

AI-Supported Conditional Progression in Gamified Higher Education: Learning, Motivation and Cognitive Sustainability

Mariarosaria Speranza

Università degli Studi Suor Orsola Benincasa, Naples, Italy

mariarosaria.speranza@studenti.unisob.na.it

Abstract: This paper presents an experimental study on gamified learning in higher education, comparing three learning conditions: traditional individual study, traditional gamification, and gamification based on AI-supported conditional progression. The study involved 142 university students, randomly assigned to the three conditions, who attended the same lecture, delivered by the course lecturer, and then worked on comparable learning content during a differentiated consolidation phase. The research adopts a cognitive ergonomics perspective, which looks beyond final performance and considers the learning experience as a whole, focusing on learning outcomes, motivation, perceived workload, and the cognitive sustainability of the path. In the control condition, students studied the materials individually, without gamified elements. In the traditional gamification condition, students moved through digital activities on a learning platform in a fluid way, without rigid constraints. In the conditional progression condition, students followed a structured digital path organised into levels, with immediate feedback, time limits, a maximum number of attempts, and advancement constraints linking each step to the previous one. Artificial intelligence functions supported the generation and organisation of the learning content, activities and feedback, always under human supervision: the researcher reviewed and validated the materials before their use in the classroom. No significant differences emerged among the three conditions in immediate learning (correct answers, $p = .922$; total score, $p = .943$) or overall perceived workload ($p = .454$), although the two gamified configurations produced different experiences. Traditional gamification showed higher perceived relevance ($p < .05$) and greater completion of required lessons than AI-supported conditional progression ($M = 3.79$ vs. 2.59 ; $p < .001$; $d = 1.79$). The latter offered more structure and guidance but shifted the focus of perceived demand toward managing progression rules, attempts and time. These results suggest that a more structured and technologically articulated design is not automatically a more effective or more sustainable one. The paper discusses implications for designing engaging, pedagogically coherent and cognitively sustainable AI-supported learning experiences, while future research should examine delayed learning and transfer and test less restrictive progression constraints across educational contexts.

Keywords: Gamification, Artificial intelligence, Higher education, Conditional progression, Motivation, Cognitive sustainability

1. Introduction

Artificial intelligence is changing how university teaching is designed. Digital platforms, instructional design tools, immediate feedback systems and gamified environments are transforming how content is organised and experienced by students. AI can help build learning paths that divide materials into progressive units, link activities to levels of difficulty, provide feedback and make progression visible. Whether this translates into educational value depends on coherent and supervised instructional design. Recent research on AI in higher education asks for a shift from technological promise to empirical verification. AI can support personalisation, assessment, feedback and the organisation of learning paths, but situated evidence is needed about what actually happens in real classrooms (Crompton and Burke, 2023; Kasneci et al, 2023). Its educational value also depends on the cognitive and operational conditions that allow students to use these systems well. Because AI in education can mean different things, studies should state clearly what the technology does inside the learning design. In this study, AI supported the design of materials, levels and feedback within a supervised gamified path. The theoretical frame is cognitive load theory: learning depends on the relationship between available mental resources and task demands (Sweller, 1988; Paas, Renkl and Sweller, 2003; van Merriënboer and Sweller, 2005). In a digital environment, students deal not only with content but also with interfaces, steps, rules, time constraints and feedback. Extraneous load can arise from interfaces, navigation and task organisation (Skulmowski and Xu, 2022), so technology-enhanced learning requires cognitively manageable structures (Speranza, Collina and Gamboz, 2025). In this frame, cognitive sustainability is the balance between the effort required, the clarity of the path and the real possibility of completing the activities. An environment rich in activities and feedback can still fail if students spend more energy managing the system than engaging with the content. Gamification is one way to organise this experience. Defined as the use of game design elements in non-game contexts (Deterding et al, 2011), it can support progression and engagement through goals, levels, tasks and feedback. Reviews and meta-analyses report promising but context-dependent effects, showing that educational outcomes depend on design, context and the indicators used to evaluate them (Hamari, Koivisto and Sarsa, 2014; Dicheva et al, 2015; Bai, Hew and Huang, 2020; Sailer and Homner, 2020). Gamification also differs in how much structure it imposes. It may add quiz-like interaction or organise the whole process through

tasks, progression rules and feedback. Conditional progression belongs to this second form. It gives the teacher more control over the sequence, but students must understand how to advance, manage attempts, interpret feedback and finish the required lessons in the time available. Motivation matters because gamified designs are usually justified on motivational grounds. The ARCS model describes motivation to learn through attention, relevance, confidence and satisfaction (Keller, 1987). These dimensions make it possible to test, rather than assume, whether a more structured and feedback-rich configuration is experienced as more interesting, useful, manageable and rewarding.

Previous research offers promising but heterogeneous evidence: AI can support personalisation and feedback, but its educational value requires empirical evaluation in authentic settings (Crompton and Burke, 2023; Kasneci et al, 2023), while gamification effects depend strongly on design and context (Hamari, Koivisto and Sarsa, 2014; Sailer and Homner, 2020). Cognitive load research also warns that digital task organisation can add extraneous demands (Skulmowski and Xu, 2022). However, evidence directly comparing fluid gamification with AI-supported conditional progression under comparable platform and content conditions remains limited.

The added value of the study lies in moving beyond a general comparison between gamified and non-gamified learning. By comparing two gamified configurations implemented on the same platform and based on comparable content, the study examines how conditional progression shapes the learning experience when AI supports the design of materials, levels and feedback under human supervision. The combined analysis of immediate learning, motivation, perceived workload and platform use provides an empirical basis for evaluating both the effectiveness and the cognitive sustainability of the learning path.

2. Aim and Research Questions

The aim of the study is to examine whether a gamified condition with AI-supported conditional progression produces differences compared with traditional gamification and non-gamified study. The comparison covers four indicators: immediate learning, motivation, perceived workload and platform use. Taken together, these indicators are used to assess the cognitive sustainability of the learning experience.

RQ1. Do the three learning conditions differ in immediate learning, measured through the number of correct answers and the total score in the Kahoot! quiz?

RQ2. Do the two gamified conditions differ in motivation, measured through the four ARCS dimensions of attention, relevance, confidence and satisfaction?

RQ3. Do the three learning conditions differ in perceived workload, measured through the six dimensions of the NASA-TLX?

RQ4. Do the two gamified conditions differ in completion of required lessons recorded in the SafetyCulture/EdApp logs?

3. Method

3.1 Study design, Participants and Context

The study used an experimental design with random assignment to three learning conditions. It took place in a real class of a Neuropsychology course at a university in Southern Italy. All participants attended the same 40-minute lecture, given by the course lecturer, on neurocognitive disorders with a focus on Alzheimer's disease. They then completed a 45-minute consolidation phase that differed according to the assigned condition. The validated database included 142 participants: 41 in the control condition, 49 in traditional gamification (GAM1) and 52 in AI-supported conditional progression (GAM2). Mean age was 20.82 years in the control group, 20.38 in GAM1 and 19.80 in GAM2, with no significant differences, $F(2, 139) = 1.252$. The sample was mostly female (92.1%), which reflects the composition of the host degree programmes in the psychological area. Participation was voluntary. Students received one additional point in the Neuroscience examination solely for taking part in the study, regardless of their performance or assigned condition. Data were analysed in aggregate form for scientific purposes. Figure 1 summarises the experimental design and outcome measures.

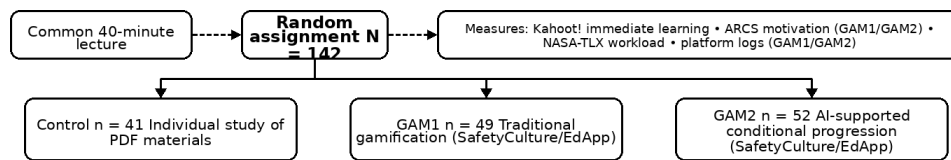


Figure 1: Experimental design and outcome measures

3.2 Experimental Conditions

In the control condition, students studied PDF materials individually, without gamification elements, progression rules or automatic feedback. In GAM1, students worked on the SafetyCulture/EdApp platform in a traditional gamification format: the digital activities were organised on the platform, and students moved through them freely, without rigid constraints. In GAM2, the same platform hosted a path with conditional progression and AI support: levels, progressive activities, advancement constraints, time limits, a maximum number of attempts and immediate feedback. In GAM2, AI was used during the design phase to help transform the source material into structured activities, level-based tasks and immediate feedback. This support made it possible to break the content into smaller units and to link them to progression rules. The researcher reviewed the generated materials, checked the consistency of questions and answers, and validated the final feedback before implementation. The comparison is therefore not between a traditional class and an autonomous intelligent tutor. It is a comparison between three ways of giving students the same content: non-gamified study, traditional gamified consolidation and AI-supported conditional progression.

3.3 Instruments

Immediate learning was assessed with a Kahoot! quiz at the end of the intervention, using two indicators: the number of correct answers, the measure most directly related to content knowledge, and the total score, which also incorporates response speed and was therefore read with more caution. Motivation was measured with a 24-item questionnaire based on the ARCS model (Keller, 1987), covering attention, relevance, confidence and satisfaction on a five-point Likert scale. The questionnaire was given only to the two gamified conditions, in two versions whose items were adapted to the type of gamification experienced. Comparisons between the aggregated subscales are appropriate, because the subscales measure the same higher order constructs; comparisons between single items have exploratory value only. Perceived workload was measured with the NASA Task Load Index (Hart and Staveland, 1988), treated as a multidimensional profile across mental demand, physical demand, temporal demand, perceived performance, effort and frustration. Both the dimension ratings and the pairwise comparison weights were analysed, so that the intensity of the workload could be distinguished from the relative importance that students gave to each dimension. Platform use was analysed through SafetyCulture/EdApp logs, available for the two gamified conditions: completion of required lessons. These logs were treated as behavioural traces to be interpreted through the theoretical frame of the study, rather than as automatic indicators of learning quality (Siemens, 2013; Wise and Shaffer, 2015).

3.4 Data Analysis

Analyses were run on the valid cases available for each measure. Univariate ANOVAs compared the three conditions on the Kahoot! indicators. Mixed ANOVAs with Group as the between-subjects factor were used for the NASA-TLX ratings and weights and for the ARCS subscales, with Greenhouse-Geisser correction where sphericity was violated and planned comparisons to follow up significant interactions. A Welch t-test compared GAM1 and GAM2 on completion of required lessons, with Cohen's d and a 95% confidence interval. The sample was a natural university class, so no a priori power analysis was possible; a posteriori sensitivity analysis showed that the design could detect at least medium effects.

4. Results

4.1 Immediate Learning

Valid cases for the Kahoot! quiz were 139. Out of a maximum of 24, the mean number of correct answers was 16.85 (SD = 4.98) in the control condition, 17.21 (SD = 3.23) in GAM1 and 17.12 (SD = 4.54) in GAM2. Univariate ANOVAs showed no significant differences among the three conditions, neither for correct answers, $F(2, 136) = 0.081$, $p = .922$, nor for total score, $F(2, 136) = 0.059$, $p = .943$. Table 1 summarises these results. The total score was in any case read with caution, because the Kahoot! metric combines correctness with response speed.

Table 1: Immediate learning outcomes measured through Kahoot!

Measure	Control M (SD)	GAM1 M (SD)	GAM2 M (SD)	F(2, 136)	p
Correct answers /24	16.85 (4.98)	17.21 (3.23)	17.12 (4.54)	0.081	.922
Total score	15027.35 (4535.38)	15165.08 (3025.55)	15316.27 (4315.55)	0.059	.943

Note: GAM1 = traditional gamification; GAM2 = AI-supported conditional progression. Valid cases = 139.

4.2 Motivation

Valid cases for the ARCS questionnaire were 47 in GAM1 and 51 in GAM2. A mixed ANOVA with Group (GAM1 versus GAM2) as the between-subjects factor and the four ARCS dimensions as the within-subjects factor, with Greenhouse-Geisser correction, showed a significant main effect of the ARCS dimensions, $F(2.754, 264.410) = 11.625$, $p < .001$, partial eta squared = .108, and a significant ARCS by Group interaction, $F(2.754, 264.410) = 5.726$, $p = .001$, partial eta squared = .056. The main effect of Group was not significant, $F(1, 96) = 0.443$, $p = .507$: overall motivation did not differ significantly between the two configurations. Planned comparisons located the difference. Attention, $t(97) = 0.377$, $p = .707$, confidence, $t(96) = -1.093$, $p = .277$, and satisfaction, $t(97) = -0.440$, $p = .661$, did not differ between conditions. Relevance, however, was significantly higher in GAM1 than in GAM2, $t(96) = 2.193$, $p < .05$, mean difference = 0.236, 95% CI [0.022, 0.450]. Table 2 reports the descriptive values. An exploratory look at single items suggested a composite picture: students in the conditional progression condition rated the personalised feedback more positively, while students in traditional gamification gave more favourable ratings on items related to interest, clarity, attention and emotional involvement. Because the item wording was adapted to each configuration, these item-level indications describe the quality of the two experiences, not a general motivational difference.

Table 2: ARCS subscales in the two gamified conditions

ARCS subscale	GAM1 M (SD)	GAM2 M (SD)	Planned comparison	p
Attention	3.79 (0.43)	3.66 (0.68)	$t(97) = 0.377$	.707
Relevance	4.09 (0.44)	3.85 (0.60)	$t(96) = 2.193$	< .05
Confidence	3.68 (0.54)	3.81 (0.62)	$t(96) = -1.093$	.277
Satisfaction	3.79 (0.46)	3.77 (0.60)	$t(97) = -0.440$	.661

Note: n (GAM1) = 47; n (GAM2) = 51. Subscales are computed on a 1-5 Likert scale. The ARCS by Group interaction was significant, $p = .001$; the Group main effect was not, $p = .507$.

4.3 Perceived Workload

The NASA-TLX ratings were analysed with a mixed ANOVA with Group as the between-subjects factor and the six dimensions as the within-subjects factor, with Greenhouse-Geisser correction. The main effect of the dimensions was significant, $F(4.018, 554.531) = 74.747$, $p < .001$, partial eta squared = .351: the dimensions were rated differently, with mental demand and effort highest in all groups. The dimension by Group interaction was not significant, $F(8.037, 554.531) = 0.468$, $p = .879$, and neither was the main effect of Group, $F(2, 138) = 0.794$, $p = .454$. Table 3 reports the descriptive values. In short, GAM2 was more structured and more constrained, but students did not experience it as globally heavier or more frustrating.

Table 3: NASA-TLX dimension ratings by learning condition

Dimension	Control M (SD)	GAM1 M (SD)	GAM2 M (SD)
Mental demand	7.28 (2.25)	7.78 (1.48)	7.84 (1.71)
Physical demand	2.42 (2.55)	2.98 (2.68)	2.90 (2.57)
Temporal demand	5.53 (3.24)	6.24 (2.52)	6.17 (2.60)
Perceived performance	5.40 (2.52)	5.98 (2.37)	5.63 (2.70)
Effort	7.00 (1.78)	6.59 (1.52)	6.69 (2.35)
Frustration	5.00 (3.38)	5.10 (2.70)	5.25 (3.18)

Note: higher scores indicate a higher perceived intensity of the dimension. No between-group effect or dimension-by-group interaction was significant.

The pairwise comparison weights of the NASA-TLX told a complementary story. They were computed for the 127 participants with complete pairwise data (39 control, 47 GAM1, 41 GAM2). A mixed ANOVA on the weights showed a significant dimension by Group interaction, $F(10, 620) = 2.30, p = .012$, partial eta squared = .036: the profile of what students considered most demanding varied across conditions. Between-group comparisons were significant for mental demand, $H(2) = 8.71, p = .013$: students in GAM2 gave significantly less weight to mental demand than students in the control condition and in GAM1 (both $p_{\text{Holm}} = .030$). Temporal pressure had the highest descriptive weight in GAM2, although the between-group comparison was not significant ($p = .198$). Overall workload intensity did not differ significantly across conditions, but its focus moved: in the conditional progression condition, what felt demanding was less the difficulty of the content and more the management of the structure and of the time of the path.

4.4 Platform use

Platform logs were available for 48 students in GAM1 and 46 in GAM2 and were read as behavioural indicators of how traversable and cognitively sustainable the two digital configurations were, not as direct measures of learning. The clearest difference concerned completed required lessons: within the same 45-minute activity period, students in GAM1 completed on average 3.79 of four required lessons, whereas students in GAM2 completed 2.59 of four required lessons. This difference was significant and strongly in favour of traditional gamification, $t(65.8) = 8.74, p < .001$, with a very large effect size, $d = 1.79$, 95% CI [1.32, 2.28]. It represents the strongest behavioural evidence from the platform data.

Table 4: Completion of required lessons in the two gamified conditions

Condition	n	Required lessons completed /4	Completion rate
GAM1 (traditional gamification)	48	3.79	94.8%
GAM2 (AI-supported conditional progression)	46	2.59	64.8%

Note: Welch $t(65.8) = 8.74, p < .001, d = 1.79$, 95% CI [1.32, 2.28]. Maximum = 4 required lessons within a 45-minute activity period.

Traditional gamification was more fluid in completion. Because both gamified conditions used the same platform and comparable content, the lower completion in GAM2 suggests that progression rules and constraints may have introduced operational friction.

5. Discussion

The AI-supported conditional progression condition did not improve immediate learning compared with traditional gamification or non-gamified study, and it did not increase the overall intensity of perceived workload. The differences appeared elsewhere: in the perceived relevance of the activity, in the focus of the workload and, most clearly, in the behavioural traces, where students under conditional progression completed fewer required lessons within a more constrained path. The more structured GAM2 path changed how students perceived and traversed the environment without improving immediate learning, reinforcing that AI-supported design should be evaluated against specific educational goals rather than technological complexity alone. The absence of learning differences should also be read in the light of the common lecture, which gave all groups a solid shared base and may have levelled the outcomes of the consolidation phase. The motivational results add a useful nuance. The significant ARCS by Group interaction shows that the two gamified configurations do not produce the same motivational profile. Attention, confidence and satisfaction did not differ significantly between conditions, while relevance was higher in traditional gamification. Students seemed to see the more fluid configuration as more directly connected to the goals of the course, perhaps because an unconstrained path makes the link between activities and content easier to see. A strongly structured path, even when rich in levels, feedback and progression, can make the educational meaning of the experience less evident (Keller, 1987). At the same time, the item-level indications suggest that students appreciated the personalised feedback of the conditional configuration. The two forms of gamification do not produce more or less of the same experience; they produce different experiences. The workload results point in the same direction. Overall workload intensity did not differ significantly across conditions, but the weights showed that students in the conditional progression condition gave less weight to mental demand, while temporal pressure became the most salient dimension of their profile. Together with the completion data, this suggests that the added structure did not make the content feel harder. It moved the attention of the students from the complexity of the material to the management of rules, steps and time. Some effects of a design do not show up as higher

self-reported workload; they show up in what students consider demanding and in how far they get. This is the core of cognitive sustainability as we use it here. Sustainability is not only effectiveness measured at the end of the activity; it is the relationship among structure, workload, motivation and completion. When a more articulated configuration does not improve learning, weakens the perceived relevance of the activity and reduces completion, the design lesson is clear: instructional guidance should support orientation without becoming operational friction. Educational value does not depend on how much technology is introduced, but on the quality of the relationship among content, platform, feedback, constraints, workload and human supervision.

These findings can also be positioned against previous research. The absence of an immediate learning advantage is consistent with reviews and meta-analyses showing that the effects of gamification are heterogeneous and strongly dependent on design, context and outcome measures (Hamari, Koivisto and Sarsa, 2014; Dicheva et al, 2015; Bai, Hew and Huang, 2020; Sailer and Homner, 2020). The shift in the workload profile in GAM2 is also compatible with evidence that digital interfaces and task organisation can generate extraneous demands even when content difficulty is unchanged (Skulmowski and Xu, 2022). From an ARCS perspective, the higher relevance observed in GAM1 supports Keller's (1987) emphasis on making the connection between activities and learning goals visible. Finally, the results reinforce recent calls to evaluate AI in higher education through specific, supervised educational uses rather than through technological novelty alone (Crompton and Burke, 2023; Kasneci et al, 2023). In our interpretation, the main difference between GAM1 and GAM2 is therefore not the presence of technology but the degree to which progression rules preserve clarity, flexibility and traversability.

6. Design Implications

The results suggest five practical indications for gamified and AI-supported learning paths in higher education. First, AI support can make the design more articulated, help build progressive levels and make immediate feedback available, but these possibilities become educationally useful only under human direction that validates content, calibrates constraints and keeps the path readable for students. Second, conditional progression must remain traversable in the time available: the number of levels and steps should be proportionate, the progression rules should be clear before the activity starts, and feedback should tell students what to do next, not only whether an answer is correct. Third, the perceived relevance of the activity deserves explicit attention: a strongly structured path should make the link between each step and the learning goals visible, so that constraints are experienced as guidance rather than as obstacles. Fourth, personalised feedback was the element students appreciated most in the AI-supported configuration and is worth preserving and refining. Finally, evaluation should combine immediate learning, motivation, workload ratings and weights, and platform logs: together they describe the sustainability of an educational design far better than any single indicator.

7. Scope and Limitations

The study was carried out inside an ordinary university teaching activity. This strengthens its ecological value, but it also defines its scope: university students, neuropsychological content, a common lecture and a short consolidation phase, in a mostly female sample from psychological degree programmes. No a priori power analysis was possible, and the sensitivity analysis indicates that small effects cannot be excluded, so non-significant results should be read as prudent indications, not as proof of equivalence. Learning was assessed immediately after the intervention; an exploratory unannounced test given about one month later to the students present that day showed no group differences, but it suffers from attendance bias and cannot be read as a controlled measure of retention. The ARCS questionnaire was given only to the gamified conditions, in versions adapted to each configuration, so item-level comparisons remain exploratory, and platform logs were available only for the gamified conditions. GAM2 should be understood as an integrated instructional configuration in which AI-supported design, conditional progression, time limits, attempt rules and personalised feedback operated together, so the findings reflect the functioning of the configuration as a whole. The findings apply most directly to short university consolidation activities in which the same content is offered through different digital configurations, without replacing the teacher's role.

8. Conclusion

This study compared a gamified configuration with AI-supported conditional progression against traditional gamification and a non-gamified condition, on four indicators: immediate learning, motivation, perceived workload and platform use. No significant differences emerged among conditions in immediate learning or

overall workload intensity. Motivation showed a more articulated profile, with relevance significantly higher in traditional gamification ($M = 4.09$ vs. 3.85 , $p < .05$), and the workload weights showed that the conditional progression condition moved the focus of the perceived demand from the content to the management of structure and time. The clearest difference was behavioural: traditional gamification supported greater completion of required lessons than AI-supported conditional progression ($M = 3.79$ vs. 2.59 , $p < .001$, $d = 1.79$). The contribution of the study is an empirical reading of AI-supported gamified learning centred on the cognitive sustainability of the experience. A more guided structure can offer design control and personalised feedback that students value, but it can also make the meaning of the activity less visible and the path harder to complete. The practical lesson is not to use less AI, but to give it a more precise pedagogical role: its success should be judged by what students learn, by how they perceive the relevance and the demands of the path, and by whether they can complete it in a meaningful way. An advantage of the AI-supported configuration is the possibility of structuring progression and personalised feedback under human supervision; its practical use, however, requires constraints to be calibrated to the time available and to learners' ability to move through the path. These findings can inform the design of short university consolidation activities and, with appropriate validation, other digital training settings in which AI supports materials and feedback without replacing pedagogical control.

Acknowledgements

The author wishes to thank Professors Nadia Gamboz and Simona Collina for their valuable supervision, support and constructive feedback throughout the doctoral research.

Ethics Declaration: The study was conducted within an ordinary university teaching activity. Participation was voluntary and all participants signed an informed consent form together with an agreement covering the processing of personal data, privacy and the ethical aspects of the research. Data were processed anonymously and analysed in aggregate form for scientific purposes. The additional examination point was awarded solely for participation and was independent of performance and assigned condition. Under these conditions, formal ethical clearance was not required for this research.

AI Declaration: Artificial intelligence tools were used in two distinct ways. Within the study itself, AI functions supported the generation and organisation of learning materials, activities, levels and feedback for the conditional progression condition, under the supervision of the researcher, as described in the Method section. In the preparation of this paper, AI-based language tools were used to support translation from Italian and language editing. The study design, data collection, statistical analyses, interpretation of results and conclusions are entirely the work of the author, who reviewed and verified all content of the final text.

References

- Bai, S., Hew, K.F. and Huang, B. (2020) "Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts", *Educational Research Review*, Vol 30, 100322. <https://doi.org/10.1016/j.edurev.2020.100322>.
- Crompton, H. and Burke, D. (2023) "Artificial intelligence in higher education: the state of the field", *International Journal of Educational Technology in Higher Education*, Vol 20, Article 22. <https://doi.org/10.1186/s41239-023-00392-8>.
- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011) "From game design elements to gamefulness: defining gamification", *Proceedings of the 15th International Academic MindTrek Conference*, Tampere, Finland, pp 9-15. <https://doi.org/10.1145/2181037.2181040>
- Dicheva, D., Dichev, C., Agre, G. and Angelova, G. (2015) "Gamification in education: a systematic mapping study", *Educational Technology & Society*, Vol 18, No. 3, pp 75-88.
- Hamari, J., Koivisto, J. and Sarsa, H. (2014) "Does gamification work? A literature review of empirical studies on gamification", *Proceedings of the 47th Hawaii International Conference on System Sciences*, pp 3025-3034. <https://doi.org/10.1109/HICSS.2014.377>.
- Hart, S.G. and Staveland, L.E. (1988) "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research", in Hancock, P.A. and Meshkati, N. (eds), *Human Mental Workload*, North-Holland, Amsterdam, pp 139-183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Kasneci, E., Sessler, K., Küchemann, S. et al (2023) "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and Individual Differences*, Vol 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Keller, J.M. (1987) "Development and use of the ARCS model of instructional design", *Journal of Instructional Development*, Vol 10, No. 3, pp 2-10. <https://doi.org/10.1007/BF02905780>.
- Paas, F., Renkl, A. and Sweller, J. (2003) "Cognitive load theory and instructional design: recent developments", *Educational Psychologist*, Vol 38, No. 1, pp 1-4. https://doi.org/10.1207/S15326985EP3801_1.
- Sailer, M. and Homner, L. (2020) "The gamification of learning: a meta-analysis", *Educational Psychology Review*, Vol 32, pp 77-112. <https://doi.org/10.1007/s10648-019-09498-w>.

- Siemens, G. (2013) "Learning analytics: the emergence of a discipline", *American Behavioral Scientist*, Vol 57, No. 10, pp 1380-1400. <https://doi.org/10.1177/0002764213498851>.
- Skulmowski, A. and Xu, K.M. (2022) "Understanding cognitive load in digital and online learning: a new perspective on extraneous cognitive load", *Educational Psychology Review*, Vol 34, pp 171-196. <https://doi.org/10.1007/s10648-021-09624-7>.
- Speranza, M., Collina, S. and Gamboz, N. (2025) "Innovative knowledge practices in technology-enhanced learning", *DigitCult*, Vol 10, No. 1, pp 47-58. <https://doi.org/10.36158/97912566920713>.
- Sweller, J. (1988) "Cognitive load during problem solving: effects on learning", *Cognitive Science*, Vol 12, No. 2, pp 257-285. https://doi.org/10.1207/s15516709cog1202_4
- van Merriënboer, J.J.G. and Sweller, J. (2005) "Cognitive load theory and complex learning: recent developments and future directions", *Educational Psychology Review*, Vol 17, pp 147-177. <https://doi.org/10.1007/s10648-005-3951-0>.
- Wise, A.F. and Shaffer, D.W. (2015) "Why theory matters more than ever in the age of big data", *Journal of Learning Analytics*, Vol 2, No. 2, pp 5-13. <https://doi.org/10.18608/jla.2015.22.2>.