Learning and Instruction 100 (2025) 102200

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

The "simple" view of learning from illustrated texts and videos

Mikko Haavisto [a,*], Janne Lepola [a], Tomi Jaakkola [b]

[a] Department of Teacher Education, University of Turku, Seminaarinkatu 1, 26100, Rauma, Finland
[b] Faculty of Education and Culture, Tampere University, Åkerlundinkatu 5, 33100, Tampere, Finland

ARTICLE INFO

Keywords:
Multimedia learning
Instructional practices
Primary education
Modality effect
Simple view of reading

STRUCTURED ABSTRACT

Background: The recent decline in children's reading skills in OECD regions poses challenges for traditional text-based learning. At the same time, teachers increasingly use videos in primary instruction. Despite these developments, limited research exists on how children's reading skills influence learning from videos versus illustrated texts in primary school classrooms.
Aims: This study investigates the roles of decoding ability and reading comprehension in learning from videos versus illustrated texts among fifth and sixth graders. It aims to determine to what degree these factors influence learning outcomes and cognitive load.
Sample: 109 children from grades 5–6 across three public primary schools.
Methods: In a within-subjects experiment, participants studied both illustrated texts and videos on two science topics. Their performance was measured through pre-, post-, and delayed tests. Mixed-effects models assessed the effect of modality and reading skills on learning outcomes and cognitive load.
Results: The children performed significantly better when learning from videos compared to illustrated texts, demonstrating higher delayed retention and lower cognitive load. There was no difference in retrieval from materials or transfer. Decoding ability and reading comprehension positively predicted learning outcomes and interacted with modality: the retention benefits of videos were more pronounced in children with lower reading skills than in those with higher reading skills.
Conclusions: The results indicate that videos are beneficial to most children across reading skill levels, especially those with weaker reading skills. This suggests that incorporating videos into primary school science instruction supports diverse learning needs associated with weaker reading skills.

1. Introduction

In the past decade, there has been an unprecedented worldwide decline in the reading skills of children. According to PISA results, today's 15-year-olds read at the level of 14-year-olds from a decade ago (OECD, 2023, p. 158). As education relies on textbook instruction, this deterioration in reading skills poses significant challenges. Luckily, at the same time, online videos have risen as an instructional approach to support children with weaker reading abilities. For example, a survey found that 90% of Spanish primary school teachers use videos weekly, primarily from YouTube (Pattier, 2021). Nevertheless, few studies have compared learning outcomes from videos and illustrated texts in primary school classroom settings (Reinwein, 2012). Moreover, despite the OECD-wide decline in reading skills, the effect of reading ability on learning from these materials has received little research attention. This study addresses these gaps by conducting experiments in fifth- and sixth-grade classrooms. In the experiments, children studied science topics from videos and illustrated texts and completed pre-, post-, and delayed tests of learning, and their reading ability was evaluated.

1.1. The simple view of reading and the modality effect

The present study conceptualizes reading ability through the simple view of reading, which posits that reading comprehension is the product of two factors: *decoding ability* and *linguistic comprehension* (Gough & Tunmer, 1986). Decoding refers to the ability to read isolated words fluently, and linguistic comprehension is the ability to interpret sentences from word information. Both components are necessary for reading comprehension, and neither is sufficient alone. Decoding without linguistic comprehension allows a person to recognize and pronounce words but leaves them unable to understand the meaning. Conversely, without decoding, the reader would not have word

* Corresponding author.
E-mail addresses: mijuhaa@utu.fi (M. Haavisto), janlep@utu.fi (J. Lepola), tomi.jaakkola@tuni.fi (T. Jaakkola).

https://doi.org/10.1016/j.learninstruc.2025.102200
Received 18 January 2025; Received in revised form 28 May 2025; Accepted 28 July 2025
Available online 14 August 2025

information to interpret writing in the first place. Therefore, in this framework, better reading comprehension predicts better learning from both videos and illustrated texts by including linguistic comprehension. However, decoding ability directly affects learning from only illustrated texts, as videos provide spoken audio that bypasses the need for decoding.

Although decoding and linguistic comprehension play distinct roles in reading comprehension, they are often correlated (Gough & Tunmer, 1986; Language and Reading Research Consortium, 2015; Lepola et al., 2016). Similarly, listening comprehension correlates with reading comprehension (Torppa et al., 2016), as they both depend on linguistic comprehension. Furthermore, recent research highlights that listening and reading comprehension processes exhibit both shared and modality-specific components, and word reading fluency contributes significantly to both reading and listening comprehension (Wolf et al., 2019). Therefore, decoding ability might also correlate with video learning outcomes.

The present study contributes to research on multimedia learning by examining the influence of reading skills on the modality effect (see meta-analyses by Ginns, 2005; Reinwein, 2012). The modality effect refers to the improved learning, when multimedia materials (e.g., videos) present information through multiple sensory modalities (e.g., pictures and spoken text utilize both sight and hearing) rather than a single modality (e.g., pictures and written text utilize only sight; Low & Sweller, 2014).

Segers et al. (2008) observed that verbal intelligence did not affect the modality effect. However, unlike broader measures like verbal intelligence and reading comprehension, decoding ability directly affects only learning from a written modality, making it the key differentiating factor for the modality effect. For this reason, the simple view of reading provides a useful framework for understanding how individual differences in reading skills affect learning from videos versus illustrated texts. The simplicity of this framework also makes it easily applicable to classrooms compared to other models, for example (a) the integrated model of text and picture comprehension (Schnotz, 2014), where written information is converted into phonological information through multiple processes, and (b) the reading systems network (Perfetti & Stafura, 2014), where decoding is embedded in a network of interdependent recursive processes.

The relationship between decoding ability and the modality effect has only been indirectly investigated by previous research. First, studies by Knoop-van Campen et al. (2018, 2019) found more evidence for the modality effect in children with dyslexia than typically developing children. Second, a prior study (Haavisto et al., 2023) using the same experimental design and materials as the present study showed that learning from illustrated texts resulted in poorer learning performance and higher cognitive load than videos. Furthermore, screen recordings from the same experiment revealed that (a) most children read more slowly than the progression speed of typical educational videos, and (b) the children's studying rate was highly uniform with videos but varied widely with illustrated texts (Haavisto & Jaakkola, 2025). These results suggest an intuitive explanation that children's weak reading skills lead to a slow reading pace, which enhances the modality effect.

The present study replicates and extends the two earlier classroom experiments mentioned above (Haavisto & Jaakkola, 2025; Haavisto et al., 2023). The studies explored the modality effect in primary school settings but had key limitations that prevented conclusions about the relationship between reading skills and the modality effect. First, the studies lacked direct measurement of children's reading comprehension and decoding. The present study assesses these reading skills and their potential impact on learning from videos and illustrated texts. Second, the prior studies were underpowered to examine interaction effects. The present study addresses this limitation with a larger sample, providing the statistical power needed to test interactions between modality and reading skills.

1.2. *Cognitive load theory and the modality effect*

Cognitive load theory (Leahy & Sweller, 2016) posits that low decoding ability increases the cognitive effort required for learning from written texts. Specifically, additional cognitive load arises from converting visual written text into verbal information within working memory (Wong et al., 2012). Thus, the benefit of learning from videos compared to illustrated texts could stem from reducing cognitive load caused by reading. This is especially true for children, as learning to read written text requires extensive effort and systematic practice (Sweller, 2020). For less-skilled readers, even graphemes may be processed as separate elements (i.e., chunks in working memory), thereby increasing cognitive load. In contrast, videos do not require decoding and may therefore be especially helpful for children with weaker decoding skills. Indeed, Reinwein and Tassé (2022) found that educated adults comprehended written multimedia messages faster than spoken ones, whereas the opposite was true for sixth graders.

Most studies on the modality effect have focused on university students (Mayer, 2020), a small, self-selected portion of the population with above-average reading abilities. This is relevant since automaticity in decoding frees cognitive resources for comprehension (Perfetti, 1985), reducing the cognitive load difference between listening and reading. While decoding is critical in the initial stages of reading, its influence diminishes as reading proficiency develops (Smith et al., 2021). The present study, conducted in public primary schools, provides a more representative sample of learners, especially regarding reading skills. In this context, a larger modality effect is expected, likely due to weaker decoding abilities.

1.3. *Studying with the material available in the classroom*

In typical primary school classrooms, children first complete exercises requiring them to find information from materials. After extensive practice with support from the materials, they take an exam without access to the materials to test their retention after a significant period of time has passed. This contrasts with many experimental studies, where children study the materials without specific exercises and then answer exam-like questions from memory, often during the same lesson, on topics they have had little time to practice (e.g., Haavisto et al., 2023; Herrlinger et al., 2017; Knoop-van Campen et al., 2018, 2019; Segers et al., 2008; Witteman & Segers, 2010). In general, multimedia learning research has focused mostly on how multimedia is presented rather than on generative activities that encourage active information processing during learning (Fiorella & Mayer, 2021). The present study addresses these limitations by mirroring real-world classroom practice: during the learning phase, students had access to the materials, but not during the delayed test. This design increases ecological validity by aligning with how knowledge is typically acquired and assessed in everyday science instruction.

Furthermore, the difference between experimental and real-world conditions poses a significant problem for generalizability, as the modality effect has been shown to be sensitive to various moderating variables (i.e., boundary conditions, Mayer & Fiorella, 2021). The replicated study by Haavisto et al. (2023) emphasized ecological validity by investigating fifth-grade science learning in authentic classroom settings, using familiar instructional materials and test question types. The present study builds on this foundation by adopting the same experimental environment, materials, and questions, while also addressing a key limitation: the earlier study's retention-based testing of new topics did not resemble classroom learning. By aligning the experiment with real-world conditions with respect to environment, learners, and procedures, the present study aims to control as many known and yet unknown moderators as possible, enhancing generalizability.

Since the primary difference between learning from illustrated texts and videos is that one is read and the other listened to, one might hypothesize that texts are better for tasks with the material available. For locating information, eye-tracking studies (Biedert et al., 2012; Simola et al., 2008) provide evidence for a faster type of reading, in which participants search for specific words without processing meaning (Brysbaert, 2019). Texts offer two advantages: readers can view multiple words simultaneously (Snell & Grainger, 2019), and eye movements are more precise and efficient than navigating videos using learner controls. This suggests that, given sufficient decoding ability, children would find information more easily from illustrated texts.

1.4. Research Questions and hypotheses

In addition to reading skills and cognitive load defined above, the present study assessed retrieval, transfer, and retention. In this study, retrieval refers to the learner's ability to locate and extract relevant information from the instructional materials to answer direct questions. The required information is explicitly presented in the available materials. Transfer refers to the learner's ability to apply acquired knowledge to novel contexts by answering open-ended questions that benefit from understanding the studied topic. Although the materials were available, they did not contain direct answers to the transfer questions. Retention refers to the learner's ability to answer direct questions based on long-term memory; it was assessed one week after studying, without access to the studied materials that contained the answers.

Research Question 1. How does studying with videos versus illustrated texts impact fifth and sixth graders' retrieval, retention, transfer, and cognitive load?

Hypothesis 1. posits that the modality effect is replicated, with videos imposing less cognitive load during initial learning. Hypothesis 2 states that illustrated texts will be superior in retrieval from materials, due to the advantages written texts provide for keyword search. No hypothesis is made concerning a modality effect in retention, as studying based on retrieval may be less beneficial with videos than with illustrated texts.

Research Question 2. To what extent do decoding ability and reading comprehension affect fifth and sixth graders' retrieval, retention, transfer, and cognitive load when studying videos versus illustrated texts?

Hypothesis 3. posits that children with weaker decoding ability or reading comprehension should benefit more from videos compared to illustrated texts across all outcome measures. This advantage is expected for three reasons: (a) weak decoding directly constrains learning through reading but not through listening, (b) weak decoding increases cognitive load during reading, and (c) decoding is a component of reading comprehension.

2. Materials and methods

2.1. Sampling and participants

The participants were 109 children (mean age 12.3 years, $SD = 0.6$) from public primary schools in Finland (Fig. 1). Three schools were selected through convenience sampling from local schools. From each school, one fifth-grade and one sixth-grade teacher were invited to participate; if a teacher declined, another at the same grade level was invited. Special education classes and classes with a language of instruction other than Finnish were excluded. Within participating classes, children who required external assistance (e.g., translation) to understand the Finnish instructions were excluded from the sample. The sample size was sufficient given the study's within-subjects design and use of mixed-effects analysis (power analyses in Section 2.5).

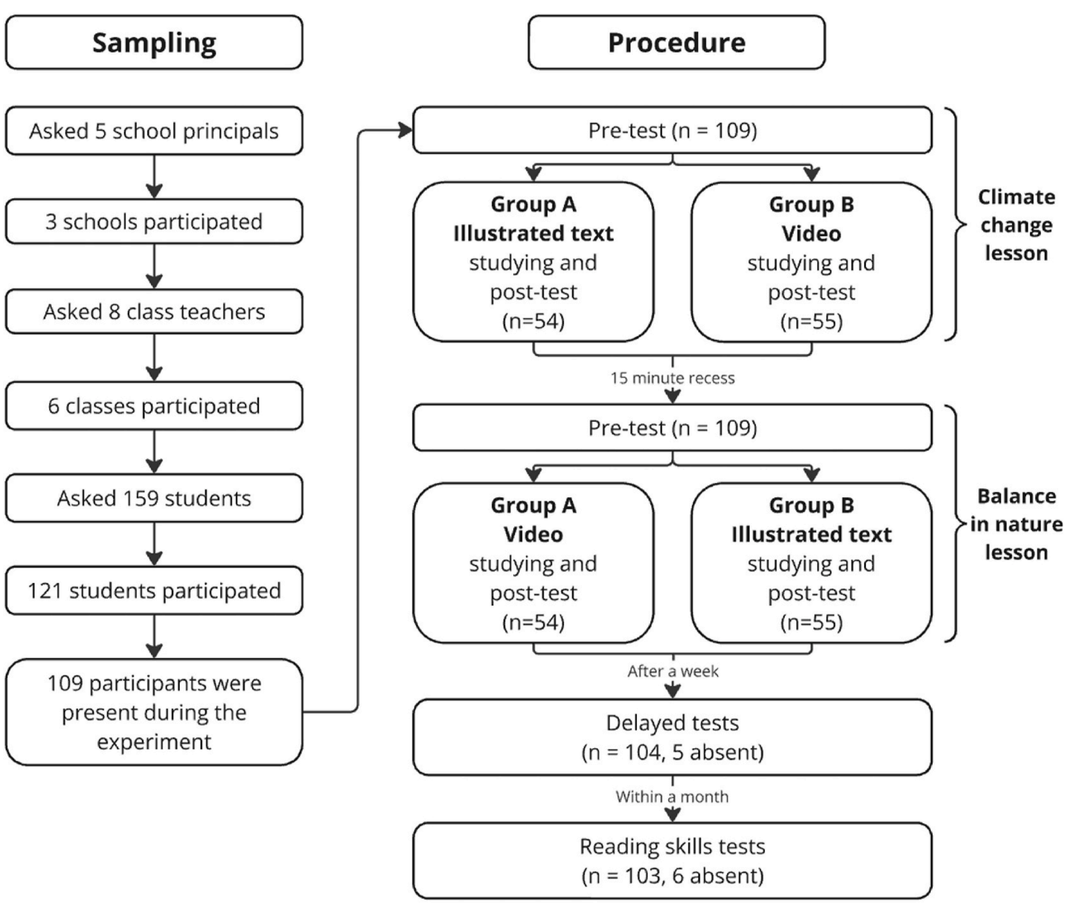

Fig. 1. Data collection.

The research was approved by the University of Turku ethics board and adhered to ethical principles for research involving children (Finnish National Board on Research Integrity TENK, 2019). Data collection was anonymous, with informed consent obtained from both the children and their guardians.

2.2. Data Collection procedure

Data collection spanned four 45-minute lessons for each participating class (Fig. 1). A confounded factorial design (Kirk, 2009) was used to assess children's learning during two science lessons conducted by the first author. Each of these two lessons began with 10 minutes dedicated to (a) randomly assigning children to research groups, (b) answering pre-test questions, and (c) explaining how to use the materials.

The remaining 35 minutes of each lesson were allocated to studying and post-testing. In the first lesson, group A studied climate change using an illustrated text, while group B used a video. After studying the material once from start to finish, the children received post-test questions, which they completed with the material available. The post-test started with a cognitive load self-assessment item. As is customary in Finnish schools, the children had a 15-minute recess after the lesson, during which they played in the schoolyard.

The second 45-minute lesson occurred after the recess. It followed the same structure as the first lesson, except that the topic was balance in nature, and this time group A used a video and group B used an illustrated text. Both modality conditions were conducted simultaneously in the same classroom to control for environmental variables.

The third lesson, conducted seven days later, assessed the children's long-term retention with delayed tests on both topics.

The fourth lesson, within one month of the experiment, assessed decoding ability and reading comprehension.

2.3. Learning materials

This study used the same video and illustrated text materials as those in Haavisto et al. (2023) Fig. 2, which addressed two science topics: climate change and balance in nature. The focus on science learning is particularly relevant, given that recent PISA results (OECD, 2023, p. 158) indicate a decline in children's science knowledge.

The materials contained identical word and picture content. For the climate change topic, both the illustrated text and the video included 474 words (for reference, English translation was 715 words) and 18 pictures; the video duration was 5 minutes and 55 seconds. The balance in nature materials included 466 words and 21 pictures; the video duration was 5 minutes and 37 seconds.

Both materials were accessed on the children's own iPads. The illustrated texts were presented as a slideshow in Google Slides, where text and pictures were displayed side by side. In the video format, the same pictures were accompanied by spoken audio (78 wpm in Finnish; for reference, English translation was 121 wpm) instead of written text and were presented on YouTube. The children had access to all learner controls (e.g., changing slides forward or backward or skipping to any slide, fast-forwarding, rewinding, scrubbing the timeline) associated with these platforms.

2.4. Measures

This study used the retrieval, retention, transfer, and cognitive load items from Haavisto et al. (2023). The post-tests and delayed tests used identical questions that assessed children's knowledge of the information presented in the materials, with the pre-test containing only a subset of these questions. The questions were of multiple types: multiple choice, pictorial, and closed-ended written questions (examples and item counts of each question type in Supplementary materials, Table S.1). These questions measured (a) retrieval in the post-test, where the children retrieved the information from the materials and (b) retention in the delayed test, as the children had to retain the information over the course of a week. Children received no feedback between the tests. In addition, the post-test included open-ended transfer questions, which evaluated the application of the instructional content—answers to these questions were not presented in the materials. The learning materials were available during the post-test, but not in the pre- or delayed tests. Table 1 presents the item-count, reliability, and descriptive statistics for each test. In all tests, one point was given for each correct answer, with the maximum score corresponding to the total number of items.

Following the initial study of the material, the participants self-assessed their level of mental effort using Paas's (1992) nine-point cognitive load scale. The reliability and validity of the scale are supported in studies by Ayres (2006) and Paas et al. (1994, 2003, 2008). The chosen measure and its timing support ecological validity: the cognitive load assessment was presented as the first question of the post-test to minimize disruption to the learning process.

In the above measures, the scoring criteria and the scorer (the first author) remained the same as in Haavisto et al. (2023). In that study, the inter-rater reliability was demonstrated in these tests with a 20% subsample (intra-class correlations: mean = .97, min = .92, max = 1.00). These intra-class correlations are in the excellent range (Koo & Li, 2016) supporting the reliability of the scoring process in the current study.

Reading comprehension was assessed using the Finnish Standardized Reading Test (Lindeman, 1998), which consisted of two informational

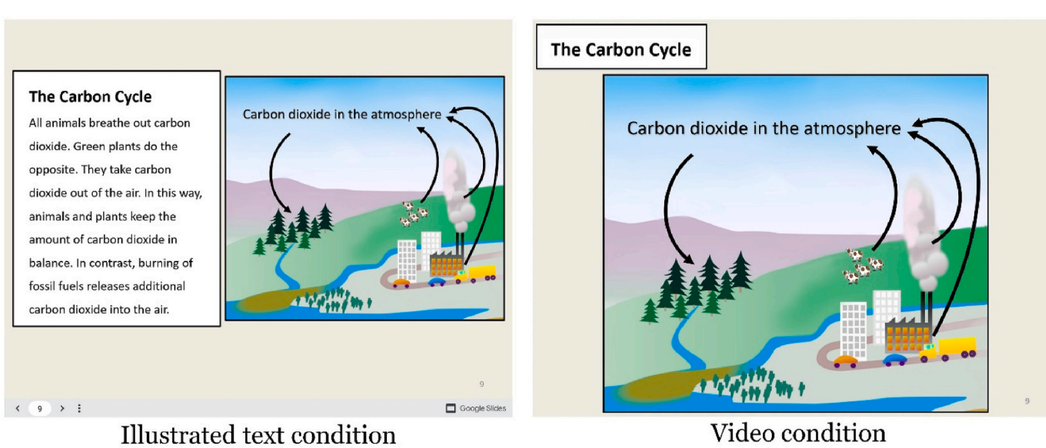

Fig. 2. An example of illustrated text and video materials.
Note. The same text that was written in the illustrated text was spoken in the video. Figure from Haavisto et al. (2023).

Table 1
Descriptive statistics and scale score reliabilities.

Test	Sample size	N of items	α	Mean	*SD*	Min	Max
CC learning cognitive load	108	1		4.22	1.38	1.00	9.00
BIN learning cognitive load	108	1		3.89	1.59	1.00	9.00
CC pre-test knowledge	109	10	.517	.117	.088	.000	.500
BIN pre-test knowledge	108	7	.168	.167	.202	.000	.429
CC post-test retrieval	109	22	.720	.764	.154	.333	1.00
BIN post-test retrieval	109	16	.777	.776	.186	.250	1.00
CC post-test transfer	109	12	.863	.221	.243	.000	.958
BIN post-test transfer	109	10	.795	.343	.235	.000	.875
CC delayed test retention	104	22	.740	.592	.177	.136	.955
BIN delayed test retention	104	16	.771	.554	.200	.063	1.00
Decoding test	103	146	.769[a]	.418[a]	.134	.144	.753
Reading comprehension test	100	24	.772	.730	.163	.125	1.00

Note. CC = climate change, BIN = balance in nature, α = Cronbach's alpha. Means, minimums, and maximums are proportions of correct answers, except for cognitive load. Raw scores can be calculated by multiplying the proportion with the number of items.

[a] Alpha was calculated from the subscores of four tests. Relatively low mean is due to time-limited tasks.

texts each accompanied by 12 multiple choice questions. Children could refer to the text for the entire duration of the test.

Decoding was measured using a validated and established test series (Nevala & Lyytinen, 2000), which included four time-limited tasks requiring the children to (a) separate individual words in a text without spaces, (b) separate syllables in words, (c) distinguish between actual words and pseudowords, and (d) recognize spelling errors. Students received one point for each correct word separation, syllable separation, pseudoword rejection, and recognized spelling error. The decoding score was the sum of correct responses across the four subtests.

Linguistic comprehension was not measured because (a) it was expected to similarly influence both modalities and (b) it overlaps with reading comprehension in the simple view of reading, contributing less unique explanatory value.

Table 1 shows that the reliability of measurements was sufficient (Taber, 2018), except for the pre-test. Pre-test statistics indicated low prior knowledge (M = 11.7%–16.7%) and low reliability. Since the same items showed sufficient reliability in the other tests, this suggests a floor effect rather than poor internal consistency. The same pattern occurred in Haavisto et al. (2023). Following their approach, the pre-test scores were excluded from the models. This decision improved both reliability and statistical power: random effects accounted for potential individual differences in prior knowledge, and including an unreliable covariate would have unnecessarily reduced power. Additionally, random assignment and the within-subjects design accounted for group differences (Section 2.6.1).

2.5. Analysis

Linear mixed-effect models were used to analyze within- and between-subject variables, and to control for and quantify all individual differences. These models efficiently used all available data, even in cases of student absences. Separate models were used for each dependent variable: cognitive load, retrieval from materials, transfer, and retention. Before analysis, these dependent variables were separately standardized within each topic to account for differences between topics. The fixed effects included the modality and decoding or reading comprehension, while the random effects were participant IDs.

Effect sizes are reported with R^2_{mar} (variance explained by fixed effects) and R^2_{con} (variance explained by fixed and random effects, Nakagawa & Schielzeth, 2013). All models accounted for heteroskedasticity using robust standard errors (White, 1980). Multicollinearity was negligible in models without interactions (VIFs <1.1) and acceptable in models with interactions (VIFs <2.55; O'Brien, 2007).

A power analysis of 1000 simulations was conducted for each model, showing sufficient statistical power, detecting Cohen's $d \geq .3$ effect sizes with 87.2% probability on average.

Analyses were conducted using R software (R Core Team, 2024) and the packages lme4 (Bates et al., 2015), performance (Lüdecke et al., 2021), clubSandwich (Pustejovsky & Tipton, 2018), and simr (Green & MacLeod, 2016).

2.6. Preliminary analyses

2.6.1. Prior knowledge

The children demonstrated low prior knowledge, answering correctly only 14.2% (SD = 7.0%) of the pre-test questions on average. A linear regression model showed no pre-test performance differences between the randomized experimental groups (p-values >.490).

2.6.2. Learning domain knowledge

Children's scores on a subset of questions included in the pre-tests show an understandable performance progression across test phases (Supplementary materials, Figure S.2). After studying, children achieved an average of 77.0% correct answers in the post-test with the material available (SD = 17.0%). A week later, their performance dropped to 54.3% correct answers in a delayed test without the material available (SD = 18.7%). Post hoc analyses (Supplementary materials, Table S.3) confirmed that the performance differences between test phases were large in effect size (R^2_{mar} = .363–.886, p-values <.001).

2.6.3. Intra-class correlations

High participant-level intra-class correlations for all dependent variables (*ICCs* .424–.546) led to the inclusion of participant random effects in mixed models. In contrast, the small classroom-level random effects were excluded (*ICCs* .013–.083).

2.6.4. Practicing for the delayed test

To reduce practicing outside of the experiment, the students were not informed about the delayed test. However, the children reported whether they had practiced the topics after the experiment. A one-tailed Welch's T-test showed that those 15 who reported practicing were not significantly better in retention than others (p = .922).

3. Results

3.1. Effect of modality on performance

Videos significantly outperformed illustrated texts in two outcome measures (Table 2): lower cognitive load, with a small effect size (R^2_{mar} = .046), and higher delayed retention, with a small effect size (R^2_{mar} = .016). Average cognitive load during the video condition was below the 'somewhat low effort' level, whereas during the illustrated text condition it exceeded that level.

Post-test transfer tasks showed no significant differences. Interestingly, videos performed no worse than texts in retrieval from materials, contrary to Hypothesis 2, which posited illustrated texts would facilitate easier location of relevant information. Notably, the minimal effect size indicates no practical difference in retrieval between modalities (R^2_{mar} = .003), with a slight trend favoring videos.

Table 2
The differences in learning and cognitive load between modalities.

Dependent variable	Video		Illustrated text		R^2_{con}	R^2_{mar}	B	SE	t
	M	SD	M	SD					
Initial study, cognitive load	3.74	1.54	4.36	1.39	.508	.046	.427	.096	4.46***
Post-test, retrieval	.779	.152	.760	.187	.509	.003	−.115	.095	−1.21
Post-test, transfer	.279	.242	.285	.252	.543	.000	.015	.092	0.16
Delayed test, retention	.598	.164	.549	.210	.556	.016	−.255	.092	−2.76**

Note. $n = 208–218$, n of groups $= 104–109$, $df = 103–108$. The retrieval and retention tests had the same items. The material was available in the post-test but not in the delayed test. Means are proportions of correct answers, except for cognitive load. Mixed-effects analysis used standardized scores.
$^{**}p < .01$, $^{***}p < .001$.

Participant-level random effects explained approximately half of the variance in all outcome measures (R^2_{con} - $R^2_{mar} = .462–.543$), showing that individual differences had a large effect on participants' performance with both video and illustrated text materials, thus aligning with previous results (Haavisto et al., 2023).

As shown in Fig. 3, videos outperformed illustrated texts in both research groups and across both topics of climate change and balance in nature. Videos also led to more consistent performance, with less individual variation in all learning metrics.

3.2. Decoding ability

As shown above, individual differences greatly influenced all outcome measures. To examine the specific role of decoding ability, it was analyzed as a fixed effect, separating it from other individual differences (Table 3). When controlling for modality with a fixed effect, decoding improved retrieval from materials, transfer, and retention with medium to large effect sizes ($\Delta R^2_{mar} = .169$, .086, and .098, respectively).

However, despite the notable variance explained by decoding, the total variance explained by the models showed minimal increases (ΔR^2_{con}-values $< .011$). This indicates that some variance previously explained by random effects was reassigned to decoding, demonstrating that random effects already effectively accounted for individual differences, such as variations in decoding ability. Consequently, it is understandable that the main effect of modality remained largely unchanged with the inclusion of decoding and their interaction.

To determine which modality is more effective for children with varying decoding abilities, interaction effects were analyzed. Table 3 shows that the interaction predicted retention with a small effect size ($\Delta R^2_{mar} = .013$): while videos improved retention overall, they were particularly beneficial for children with lower decoding abilities.

This interaction is illustrated in Fig. 4, where each participant's standardized retention score in the illustrated text condition is subtracted from their corresponding score in the video condition. Among children in the lowest 25% of decoding ability, retention was substantially higher in the video condition. In contrast, for children in the highest 25%, the difference between modalities was minimal. Although other interactions were not significant, a similar pattern was observed in transfer performance, with a trend approaching significance ($p = .071$), also favoring videos for children with weaker decoding abilities.

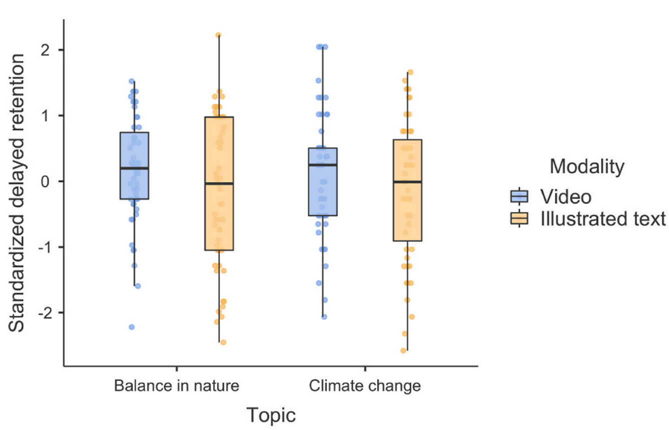

Fig. 3. Standardized delayed retention across modalities and topics.

Decoding did not influence cognitive load when modality was controlled, nor did the interaction between modality and decoding significantly predict cognitive load. These results are surprising: they contradict Hypothesis 3, which predicted children with better decoding ability would experience less effort learning from texts (Section 1.2).

Exploratory analyses provided further insights into the relationship between modality, decoding, and retention. An interaction between modality and gender revealed that videos were more effective for boys in retention ($\Delta R^2_{mar} = .014$, $p = .012$). Similarly, grade and modality interacted, with fifth graders benefiting more from videos than sixth graders in retention ($\Delta R^2_{mar} = .010$, $p = .041$). These results are reasonable, as girls exhibited higher decoding than boys ($R^2 = .084$, $p = .004$), and sixth graders had somewhat higher decoding than fifth graders, although this difference was not significant ($p = .154$).

3.3. Reading comprehension

The same mixed-effects models described in Section 3.2 were reanalyzed, but reading comprehension replaced decoding (Supplementary materials, Table S.4). After controlling for modality, reading comprehension improved retrieval from materials, transfer, and retention with medium to large effect sizes ($\Delta R^2_{mar} = .222$, .115, and .276, respectively; p-values $< .004$). These effects repeated the same patterns observed for decoding but exhibited much greater effect sizes. Like decoding, reading comprehension interacted with modality, though the interaction was weaker: the retention benefits of video were more pronounced among children with lower reading comprehension ($\Delta R^2_{mar} = .009$, $p = .036$). This is illustrated in Fig. 5; videos especially benefited children in the lowest 25% of reading comprehension, with smaller differences across higher comprehension levels. Additionally, similar to decoding, reading comprehension did not affect cognitive load, and other interactions were not significant.

These results highlight that while children's reading comprehension has qualitatively similar but stronger effects than decoding on learning outcomes, decoding has a more decisive role in determining which modality is more beneficial. These findings are consistent with the simple view of reading, which conceptualizes decoding as a subcomponent of reading comprehension. As such, reading comprehension provides a broader advantage for learning from both illustrated texts and videos, whereas decoding specifically enhances learning in the illustrated text condition.

Finally, the effect of reading comprehension was evaluated after controlling for modality, decoding, and the interaction between modality and decoding with fixed effects. Reading comprehension emerged as a significant predictor for retrieval, transfer, and retention ($\Delta R^2_{mar} = .100$, .047, .178, respectively; p-values $< .022$; Supplementary materials, Table S.5). In these models, the effect of modality remained significant

Table 3
Effect of modality (MOD) and decoding (DEC) on cognitive load (CL), retrieval (RET), transfer (TRA), and delayed retention (DEL).

Model	R^2_{con}	R^2_{mar}	Coefficient	*B*	*SE*	*t*	*df*	*p*
CL ~ MOD + DEC	.493	.042	MOD	.400	.098	4.07	101.4	<.001
			DEC	−.028	.082	−0.34	37.8	.737
CL ~ MOD * DEC	.489	.042	MOD	.400	.099	4.06	100.4	<.001
			DEC	−.030	.094	−0.32	37.2	.749
			MOD:DEC	.005	.086	0.05	37.2	.957
RET ~ MOD + DEC	.520	.172	MOD	−.092	.096	−0.96	102.0	.334
			DEC	.412	.060	6.91	38.0	<.001
RET ~ MOD * DEC	.518	.172	MOD	−.092	.096	−0.96	101.0	.341
			DEC	.388	.060	6.48	38.0	<.001
			MOD:DEC	.047	.090	0.52	38.0	.607
TRA ~ MOD + DEC	.546	.086	MOD	.004	.095	0.05	102.0	.964
			DEC	.300	.083	3.61	38.0	<.001
TRA ~ MOD * DEC	.554	.092	MOD	.004	.095	0.05	101.0	.964
			DEC	.223	.095	2.35	38.0	.024
			MOD:DEC	.153	.082	1.86	38.0	.071
DEL ~ MOD + DEC	.556	.114	MOD	−.244	.097	−2.53	97.0	.013
			DEC	.316	.083	3.82	36.7	<.001
DEL ~ MOD * DEC	.579	.127	MOD	−.243	.094	−2.59	96.0	.011
			DEC	.201	.089	2.26	36.7	.030
			MOD:DEC	.230	.099	2.32	36.7	.026

Note. Modality was coded illustrated text = 1, video = 0. Other variables were standardized. Colon denotes interaction. All models included random effects for participants.

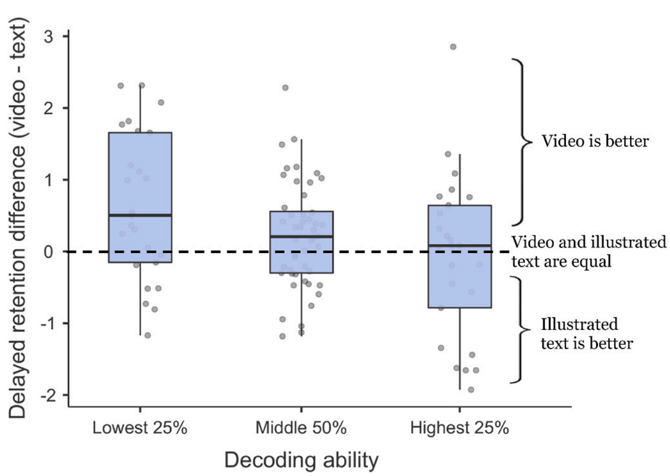

Fig. 4. Standardized delayed retention difference score (video minus illustrated text) across decoding ability percentiles.

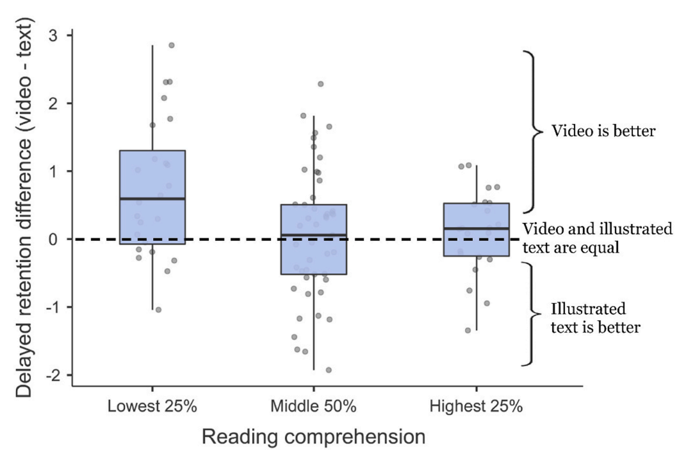

Fig. 5. Standardized delayed retention difference score (video minus illustrated text) across reading comprehension percentiles.

with respect to retention and cognitive load (p-values <.012). Notably, the interaction between decoding and modality also remained significant ($p = .034$) even after accounting for reading comprehension. This finding again highlights that decoding specifically supports learning from illustrated texts, while reading comprehension predicts learning from both modalities. Lastly, it was observed that the correlation between reading comprehension and decoding was high ($r = .496$, Supplementary materials, Table S.6), aligning with empirical research and the theoretical framework of the simple view of reading (Section 1.1).

4. Discussion

The purpose of this within-subject experimental study was to compare the effectiveness of videos and illustrated texts for learning science topics and to examine how reading skills influence the relative effectiveness of these learning materials. A total of 109 fifth and sixth graders participated, and their retrieval, transfer, retention, and cognitive load were evaluated using pre-, post-, and delayed tests. The experiment was conducted in authentic classroom settings with materials closely resembling typical primary school resources. The study revealed several key findings: (a) learning from videos led to better retention and lower cognitive load compared to illustrated texts, (b) videos were as effective as illustrated texts for retrieval of information from the materials, and (c) reading abilities interacted with modality, influencing retention: the benefit of videos was more pronounced in children with weaker decoding skills.

4.1. Videos enhance retention and reduce cognitive load

Consistent with Hypothesis 1 and prior multimedia research (Ginns, 2005; Haavisto et al., 2023; Mayer, 2020; Reinwein, 2012), videos were more effective than illustrated texts in science learning among fifth and sixth graders. The modality effects align with cognitive load theory (Leahy & Sweller, 2016; Sweller, 2020), which posits that presenting information through both visual and auditory channels reduces cognitive load and enhances learning.

Herrlinger et al. (2017) similarly observed a modality effect for fourth graders but suggested it may have been due to the materials being studied for only 8 minutes. They argued that this limited time may have disproportionately hindered learning from illustrated texts, as the young students lacked sufficient time to select, organize, and integrate words and pictures, as described by the cognitive theory of multimedia learning (Mayer, 2020). The present study addressed this limitation by allowing 35 minutes for studying in both conditions. The modality effect persisted, demonstrating its robustness among younger students in a lesson-length setting.

The findings descriptively suggest that videos promote more consistent engagement with the learning material, as students exhibited less individual variation across learning metrics with videos compared to illustrated texts. This consistency may stem from videos uniformly guiding and scaffolding students through the content, whereas reading rates can vary widely among individuals, affecting the study of illustrated texts (Haavisto & Jaakkola, 2025).

The transient information effect, as described by Singh et al. (2012), predicts that videos impose higher cognitive load because of their fleeting content that demands additional working memory resources to process and integrate disappearing information. However, in this study, videos imposed less cognitive load than illustrated texts. Furthermore, Kim and Phillips (2014) emphasized that comprehension of transient information relies on sustained attention and inhibitory control, making individual differences a key factor. Yet, random effects controlled for these differences in the present study. Thus, the cognitive load differences between modalities do not seem to be explained by differences in transience or attentional demands.

4.2. Videos and illustrated texts are comparably effective for retrieval

Contrary to Hypothesis 2, videos were as effective as illustrated texts for retrieval tasks with the material available. Texts were expected to have an advantage due to easier keyword searching and scanning for specific information (Section 1.3). However, the results showed a slight, albeit not statistically significant, benefit for videos in retrieval. This combination of no differences in retrieval and a benefit of videos in retention indicates that the children not only focused on fact-finding in the video condition but also on broader learning. Possibly, they had to engage with more content to find specific information, enhancing retention. Prior research supports this, as Haavisto & Jaakkola (2025) showed that learners using videos covered more content than those using illustrated texts within the same study time.

A key difference between the previous study (Haavisto et al., 2023) and the present study was that children had the material available during the post-test, unlike in the previous study. Despite this, a similar pattern of results was replicated: a modality effect was observed for retention and cognitive load, but not for transfer. However, the modality effects were weaker across all measures in the present study, suggesting that having the material available reduces the benefits of videos over illustrated texts.

4.3. Comparing outcomes across studies with different material availability

Practicing with the material available is grounded on evidence-based benefits of generative active processing in multimedia learning (Fiorella & Mayer, 2021; Fyfield et al., 2022). As such, it is valuable to compare studying based on retrieval from materials in the present study with the retention-based studying in the replicated study (Haavisto et al., 2023). Pre-test scores support this comparison, showing a minimal 2.3% difference in prior knowledge between the studies.

Unsurprisingly, the present study observed 26.3% higher post-test scores with the material available. More compellingly, participants scored 8.3% higher in the delayed retention test (without material in both studies). Retrieving relevant knowledge during the post-test likely reinforced the learning evaluated in the delayed test. Cognitive load during initial studying was 13.6% lower than in Haavisto et al. (2023), possibly because memorizing content for the retention test demands more effort than studying with the expectation of accessing the materials later. These findings are relevant for real-world classroom contexts, where children do not commonly complete exam-like tests for new topics. Such tests could increase stress and waste time by asking children to recall information they have not yet learned.

However, studying with the material available was not entirely beneficial, as transfer performance was 26.4% worse in the present study. This may be because children (a) spent more time on retrieval tasks, leaving less time for subsequent transfer tasks, or (b) attempted to locate answers in the materials, which proved ineffective since the transfer questions were designed to extend beyond the provided information. Overall, the findings reveal both advantages and drawbacks of studying based on retrieval from materials.

4.4. The advantage of videos increases with lower reading skills

The benefit of videos was even more pronounced in children with weaker decoding skills, supporting Hypothesis 3. This finding adds a novel contribution to multimedia research on primary school children learning science topics in classroom settings.

The results regarding decoding and reading comprehension highlight the distinct roles of children's reading skills in learning from videos and illustrated texts. Overall, reading comprehension was a stronger predictor of learning outcomes than decoding, which is reasonable as reading comprehension is a product of decoding and linguistic comprehension (Section 1.1). However, the weaker interaction between modality and reading comprehension suggests that reading comprehension enhances learning across modalities. In contrast, the stronger interaction between modality and decoding supports the notion that decoding is a key factor for the modality effect. Decoding directly constrains learning from written texts but not from spoken text. Furthermore, decoding strongly correlated with reading comprehension, meaning higher decoding also predicts better learning from videos.

The above findings align with previous research (Section 1.1) and extend it by directly showing how individual differences in decoding ability influence learning from videos versus illustrated texts in primary school students. Knoop-van Campen et al. (2018, 2019) suggest that children with dyslexia benefit from spoken multimedia. The present study demonstrated that, more broadly, among fifth and sixth grade children with varying reading levels, children with weaker decoding skills remembered more of what they learned from videos, while the benefit was less for those with average decoding and almost nonexistent for those with high decoding. Thus, videos were not harmful for any group of readers and led to more equitable retention across student skill levels.

Previous studies on the modality effect have also identified reading comprehension as a predictor of improved multimedia learning in primary school. Herrlinger et al. (2017) reported that fourth graders' reading comprehension explained a large part of variance in retention and transfer scores ($R^2 = .28$–$.37$). Similarly, Witteman and Segers (2010) found that better reading comprehension among sixth graders predicted better overall learning from written and spoken multimedia. However, the influence of reading comprehension on learning between the modalities was not examined in either study, overlooking a key interaction that was found to explain learning differences in the present study.

Interestingly, adding decoding and reading comprehension into the statistical model did not remove the significance of the modality. This suggests an additional benefit of videos beyond bypassing the conversion of written word forms into a phonological code in working memory. This is especially true with respect to cognitive load. Decoding was not related to cognitive load directly, nor through interaction with modality, whereas modality had the strongest effect on cognitive load out of all dependent variables. This is surprising since interpreting written text is considered more cognitively demanding for children with low decoding ability (Section 1.2). Instead, the results suggest that variability in children's decoding ability does not explain the reduced cognitive load in videos.

4.5. Limitations and directions for future research

This study did not experimentally test the effect of material availability on learning. Although studying based on retrieval from materials was compared to retention-based studying in a highly similar research setting (Haavisto et al., 2023), future research should directly manipulate material availability to confirm its impact.

The use of learner controls in the materials was also not analyzed. Prior work that analyzed screen recordings of children studying the same materials without specific questions (Haavisto & Jaakkola, 2025) suggests limited navigation during studying. However, in the present study, the retrieval questions directly required learners to browse the instructional content, which likely increased navigation. Future research should examine how navigation is used when materials are available during testing.

Additionally, the study did not assess the impact of linguistic comprehension on the modality effect. By measuring only two components of the simple view of reading, it was not possible to empirically validate whether the relationships between the components align with the theoretical model. Future work should test whether these relationships hold in multimedia contexts.

Lastly, this study focused on primary school science learning. It is plausible that the findings extend to other domains involving expository texts with diagrams and pictures, and particularly to younger students due to their developing reading skills. Further research is needed in these areas.

4.6. Conclusion

The aim of this study was to address gaps in prior research on multimedia learning, which has (a) primarily focused on retention-based studying, (b) predominantly studied groups with higher reading skills, and (c) overlooked measuring the influence of reading skills on the modality effect. This study demonstrates that videos are a more efficient format for learning science topics than illustrated texts, leading to better long-term retention with less mental effort within the same study duration. As expected, the retention advantage of videos was more pronounced among children with lower reading skills. However, contrary to expectations, videos were as effective as illustrated texts for retrieval from materials.

Prior research highlights the sensitivity of the modality effect to contextual changes. This study investigated the modality effect within the context of studying with the material available in the classroom. The findings demonstrate that the modality effect persists for most children, despite individual differences. Notably, videos are a viable instructional approach to address individual science learning needs, particularly for weaker readers. Theoretically, these findings offer perspectives on the aptitude-treatment interactions documented among beginning readers (Connor, 2011).

The study contributes to multimedia learning and cognitive load theory by demonstrating that the reduction in cognitive load and the improvement in learning with videos are separate effects, aligning with Haavisto et al. (2023). Additionally, while reading skills partially explain the benefits of videos, the lower cognitive load associated with videos is not due to bypassing the need to decode written text into phonological information. Instead, a significant role is played by advantages that are independent of reading skills and other individual differences—one possible example being the automatic progression of videos, which reduces the need for active cognitive control during studying.

The findings have practical significance for educators. First, the materials and scenarios closely reflect real-world classroom environments, enhancing generalizability. Second, selecting between modalities and study methods is a practical decision that teachers can make with minimal additional resources, especially given the widespread availability of online videos. Third, decoding tests are common and easy to implement, allowing teachers to identify students who are most likely to benefit from video-based learning.

CRediT authorship contribution statement

Mikko Haavisto: Writing – original draft, Methodology, Formal analysis, Writing – review & editing, Visualization, Investigation, Conceptualization. **Janne Lepola:** Conceptualization, Writing – review & editing. **Tomi Jaakkola:** Writing – review & editing, Conceptualization, Methodology.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used GPT-4o to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of interest

Declarations of interest: none.

Acknowledgements

The authors are deeply grateful to the participating children and teachers. This work was supported through a salaried doctoral researcher position funded by the Doctoral Programme on Learning, Teaching and Learning Environments Research (OPPI), University of Turku Graduate School.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2025.102200.

Data availability

The authors do not have permission to share data.

References

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16*(5), 389–400. https://doi.org/10.1016/j.learninstruc.2006.09.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Biedert, R., Hees, J., Dengel, A., & Buscher, G. (2012). A robust realtime reading-skimming classifier. In *Proceedings of the symposium on eye tracking research and applications* (pp. 123–130). https://doi.org/10.1145/2168556.2168575

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language, 109*, Article 104047. https://doi.org/10.1016/j.jml.2019.104047

Connor, C. M. (2011). Child characteristics–instruction interactions: Implications for students' literacy skills development in the early grades. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 3). Guilford Publications.

Finnish National Board on Research Integrity TENK. (2019). The ethical principles of research with human participants and ethical review in the human sciences in Finland. *Finnish National Board on Research Integrity TENK Guidelines 2019*. No. 3 https://tenk.fi/sites/default/files/2021-01/Ethical_review_in_human_sciences_2020.pdf.

Fiorella, L., & Mayer, R. E. (2021). The generative activity principle in multimedia learning. In R. E. Mayer, & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 339–350). Cambridge University Press. https://doi.org/10.1017/9781108894333.036.

Fyfield, M., Henderson, M., & Phillips, M. (2022). Improving instructional video design: A systematic review. *Australasian Journal of Educational Technology, 38*(3). https://doi.org/10.14742/ajet.7296. Article 3.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*(4), 313–331. https://doi.org/10.1016/j.learninstruc.2005.07.001

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Green, P., & MacLeod, C. J. (2016). Simr: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Haavisto, M., & Jaakkola, T. (2025). Fifth graders study videos faster than illustrated texts: Studying rate, review strategies, and learning implications. *PsyArXiv*. https://doi.org/10.31234/osf.io/mztpw_v1.

Haavisto, M., Jaakkola, T., & Lepola, J. (2023). Video outperforms illustrated text: Do old explanations for the modality effect apply in a learner-paced fifth-grade classroom context? *Computers & Education, 199*, 104775. https://doi.org/10.1016/j.compedu.2023.104775

Herrlinger, S., Höffler, T. N., Opfermann, M., & Leutner, D. (2017). When do pictures help learning from expository text? Multimedia and modality effects in primary schools. *Research in Science Education, 47*(3), 685–704. https://doi.org/10.1007/s11165-016-9525-y

Kim, Y.-S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly, 49*(3), 269–281. https://doi.org/10.1002/rrq.74

Kirk, R. E. (2009). Experimental design. In R. Millsap, & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 23–45). SAGE Publications Ltd. https://doi.org/10.4135/9780857020994.n2.

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2018). The modality and redundancy effects in multimedia learning in children with dyslexia. *Dyslexia, 24*(2), 140–155. https://doi.org/10.1002/dys.1585

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2019). Modality and redundancy effects, and their relation to executive functioning in children with dyslexia. *Research in Developmental Disabilities, 90*, 41–50. https://doi.org/10.1016/j.ridd.2019.04.007

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*(2), 151–169. https://doi.org/10.1002/rrq.99

Leahy, W., & Sweller, J. (2016). Cognitive load theory and the effects of transient information on the modality effect. *Instructional Science, 44*(1), 107–123. https://doi.org/10.1007/s11251-015-9362-9

Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., & Niemi, P. (2016). Early Oral Language Comprehension, Task Orientation, and Foundational Reading Skills as Predictors of Grade 3 Reading Comprehension. *Reading Research Quarterly, 51*(4), 373–390. https://doi.org/10.1002/rrq.145

Lindeman, J. (1998). ALLU. Ala-asteen lukutesti. Tekniset tiedot. *Turun yliopisto, Oppimistutkimuksen keskus*.

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 227–246). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.012.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781316941355

Mayer, R. E., & Fiorella, L. (2021). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer, & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 243–260). Cambridge University Press. https://doi.org/10.1017/9781108894333.025.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nevala, J., & Lyytinen, H. (2000). Sanaketjutesti. *Niilo Mäki Instituutti & Jyväskylän yliopiston lapsitutkimuskeskus*.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

OECD. (2023). *PISA 2022 results (volume I): The State of Learning and Equity in Education*. OECD. https://doi.org/10.1787/53f23881-en

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Paas, F., Ayres, P., & Pachman, M. (2008). Assessment of cognitive load in multimedia learning: Theory, methods, and applications. In D. H. Robinson, & G. Schraw (Eds.), *Recent innovations in educational technology that facilitate student learning* (pp. 11–35). Information Age Publishing.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F., van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*(1), 419–430. https://doi.org/10.2466/pms.1994.79.1.419

Pattier, D. (2021). Teachers and YouTube: The use of video as an educational resource. *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education, 16*(1). https://doi.org/10.6092/issn.1970-2221/11584. Article 1.

Perfetti, C. (1985). *Reading ability*. Oxford University Press.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Pustejovsky, J. E., & Tipton, E. (2018). Small-Sample methods for cluster-robust variance estimation and Hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics, 36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/.

Reinwein, J. (2012). Does the modality effect exist? And if so, which modality effect? *Journal of Psycholinguistic Research, 41*(1), 1–32. https://doi.org/10.1007/s10936-011-9180-4

Reinwein, J., & Tassé, S. (2022). Modality effects examined by means of an online sentence-picture comparison task. *Journal of Psycholinguistic Research, 51*(3), 521–542. https://doi.org/10.1007/s10936-022-09849-9

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.006.

Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology, 22*(3), 375–387. https://doi.org/10.1002/acp.1413

Simola, J., Salojärvi, J., & Kojo, I. (2008). Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cognitive Systems Research, 9*(4), 237–251. https://doi.org/10.1016/j.cogsys.2008.01.002

Singh, A.-M., Marcus, N., & Ayres, P. (2012). The transient information effect: Investigating the impact of segmentation on spoken and written text. *Applied Cognitive Psychology, 26*(6), 848–853. https://doi.org/10.1002/acp.2885

Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading Psychology, 42*(3), 214–240. https://doi.org/10.1080/02702711.2021.1888348

Snell, J., & Grainger, J. (2019). Readers are parallel processors. *Trends in Cognitive Sciences, 23*(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research & Development, 68*(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Torppa, Georgiou, Lerkkanen, Niemi, Poikkeus, & Nurmi. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly, 62*(2), 179. https://doi.org/10.13110/merrpalmquar1982.62.2.0179

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838. https://doi.org/10.2307/1912934

Witteman, M. J., & Segers, E. (2010). The modality effect tested in children in a user-paced multimedia environment. *Journal of Computer Assisted Learning, 26*(2), 132–142. https://doi.org/10.1111/j.1365-2729.2009.00335.x

Wolf, M. C., Muijselaar, M. M. L., Boonstra, A. M., & de Bree, E. H. (2019). The relationship between reading and listening comprehension: Shared and modality-specific components. *Reading and Writing, 32*(7), 1747–1767. https://doi.org/10.1007/s11145-018-9924-8

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction, 22*(6), 449–457. https://doi.org/10.1016/j.learninstruc.2012.05.004

Mikko Haavisto[a] is a Doctoral Researcher with research interests in learning applications of multimedia in classroom environments.

Janne Lepola[a], a Professor of Educational Psychology, researches the development of early comprehension skills and teacher-child interaction.

Tomi Jaakkola[b], an Associate Professor of Teacher Education, studies learning and interest in technology-enhanced STEM education