Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From Design to Use: A Focus on Engagement with Optional Learning Resources

Amedeo Viccari

amedeo.viccari@uni-graz.at

University of Graz

Research Article

Keywords: Game-Based Learning, Learner Engagement, Knowledge Acquisition, Educational Technology, Iterative Redesign, Video Games

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11664749/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

From Design to Use: A Focus on Engagement with Optional Learning Resources

Abstract

Optional instructional resources can enrich interactive learning environments, but their availability does not ensure that learners will use them or benefit from them. *Limes*, a digital environment for learning Roman history, includes an optional Codex with supplementary historical information. Based on findings from prior studies, the *Limes* Master Edition was developed through an evidence-informed, iterative redesign of the Codex to encourage engagement. In a preregistered, nonrandomized cross-study comparison pooling three *Limes* data collections (N = 200), we examined Codex engagement and posttest knowledge across Version (Original vs. Master Edition) and Format (Game vs. Text). Preregistered 2 × 2 ANCOVAs showed no overall higher level of Codex engagement in the Master Edition. For clicks, however, the Version × Format interaction was significant: Text participants clicked more than Game participants in both versions, but this difference was smaller in the Master Edition, primarily because Master Text participants clicked less than Original Text participants. The interaction for Codex time was not significant. Adjusted posttest knowledge was lower in the Master Edition sample. Within the Master Edition, greater Codex time and more clicks were associated with higher posttest knowledge, replicating the pattern observed in the earlier studies. Exploratory analyses further tested whether accessing individual Codex entries was associated with correctly answering corresponding knowledge questions; nine of 29 predefined associations remained significant after multiple-comparison correction. These findings show that redesigning optional instructional resources may not necessarily increase their use and underscore the importance of examining learners' actual engagement with them.

Keywords: Game-Based Learning; Learner Engagement; Knowledge Acquisition; Educational Technology; Iterative Redesign, Video Games

From Design to Use: What Engagement with Optional Learning Resources Reveals About Learning

Digital learning environments can provide learners with rich instructional content while allowing them to decide when, how, and whether to engage with particular resources. In game-based learning environments, instructional information often coexists with narrative, exploration, goals, feedback, and other interactive features (Wouters et al., 2013). Consequently, educational value may depend not only on whether relevant information is available, but also on whether learners actually use it (Schumacher et al., 2026; von Keyserlingk et al., 2026). This perspective is consistent with research emphasizing specific design features rather than treating the medium itself as the active ingredient. Clark et al. (2016), for example, found learning benefits not only for games compared with nongame conditions but also for augmented game designs compared with standard versions, highlighting the role of design beyond medium. Similarly, Plass et al. (2015) conceptualized game-based learning as a designed experience in which cognitive, behavioral, affective, and sociocultural engagement is shaped by particular game features.

This issue is closely related to learner control. Learner-controlled environments allow learners to select and organize instructional content (Scheiter, 2021) but such control also places cognitive and metacognitive demands on them. Scheiter (2021) argues that learner control is most effective when learners have sufficient prior knowledge and receive support for orienting and regulating their learning. Consistent with this view, Karich et al. (2014) found little overall benefit of learner control and noted that learners may skip important instructional material. Thus, making information optional creates a central design challenge: availability alone does not ensure meaningful engagement.

One response to this challenge is to guide learners' attention toward relevant information. The FoRe-Squares model proposes that technology-augmented instruction is more likely to support learning when it provides sufficient focus support and reward for active, goal-directed processing (Eitel et al., 2025). Focus support refers to cognitive and metacognitive support that directs processing toward the most important elements of instruction. Prompts can serve a related function by directing attention or encouraging learners to engage in particular cognitive or metacognitive activities. In a recent systematic review and meta-analysis of 68 experimental studies, Thomann & Deutscher (2025) found that prompts in digital learning environments can support learning achievement, while also showing that their effectiveness varies according to characteristics such as timing, personalization, context, and presentation. Importantly, prompts should therefore not be understood as a universally effective solution. Design features may make relevant information more noticeable or encourage its use, but whether learners subsequently engage with that information remains an empirical question.

Research in game-based learning further illustrates the importance of examining learners' actual interactions with instructional resources. In *Crystal Island*, time spent interacting with books and research articles was positively associated with learning gains, while learners' interactions with instructional materials also differed depending on whether the information was relevant to the assessed content (Dever et al., 2020). Dever et al. (2022) extended this process-oriented perspective by examining sequences of interactions with instructional materials using log-file and eye-tracking data. Together, these studies show how process data can reveal which resources learners interact with, how they engage with them over time, and how these patterns relate to learning.

The present research addresses this problem using *Limes* (Viccari et al., 2024), an interactive environment for learning about Roman history. *Limes* includes a Codex containing optional historical

information on historical figures, places, and events. Previous research consistently indicated that engagement with this resource was associated with knowledge acquisition. First, in a study comparing the videogame *Limes* with a text-based version, additional analyses revealed that greater Codex engagement was associated with higher posttest knowledge, and text participants accessed the Codex more frequently and spent more time with it than game participants (Study 1, Viccari, Kienitz, et al., 2026). Second, in a study which examined the effects of sound in the game-based version of *Limes*, Codex interactions and time spent with Codex content significantly predicted posttest knowledge after accounting for relevant learner characteristics (Study 2, Viccari, Lindner, et al., 2026). These findings motivated an evidence-informed redesign of *Limes*, resulting in the Master Edition (Study 3), evaluated in the present work. Because Codex engagement was associated with higher knowledge scores but was lower in the game-based format, the redesign aimed to encourage learners to access optional historical information more frequently, particularly during gameplay. Hereafter, Studies 1–3 refer to the three studies, whereas Original version (used in Studies 1 and 2) and Master Edition (used in Study 3) refer to the respective versions of *Limes*. In the game-based Master Edition, the Codex was made more prominent by adding an animated notification and end-of-level prompts encouraging learners to revisit its content. These additional guidance features were specific to the game-based format and were not introduced in the text-based format. Logging was also extended from aggregate measures to entry-level records (Viccari, 2026).

The Present Study

The present study tested whether participants playing the Master Edition had higher Codex engagement and posttest knowledge than those playing the original game. We pooled eligible participants from *Studies 1–3*: Studies 1 and 2 used the Original version, whereas Study 3 used the Master Edition. This yielded a 2 × 2 design with Version (Original vs. Master Edition) and Format (Game vs. Text) as between-subjects factors. Importantly, Version was not randomly assigned, as the Original version and Master Edition were used in separate data collections. Therefore, the comparison between versions represents a cross-study comparison rather than a randomized experimental manipulation of the redesign. Hypotheses, measures, covariates, exclusion rules, and statistical analyses were preregistered (preregistration: https://aspredicted.org/3ny54j.pdf).

Hypothesis 1 (H1). We hypothesized that participants using the Master Edition would engage more with Codex content than participants using the Original version. The Master Edition was designed to increase the visibility of Codex content and guide learners' attention toward optional explanatory information, consistent with the focus-support perspective (Eitel et al., 2025).

Hypothesis 2 (H2). We expected Version to moderate the effect of Format on Codex engagement, such that the difference between Game and Text participants would be smaller in the Master Edition. This prediction was motivated by *Study 1*, in which Text participants engaged more extensively with the Codex than Game participants. The game-based Master Edition was designed in part to address this difference through greater Codex visibility and prompts.

Hypothesis 3 (H3). Within the Master Edition, we hypothesized that greater Codex engagement would be associated with higher posttest knowledge, replicating the findings of *Studies 1 and 2* to the redesigned environment.

Hypothesis 4 (H4). We hypothesized that participants using the Master Edition would achieve higher posttest knowledge scores than participants using the Original version. The redesign was expected to support knowledge acquisition by making explanatory information more visible and

directing learners toward content-rich material, consistent with evidence that engagement with relevant informational materials can support learning (Dever et al., 2020).

In addition to these confirmatory hypotheses, the more detailed entry-level behavioral data available in *Study 3* allowed us to address two preregistered exploratory questions. We examined which Codex entries were accessed most and least frequently and whether accessing specific entries was associated with correctly answering the corresponding knowledge-test items.

Methods

Participants and Design

The pooled dataset comprised $N = 200$ unique university students from *Studies 1–3* (Viccari, Kienitz, et al., 2026; Viccari, Lindner, et al., 2026). Participants were enrolled in a range of university programs, including psychology, teacher education, and sport science. The final sample included 162 women, 36 men, and two nonbinary/diverse participants. Participants were between 18 and 45 years of age ($M = 22.56$, $SD = 4.55$).To ensure comparability across studies, only participants who had received the learning instruction (in Study 1, partipants were either instructed to focus on learning or on having fun) and used *Limes* without sound (in Study 2, participants either engaged with Limes with or without sound) were included. *Study 1* contributed $n = 69$ participants using the *Original version* (36 Game, 33 Text). *Study 2* included 56 participants in the no-sound Game condition. However, 36 of these participants had also taken part in Study 1. To avoid counting them twice, only the remaining 20 participants were added to the pooled sample. *Study 3*, which used the *Master Edition*, contributed $n = 111$ participants (56 Game, 55 Text). Appendix Table A1 summarizes the source studies and construction of the pooled sample.

The pooled data formed a 2 × 2 design with *Version* (Original vs. Master Edition) and *Format* (Game vs. Text). Cell sizes were $n = 56$ for Original–Game, $n = 33$ for Original–Text, $n = 56$ for Master–Game, and $n = 55$ for Master–Text.

Participants were recruited from university populations and generally received course credit. During a limited recruitment period in Study 1, participants could alternatively receive €20. Eligibility criteria included being at least 18 years old, sufficient German proficiency, not being enrolled in a history degree program, and no uncorrected severe visual impairment. All participants provided informed consent.

The preregistered power analysis targeted the main effect of Version on Codex engagement in a 2 × 2 ANCOVA. Previous studies with Limes had identified Codex engagement as a relevant learning behavior, providing a rationale for the redesign. Because the redesign directly targeted this behavior through increased visibility and additional prompts, an effect on Codex engagement was expected. However, no sufficiently comparable estimate of the redesign effect was available. We therefore used a medium effect of $f = 0.25$ as a planning assumption. With $\alpha = .05$ and power = .80, G*Power 3.1 (Faul et al., 2009) indicated a minimum required sample of $N = 128$.

The preregistration reported a maximum available pooled sample of $N = 238$. However, after reviewing the participant records and correcting the sample counts, the final sample consisted of $N = 200$ unique participants. As described above, 36 participants in Study 2 had also taken part in Study 1 and were counted only once. The preliminary estimate of 71 participants from Study 1 mistakenly included two participants from the Text condition who had already been excluded from the original study, leaving 69 eligible participants. The final sample remained above the preregistered minimum,

and the preregistered hypotheses, measures, covariate-selection procedure, and statistical models were otherwise unchanged.

Experimental Manipulation and Materials

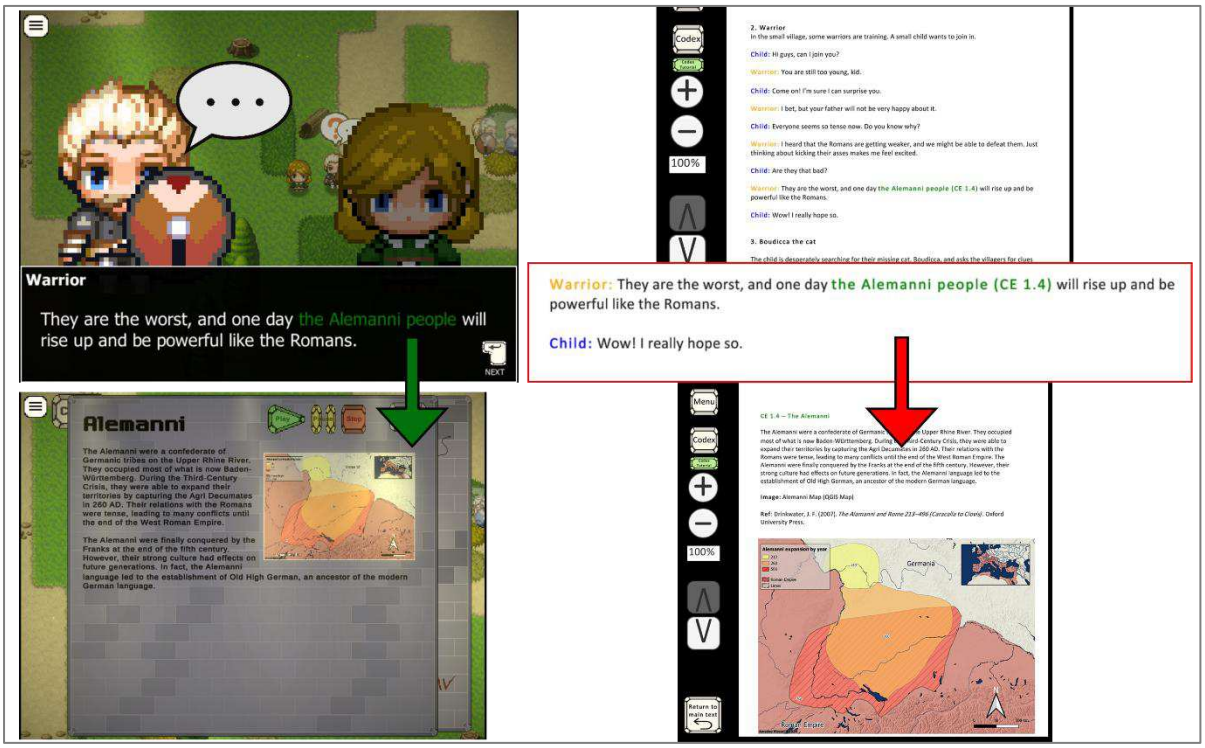
Limes is a digital learning environment designed to teach factual knowledge about Roman history through a historically grounded narrative set along the Roman frontier during the third century CE. The narrative unfolds across five levels containing a main storyline, optional secondary missions, and supplementary Codex entries.

The Codex comprises 36 short entries covering historical figures, events, places, military structures, and everyday life. In the German version, entries averaged 148 words (range = 59–265; 6,328 words in total). Entries were unlocked through interactions with characters or objects and could subsequently be accessed optionally.

Limes was implemented in Game and Text formats. Both presented the same narrative, five-level progression, core learning content, and Codex material, while differing in how participants interacted with the environment (see Figure 1).

Figure 1

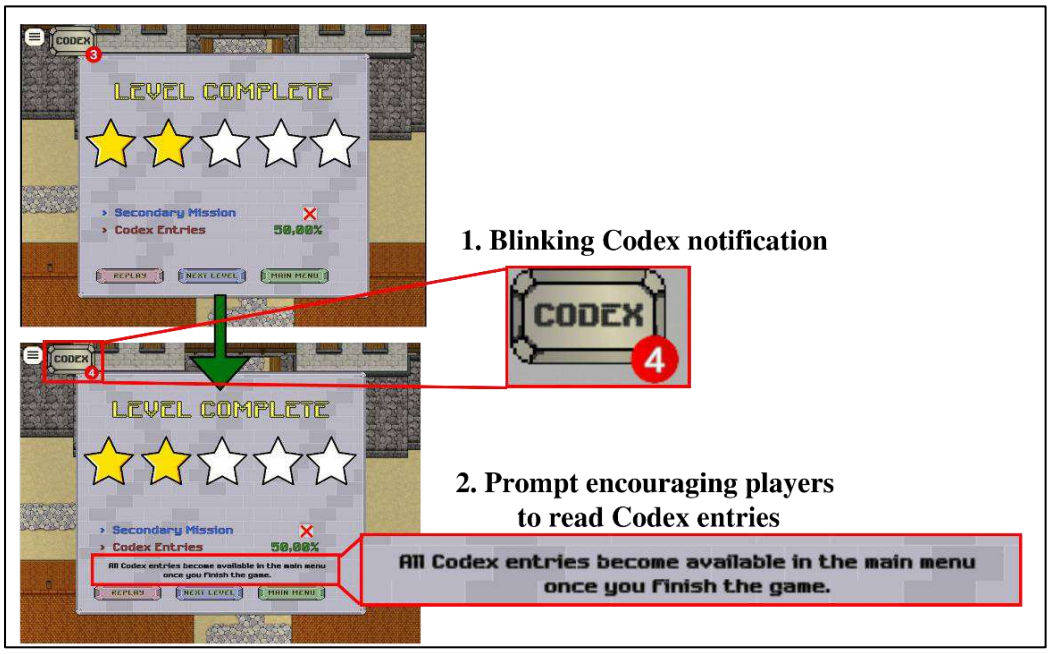
Corresponding narrative and Codex content in the original game-based and text-based versions of Limes.



Note. The left panels show the game-based format and the right panels the text-based format. The upper panels present corresponding narrative content, and the lower panels show the associated Codex entry.

In the Original version, behavioral logging captured aggregate Codex time and clicks. In the game-based Master Edition, two guidance features were added: an animated Codex notification and an end-of-level prompt encouraging participants to revisit unlocked entries (see Figure 2). The Master Edition also extended logging to individual Codex entries.

Figure 2



Codex guidance features introduced in the game-based Limes Master Edition.

Measures

Pre-Test Background Variables and Covariates. Prior factual knowledge of Roman history was assessed at pretest using five items from the knowledge test (score range: 0–5). Prior interest in history was assessed with three items adapted from Schumacher et al. (2025), rated on a 7-point Likert scale from 1 (not at all interested) to 7 (very interested) and averaged into a composite score (α = .73). Self-reported prior historical knowledge was assessed with three items covering knowledge of Roman history, ancient history, and general history, also adapted from Schumacher et al. (2025). Items were rated on a 7-point Likert scale from 1 (no knowledge at all) to 7 (a great deal of knowledge) and averaged into a composite score (α = .80). Participants additionally reported their average weekly gaming hours. These four variables were preregistered as candidate covariates.

Knowledge Test. Factual knowledge of Roman history was assessed after the learning phase using a self-developed 25-item single-choice test covering content presented in *Limes*. Each item had four response alternatives and an "*I do not know*" option; correct responses were scored 1 and all others 0, yielding a total score from 0 to 25. The same test and scoring procedure were used across versions and formats. Descriptive statistics are reported in Table 1.Internal consistency of the 25-item knowledge test was acceptable overall (Cronbach's α = .77), as well as in the Original Game (α = .80), Original Text (α = .71), Master Edition Game (α = .72), and Master Edition Text (α = .78) conditions.

Codex Engagement. Engagement with the optional Codex material was assessed using behavioral log data. Two indicators were available across versions and formats: total Codex time and clicks, with clicks reflecting entry openings. Codex time was recorded in seconds for the statistical analyses. Descriptive statistics are reported in Table 1.

Entry-Level Codex Engagement. In Study 3, the Master Edition additionally recorded which of the 36 Codex entries participants accessed, how long they spent on each entry, and how often they clicked on them. These data were used in the preregistered exploratory analyses to examine which

entries were accessed most often and whether accessing a specific entry was associated with correctly answering the corresponding knowledge-test question.

All measures used in our study are reported in the Online Supplement.

Table 1

Descriptive Statistics for Knowledge Test Performance and Codex Engagement by Version and Format

Measure	Version	Format	*M*	*SD*
Posttest knowledge	Overall	—	13.88	4.41
	Original	Overall	14.40	4.51
	Original	Game	13.84	4.77
	Original	Text	15.36	3.90
	Master Edition	Overall	13.45	4.31
	Master Edition	Game	12.30	3.92
	Master Edition	Text	14.62	4.40
Codex time (min)	Overall	—	12.40	10.50
	Original	Overall	13.09	10.37
	Original	Game	9.92	9.58
	Original	Text	18.45	9.53
	Master Edition	Overall	11.85	10.61
	Master Edition	Game	9.73	8.98
	Master Edition	Text	14.01	11.74
Codex clicks	Overall	—	23.07	16.58
	Original	Overall	23.30	15.40
	Original	Game	15.70	13.21
	Original	Text	36.21	8.92
	Master Edition	Overall	22.87	17.53
	Master Edition	Game	17.71	14.44
	Master Edition	Text	28.13	18.92

Procedure

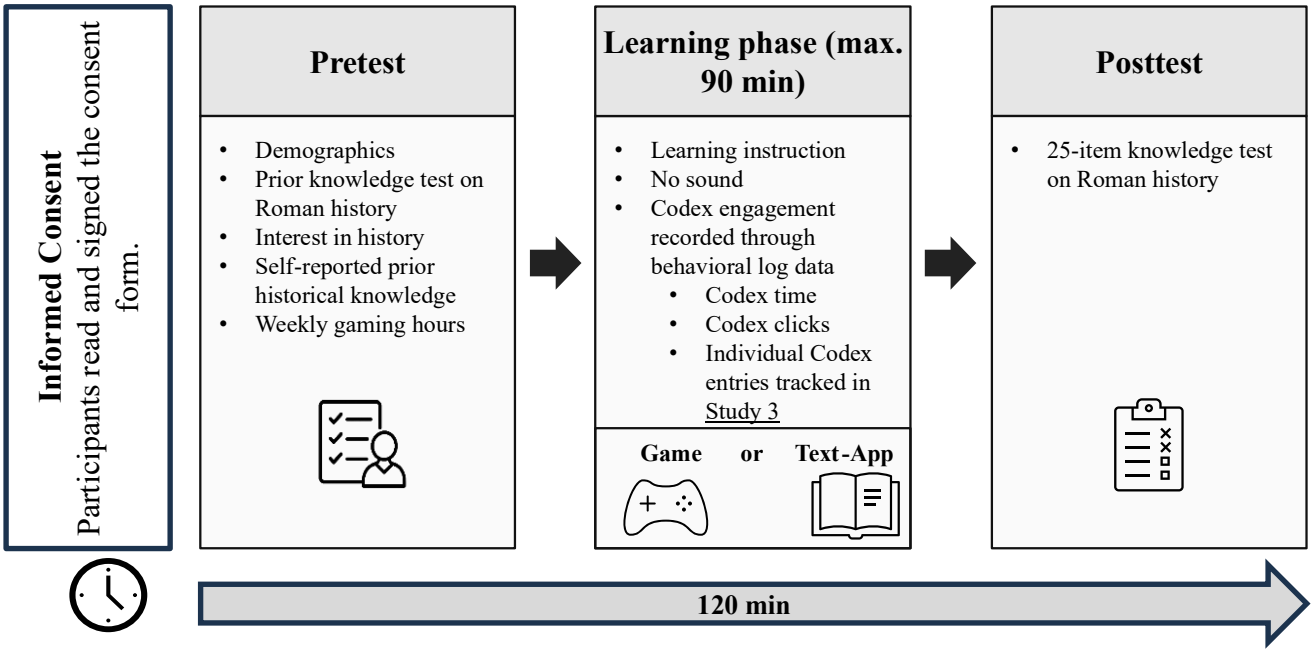
The studies were conducted in a laboratory at a German university using closely aligned procedures (see Figure 3). After being informed about the study and data protection and providing consent, participants completed demographics and pretest measures in Unipark via Firefox on university computers.

Participants then received the learning instruction and entered their assigned *Limes* condition. Studies 1 and 2 used the Original version and Study 3 the Master Edition; Studies 1 and 3 included Game and Text, whereas Study 2 used Game only. For the present analyses, we included only participants from the learning-instruction condition in Study 1 and the no-sound condition in Study 2. The corresponding standalone application was launched, and participants had up to 90 minutes in *Limes*. Codex engagement was logged throughout, with entry-level logging additionally available in

Study 3. Participants then completed the posttest in Unipark, including the 25-item knowledge test. Sessions lasted up to 120 minutes.

Figure 3

Common procedure and measures across the three contributing Limes studies



Analyses

We conducted 2 × 2 ANCOVAs with Version (Original vs. Master Edition) and Format (Game vs. Text) as between-subjects factors to examine differences in Codex engagement and posttest knowledge. Candidate pretest covariates were included when they were statistically significantly associated with the respective dependent variable, as preregistered. For H3, separate multiple regressions within the Master Edition sample examined associations between Codex engagement and posttest knowledge while controlling for Format and relevant covariates. Exploratory entry-level analyses examined access to individual Codex entries and their associations with corresponding knowledge-test items. All significance tests were evaluated at $\alpha = .05$.

Results

Preliminary Analyses

Across the full sample, participants spent an average of 1.91 minutes on the instruction screen, 65.18 minutes in the *Limes* digital environments, and 10.24 minutes completing the posttest (see Table 2).

Baseline differences across the four Version × Format conditions were examined using Kruskal–Wallis tests. Age differed significantly across the four Version × Format groups, $\chi^2(3) = 12.68$, $p = .005$. No significant differences were found for weekly gaming hours, $\chi^2(3) = 0.39$, $p = .942$; self-reported prior historical knowledge, $\chi^2(3) = 3.10$, $p = .376$; interest in history, $\chi^2(3) = 1.01$, $p = .799$; or pretest factual knowledge, $\chi^2(3) = 3.67$, $p = .300$. Because some gender categories contained very few participants across conditions, Fisher's exact test with Monte Carlo simulation was used. Gender distribution did not differ significantly across conditions, $p = .234$. The preregistered covariates were

pretest factual knowledge, weekly gaming hours, interest in history, and self-reported prior historical knowledge. Descriptive statistics for the main outcome variables are reported in Table 3.

Table 2

Time Spent Across the Experimental Procedure by Version and Format

	Original Game (n = 56)		**Original Text (n = 33)**		**Master Game (n = 56)**		**Master Text (n = 55)**	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Instruction Screen Time	1.63	0.47	1.91	0.57	1.94	0.52	2.17	1.04
Main Content Time	63.97	13.27	45.38	9.30	56.99	10.77	41.59	8.94
Codex Time	9.92	9.58	18.45	9.53	9.73	8.98	14.01	11.74
Total Time in *Limes*	73.89	13.84	63.83	16.24	66.67	12.71	55.60	16.22
Posttest Total Time	9.33	2.31	10.69	2.46	10.46	2.22	10.67	3.00

Note. All times are reported in minutes. Main Content Time refers to time spent engaging with the main storyline and secondary content in the game or text application, excluding Codex content. Total Time in *Limes* comprises Main Content Time and Codex Time.

Table 3

Means and Standard Deviations for Outcome Variables

	Original Game (n = 56)		**Original Text (n = 33)**		**Master Game (n = 56)**		**Master Text (n = 55)**	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Posttest Knowledge (0–25)	13.84	4.77	15.36	3.90	12.30	3.92	14.62	4.40
Codex Clicks (Count)	15.70	13.21	36.21	8.92	17.71	14.44	28.13	18.92
Codex Time (Minutes)	9.92	9.58	18.45	9.53	9.73	8.98	14.01	11.74

We inspected the correlations between the preregistered candidate covariates and the dependent variables (see Table 4). Covariates were included in the respective models when they showed a significant association with the dependent variable, in accordance with the preregistered analysis plan.

Table 4

Correlations Among Study Variables

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1. Interest in History (Pretest)	—						
2. Self-Reported Historical Knowledge (Pretest)	.57***	—					
3. Pretest Knowledge (5 Items)	.26***	.39***	—				
4. Weekly Gaming Hours (Pretest)	.12	.16*	.00	—			
5. Posttest Knowledge (25 Items)	.35***	.27***	.15*	.11	—		
6. Codex Clicks	.16*	.06	-.08	.02	.53***	—	
7. Codex Time (Minutes)	.15*	-.02	-.11	-.03	.55***	.74***	—

Note. *Correlation is significant at the .05 level (two-tailed); **Correlation is significant at the .01 level (two-tailed); ***Correlation is significant at the .001 level (two-tailed).

Hypothesis Testing

Overall Codex Engagement: Original Version vs. Master Edition (Hypothesis 1) For Codex time, the main effect of Version was not significant, $F(1, 195) = 2.53$, $p = .114$, $\eta^2_p = .013$. Adjusted means were 14.17 min for the Original version and 11.87 min for the Master Edition. For Codex clicks, the main effect of Version was also not significant, $F(1, 195) = 2.00$, $p = .159$, $\eta^2_p = .010$, with adjusted means of 25.9 and 22.9 clicks, respectively. Pretest interest in history was a significant

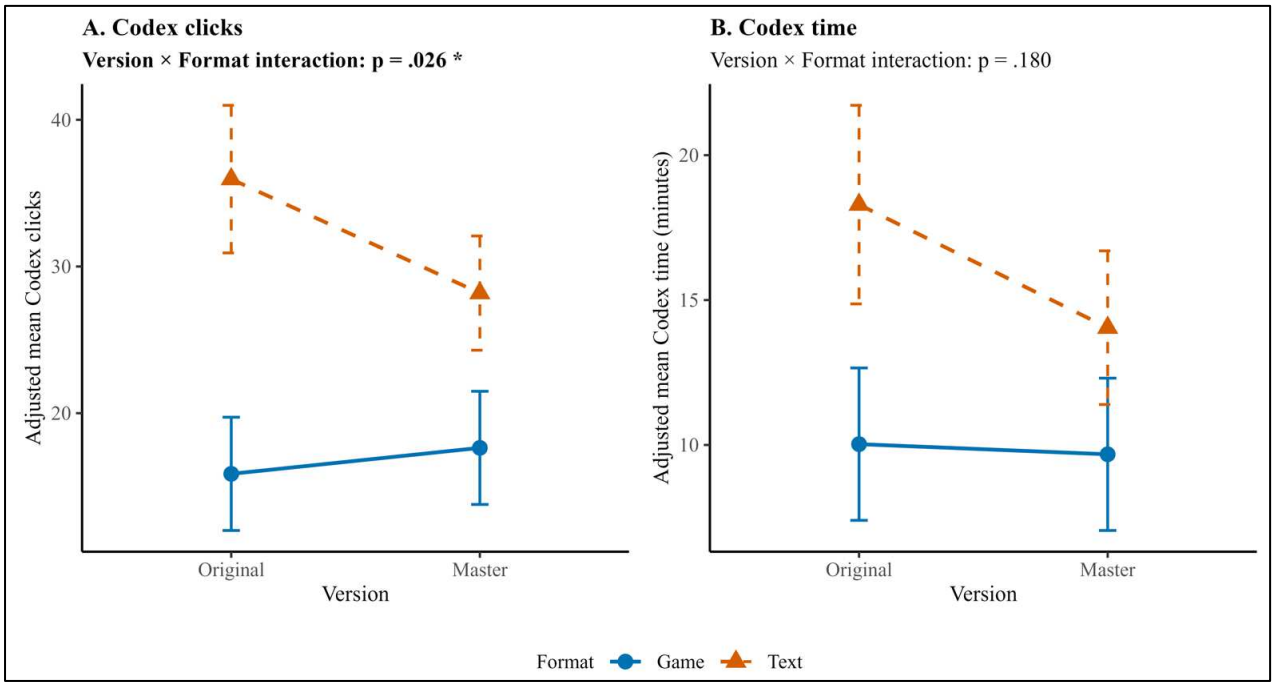
covariate for both Codex time, $F(1, 195) = 4.45$, $p = .036$, $\eta^2_p = .022$, and the Codex click, $F(1, 195) = 5.44$, $p = .021$, $\eta^2_p = .027$. Hypothesis 1 was therefore not supported for either measure of Codex engagement.

Moderation of the Game–Text Difference in Codex Engagement (Hypothesis 2). For Codex time, the Version × Format interaction was not significant, $F(1, 195) = 1.81$, $p = .180$, $\eta^2_p = .009$. Adjusted means were 10.03 min (Game) and 18.30 min (Text) in the Original version, compared with 9.68 min (Game) and 14.05 min (Text) in the Master Edition (Figure 4). The simple effect of Format was significant in both the Original version, $p < .001$, and the Master Edition, $p = .022$. Thus, Hypothesis 2 was not supported for Codex time.

For Codex clicks, the Version × Format interaction was significant, $F(1, 195) = 5.03$, $p = .026$, $\eta^2_p = .025$. Adjusted means were 15.9 clicks (Game) and 36.0 clicks (Text) in the Original version, compared with 17.6 clicks (Game) and 28.2 clicks (Text) in the Master Edition (Figure 4). The Game–Text difference therefore decreased from 20.1 to 10.6 clicks. The simple effect of Format was significant in both versions, $ps < .001$. Thus, even though the difference between game and text was smaller in the Master edition, which we had outlined in Hypothesis 2, the smaller Game–Text difference primarily reflected fewer clicks in the Master Text condition rather than more clicks in the Master Game condition.

Figure 4

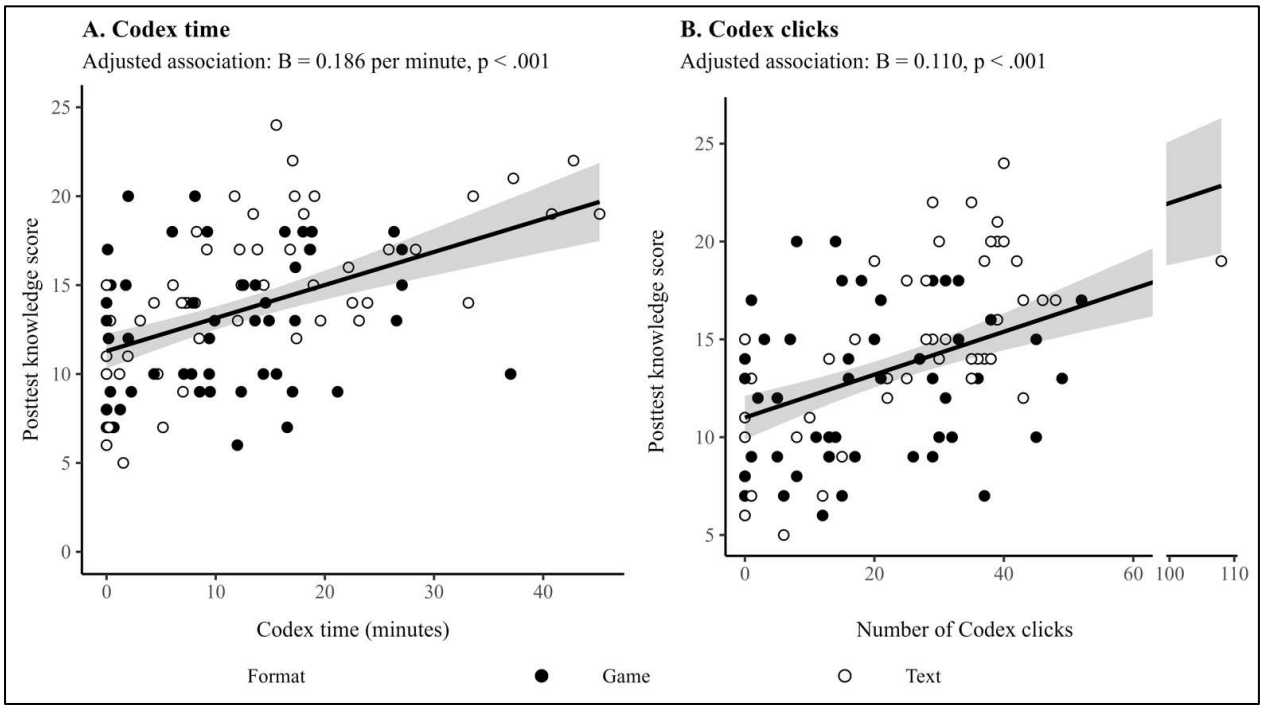
Version × Format Effects on Codex Engagement



Association Between Codex Engagement and Knowledge in the Master Edition (Hypothesis 3). Greater Codex time was significantly associated with higher posttest knowledge, $B = 0.186$, $p < .001$. Greater Codex clicking was likewise significantly associated with higher posttest knowledge, $B = 0.110$, $p < .001$ (Figure 5). The models controlled for Format, weekly gaming hours, pretest interest in history, and self-reported prior historical knowledge. Thus, Hypothesis 3 was supported for both measures of Codex engagement.

Figure 5

Associations Between Codex Engagement and Posttest Knowledge in the Master Edition



Effects on Knowledge Acquisition (Hypothesis 4). The ANCOVA revealed a significant main effect of Version on posttest knowledge, $F(1, 193) = 4.47$, $p = .036$, $\eta^2_p = .023$. Contrary to the prediction, adjusted knowledge scores were higher in the Original version ($M = 14.6$) than in the Master Edition ($M = 13.4$). There was also a significant main effect of Format, with higher adjusted posttest knowledge in the Text than in the Game format, $F(1, 193) = 9.85$, $p = .002$, $\eta^2_p = .049$. The Version × Format interaction was not significant, $F(1, 193) = 0.79$, $p = .376$, $\eta^2_p = .004$, indicating that the Format difference did not significantly vary between the Original version and Master Edition (see Appendix Figure A2). Pretest interest in history was a significant covariate, $F(1, 193) = 12.49$, $p < .001$, $\eta^2_p = .061$. Thus, Hypothesis 4 was not supported.

Exploratory Analysis

Single Codex-Entry Analyses. The Master Edition redesign introduced entry-level logging of Codex interactions, making it possible to examine whether accessing individual Codex entries was associated with correctly answering the corresponding knowledge-test questions. A separate logistic regression was conducted for each Codex-entry/question pair. The 25 knowledge-test questions produced 29 pairs because four questions were linked to two Codex entries each[1]. Of the 29 associations, 13 were significant before correction ($p < .05$), ranging between 2.51 and 19.11. After Benjamini–Hochberg false discovery rate (FDR) correction across the 29 tests, nine remained significant and are reported in Table 5. Full results for all 29 analyses are reported in Appendix A3.

[1] Four knowledge-test questions were linked to two Codex entries rather than one. Specifically, Question 1 was linked to CE 1.1 (Alexander Severus) and CE 2.3 (Crisis of the Third Century); Question 2 to CE 1.2 (Hannibal) and CE 1.3 (The Carthaginians); Question 10 to CE 2.7 (Spartacus) and CE 2.8 (Third Servile War); and Question 20 to CE 3.4 (Raetia) and CE 3.5 (Roman Province).

Table 5

FDR-Significant Associations Between Codex-Entry Access and Corresponding Knowledge-Test Questions

Question	Codex entry	Correct if accessed (%)	Correct if not accessed (%)	Adjusted OR	FDR-adjusted *p*
5	CE 1.4 – The Alemanni	76.3%	33.3%	6.12	< .001
9	CE 5.1 – Hadrian's Wall	55.3%	22.2%	4.75	.005
10	CE 2.8 – Third Servile War	48.0%	21.7%	3.30	.035
15	CE 4.8 – The Suebi	55.6%	14.3%	4.79	.019
16	CE 4.5 – Pit-house	91.7%	33.9%	19.11	< .001
17	CE 4.6 – Arminius	70.2%	38.1%	3.34	.027
18	CE 4.7 – Emperor Augustus	93.2%	51.5%	11.10	.004
19	CE 1.8 – Queen Zenobia	55.0%	26.0%	3.41	.019
24	CE 3.6 – Emperor Nero	62.7%	23.5%	3.79	.023

Note. Percentages indicate the proportion of participants who answered the corresponding knowledge question correctly, depending on whether they accessed the Codex entry. *OR* = adjusted odds ratio comparing participants who accessed the entry with those who did not. Values above 1 indicate higher odds of a correct answer among participants who accessed the entry. *ORs* were adjusted for Format, weekly gaming hours, pretest interest in history, and self-reported prior historical knowledge. *FDR* = false discovery rate; *p* values were adjusted for multiple comparisons. *CE* = Codex entry; the first number identifies the Limes level and the second identifies the entry within that level.

Discussion

The present study examined the Master Edition of Limes as an evidence-informed redesign of a digital learning environment. Building on findings from two earlier studies, the redesign aimed to encourage engagement with optional historical information and support knowledge acquisition. The Original version and Master Edition were evaluated in separate data collections rather than through random assignment to versions. Consequently, the observed differences between versions cannot be attributed solely to the redesign.

Contrary to Hypothesis 1, neither Codex time nor Codex clicks were significantly higher overall in the Master Edition. For Hypothesis 2, the smaller Game–Text difference in clicking primarily reflected fewer clicks in the Master Text condition than in the Original Text condition, rather than greater clicking among Master Game participants. These findings highlight a challenge of iterative educational technology development: redesigning optional learning resources to make them more visible does not necessarily translate into greater engagement. From the perspective of the FoRe-Squares model, focus support alone may be insufficient when learners are not sufficiently motivated to initiate and maintain active, goal-directed processing (Eitel et al., 2025). Similarly, Thomann and Deutscher (2025) found that the effectiveness of digital prompts varied with their design, with stronger effects for more targeted approaches. The non-adaptive prompts introduced in the Master Edition Game may therefore have been insufficient to encourage greater Codex engagement.

The clearest finding concerned Hypothesis 3. Within the Master Edition, greater Codex time and more Codex clicks were associated with higher posttest knowledge after controlling for Format, weekly gaming hours, pretest interest in history, and self-reported prior historical knowledge. Importantly, this finding is consistent with the two previous *Limes* studies, both of which showed that greater Codex engagement was associated with higher posttest knowledge (Study 1: Viccari, Kienitz, et al., 2026; Study 2: Viccari, Lindner, et al., 2026). The present findings therefore provide converging

evidence across three studies that greater engagement with the Codex is consistently associated with higher knowledge scores.

Hypothesis 4 was not supported. Contrary to the prediction, adjusted posttest knowledge was lower among Master Edition participants than among participants using the Original version. This difference should be interpreted cautiously, as the samples came from separate data collections. Even though the participants did not significantly differ in any assessed pretest characteristics, except for their age distributions, other differences in participant groups or study contexts may have contributed to the observed result (Waddington et al., 2023).

Although the effects of the redesign on overall engagement and knowledge were limited, the more detailed logging implemented in the Master Edition enabled a more fine-grained analysis of how learners used individual Codex entries. Of the 29 predefined Codex-entry/question associations, nine remained significant after correction for multiple testing. For these topics, learners who accessed the relevant Codex entry were more likely to answer the corresponding knowledge question correctly. This is consistent with Dever et al. (2020) who found that spending more time with books and research articles in a game-based learning environment was associated with greater learning gains.

Limitations and Directions for Future Research

Several limitations should be acknowledged. First, the comparison between versions relied on pooled data from separate studies. Although the datasets were harmonized and preregistered covariates were included when appropriate, unmeasured differences between the studies or samples may still have influenced the results. A stronger test of the redesign would therefore randomly assign participants within the same study to either the Original version or the Master Edition.

Second, clicks and time capture only observable aspects of engagement and do not show how learners processed the information. Bergdahl et al. (2024) found that higher-education learning-analytics research relies predominantly on behavioral indicators such as clicks and time on task, while cognitive engagement and deeper processing are measured much less often. Opening a Codex entry or spending more time on it therefore does not necessarily mean that the information was carefully read or understood. This limitation is also illustrated by Dever et al. (2022) who noted that log data showing that an instructional resource was opened were not sufficient to establish that learners had actually read it. Future research could therefore examine more detailed patterns of Codex use, including entry-specific dwell time, revisits, navigation sequences, and transitions between gameplay and informational resources (Dever et al., 2022). These behavioral traces could also be combined with multimodal measures to provide a more detailed picture of learners' cognitive, affective, and motivational processes during learning (Cloude et al., 2024; Zhao et al., 2026).

Conclusions

Taken together, the findings illustrate the challenges of evidence-informed, iterative redesign: making optional learning resources more visible is not necessarily accompanied by greater use or better learning outcomes. In *Limes*, the Master Edition was developed based on findings from earlier studies, but participants using it did not show higher overall Codex engagement or posttest knowledge than those using the Original version in the cross-study comparison. At the same time, greater Codex use was consistently associated with higher knowledge scores, and the more detailed logging introduced in the Master Edition enabled the identification of similar associations for several specific Codex entries and their corresponding knowledge questions. The central implication is that the educational value of optional learning resources may depend not only on their availability or prominence, but also

on how learners use them. Digital learning design should therefore consider not only how information is provided, but also how learners can be supported in making effective use of it.

References

Bergdahl, N., Bond, M., Sjöberg, J., Dougherty, M., & Oxley, E. (2024). Unpacking student engagement in higher education learning analytics: A systematic review. *International Journal of Educational Technology in Higher Education*, *21*(1), 63. https://doi.org/10.1186/s41239-024-00493-y

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research*, *86*(1), 79–122. https://doi.org/10.3102/0034654315582065

Cloude, E. B., Dindar, M., Ninaus, M., & Kiili, K. (2024). Synchrony Between Facial Expressions and Heart Rate Variability During Game-Based Learning: Insights from Cross-Wavelet Transformation. In R. Ferreira Mello, N. Rummel, I. Jivet, G. Pishtari, & J. A. Ruipérez Valiente (Eds.), *Technology Enhanced Learning for Inclusive and Equitable Quality Education* (pp. 90–104). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72315-5_7

Dever, D. A., Amon, M. J., Vrzáková, H., Wiedbusch, M. D., Cloude, E. B., & Azevedo, R. (2022). Capturing Sequences of Learners' Self-Regulatory Interactions With Instructional Material During Game-Based Learning Using Auto-Recurrence Quantification Analysis. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.813677

Dever, D. A., Azevedo, R., Cloude, E. B., & Wiedbusch, M. (2020). The Impact of Autonomy and Types of Informational Text Presentations in Game-Based Environments on Learning: Converging Multi-Channel Processes Data and Learning Outcomes. *International Journal of Artificial Intelligence in Education*, *30*(4), 581–615. https://doi.org/10.1007/s40593-020-00215-1

Eitel, A., Krebs, M.-C., & Schöne, C. (2025). Introducing the FoRe-Squares Model: Focus Support and Reward as Key Predictors of Effective Technology-Augmented Instruction. *Educational Psychology Review*, *37*(3), 67. https://doi.org/10.1007/s10648-025-10045-z

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Karich, A. C., Burns, M. K., & Maki, K. E. (2014). Updated Meta-Analysis of Learner Control Within Educational Technology. *Review of Educational Research*, *84*(3), 392–410. https://doi.org/10.3102/0034654314526064

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Scheiter, K. (2021). The Learner Control Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (3rd ed., pp. 418–429). Cambridge University Press. https://doi.org/10.1017/9781108894333.043

Schumacher, A., Kammerer, Y., Scharinger, C., Gottschling, S., Hübner, N., Tibus, M., Kasneci, E., Appel, T., Gerjets, P., & Bardach, L. (2025). How do intellectually curious and interested people learn and attain knowledge? A focus on behavioral traces of information seeking. *European Journal of Personality*, 08902070241309124. https://doi.org/10.1177/08902070241309124

Schumacher, A., Kammerer, Y., Scharinger, C., Gottschling, S., Hübner, N., Tibus, M., Kasneci, E., Appel, T., Gerjets, P., & Bardach, L. (2026). How do intellectually curious and interested people learn and attain knowledge? A focus on behavioral traces of information seeking. *European Journal of Personality*, *40*(1), 87–110. https://doi.org/10.1177/08902070241309124

Thomann, H., & Deutscher, V. (2025). Scaffolding through prompts in digital learning: A systematic review and meta-analysis of effectiveness on learning achievement. *Educational Research Review*, *47*, 100686. https://doi.org/10.1016/j.edurev.2025.100686

Viccari, A. (2026). *OLD. Limes Master Edition: Using Behavioural Data to Refine Game Mechanics in Research*. https://doi.org/10.31234/osf.io/j5s6b_v1

Viccari, A., Göllner, R., Hahn, J.-U., & Bardach, L. (2024). Limes (the Roman Frontier): Developing a Video Game for History Learning. *European Conference on Games Based Learning*, *18*(1), 839–847. https://doi.org/10.34190/ecgbl.18.1.2833

Viccari, A., Kienitz, A., Eitel, A., Hahn, J.-U., Göllner, R., Lindner, M., & Bardach, L. (2026). *Playing, Reading, Learning, and Having Fun? Knowledge, Interest, and Emotions in Video Game and Text-based Environments* (Fa28t_v3). PsyArXiv. https://osf.io/preprints/psyarxiv/fa28t_v3/

Viccari, A., Lindner, M. A., Kienitz, A., Hahn, J.-U., Göllner, R., Emslander, V., & Bardach, L. (2026). The sound of the roman empire: Effects of playing a history video game with and without sound. *Trends in Neuroscience and Education*, *43*, 100290. https://doi.org/10.1016/j.tine.2026.100290

von Keyserlingk, L., Schwerter, J., Kunina-Habenicht, O., & Lauermann, F. (2026). Motivation and self-regulated learning as predictors of learning resource use in learning management systems in higher education. *Educational Psychology*, *46*(3), 437–457. https://doi.org/10.1080/01443410.2025.2564891

Waddington, H. S., Villar, P. F., & Valentine, J. C. (2023). Can Non-Randomised Studies of Interventions Provide Unbiased Effect Estimates? A Systematic Review of Internal Replication Studies. *Evaluation Review*, *47*(3), 563–593. https://doi.org/10.1177/0193841X221116721

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Zhao, L., Raković, M., Lee, G. A., Abuzayyad-Nuseibeh, H., Cloude, E. B., & Bardach, L. (2026). Capturing Momentary Motivational States Using Detector-Driven In-Situ Questionnaire. In E. G. Blanchard, G. Chen, M. Chi, & S. Isotani (Eds.), *Artificial Intelligence in Education. Late Breaking Results, WideAIED, Practitioners, Industry and Policies, Blue Sky, Doctoral Consortium, FoL Workshops and Tutorials, FoL Invited Papers* (pp. 538–544). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-29788-4_77

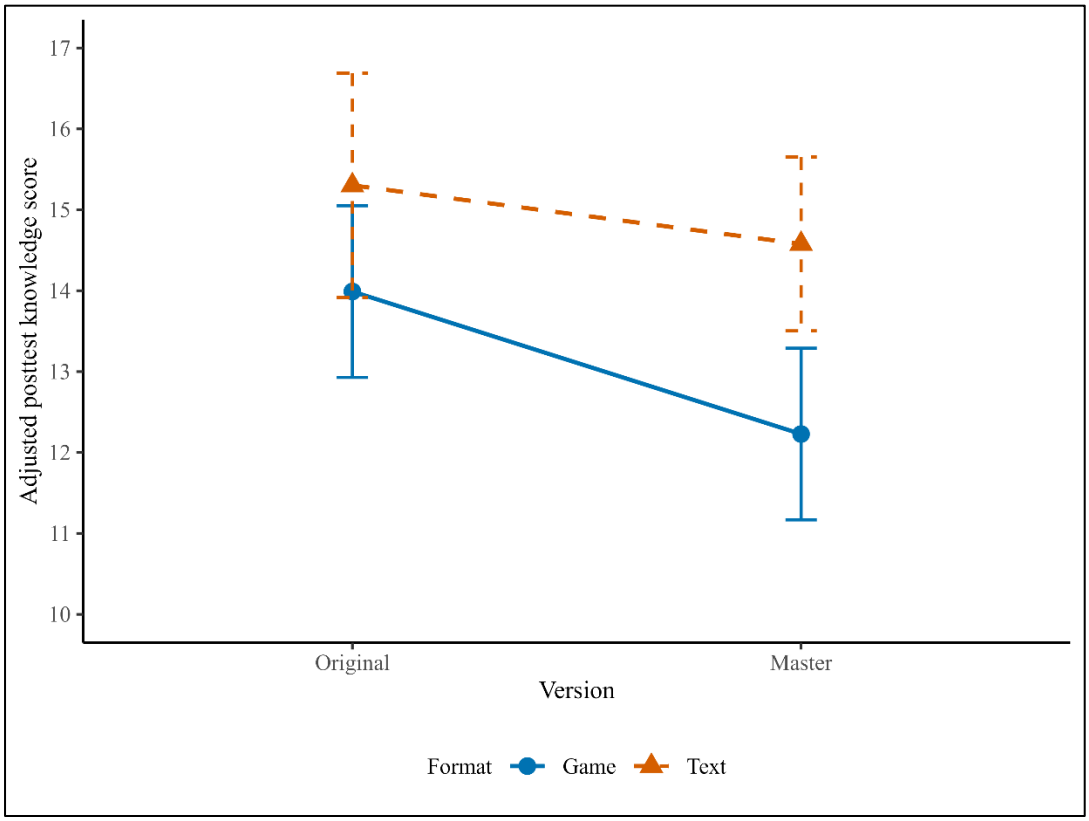
Appendix Table A1

Overview of Source Studies and Construction of the Final Pooled Sample

Source study	Original design	Raw *N*	Final original-study sample after exclusions	Portion relevant to present analysis	Overlap handling	Unique contribution to pooled sample
Study 1	2 × 2 Format × Instruction	143	**N = 140:** Game–Fun = 36; Game–Learning = 36; Text–Fun = 35; Text–Learning = 33	Learning only: Game = 36; Text = 33; total = 69	No removal at this stage; Study 1 records retained as the primary record when also represented in Study 2	**69**
Study 2	Sound vs. No Sound; game format	112	**N = 111:** Sound = 55; No Sound = 56	No-Sound condition = 56	36 of the 56 were already represented in Study 1 and were not counted again; 20 were unique	**20**
Study 3	Master Edition, Game vs. Text	114	**N = 111:** Game = 56; Text = 55	Entire retained sample: Game = 56; Text = 55	No Study 1/2 overlap requiring removal	**111**
Final pooled sample	2 × 2 Version × Format analytical structure	—	—	Original–Game = 56; Original–Text = 33; Master–Game = 56; Master–Text = 55	Unique participants only	**200**

Appendix Table A2

Adjusted Posttest Knowledge by Version and Format



Appendix Table A3

Associations Between Codex-Entry Access and Corresponding Knowledge-Test Questions

Question	Codex entry	Correct if accessed (%)	Correct if not accessed (%)	Adjusted OR	Raw p	FDR-adjusted p
Q1	CE 1.1 – Alexander Severus	96.2	83.9	4.75	.056	.116
Q1	CE 2.3 – Crisis of the Third Century	93.0	92.3	1.40	.675	.753
Q2	CE 1.2 – Hannibal	18.8	13.0	1.45	.500	.630
Q2	CE 1.3 – The Carthaginians	18.6	13.7	1.29	.639	.753
Q3	CE 1.3 – The Carthaginians	54.2	54.9	0.94	.885	.917
Q4	CE 1.4 – The Alemanni	81.4	74.5	1.67	.293	.425
Q5	CE 1.4 – The Alemanni	76.3	33.3	6.12	< .001	< .001
Q6	CE 1.7 – Sassanid Empire	45.2	20.8	2.51	.043	.104
Q7	CE 2.4 – Roman Fort	93.7	83.9	2.28	.251	.393
Q8	CE 2.4 – Roman Fort	93.7	83.9	3.43	.103	.199
Q9	CE 5.1 – Hadrian's Wall	55.3	22.2	4.75	< .001	.005
Q10	CE 2.7 – Spartacus	44.4	23.2	2.58	.047	.104
Q10	CE 2.8 – Third Servile War	48.0	21.7	3.30	.011	.035
Q11	CE 3.1 – Limes	71.2	64.9	1.15	.757	.813
Q12	CE 3.2 – Julius Caesar	83.3	53.2	3.22	.027	.071
Q13	CE 3.2 – Julius Caesar	35.4	27.4	1.22	.654	.753
Q14	CE 4.1 – The Juthungi	78.1	70.3	1.60	.347	.479
Q15	CE 4.8 – The Suebi	55.6	14.3	4.79	.003	.019
Q16	CE 4.5 – Pit-house	91.7	33.9	19.11	< .001	< .001
Q17	CE 4.6 – Arminius	70.2	38.1	3.34	.007	.027
Q18	CE 4.7 – Emperor Augustus	93.2	51.5	11.10	< .001	.004
Q19	CE 1.8 – Queen Zenobia	55.0	26.0	3.41	.004	.019
Q20	CE 3.4 – Raetia	36.9	26.7	1.84	.181	.318
Q20	CE 3.5 – Roman Province	36.7	28.0	1.63	.258	.393
Q21	CE 1.2 – Hannibal	35.9	15.2	3.31	.020	.059
*Q22	CE 5.5 – Germanic Food	100.0	87.0	1.16×10^{8}[a]	.993	.993
Q23	CE 1.5 – Boudicca	55.7	42.5	1.74	.187	.318
Q24	CE 3.6 – Emperor Nero	62.7	23.5	3.79	.005	.023
Q25	CE 3.3 – Roman Republic	20.8	12.9	1.48	.489	.630

Note. OR = adjusted odds ratio; *FDR* = false discovery rate. Percentages show the proportion of participants who answered the corresponding knowledge-test question correctly among those who did and did not access the relevant Codex entry. *ORs* were estimated using logistic regression and indicate the association between Codex-entry access and a correct response, controlling for Format, weekly gaming hours, pretest interest in history, and self-reported prior historical knowledge. Benjamini–Hochberg correction was applied across the 29 tests. *The model for Q22/CE 5.5 showed separation-like behavior and an extreme odds ratio; this *OR* should not be interpreted.