learning Analytics: A User-Centered Dashboard for LMS Platforms (Doctoral dissertation, Polytechnique Montréal).

Shao, W. (2026, March). Interpretable Ensemble Learning for Network Traffic Anomaly Detection: A SHAP-based Explainable AI Framework for Embedded Systems Security. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-4). IEEE.

McGuire, M. (2026). Feedback, reflection and psychological safety: rethinking assessment for student well-being in higher education. Assessment & Evaluation in Higher Education, 51(3), 532-556.

Huang, J., Xu, T., Yin, J., & Yang, J. (2026). Data Quality Monitoring and Intelligent Scheduling Optimization in Financial Data Workflow Automation. Available at SSRN 7179338.

Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.

Sultana, R. (2025). Artificial intelligence in data visualization: Reviewing dashboard design and interactive analytics for enterprise decision-making. International Journal of Business and Economics Insights, 5(3), 01-29.

Hong, Z., Chen, H., & Xu, T. (2026). Real-Time Capacity Risk Forecasting for Low-Latency Financial Trading Cloud Platforms. Available at SSRN 7118180.

Li, J., Ma, R., & Gu, Y. (2026). From Audit Requirements to Computable Software Architecture: A Rule-Engine and Evidence-Indexing Method for Trusted Digital Infrastructure.

Goh, T. T. (2026). Learning management system log analytics: the role of persistence and consistency of engagement behaviour on academic success. Journal of Computers in Education, 13(1), 283-306.

Wu, J., Wu, D., Zhang, J., & Peng, Y. (2026). Generative AI Feedback in Junior High School Artistic Creation Evaluation. Available at SSRN 7244838.

Shao, W. (2026). Design and Implementation of an Open-Source Security Framework for Cloud Infrastructure. arXiv preprint arXiv:2604.03331.

Brandmo, C., & Gamlem, S. M. (2025, May). Students' perceptions and outcome of teacher feedback: a systematic review. In Frontiers in education (Vol. 10, p. 1572950). Frontiers Media SA.

Xiong, W., & Guo, Y. (2026). A Study on the Phenomenon of Abnormal Load Concentration and Its Cross-system Transmission Characteristics in High-density Building Operations. Available at SSRN 7120661.

Wu, J., Zhang, J., Wu, D., & Peng, Y. (2026). Visual Transformation Mechanisms for Integrating Digital Cultural Heritage Resources into Junior High School Art Curricula.

Abouelenein, Y. A. M., Selim, S. A. S., & Aldosemani, T. I. (2025). Impact of an adaptive environment based on learning analytics on pre-service science teacher behavior and self-regulation. Smart Learning Environments, 12(1), 8.

Huang, J., Yin, J., Yang, J., & Xu, T. (2026). Construction of Audit Evidence Chain and SOX Control Verification Mechanism for Transaction-level Financial Data in High Volume E-commerce Platform. Available at SSRN 7179259.

Zhang, Z., Wang, J., Li, Z., Wang, Y., & Zheng, J. (2025). Anncoder: A mti-agent-based code generation and optimization model. Symmetry, 17(7), 1087.

Echeverria, V., Nieto, G. F., Zhao, L., Palominos, E., Srivastava, N., Gašević, D., ... & Martinez‐Maldonado, R. (2025). A learning analytics dashboard to support students' reflection on collaboration. Journal of Computer Assisted Learning, 41(1), e13088.

Lin, T., Spivack, Z., & Chen, A. (2026). Monitoring Teaching Quality and Validating Educational Equity in Public Art Education. Available at SSRN 6512778.

Zhang, G., Xu, Y., Zhang, M., Wang, S., & Lin, N. (2021). The brain network in support of social semantic accumulation. Social cognitive and affective neuroscience, 16(4), 393-405.

Leibur, T., & Saks, K. (2025, October). Leveraging learning analytics to support teachers' professional development: insights from a digital application. In Frontiers in Education (Vol. 10, p. 1639217). Frontiers Media SA.

Yin, J., Rao, H., & Huang, Y. (2026, March). Dynamic Modeling and Heterogeneity Analysis of Platform User Behavior Time Series. In 2026 International Conference on AI in Education Technology and Applications (AIETA) (pp. 30-33). IEEE.

Andargie, A., Amogne, D., & Tefera, E. (2025). Effects of project-based learning on EFL learners' writing performance. PloS one, 20(1), e0317518.

Zhang, Z., Gao, Y., & Tong, Y. (2026). CausalML4CX: A Causal Inference and Machine Learning Framework for Customer Experience Optimization with Application in Banking.

Rüdian, S., Podelo, J., Kužílek, J., & Pinkwart, N. (2025, March). Feedback on feedback: student's perceptions for feedback from teachers and few-shot LLMs. In Proceedings of the 15th international learning analytics and knowledge conference (pp. 82-92).

Meier, F. L., Jäger, L., Senn, O., Markun, S., & Burgstaller, J. M. (2025). Effectiveness of interactive dashboards as audit and feedback tools in primary care: A systematic review. Plos one, 20(6), e0327350.

Qi, C., & Qiao, X. (2026). Using AI to Monitor AI: Automated Operations Through Log-Driven Intelligence. Available at SSRN 6795840.

ETHICAL DECLARATION

Conflict of interest: No declaration required. **Financing:** No reporting required. **Peer review:** Double anonymous peer review.