Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

From reconstruction to remembering: Learning, perceived realism, and narrative engagement in a company history video with AI-generated sequences

Andrea Laudadio[1] and Maria Ermelinda De Carlo[2*]

[1] TIM S.p.A., Rome, Italy; email: andrea@laudadio.it
[2] Università Telematica Pegaso, Naples, Italy; email: mariaermelinda.decarlo@unipegaso.it
* Corresponding author: mariaermelinda.decarlo@unipegaso.it

Abstract: Corporate training seldom examines how established employees learn about their organization's history. This study investigated collective viewing of a 33-minute historical video produced for TIM, an Italian telecommunications company, including approximately 14 minutes of AI-generated reconstructions. In a single-group, three-wave field study, 57 employees responded at baseline, 47 after viewing, and 43 at follow-up (33–60 days later). Measures included factual knowledge, perceived realism, narrative engagement, interest, organizational identification, and pride. Among matched respondents, accuracy on taught items increased from 58.1% to 83.0%; pooled control-item accuracy remained stable, although control-item difficulty differed between forms. Of the immediate knowledge gain, 68% was retained among the 42 respondents completing all waves. Interest, identification, and pride did not change reliably. Perceived realism correlated with narrative engagement (r = .56), but associations with knowledge and attitudes were not robust to an influential observation. Post-viewing interest was associated with self-reported voluntary behaviors at follow-up. The findings support assessing experience, factual learning, and retention separately. The study cannot isolate the educational contribution of the AI-generated scenes.

Keywords: organizational history; generative AI; perceived realism; narrative engagement; training evaluation

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Corporate history can inform learning beyond the induction of newcomers: it situates employees' everyday work within technological and organizational change. Organizational narratives also shape identity, however, and can present a selective interpretation of the past (Klein & Weaver, 2000; Aeon & Lamertz, 2021). Historical training therefore raises both an instructional question—what is learned and

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

retained?—and an epistemic one—how is a convincing reconstruction distinguished from evidence?

Generative AI makes it possible to illustrate historical episodes for which moving images do not exist. Research on AI-generated instructors has found no consistent improvement in learning over human-produced videos, and sometimes less favorable engagement (Arkün-Kocadere & Çağlar-Özhan, 2024; Netland et al., 2025). Reconstructed scenes are different from synthetic lectures: they support a historical narrative while risking confusion between visual plausibility and documentary authenticity (Doss et al., 2023).

Narratives may aid understanding and recall (Mar et al., 2021), and multimedia learning theory identifies benefits when visual information is relevant and cognitively manageable (Mayer, 2024). Yet perceived realism is multidimensional and need not signal accuracy or educational effectiveness (Cho et al., 2014). Narrative engagement includes attention, understanding, presence and emotion (Busselle & Bilandzic, 2009); interest may motivate later voluntary engagement (Hidi & Renninger, 2006). Crucially, satisfaction, demonstrated learning and delayed retention are distinct evaluation outcomes (Alliger et al., 1997; Sitzmann et al., 2008).

This field study examines a collective screening of corporate-history video containing AI-generated reconstructions. It asks whether factual knowledge improves specifically on narrated content and persists after several weeks; whether perceived realism accompanies engagement or measured outcomes; and whether interest, identification and pride change or relate to subsequent voluntary behaviors. The study tests associations and within-person differences, not the causal benefit of AI-generated scenes.

Perceived realism warrants particular attention because it is not equivalent to either pictorial detail or historical veracity. A reconstruction may seem credible while incorporating speculative scenery, gestures or dialogue. In narrative research, engagement is also multidimensional: focused attention and comprehension may not track emotional involvement or the sensation of entering the story (Busselle & Bilandzic, 2009). These distinctions motivate an evaluation that treats realism and engagement as features of the experienced presentation rather than outcome measures. Similarly, knowledge acquisition need not imply a new sense of organizational belonging, especially when an audience has extensive prior experience of the organization. The framework therefore contrasts performance on narrated content with an internal reference, measures identity-related attitudes independently, and examines whether self-reported engagement persists beyond the screening. None of these comparisons identifies what participants would have learned from an otherwise identical video without the AI-generated sequences.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

1.1. Research expectations

H1: The T0–T1 gain on taught historical items exceeds any change on untaught control items. H2: Knowledge and any interest gains are at least partly retained at T2; identification and pride are examined without directional predictions. H3a: Perceived realism is associated with engagement; H3b: these judgments explain additional T1 outcome variance after baseline adjustment. H4a: T1 interest explains variance in knowledge and retention; H4b: it is associated with voluntary activities at follow-up. An exploratory question concerns reported recollection, discussion and further information-seeking, without equating them with demonstrated historical appropriation.

2. Materials and Methods

2.1. Design and participants

A single-group study measured respondents before viewing (T0, July 29, 2026), after viewing (T1, July 31–August 3), and after 33–60 days (T2; median 40 days). Of 140 TIM employees invited, 65 attended; 57 completed T0, 47 T1 and 43 T2, with 42 completing all waves. Most baseline respondents had over 15 years of tenure (82.5%) and were aged 51–60 (71.9%). Participation was voluntary and unpaid, and employees were informed that the questionnaire formed part of a scientific study. The study was not preregistered and was not submitted to an ethics committee.

The knowledge instrument comprised two crossed parallel forms (A and B), with ten taught and two untaught control items each: A–B–A or B–A–B across T0–T2. Forms covered matching historical periods; their order was randomized and counterbalanced, limiting exact-item repetition at T1. T2 repeated the T0 form, so pretest sensitization remains possible. The two control items per form differed in difficulty and are a limited internal reference rather than a comparison condition (Haccoun & Hamtiaux, 1994; Shadish et al., 2002). Waves were matched by participant-generated codes; only 24 of 43 T2 codes were identical across waves. Incomplete codes were reconciled by similarity and elimination, and a sensitivity analysis used 30 T0–T2 pairs matching on at least three components. The analyses used participant labels rather than names (Kearney et al., 1984).

Attrition was assessed against baseline responses: of 22 unadjusted comparisons between those retained and lost, one was nominally significant for greater prior familiarity among those lost at T2 ($p = .045$). The non-random attendance and attrition mean that the sample cannot be treated as representative of TIM employees. Data were collected using online questionnaires, with the post-viewing survey opened the day after screening and the follow-up distributed in two waves; 22 T2 responses were collected on days 33–40 and 21 on days 53–60. These different intervals were considered in the robustness analyses.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

The longitudinal linkage deserves particular caution. Each self-generated code combined the first two letters of each parent's first name, day of birth, and first two letters of the city of birth. No participant names were collected, but the derived code does not guarantee complete anonymity. At T1, 42 of 47 codes matched exactly, whereas at T2 only 24 of 43 did. The matching procedure for other records introduced uncertainty about whether an apparent within-person difference reflected the same individual. Accordingly, the principal longitudinal estimates are reported together with the more securely linked subset. Because employees participated in a company setting and one author was employed by TIM, favorable ratings may also have been influenced by social desirability. These circumstances, as well as the absence of an ethics-committee review, are disclosed rather than retrospectively represented as formal ethical approval.

2.2. Intervention and measures

Participants collectively viewed "TIM: Connessioni," a 33-minute video about the development of Italian telecommunications from Antonio Meucci to mobile networks. A live-action host narrated the story, while approximately 14 minutes of AI-generated scenes illustrated otherwise unfilmed episodes, including telephone exchanges, line workers and public phone booths. The visual reconstructions were not archival footage. The AI origin had been communicated; at T1, 46 of 47 respondents said they knew or inferred it by the end of viewing.

Figure 1. Stills from the video "TIM: Connessioni": the host filmed in live action and seven AI-generated sequences

 The host, an actor-narrator, inside the TIM museum (real footage)	 A dramatization of Antonio Meucci's telephone experiments (AI)
 Linemen at work (AI)	 The switchboard operators (AI)
 Leonardo Chiariglione arrives at work (AI)	 People lining up at the phone booths on the night of the 1982 World Cup victory (AI)
 Mauro Sentinelli announces the launch of the prepaid card (AI)	 A technician installs Alice ADSL in an apartment (AI)

Note. The label in parentheses gives the origin of each image: real footage or AI-generated sequence.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

The narrative was organized around museum-like objects handled by a live-action host: telephone apparatus, photographs, portraits and a token introduced successive historical episodes. Its chronology included early telephony, operators and network workers, SIP, telephone booths, research at CSELT, mobile telecommunications and broadband. Transitions often animated a period object or photograph into a synthetic scene while the host continued narrating. Thus, generative material served as situated illustration embedded within an otherwise live-action, voice-led story, not as a synthetic teacher delivering a lecture. The same intervention was presented to all participants; there was no comparison version, and the study did not experimentally vary realism, narration, historical content or the presence of colleagues. The distinction between the filmed presenter and the generated historical scenes was retained in the figure captions.

Factual knowledge was scored from 10 multiple-choice items per form (0–10); the corresponding untaught reference used two items. Internal consistency for taught items was KR-20 = .77–.80 at baseline, .88 for form A and .40 for form B at T1, and .76/.67 at T2. The low T1 reliability of form B and an item with a recurrent distractor response were checked in sensitivity analyses.

At T1, narrative engagement was assessed using a translated 12-item Narrative Engagement Scale (total α = .71; Busselle & Bilandzic, 2009); its processing and transportation composites were uncorrelated (r = −.07). Perceived realism combined two seven-point ratings of visual credibility and plausibility (r = .92); these ratings did not independently establish factual accuracy. Across all waves, scales measured interest in company history (four items; α = .86–.87), identification (six items; α = .89–.91), and state pride (three items; α = .85–.93). The interest scale did not separate situational from enduring interest. At follow-up, participants reported conversation, spontaneous recollection and independent information searches. Interest items were adapted from Linnenbrink-Garcia et al. (2010), identification from Mael and Ashforth (1992), and pride was based on items inspired by Gouthier and Rhein (2011).

Besides the knowledge and central attitudinal measures, baseline questionnaires asked about prior familiarity with company history and attitudes toward AI (AIAS-4; Grassini, 2023), while T1 measured satisfaction, co-viewing and discomfort. The history narrative included episodes concerning operators, installation technicians, technological research and mobile services. Viewers were explicitly invited to connect organizational history with future development. Since the videos had not been historically verified at the level of every visual detail, the generated scenes were treated in the study as plausible representations; competence in distinguishing documentary sources from reconstruction was not tested.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

The engagement measure comprised narrative understanding, attentional focus, narrative presence and emotional engagement. Processing (the first two dimensions) showed pronounced ceiling effects, with 70% at the scale maximum, while presence and emotional engagement retained more variation. Perceived realism was the average of ratings for visual credibility and plausibility; a factual-fidelity item was available in only one questionnaire version (n = 24) and was therefore descriptive rather than part of the index. Satisfaction and discomfort were measured separately. These distinctions matter for interpretation: high perceived realism does not constitute source verification, and the topic-interest scale, measured at all three waves, does not isolate a momentary state of situational interest from a stable disposition toward company history. The follow-up behavioral items captured self-report, not independently observed action.

2.3. Statistical analysis

Within-person paired tests compared proportions correct for taught versus control items; generalized estimating equations (GEE), between-sequence comparisons on identical forms and a person-intercept mixed model provided checks. Retained gain was estimated among complete cases with bootstrap confidence intervals. Holm (1979) correction was applied to three primary T0–T1 comparisons. Null findings were additionally examined with equivalence tests using post hoc bounds (Lakens, 2017), which must be interpreted exploratorily. Baseline-adjusted regressions examined realism and engagement; follow-up behaviors were analysed with Firth logistic models because of sparse cells (Firth, 1993; Heinze & Schemper, 2002). Bootstrap intervals and leave-one-out checks assessed robustness. Exploratory associations were not multiplicity-adjusted. One genuine influential respondent rated realism at the minimum and answered "Don't know" to all ten T1 taught items; models involving realism are presented with and without this case.

For the primary hypothesis, the estimand was the within-person difference between the change on taught items and the change on untaught items. It is more informative than the taught-item gain alone, but imperfect when the internal reference consists of only two questions with unequal difficulty across forms. Exact-item comparisons between the two counterbalancing sequences and item-level GEE models therefore offer complementary evidence. The repeated-measures mixed model included a participant intercept and form. For associations involving perceived realism, regressions adjusted for the corresponding T0 outcome when available, with percentile bootstrap intervals from 5,000 resamples; leave-one-out estimates assessed whether one participant determined the inference. Behavioral odds ratios used Firth penalization because some response categories were small. These association tests remain exploratory, and the equivalence bounds for the attitude outcomes (dz = ±0.5) were selected after data collection.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

3. Results

Table 1. Descriptive outcomes by survey wave (all respondents available at each wave).

Measure	T0 (n = 57)	T1 (n = 47)	T2 (n = 43)
Taught items, correct (%)	59.5	83.0	74.2
Control items, correct (%)	65.8	64.9	64.0
Interest in company history (1–5)	4.40 (0.72)	4.52 (0.64)	4.43 (0.59)
Organizational identification (1–5)	4.23 (0.63)	4.16 (0.74)	4.12 (0.73)
State organizational pride (1–7)	5.89 (0.95)	6.00 (1.11)	5.84 (1.17)

Note. Attitudinal scale values are M (SD). Descriptives use all available cases; inferential comparisons below use matched subsamples and therefore have different T0 values.

3.1. Learning and retention

For 47 matched T0–T1 respondents, taught-item accuracy increased from 58.1% to 83.0% (mean gain 2.49/10, 95% CI [1.60, 3.38]; $t(46) = 5.61$; $p < .001$; $dz = 0.82$). Control accuracy changed from 66.0% to 64.9% ($p = .875$). The primary taught-versus-control interaction was +26.0 percentage points (95% CI [11.2, 40.7]; $t(46) = 3.54$; Holm-adjusted $p = .003$; $dz = 0.52$); response-level GEE was consistent (OR = 3.46, 95% CI [1.89, 6.35]). However, the interaction differed by sequence (+45.8 versus +5.2 points), and differences in control-item difficulty weaken the specificity claim. On identical items across independent sequences, post-viewing knowledge exceeded baseline for both forms (Hedges's $g = 1.09$ and 0.92).

At T2, 43 matched respondents scored 74.2% on taught items versus 56.3% at T0 (mean gain 1.79/10; 95% CI [1.02, 2.56]; $t(42) = 4.69$; $p < .001$; $dz = 0.72$); controls were 64.0% at both waves. Among 42 complete cases, 68% of the immediate knowledge gain persisted (bootstrap 95% CI [45%, 100%]). The adjusted mixed model estimated gains of 2.40 points at T1 and 1.58 at T2. Retention was similar in the 30 more securely matched T0–T2 pairs ($dz = 0.74$). The apparent T1–T2 decline also depended on form sequence, and re-use of the T0 form at T2 may inflate the retained-gain estimate.

The stability of the pooled control accuracy masks divergent changes by counterbalancing sequence: the two-item control score shifted by −0.50 in one sequence and +0.48 in the other. Specificity therefore appeared clearly on one form but only borderline on the other when equal items were compared between independent respondents. The main H1 result should consequently be described as converging evidence of content-related learning rather than proof that every possible

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

test form would show the same taught-versus-control differential. At follow-up, the mixed model estimated form B to be harder by approximately 0.64 points (p = .052). The nominal decline from T1 to T2 was −0.83 points (95% CI [−1.65, −0.01]; p = .047), but differed sharply by form sequence. Its Bayes factor was approximately 1.1, offering little independent evidence for a general decline. Both follow-up mailing periods nevertheless yielded scores above baseline.

3.2. Experience, attitudes and follow-up activity

Initial interest, identification and pride were high (means 4.40/5, 4.23/5 and 5.89/7). Interest increased by only 0.13 at T1 (uncorrected p = .049; Holm-adjusted p = .098) and returned close to baseline at T2; pride and identification showed no reliable corrected change. The study had limited power to detect smaller attitudinal effects. Engagement averaged 5.48/7 and perceived realism 5.99/7; 42.6% rated realism at the maximum.

Realism correlated with narrative engagement (r = .56; bootstrap 95% CI [.36, .70]; r = .53 without the influential case), especially its emotional/transportation component (r = .46). Baseline-adjusted associations of realism with knowledge, pride and interest did not survive omission of the influential case. Interest at T1 was unrelated to knowledge at T1 and T2 (both β = .14; p = .410 and .339). However, higher interest was associated with reported conversations, information seeking and spontaneous recollection (penalized ORs = 2.29, 4.57 and 5.92 per SD, respectively; all exploratory, with wide intervals). Adjusted for baseline interest, the rank correlation with the number of reported behaviors was ρ = .36 (p = .021).

At follow-up, 39/43 (90.7%) reported talking about the experience, 37/43 (86.0%) thinking back to scenes, and 17/43 (39.5%) seeking further information. Realism was associated with spontaneous recollection (ρ = .51, p < .001), not with the other two behaviors. Sensitivity checks showed that the taught-item gain persisted after excluding the problematic form-B item (dz = .82) and was larger without the influential observation (dz = .99). These self-reports indicate subsequent engagement, not independently verified learning or critical appropriation.

Further exploratory checks showed that perceived realism was correlated with co-viewing (r = .65), satisfaction (r = .53) and baseline attitude toward AI (r = .41). Only four of 47 respondents noticed visual flaws, and mean discomfort was 2.26/7. Correlations of realism with T1 interest and pride appeared positive in full-sample regressions, but decreased to nonsignificant levels after exclusion of the influential case (β = .15 and .18). No evidence of a stable indirect pathway through narrative engagement was found. These results reinforce the distinction between a favorable subjective experience and evidence of specific knowledge acquisition.

The sparse binary contrasts warrant particular caution in H4b. The 95% profile-likelihood confidence intervals around the penalized odds ratios were [1.14, 5.17] for

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

conversation, [1.26, 32.14] for information searching and [2.11, 39.11] for thinking back to scenes. Only four respondents reported no conversation and six reported never thinking back, making the estimated magnitudes highly imprecise. Rank-order results agreed in direction: interest at T1 correlated with conversations (ρ = .44), seeking information (ρ = .40) and recollection (ρ = .47). However, higher interest did not explain extra variation in knowledge at either posttest or follow-up. A coherent interpretation is thus an association with self-reported continuation of engagement, rather than an effect on test performance. These analyses were not adjusted for multiple comparisons and should not be promoted to confirmatory evidence.

Overall, H1 was supported in the pooled analysis, subject to form-specific caveats; H2 was supported for knowledge retention, not sustained interest. H3a was supported, whereas H3b was not robust to sensitivity checks. H4a was not supported, while H4b received exploratory, imprecise support for self-reported voluntary behaviors.

4. Discussion

This study distinguishes three outcomes that can otherwise be conflated: a positively rated experience, learning of narrated historical facts, and persistence of acquired scores several weeks later. The improvement is compatible with narrative and multimedia-learning research (Mar et al., 2021; Mayer, 2024), but the absence of an alternative version of the video prevents attribution to the AI-generated scenes. The taught-versus-control comparison provides a within-study reference, yet the two control items per form, unequal form difficulty and crossover-specific effects demand caution. Repeated exposure to baseline questions may also have directed attention to content and supported recognition at follow-up (St. Hilaire et al., 2024).

Realism was associated with narrative engagement but not robustly with knowledge or changed attitudes. Its high ceiling and the correlation between visual credibility and plausibility suggest a broad appraisal of the experience rather than an independent estimate of visual or historical truth. Judgments of realism therefore cannot replace learning tests. Nor does the absence of reliable shifts in interest, identification or pride establish that such outcomes are unchangeable: participants had long organizational tenure and high initial scores, leaving limited measurable room for change.

Post-viewing interest accompanied later conversations, curiosity and recollection, in line with the possibility that interest encourages further engagement (Hidi & Renninger, 2006). The observations establish neither a causal pathway nor appropriation in Wertsch's (1998) sense of making a historical narrative one's own. The links were exploratory, based on small and imbalanced categories, with possible common-method and social-desirability bias.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

From an educational-design perspective, generative reconstructions should be presented as interpretations supported, where possible, by identifiable historical sources, never as archival recordings. Specific source-critical guidance may be preferable to an undifferentiated "AI-generated" label when visuals depict invented details (Butler et al., 2009; Doss et al., 2023). Evaluation should combine knowledge measures, preferably with a matched comparison condition and stronger reference items, with delayed follow-up and appropriately qualified behavioral indicators.

AI-generated dramatizations call for explicit differentiation between documentary evidence and interpretative visual choices, particularly when emotionally compelling scenes may obscure uncertainty about historical sources. In the present evaluation, the disclosure of AI production was nearly universal, yet without an uninformed comparison group it remains unknown whether awareness affected realism ratings, learning or recollection. The design also cannot separate the effect of watching with colleagues from the influence of narration or video content.

A further distinction concerns memory for a narrated historical proposition versus memory for a synthetic picture. Participants answered factual multiple-choice questions, not tasks testing whether they could identify documentary evidence, infer source uncertainty, or reject visual details invented by generative tools. The observed knowledge gain therefore cannot establish improved historical reasoning or critical media literacy. An instructional extension could connect each reconstruction to dated primary sources and ask learners which parts of a scene follow directly from evidence and which are plausible artistic decisions. Future studies might compare otherwise equivalent narrated videos with documentary images, schematic animation and generative reconstruction; they could also manipulate explicit source annotations, measure comprehension and source discrimination, and use delayed tests with new items. The present field design supplies a useful baseline for those comparisons, not evidence that one representational technique is superior.

The absence of attitudinal change should be interpreted against high baseline affiliation and long organizational tenure. Findings concerning realism and engagement describe this live-action narrated video with AI-generated historical scenes, not synthetic-instructor lectures or fully AI-generated films. The positive association between realism and spontaneous recollection several weeks later suggests a potentially relevant design question, but it cannot establish whether realistic imagery caused memory persistence or simply covaried with pre-existing interest and appraisal of the screening.

4.1. Limitations

The study used a small, self-selected sample from one company and no comparison group. The limited and unequally difficult control items, low T1 reliability of form B, variable T2 timing, pretest sensitization, and imperfect

participant-code matching restrict inference. Knowledge items assessed factual recognition rather than conceptual understanding, transfer or capacity to distinguish invented imagery from documented history. Realism and engagement were measured concurrently, and neither realism nor collective viewing was manipulated. Voluntary activities were self-reported; several analyses were exploratory and uncorrected for multiplicity, and some estimates depended on one observation. The first author was employed by TIM, potentially increasing social-desirability bias; the study was not reviewed by an ethics committee.

5. Conclusions

Following a narrative video featuring AI-generated historical scenes, employees showed an increase in recognition of taught facts and retained part of that gain after 33–60 days. Perceived realism accompanied engagement and later recollection, but was not a robust indicator of factual learning or attitudinal change. These observational results support distinguishing learning experience, knowledge and retention in educational evaluation; they do not demonstrate the specific educational effectiveness of AI-generated reconstructions. Future controlled studies should compare visual representations, assess historical source literacy and test longer-term comprehension.

References

Aeon, B., & Lamertz, K. (2021). Those who control the past control the future: The dark side of rhetorical history. Organization Studies, 42(4), 575–593. https://doi.org/10.1177/0170840619844284

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. Personnel Psychology, 50(2), 341–358. https://doi.org/10.1111/j.1744-6570.1997.tb00911.x

Arkün-Kocadere, S., & Çağlar-Özhan, Ş. (2024). Video lectures with AI-generated instructors: Low video engagement, same performance as human instructors. The International Review of Research in Open and Distributed Learning, 25(3), 350–369. https://doi.org/10.19173/irrodl.v25i3.7815

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. Media Psychology, 12(4), 321–347. https://doi.org/10.1080/15213260903287259

Butler, A. C., Zaromb, F. M., Lyle, K. B., & Roediger, H. L., III. (2009). Using popular films to enhance classroom learning: The good, the bad, and the interesting. Psychological Science, 20(9), 1161–1168. https://doi.org/10.1111/j.1467-9280.2009.02410.x

Cho, H., Shen, L., & Wilson, K. (2014). Perceived realism: Dimensions and roles in narrative persuasion. Communication Research, 41(6), 828–851. https://doi.org/10.1177/0093650212450585

Doss, C., Mondschein, J., Shu, D., Wolfson, T., Kopecky, D., Fitton-Kane, V. A., Bush, L., & Tucker, C. (2023). Deepfakes and scientific knowledge dissemination. Scientific Reports, 13, Article 13429. https://doi.org/10.1038/s41598-023-39944-3

Firth, D. (1993). Bias reduction of maximum likelihood estimates. Biometrika, 80(1), 27–38. https://doi.org/10.1093/biomet/80.1.27

Gouthier, M. H. J., & Rhein, M. (2011). Organizational pride and its positive effects on employee behavior. Journal of Service Management, 22(5), 633–649. https://doi.org/10.1108/09564231111174988

Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. Frontiers in Psychology, 14, Article 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Haccoun, R. R., & Hamtiaux, T. (1994). Optimizing knowledge tests for inferring learning acquisition levels in single group training evaluation designs: The internal referencing strategy. Personnel Psychology, 47(3), 593–604. https://doi.org/10.1111/j.1744-6570.1994.tb01739.x

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. Statistics in Medicine, 21(16), 2409–2419. https://doi.org/10.1002/sim.1047

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. Educational Psychologist, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 6, 65–70. https://www.jstor.org/stable/4615733

Kearney, K. A., Hopkins, R. H., Mauss, A. L., & Weisheit, R. A. (1984). Self-generated identification codes for anonymous collection of longitudinal questionnaire data. Public Opinion Quarterly, 48(1B), 370–378. https://doi.org/10.1093/poq/48.1b.370

Klein, H. J., & Weaver, N. A. (2000). The effectiveness of an organizational-level orientation training program in the socialization of new hires. Personnel Psychology, 53(1), 47–66. https://doi.org/10.1111/j.1744-6570.2000.tb00193.x

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. Social Psychological and Personality Science, 8(4), 355–362. https://doi.org/10.1177/1948550617697177

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. Educational and Psychological Measurement, 70(4), 647–671. https://doi.org/10.1177/0013164409355699

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. Journal of Organizational Behavior, 13(2), 103–123. https://doi.org/10.1002/job.4030130202

Mar, R. A., Li, J., Nguyen, A. T. P., & Ta, C. P. (2021). Memory and comprehension of narrative versus expository texts: A meta-analysis. Psychonomic Bulletin & Review, 28(3), 732–749. https://doi.org/10.3758/s13423-020-01853-1

Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. Educational Psychology Review, 36(1), Article 8. https://doi.org/10.1007/s10648-023-09842-1

Netland, T., von Dzengelevski, O., Tesch, K., & Kwasnitschka, D. (2025). Comparing human-made and AI-generated teaching videos: An experimental study on learning effects. Computers & Education, 224, Article 105164. https://doi.org/10.1016/j.compedu.2024.105164

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. Journal of Applied Psychology, 93(2), 280–295. https://doi.org/10.1037/0021-9010.93.2.280

St. Hilaire, K. J., Chan, J. C. K., & Ahn, D. (2024). Guessing as a learning intervention: A meta-analytic review of the prequestion effect. Psychonomic Bulletin & Review, 31(2), 411–441. https://doi.org/10.3758/s13423-023-02353-8

Wertsch, J. V. (1998). Mind as action. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195117530.001.0001