Successful learning with whiteboard animations – a question of their procedural character or narrative embedding?

Schneider, Sascha ; Krieglstein, Felix ; Beege, Maik ; Rey, Günter Daniel

Abstract: Although whiteboard animations are increasingly used for educational purposes, there is little empirical evidence as to why such animations can enhance learning. To specify essential elements, their dynamic visual presentation, as well as their narrative embedding, were found to be theortically important. In a first Experiment (N = 133) with a 2 (presentation mode: static pictures vs. progressive drawing) x 2 (narrative context: with vs. without a narrative) between-subject factorial design, motivational, cognitive, affective variables, as well as learning outcomes, of secondary school students were measured. Results revealed that progressive drawing, as well as a narrative context, are mostly associated with an increase in learning-relevant variables. In a second experiment with the same sample and the same experimental design but a different whiteboard animation, results from Experiment 1 generalize to another learning content. Again, a progressive drawing, as well as a narrative context within whiteboard animation, fostered learning relevant variables as well as learning outcomes. Results are discussed considering the cognitive theory of multimedia learning, the contiguity effect as well as the instructional design theory of anchored instruction.

DOI: https://doi.org/10.1016/j.heliyon.2023.e13229

Originally published at:
Schneider, Sascha; Krieglstein, Felix; Beege, Maik; Rey, Günter Daniel (2023). Successful learning with whiteboard animations – a question of their procedural character or narrative embedding? Heliyon, 9(2):e13229.
DOI: https://doi.org/10.1016/j.heliyon.2023.e13229

Heliyon 9 (2023) e13229

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Successful learning with whiteboard animations – A question of their procedural character or narrative embedding?

Sascha Schneider [a,*], Felix Krieglstein [b], Maik Beege [c], Günter Daniel Rey [b]

[a] Educational Technology, Institute of Education, University of Zurich, Switzerland
[b] Psychology of Learning with Digital Media, Institute for Media Research, Chemnitz University of Technology, Germany
[c] Digital Media in Education, Department of Psychology, Freiburg University of Education, Germany

ARTICLE INFO

Keywords:
Whiteboard animations
Cognitive theory of multimedia learning
Presentation mode
Narrative context
Learning outcomes

ABSTRACT

Although whiteboard animations are increasingly used for educational purposes, there is little empirical evidence as to why such animations can enhance learning. To specify essential elements, their dynamic visual presentation, as well as their narrative embedding, were found to be theortically important. In a first Experiment ($N = 133$) with a 2 (presentation mode: static pictures vs. progressive drawing) x 2 (narrative context: with vs. without a narrative) between-subject factorial design, motivational, cognitive, affective variables, as well as learning outcomes, of secondary school students were measured. Results revealed that progressive drawing, as well as a narrative context, are mostly associated with an increase in learning-relevant variables. In a second experiment with the same sample and the same experimental design but a different whiteboard animation, results from Experiment 1 generalize to another learning content. Again, a progressive drawing, as well as a narrative context within whiteboard animation, fostered learning relevant variables as well as learning outcomes. Results are discussed considering the cognitive theory of multimedia learning, the contiguity effect as well as the instructional design theory of anchored instruction.

1. Introduction

Looking at classroom contexts in which teachers write instructions on black- or whiteboards, digital animations in which instructions are dynamically drawn on something resembling a whiteboard become increasingly popular. In addition to institutional contexts (e.g. Ref. [1]), such whiteboard animations are often used in informal instructional contexts (for example on *YouTube* [2]). Previous research suggests that whiteboard animations offer a learning advantage over other instructional media such as narrated slides or text presentations. However, such media comparisons must be viewed critically in terms of their informative value, as no precise conclusions can be drawn about the specific design. In this context, concrete design recommendations based on empirical support for whiteboard animations and how viewers perceive such animations are rare (e.g. Ref. [3]). In consequence, recommendations for designing whiteboard animations can only be derived from research concerning the use of general animations in instructional settings (e.g. Ref. [4]). Based on the literature on dynamic media, a dynamic presentation (i.e., a time-varying visual presentation of information) and a narrative context (i.e., using a story to embed information) were identified as essential components in terms of the visual as well as verbal design of whiteboard animations. In an experimental series, these two components were

* Corresponding author.
E-mail address: sascha.schneider@uzh.ch (S. Schneider).

https://doi.org/10.1016/j.heliyon.2023.e13229
Received 13 June 2022; Received in revised form 9 January 2023; Accepted 21 January 2023
Available online 26 January 2023

examined under controlled conditions and further explained by cognitive, motivational, and affective variables to provide specific theoretical and practical implications. The results of these experiments will help instructional designers to decide on the specific conditions to use whiteboard animations in instructional contexts but also help research to shed light on the effects of dynamic features in learning media.

2. Literature review

2.1. Learning with animations

Animations are often used in teaching and learning to represent dynamic processes (e.g., biological or mechanical processes), which would be difficult to show with static images (for a recent review, see Ref. [4]). Within an animation, pictorial frames are artificially strung together, so that learners perceive a continuous movement over time ([5,6]). Hereby, animations help learners to build up a mental model of the learning content, whereby in static media, transitions between pictures must be filled by the learner causing cognitive resources ([7]). In this vein, animations are characterized by two important features – their visuospatial and spatiotemporal organization (e.g. Ref. [8]). Hereby, dynamic visualizations can portray the spatial arrangement of graphic entities and their organization (e.g., components of a clockwork [9]). In terms of the spatiotemporal organization, animations offer the possibility to show the chronological sequence of events (e.g., chemical processes within the photosynthesis [10]). Such changes in time and space cannot be adequately represented in static pictures what justifies the instructional benefit of animations.

In contrast to educational videos, self-created animations focus on relevant aspects of the learning material since rather irrelevant aspects can be omitted and important aspects can be emphasized more ([6]). Several meta-analyses confirmed that dynamic visualizations can offer a learning benefit in comparison to static pictures ([5,11,12]). However, found effects were small and subject to methodological weaknesses (e.g., missing control for participants characteristics [13]). Mixed findings on the learning benefits of animations also manifest themselves in a large number of non-significant pair-wise comparisons between animations and static pictures (e.g., 83 out of 140 effect sizes in the meta-analysis by Berney & Bétrancourt [5]). A review from Castro-Alonso et al. ([14]) also indicated that studies comparing the instructional effectiveness of static with dynamic visualizations do not control for important moderating variables such as appeal, realism, or interaction and are more like an "apple and oranges" comparison. Nevertheless, a recent meta-analysis from Castro-Alonso et al. [11] found a positive effect of $g+ = 0.23$ indicating that animations are more effective for learning than static pictures. Interestingly, animations unfold a positive effect on learning only when 59% or fewer females participated in the respective primary study suggesting a gender imbalance. Since many studies rely on samples with unequal gender distribution, this factor could have an important influence when comparing static and dynamic visualizations. In general, animations seem to be particularly conducive when learning the STEM domains science, technology, engineering, and mathematics ([5,11,12]).

2.2. The special case of whiteboard animations

Whiteboard animations as a subtype of animations are an increasingly used tool for educational purposes. In general, they can be defined as "videos that depict the process of drawing a finished picture, usually on a whiteboard or something resembling a whiteboard" ([15], p. 103). Hereby, the learning content is presented through hand-drawn illustrations accompanied by a voice commentary ([16]). In contrast to conventional animations, whiteboard animations are characterized by their uplifting character indicating that the contents are built over a certain period of time. This could make it easier for the learner to follow the progress of the animation. Hereby, the knowledge acquisition is also supported by the narration which underpins the animated contents with spoken explanations.

However, empirical evidence for the beneficial effect of whiteboard animations in multimedia learning research is still quite rare. The most cited experimental study by Türkay ([15]) compared four different instructional conditions: whiteboard animation, slideshow, audio, and a text containing the same learning content. The results show that whiteboard animations positively influence retention performances, engagement, and enjoyment. However, comparing different forms of medial representation harbors the danger of confounded comparisons. For instance, whiteboard animations including voice-overs address both channels of information processing, whereas slides, audio or text do not (see *modality principle* [17]). Due to the novelty of the learning medium, Türkay ([15]) also states that curiosity and a higher level of attention might explain the found results.

Whiteboard animations can also be used for classroom instruction. For example, Li et al. [16] presented whiteboard animations in the context of a flipped classroom. Hereby, learners were allowed to watch an animation before attending a tutorial. In general, students who watched the animation before the tutorials performed better on the learning test than those who did not. In a study by Suhroh et al. [18], whiteboard animations were implemented in a project-based learning setting. Hereby, students were asked to create such an animation as a part of a presentation while the control group only relied on traditional media. Again, there was a positive effect in favor of whiteboard animations showing that students demonstrated better presentation skills in the posttest. Whiteboard animations also unfold positive effects in vocabulary and grammar learning as shown in a pretest – posttest experimental study by Syafrizal et al. ([1]). Overall, whiteboard animations were shown to facilitate learning. However, clear empirical evidence on why such animations might help is still missing.

2.3. Cognitive and motivational processes when learning with animations

Cognitive processes involved when learning with animations can be explained by the basic assumptions of the *Cognitive Load Theory*

(CLT [19,20]). Central for learning with animations is hereby the interplay of the working and long-term memory ([21]). While new information is processed in the working memory system, it has to be transferred to long-term memory to be stored for later retrieval phases. Working memory, however, has a limited capacity suggesting that only a small amount of information can be processed simultaneously ([22,23]). In addition, information can only remain up to 30 s before it is forgotten again ([24]). However, under the premise that the information is not repeated again. The load imposed on the learner's working memory is composed of three additive types – intrinsic, extraneous, and germane cognitive load. Intrinsic cognitive load (ICL) results from the learning tasks' complexity and is also influenced by the learner's domain-specific prior knowledge. In contrast, extraneous cognitive load (ECL) is caused by the presentation form of the learning material ([20]). The role of the germane cognitive load (GCL), defined as load which results from learning activities such as schema acquisition and automatization, is still under debate as recent findings suggest that it is not a load per se but rather has an allocation function by allocating cognitive resources to learning-relevant activities ([19]).

Beyond that, cognitive load, which is also named *mental load* in other psychology disciplines ([25]), is closely intertwined with the construct mental effort ([26]). In general, mental load is experienced rather passively by the learner. Accordingly, the ICL and ECL, which are elicited by characteristics of the learning task, are perceived passively as they cannot be influenced. In contrast, learners can only influence their learning process by actively investing his or her cognitive resources for activities relevant for learning, which is often called GCL or active load (for an overview see Ref. [27]). Following these assumptions, animations with their dynamic content and transient information (i.e., the transient information effect [28–31]) can be also detrimental to learning, because information must be processed in a comparatively quick way. However, since not only learning-relevant information (i.e., intrinsic cognitive load) but also learning-irrelevant information (i.e., extraneous cognitive load) must be processed during watching the animation, this can quickly lead to an overload of the working memory and learning activities (i.e., germane cognitive load) can be prevented. Learning-relevant information has additionally to be stored in the working memory before new information appears ([32]). In order to avoid negative influences of transient information, animations have to be designed in a way that helps learners to be able to cope better with the cognitive burdens. For example, the *segmentation principle* (for a meta-analysis, see Ref. [33]), which describes that the division of learning materials into learner-specific segments can lead to an improvement in learning, can help to mitigate the negative effects of transient information ([32]). Typically, pauses are implemented in the animation to give learners sufficient time for cognitive processing. This divides the animation into individual segments and helps learners to manage to rather high cognitive load (e.g., [34]). However, when looking at whiteboard animations, such pauses might also have negative effects since the storyline might be interrupted.

In recent years, researchers have increasingly shown that motivation to engage with a learning object is a crucial key mediator for the success of multimedia learning ([35]). According to Ryan and Deci ([36]), motivation can be defined as people's inherent drive to start and engage in activities. This drive can be internally motivated (i.e., intrinsic motivation) or externally evoked (i.e., extrinsic motivation). Moreno and Mayer [37], for example, proposed an extended theory on multimedia learning, the *Cognitive-Affective Theory of Learning with Media* (CATLM), by incorporating motivational factors. The idea behind is that the individual's motivation affects the cognitive engagement which in turn have an influence on learning success. Motivation can, therefore, improve the way information is selected, organized, and integrated into a coherent mental model. Mayer ([38]) qualified the hypothesis by the addition that sufficient cognitive resources must first be available for the learning-promoting effect of motivation. In general, it is assumed that motivational factors can enhance the learner's cognitive engagement to deal with the task which might lead to better learning ([39]).

Whiteboard animations are also said to have motivational effects. In this vein, Türkay ([15]) could prove the positive effect of this learning medium on enjoyment and attention, which can be explained by its inherent design. Thus, the rather entertaining drawings prime affective and motivational processes supporting the emotional design hypothesis (i.e., design factors that increase positive and activating affective states, e.g. Ref. [40]). Another explanation for the motivational effects of whiteboard animations is the rather unconventional form of information presentation maintaining student's attention and engagement.

2.4. The peculiarities of whiteboard animations for learning

Due to the lack of experimental studies regarding whiteboard animations, it is still unclear which factors lead to whiteboard animations being an effective learning tool. On the one hand, it might be because of the progressive drawing of the plot, which helps learners to better understand the individual processes and the costs for the cognitive processing of the individual steps can be outsourced to the learning material. On the other hand, it might be because of the underlying story, which motivates learners to actively process information and provides a schema that helps to structure the learning process and storage.

2.5. The learning-promoting effect of progressive build-up

Dynamic learning media such as animations are characterized in particular by their progressive build-up over time. Given a learning situation with whiteboard animation, the contents build up over a certain period of time and are then replaced by new ones. In contrast, static pictures represent a factual situation without any movement. When designing learning materials, an instructional designer often faces the question if dynamic or static visualizations will be more effective. Overall, dynamic visualizations seem to be more beneficial for learning what is underlined with meta-analytical findings (e.g. Ref. [11]). But what is the reason for this advantage in favor of dynamic visualizations using whiteboard animations as an example? Generally, the progressive build-up promotes the learners to mentally visualizing the contents resulting in a lower cognitive load while learning ([12]). In contrast, static pictures inevitably force the learner to infer the portrayed process from the picture series ([5]). Even though dynamic visualizations fill the gaps between static graphics, the continuous temporal changes pose different challenges to information processing ([41]). These demands usually lead to learners using their visual attention quite selectively ([42]). One way to compensate for the lack of attention during learning dynamic visualizations is attention guidance. Hereby, this strategy supports the learner in extracting relevant information from the learning material ([43]). In line with the cueing principle (for a meta-analysis see Ref. [44]), visual cues such as arrows, colored circles or lines, can be implemented into the animation for highlighting learning-relevant aspects ([45]). Unlike cueing conventional dynamic visualizations, whiteboard animations offer a new approach to attention guidance. Hereby, the progressive build-up within a whiteboard animation guides the learner's attention on important contents – namely those that are currently drawn by the hand.

Furthermore, it was assumed that the successive appearance of the learning content could promote the process of building mental models. For instance, Lowe et al. ([46]) found empirical evidence for the learning-beneficial effect of a successive presentation. In their study, participants watched a set of images depicting a kangaroo hop in either a dynamic, successive, or simultaneous format. After viewing the participants were asked to arrange a random sequence of the pictures in the correct order. Hereby, participants who have previously seen the successive format performed best in this test indicating that participants benefitted from the emphasis on configuration differences within the frames. In contrast, a recent study by H'mida et al. ([47]) confirmed that a continuous presentation via video of a complex judo skill resulted in better learning performances than several static pictures formats. The positive effect of a continuous presentation also led to a reduced perception of cognitive load while learning compared to a static-simultaneous-permanent pictures condition (all images were presented at the same time and remain on the screen) and a static-sequential-transient pictures condition (images appear and disappear one after the other) in which learners need to fill transitions mentally to create a coherent model. Moreover, learners in the continuous video condition reported the highest intrinsic motivation.

Similarly, drawn pictures in a whiteboard animation can be seen as components of a successive build-up. Supplemented by the verbal commentary, the gradual build-up through drawing could lead to a deeper understanding since two channels of information processing (visual and verbal information) can be used to build a coherent mental model as proposed by CTML ([38]). For this, the verbal information presented in whiteboard animation is also an essential part of the successful building of mental models.

2.6. *Learning through stories and narrativity*

Storytelling is a widely used method used in language education (e.g. Ref. [48], moral education (e.g. Ref. [49]), school education (e.g. Ref. [50]), or educational video games (e.g. Ref. [51]). However, only a few studies focused on the examination of stories themselves. Stories consist of a narrative described as a sequence of events, which build on each other and thus are interdependent (e.g. Refs. [52,53]). Moreover, narrative include the presence of relationships between events and/or characters supporting meaning construction and invite to actively process information ([54]). In conclusion, chronicles, where the construction of events is not completed, scientific reports or lectures, where often no relational structure, narrating voice, and a conclusion is presented, are not narratives ([54,55]). Some researchers used the terms *plot* and *story* interchangeably. Förster ([56]) highlighted their difference by explaining that a story is a narration that is set in a timed sequence, while a plot is a special narrative that emphasizes the causality of events. In this paper, we use the word *narrative* as a timeline of events connected by causality and use this word interchangeably with *story*.

While there is a grown body of research on the comparison of narrative versus expository/explanatory/informative texts (e.g. Refs. [57,58]), this study focused on the analysis of story elements as narrative rather than a narrative as a text genre. In this vein, Mandler and Johnson ([59]) described a set of rules to describe how a list of events can be transformed into a story. This story grammar principle includes rules like setting a narrative form, inserting names and characters into non-personalized sentences, assigning an intention to act to characters or relating characters to each other. Brunauer ([60]) furthermore proposes a narrative structure for learning, describing that a character should include a setting, agent, act, means, goal, and trouble. Wu ([61]) concluded that a story always includes a main character, a changing moment, and a theme or the main question. However, these are only basic elements that all stories have in common.

Narratives are also described as a cognitive tool that can enhance learning ([50,62]) and are based on two cognitive theories. According to the schema theory by Bartlett ([63]), people store information in a narrative form to be able to quickly assess memory in new situations. This net of information is called schema, while such schemata include components, attributes, prototypes, and relationships. According to the script theory by Schank and Abelson ([64]) people organize knowledge around stereotypic situations that often have a procedural character, such as riding a train. For these situations, the term *script* was used to help people to reduce the costs of actively processing information since routinely trained memory does not burden the working memory system. Thus, pre-prepared scripts such as a narrative can help to build such a script and thus, might help to motivate people to actively process information and foster learning.

There is not only theoretical support for the idea that narratives support learning but also empirical evidence (e.g. Refs. [49,54,65]). Narratives were not only shown to motivate learners ([66]) but also increase their listening skills ([67–69]). Comparing expository texts with narratives, researchers showed learning with stories has simpler grammatical structures, is more familiar to spoken language, and provides less cognitive demand ([70]). Narratives were shown to facilitate learning and motivate learners in various media contexts, such as learning with videos ([71]) or playing videos games (e.g. Refs. [72,73], for an overview see [74]). Especially in the comparison between narrative and expository texts, narratives seems to have a positive effect on learning (e.g. Refs. [75–80]). Narratives also not only enhanced memory processes but also facilitated comprehension of learning topics [81] and thus, can serve as a cognitive scaffolding for transferring knowledge to new problem tasks ([73]). In addition, a narrative contains elements necessary for a learner to stay emotionally engaged since emotionally charged words are usually used to describe a story ([82]).

Narratives can shift towards both more activating emotions (i.e., higher arousal) and more positive feelings (i.e., higher valence) when added to dynamic media such as video games ([83]) or films ([84]). Clark ([85]) explained the learning facilitating by saying that it situates and anchors learning in a context. However, there are also some studies showing no differences between stories and information texts (e.g. Ref. [61]) suggesting there might be moderators for the activation of the learning potential of using narratives requiring a closer examination.

2.7. Research question and hypotheses

Both the visual and verbal presentation of whiteboard animations are essential factors that might accelerate positive motivational, affective and cognitive effects. On the one hand, this study examines if the unique procedural presentation mode of whiteboard animations, namely the progressive construction of a composite illustration, helps better understanding the learning content. According to research on the comparison between static and dynamic visualizations ([12]), it is supposed that such a progressive build-up fosters learning by building a better and more automatized mental model ([46]) decreasing knowledge retrieval time. In addition, the progressive build-up was shown to decrease cognitive load and foster motivational processes while learning ([47]). For this, the following hypotheses were postulated:

Learners exposed to a progressive build-up of a described scenario …

H1a. achieve higher learning scores …

H1b. report less cognitive load …

H1c. report more mental effort …

H1d. indicate a higher motivation …

Than learners not exposed to a progressive build-up in whiteboard animations.

On the other hand, this study explores how the usage of stories in such an animation can help to organize knowledge and thus, foster learning. According to empirical studies, narratives were shown to enhance learning (e.g. Refs. [49,71]), motivate learners ([66]), provide less cognitive demand but also a higher cognitive involvement ([70]), and increase feelings of valence and activation (e.g. Ref. [83]) compared to mainly informational media. According to the schema theory by Bartlett ([63]) and the script theory by Schank and Abelson ([64]), people storing information in a narrative form can quickly assess memory in new situations while this routinely trained memory does not burden the working memory system. For this, the following hypotheses were postulated:

Learners exposed to a narrative in whiteboard animations …

H2a. achieve higher learning scores …

H2b. report less cognitive load …

H2c. report more mental effort …

H2d. indicate a higher motivation …

H2e. perceived more positive, activating emotions …

Than learners not exposed to a narrative in whiteboard animations.

Since both a progressive build-up as well as a narrative are assumed integral parts of whiteboard animations, their interaction will be examined explanatorily.

3. Experiment 1

3.1. Method

Design and participants. The research hypotheses were examined using a 2 (presentation mode: static pictures vs. progressive drawing) x 2 (narrative context: with vs. without a narrative) between-subject factorial design. Based on the effect sizes revealed in previous experiments in the area of the comparison between static and dynamic media (e.g. Refs. [86,87]) and trying to avoid searching for small effects that are of little practical significance, a power analysis was conducted with *G*Power* ([88]) using an effect size of $\eta_p^2 = 0.06$, a test power of $1 - \beta = 0.80$, and an error probability of $\alpha = 0.05$. This test revealed a minimum number of participants of $N = 128$. The experiment was conducted at a secondary school in Greiz, Germany. Ethical approval for this study was given by the ethics committee of the faculty of humanities at Chemnitz University of Technology (EPHF_09,022,021_004). Participants were asked for informed consent prior to their participation in the experiment. Each class participated as part of a double lesson so that the first experiment could be conducted in the first lesson and the second experiment in the second lesson. However, students were randomly assigned to the conditions for each experiment separately. Overall, 133 students (65.4% females; age: $M = 17.62$; $SD = 1.09$) took part in this experimental series. They were recruited from grade eleven until thirteen with three different profiles: economics ($N = 26$), design and media technology ($N = 50$), health and social affairs ($N = 57$). The domain-specific prior knowledge of the learning topic (for more details, see Measures) can be considered as very low since no person was able to give an appropriate answer to the open-answer question ("What statements can be derived about the Pygmalion effect by Robert Rosenthal?"). Group sizes were

(caption on next page)

Fig. 1. Overview of the nine, fully drawn pictures explaining the Pygmalion effect in Experiment 1.

controlled after each lesson so that they could be assigned to one condition of the experimental design in a counterbalanced way: 32 students received a whiteboard animation with progressive drawings and a narrative, 33 students watched the animation with progressive drawings but without a narrative, 33 students viewed the animation without progressive drawings but with a narrative, and 35 students were assigned to the condition where the animation did not contain a progressive drawing and no narrative.

Materials. To examine the progressive build-up and the narrative of whiteboard animations, an already created whiteboard animation from a YouTube channel ([89]) was used as a basis for all conditions. This animation covers the topic "The Pygmalion Effect" (for an overview see [90]). The Pygmalion effect is the phenomenon where higher expectations lead to higher performance. The effect has also become known as the Rosenthal experiment. This topic was selected because (1) it is not part of the school curriculum and (2) the complexity of the learning material is high enough to push learners in grades 11 to 13 to the edge of their working memory capacity. In line with the assumption of Tuovinen ([91]), who concludes that if learners have little prior knowledge, the presentation form should be given high priority to minimize cognitive load. Consent to use and manipulate said animation as part of a university study was collected beforehand. The original whiteboard animation is 4 min and 30 s long and comprises nine painted pictures (for an overview, see Fig. 1), which are progressively drawn during the video (see Fig. 2). The colored drawing is shown without a hand but in a progressive way.

In order to manipulate the presentation mode of the animation, the final appearances of the nine painted pictures were used to be shown in the conditions without a progressive build as a static image during the animation (see Fig. 1). For each of the image series, the fully constructed, final still image of a whiteboard animation scene was taken and merged with the same off-text of its animated counterpart, so that viewers of this version received a consistent cognitive load compared to those who viewed a fully animated version.

In order to manipulate the narrative context of the animations, the off-texts of both videos were re-dubbed to alienate the original version on the one hand, but also to create two versions in an informational and a narrative execution. For the narrative version, the existing off-text was first reformulated and shaped through using the following operands according to the story grammar principle ([59]).

1. Introduce a setting or theme.
2. Determine the form of narration (e.g., lyrical I, third person).
3. Add names/persons to non-personalized sentences.
4. Assign an intention to names/persons to create a plot.
5. Relate names/persons (e.g., teacher - student).
6. Create a dilemma/problematic moment and solve it.

The operands were chosen to make the reformulations of an originally factual text into a narrative form understandable and repeatable. However, the animation is, of course, nevertheless provided with a temporal sequence and is also partly based on a sequence of events, so it cannot be understood entirely as a condition without narrative. Table 1 shows examples illustrating the implementation of the above operands in comparison with the off-texts of the informational context. In the next step, the animations were adapted in their length to the newly spoken off-text, whereby no content-related scene was shortened or removed. While animations with a narrative context lasted 4 min and 35 s, animations with an informational context lasted 4 min and 30 s. This resembles a typical animation length and is longer than earlier research on animations (e.g. Ref. [92]).

Measures. For all measurements with more than one item, internal consistencies were calculated with the help of Cronbach's alpha ([93]). Hereby, a minimum value of 0.70 was considered satisfactory ([94]).

Prior knowledge. As the learner's domain-specific prior knowledge plays a crucial role when investigating the effectiveness of learning materials (e.g. Ref. [95]), it was measured with one open-answer question: "What statements can be derived about the Pygmalion effect by Robert Rosenthal?" Two independent raters were consulted to rate the answers based on a pre-set list of correct answers. A maximum of ten points could be reached. However, none of the participants' answers was sufficient to reach at least one point. Therefore, no inter-rater reliability could be calculated.

Cognitive load. The German-speaking self-report scales from Krell ([25]) were used to measure how learners perceived the task-inherent load (mental load) and managed the task-inherent load (mental effort). Hereby, subjective measures are frequently used in cognitive load research compared with physiological or dual-task methods ([96]). In line with recent findings from Klepsch and Seufert ([27]), the construct mental load is rather a passive part representing the amount to which the task complexity can burden working memory capacity. In contrast, mental effort represents the actual working memory capacity used to master a task. Specifically, the scales for mental load (five items, $\alpha = 0.84$; e.g., "The contents of the tasks were complicated.") and mental effort (six items, $\alpha = 0.82$; e.g., "I have tried hard to answer the tasks correctly.") were used for this study. Each item was rated on a 5-point scale ranging from (1) "does not apply at all" to (5) "applies fully".

Motivation. Learner's motivation while learning was measured with the *Situational Motivation Scale* from Guay et al. ([97]). Concretely, the facets of intrinsic motivation (four items, $\alpha = 0.88$; e.g., "Because I think that this activity is interesting.") and extrinsic regulation (four items; $\alpha = 0.84$; e.g., "Because I am supposed to do it"), as sub-concepts of motivation ([36]) were used to gain deeper insights under which motives the students have worked with the learning materials. The items were evaluated by the participants on a 7-point scale ranging from (1) "does not apply at all" to (7) "applies fully".

(caption on next page)

Fig. 2. Example of a build-up of one picture in the animation on the Pygmalion effect in Experiment 1.

Table 1
Examples of the off-text according to pre-defined operants by narrative context in Experiment 1.

	Narrative context	
Operant	Narrative	Informative
Introduce a setting or theme.	Andreas, the coach of a sports team, observes his team on the first day.	A coach observes young people of a sports team on the first day.
Determine the form of narration	He, His	The trainer
Add names/persons to non-personalized sentences.	Students Chris and Joe are new members of his team.	Person A and person B are new members of the team.
Assign an intention to names/persons to create a plot.	Chris reminds Andreas of Michael Jordon. Joe reminds him instead of an unfriendly relative.	Person A reminds the coach of a well-known athlete. Person B reminds the trainer instead of an unfriendly person.
Relate names/persons (e.g. teacher - student).	Coach Andreas' expectations influence the actions of the two students.	The expectations of the trainer influence the expectations of the acting persons.
Create a dilemma/problematic moment and solve it.	Chris is sure to reel in the win at the end of the season. Joe no longer believes the season will end successfully.	

Emotional states. As learning with multimedia learning materials may influence the learner's emotions, their affective states were measured with the *Positive Affect, Negative Affect and Valence Short Scales* (PANAVA-KS) Questionnaire from Schallberger ([98]). Specifically, the sub-components positive affect (PA) and negative affect (NA) were collected with the help of bipolar items. For example, the contrary items tired – wide awake (positive affect) and calm – nervous (negative affect) should be evaluated on a 5-point scale. Participants were asked to fill in the questionnaire before (PA: $\alpha = 0.80$; NA: $\alpha = 0.72$) and after learning (PA: $\alpha = 0.82$; NA: $\alpha = 0.77$) in order to identify affective changes during the learning process.

Learning outcomes. In order to measure participants' learning performance, retention and transfer tasks were prepared as common measures in multimedia learning ([99]). Concerning retention, which can be defined as remembering information explicitly mentioned in the learning material ([99]), five multiple-choice questions were formulated ($\alpha = 0.704$). Each retention question was given with four answer possibilities. Each answer has to be rated on its correctness. The number of correct answers differed among all tasks, however, at least one answer was correct. For instance, the question "Which statements about the Pygmalion effect are correct?" was presented with the answer options: "a) Low expectations lead to more motivation and higher performance", "b) Certain actions affect a person's self-image", "c) One's expectations about a person's abilities influence one's actions toward that person", and "d) Higher expectations lead to higher performance." In this example, the answer options b), c) and d) were correct and consequently had to be marked. Another point was given when answer option a) was not marked as correct. Following this example, each correct crossing or omission of a crossing was rewarded with one point resulting in a maximum of four points per question. Overall, 20 points could be achieved in the retention test.

Transfer refers to understanding ([99]) so that learned information can be applied in new scenarios. Overall, five multiple-choice questions including four answer possibilities were given ($\alpha = 0.738$). Again, each question was given with four answer possibilities. For example, the question "An employee of a company believes that he will soon be promoted. According to the Pygmalion effect, how are his superiors likely to treat him?" was given with the answer possibilities: "a) They ignore him and also do not pay attention to his successes at work"; "b) They value his achievements"; "c) They give him helpful tips on how to improve"; and "d) They give him helpful improvement tips, but do not appreciate his work". Hereby, the answers b) and d) were correct and had to be marked. Each correct crossing or crossing-out was rewarded with one point resulting in a maximum of four points per question. Overall, 20 points could be achieved in the transfer test.

Procedure. The experiment was conducted in a computer lab of a school with 25 workstations. Students were randomly assigned to one of the four experimental groups by drawing lots, and controlling for the number of participants in each group after each lesson (i.e., block randomization). Computers had been prepared by opening the first questionnaire before each experiment started. Before the experiment started and after the lesson started, all students were instructed to follow the instructions on their screens, read all information carefully, and raise their hands if any question occurs. All students completed the three parts (i.e. a first questionnaire, a learning website, and a second questionnaire) of the experiment individually. In the first questionnaire, demographic information was gathered and prior knowledge, as well as positive affect and negative affect, was measured to secure baseline before the learning environment. On the learning website, students had to watch the whiteboard animation on a web page. This learning web page differed according to the experimental group. In the second questionnaire, all dependent variables were measured in the following order: (1) Emotional, motivational, and mental load and mental effort scales, (2) learning tasks, (3) and demographic data. After students reached the last page, they were instructed to sit quietly at their computer desks. Overall, the experiment lasted 35–45 min. Overall, the full sample size, all data exclusions (if any), all manipulations, and all measures are reported in this manuscript.

4. Results

In order to calculate possible group differences, multivariate analyses of variance (MANOVAs) and univariate analyses of variance (ANOVAs) were conducted. Follow-up ANOVAs were only used if the previously calculated MANOVA showed significant effects. The independent variables presentation mode (static pictures vs. progressive drawing) and narrative context (with vs. without a narrative) were used for all variance analyses. Before the data analysis, the influence of possible confounding variables was investigated. Hereby, prior knowledge was not included as a covariate since no participant reported any prior knowledge. The four treatment groups also did not differ in terms of age, p = [.127; .588]. Chi-squared tests indicated that there were no other differences with respect to gender, p = [0.358; 0.861], class level, p = [0.944; 0.976], and profile, p = [0.399; 0.754]. Consequently, these variables were not considered as covariates for the following analyses. Descriptive values of all dependent variables including mean scores and standard deviations are shown in Table 2.

Learning outcomes. In order to check for learning differences, a MANOVA was conducted with the two dependent variables retention and transfer. Significant main effects were found for the presentation mode, Wilk's $\Lambda = 0.83$, $F(2, 128) = 13.14$, $p < .001$, ηp2 = 0.17, and the narrative context, Wilk's $\Lambda = 0.77$, $F(2, 128) = 19.54$, $p < .001$, ηp2 = 0.23. The interaction also showed a significant effect, Wilk's $\Lambda = 0.64$, $F(2, 128) = 36.75$, $p < .001$, ηp2 = 0.37. Based on the significant results, follow-up ANOVAs were carried out.

In terms of retention, a progressive drawing led to better retention performances, $F(1, 129) = 25.72$, $p < .001$, ηp2 = 0.17. However, manipulating the narrative context was not associated with significant differences ($p = .151$, ηp2 = 0.02). The significant interaction, $F(1, 129) = 19.71$, $p < .001$, ηp2 = 0.13, revealed that the beneficial learning effect of progressive drawing changes in dependence of the narrative context. When a progressive drawing was presented, a narrative was better than an informative context ($p = .021$; $\eta^2 = 0.06$). In contrast, when static pictures were presented, the informative context without narration led to better retention performances ($p < .001$; $\eta^2 = 0.21$). For an overview, see Fig. 3. With respect to transfer, the significant main effect for presentation mode, $F(1, 129) = 13.56$, $p < .001$, $\eta_p^2 = 0.10$, indicated that progressive drawing resulted in better transfer scores. Moreover, a narrative led to better transfer scores, $F(1, 129) = 16.87$, $p < .001$, $\eta_p^2 = 0.12$. The interaction of both factors also reached significance, $F(1, 129) = 11.43$, $p = .001$, $\eta_p^2 = 0.08$. Thus, a narrative context led to better transfer performances when static pictures were used ($p < .001$; $\eta^2 = 0.26$). In contrast, no differences occurred when the learning material was presented with progressive drawing ($p = .285$). For an overview, see Fig. 4. Overall, the hypotheses H1a and H2a can be confirmed with the exception that a narrative within the whiteboard animation did not lead to better retention scores.

Cognitive processes. A MANOVA with the dependent variables mental effort and mental load showed a significant main effect for the presentation mode, Wilk's $\Lambda = 0.81$, $F(2, 128) = 15.35$, $p < .001$, $\eta_p^2 = 0.19$, and narration, Wilk's $\Lambda = 0.80$, $F(2, 128) = 15.79$, $p < .001$, ηp2 = 0.20, but no significant interaction, $p = .793$, ηp2 = 0.004. Consequently, only the main effects can be analyzed for the following ANOVAs.

With respect to mental load, static pictures resulted in a significant higher perceived mental load, $F(1, 129) = 28.21$, $p < .001$, $\eta_p^2 = 0.18$. Furthermore, a narrative context within the learning material was associated with a significant lower perceived mental load, $F(1, 129) = 13.63$, $p < .001$, $\eta_p^2 = 0.09$. Based on these results, the hypotheses H1c and H2c can be confirmed. Concerning mental effort, only a significant main effect for narration was found, $F(1, 129) = 23.57$, $p < .001$, $\eta_p^2 = 0.16$, whereby a narrative increased the invested mental effort. The main effect for presentation mode, $p = .469$, $\eta_p^2 = 0.004$, failed to reach significance. Accordingly, hypothesis H1d must be rejected while hypothesis H2d can be confirmed.

Motivational processes. Based on the proposed scales of intrinsic motivation and external regulation by Guay et al. ([97]), a MANOVA was conducted. Hereby, a significant main effect for the presentation mode occurred, Wilk's $\Lambda = 0.91$, $F(2, 128) = 6.38$, $p = .002$, ηp2 = 0.09. The main effect for the narrative context, $p = .382$, $\eta_p^2 = 0.015$, as well as for the interaction, $p = .569$, ηp2 = 0.009, were not significant.

Regarding the intrinsic motivation, it could be shown that progressive drawing resulted in higher perceptions of this construct, F (1,

Table 2
Mean scores of all dependent variables and their corresponding standard deviations by experimental groups of Experiment 1.

	Experimental groups							
	Progressive drawing				Static pictures			
	With a narrative (N = 32)		Without a narrative (N = 33)		With a narrative (N = 33)		Without a narrative (N = 35)	
Type of scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mental load	2.33	0.76	2.76	0.60	2.98	0.64	3.48	0.90
Mental effort	5.11	0.74	4.50	0.93	5.32	0.70	4.51	0.97
Intrinsic motivation	5.02	1.13	4.93	1.26	4.05	1.28	4.39	1.23
External regulation	5.02	1.38	5.24	1.37	4.83	1.42	5.21	1.48
Positive activation (a priori)	3.92	0.85	3.88	0.76	4.03	0.82	3.86	0.75
Positive activation (post-hoc)	4.97	1.23	4.02	1.06	5.09	1.17	4.23	1.20
Negative activation (a priori)	3.15	0.69	3.10	0.66	3.04	0.75	3.26	0.99
Negative activation (post-hoc)	2.57	1.15	2.55	1.19	2.30	1.04	2.51	1.30
Retention	15.13	2.93	13.61	2.99	10.30	2.80	13.29	2.97
Transfer	15.94	2.20	15.58	2.86	15.78	2.96	12.06	3.31
Test time (in min)	9.38	1.77	10.00	1.87	9.43	1.70	9.57	1.70
Retention efficiency	1.66	0.43	1.42	0.46	1.13	0.39	1.43	0.40
Transfer efficiency	1.76	0.42	1.63	0.50	1.74	0.50	1.30	0.44

Note. The scales of mental effort and mental load each ranged from 1 to 5. The scale of intrinsic motivation and external regulation each ranged from 1 to 7. Each scale of positive and negative activation each ranged from 1 to 5. The learning scores for retention and transfer ranged each from 0 to 20 points. All efficiency scores represent the corresponding learning score per invested minute in the learning test.

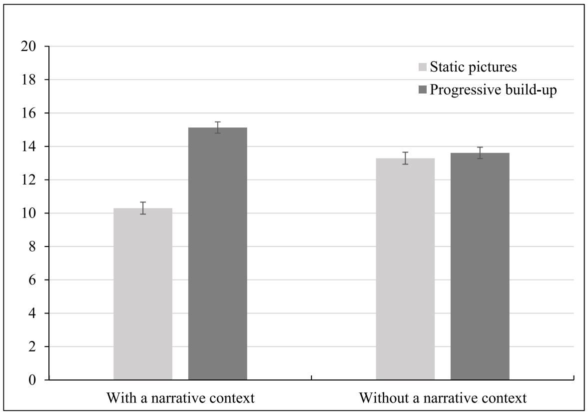

Fig. 3. Retention scores by experimental groups of Experiment 1. *Note*. Mean scores and corresponding standard errors by experimental group. The scale of retention ranged from zero to twenty points.

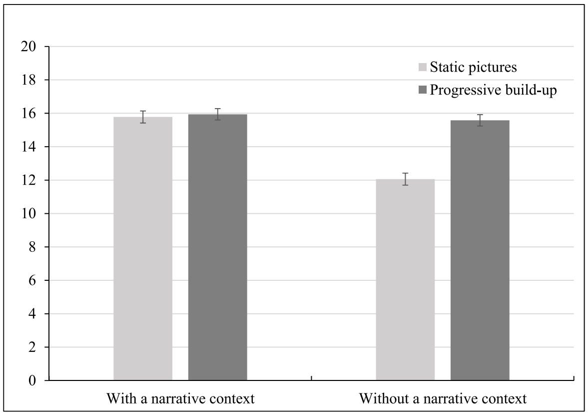

Fig. 4. Transfer scores by experimental groups of Experiment 1. *Note*. Mean scores and corresponding standard errors by experimental group. The scale of transfer ranged from zero to twenty points.

129) = 12.42, $p = .001$, $\eta_p^2 = 0.09$. In terms of external regulation, the effect of presentation mode failed to reach significance, $p = .657$, $\eta_p^2 = 0.002$. In sum, hypothesis H1e can be confirmed for intrinsic motivation while hypothesis H2e must be rejected.

Affective changes. A mixed MANOVA was calculated with time as an additional within-subject factor and the positive affect and negative affect scores as dependent variables. The analysis revealed a significant main effect for time, Wilk's $\Lambda = 0.45$, F (2, 128) = 78.00, $p < .001$, $\eta_p^2 = 0.55$) and a significant interaction effect of time and narration, Wilk's $\Lambda = 0.90$, F (2, 128) = 7.27, $p = .001$, $\eta_p^2 =$ 0.10). In contrast, the interaction between time and presentation mode, $p = .274$, $\eta_p^2 = 0.02$, as well as the interaction between time, presentation mode, and narrative context were not significant, $p = .856$, $\eta_p^2 < 0.01$.

Regarding the effect of time, both the positive affect, F (1, 129) = 39.05, $p < .001$, $\eta_p^2 = 0.23$, as well as the negative affect, F (1, 129) = 112.08, $p < .001$, $\eta_p^2 = 0.47$, reached significance. Watching a whiteboard animation led to an increase in the scores of positive affect and a decrease in the scores of negative affect.

In terms of positive affect, a significant interaction between time and narrative context was found, F (1, 129) = 14.61, $p < .001$, η_p^2 = 0.10. When the whiteboard animation included a narrative context, $p = <.001$; $\eta_p^2 = 0.11$, positive affect increased more than watching the whiteboard animation without a narrative context, $p = .004$; $\eta_p^2 = 0.03$. With respect to negative affect, no significant interaction between time and narrative context occurred, $p = .967$, $\eta_p^2 < 0.01$.

5. Experiment 2

The promising results of the first experiment may also have been partly caused by the novelty effect (e.g. Ref. [100]). Since various empirical studies support the stability of the novelty effect (e.g. Refs. [101,102]), it could be considered as a possible explanation for the results of Experiment 1. In order to determine whether the found effects can be generalized, we replicated the scenario in a second experiment. This is particularly important because replications in educational psychology are still rare although they are the key to high-quality research (e.g. Ref. [103]). Moreover, the manipulation of the narrative context in the first experiment is confounded because both animations include a sequence of events. In order to strengthen the generalizability, the second experiment also included

(caption on next page)

Fig. 5. Overview of the ten, fully drawn pictures explaining the eight stages of psychological development in Experiment 2.

another topic which can be better used to manipulate a narrative context. For this, Experiment 2 was conducted in order to be able to generalize findings from Experiment 1 to a different content, but using the same sample.

5.1. Method

Participants and design. For the second experiment, the same design and therefore the same power analysis was used as in Experiment 1. As described in Experiment 1, the same participants engaged for both experiments so that demographic data remained the same. The domain-specific prior knowledge of the learning topic (for more details, see Measures) can be considered as very low since no person was able to give an appropriate answer to the open-answer question ("What key central concepts come to mind about Erik Erikson's stage model of psychosocial development?").

Participants were randomly assigned to one condition of the experimental design: 32 students received a whiteboard animation with progressive drawings and a narrative, 33 students watched the animation with progressive drawings but without a narrative, 33 students viewed the animation without progressive drawings but a narrative, and 35 students were assigned to the condition where the animation did not contain a progressive drawing and no narrative.

Materials. Again, an already created whiteboard animation from the same YouTube channel ([104]) was used as a basis for all conditions. This animation covers the topic "The eight stages of psychological development" (for an overview see [105]). This theory describes the full human life cycle as a relationship between the psychological and social dimensions of life. This topic was selected because of the same reasons stated in Experiment 1: (1) it is not part of the school curriculum, and (2) the complexity of the learning material is high enough to push learners in grades 11 to 13 to the edge of their working memory capacity. Fig. 5 shows an example of the main structure of an animation. Consent to use and manipulate said animation as part of a university study was collected beforehand. The original whiteboard animation is 4 min and 30 s long and comprises ten painted pictures, which are progressively drawn during the video. The black-and-white drawing is shown with a hand in a progressive way.

This whiteboard animation was modified under the same conditions as in Experiment 1. First, the presentation mode was manipulated by creating a version that shows a continuous build-up of the image and a version that shows the final image of each scene as a static image meanwhile. Second, the narrative context was changed according to the rules described in Experiment 1, creating a version with an informative presentation of the topic and a version with a narrative presentation of the topic. For example, the sentence "When you're a toddler, you wonder if the world is trustworthy and if it's safe" was transformed into "As a young child, Max wonders if he can trust the world he lives in and if it is safe." While animations with a narrative context lasted 4 min and 41 s, animations with an informational context lasted 4 min and 56 s.

Measures. The same instruments as in Experiment 1 were used to measure intrinsic motivation ($\alpha = 0.95$), external regulation ($\alpha = 0.88$), mental load ($\alpha = 0.83$), mental effort ($\alpha = 0.83$), as well as positive affect (PA; $\alpha = 0.84$), and negative affect (NA; $\alpha = 0.83$). Since a new learning topic was presented during the animation, all learning-related measures (i.e., prior knowledge, retention, and transfer scores) were newly created or calculated.

Prior knowledge. Prior knowledge was measured with one open-answer question: "What key central concepts come to mind about Erik Erikson's stage model of psychosocial development?" Based on a pre-set list of correct answers, a maximum of eight points could be reached. Two independent raters were used to rate the answers. However, none of the participants' answers was sufficient to reach at least one point. Therefore, no inter-rater reliability could be calculated.

Learning outcomes. In order to measure participants' learning performance, retention and transfer tasks were prepared again. Concerning retention, five multiple-choice questions were formulated ($\alpha = 0.704$). Again, each retention question was given with four answer possibilities and each answer has to be rated on its correctness. The number of correct answers differed among all questions, however, at least one answer was correct. For instance, the question "Which statements are true regarding the first stage of development according to Erikson's theory?" was presented with the answer options: "a) You learn if you can trust someone.", "b) The phase is in the child age of 1–2 years.", "c) When you experience fear, you develop anger and hostility.", and "d) The key to development here are the grandparents." In this example, the answer options a) and b) were correct and consequently had to be marked. Another two points were given when answer options c) and d) were not marked as correct. Following this example, each correct crossing or omission of a crossing was rewarded with one point resulting in a maximum of four points per question. Overall, 20 points could be achieved in the retention test.

In order to measure transfer, five multiple-choice questions including four answer possibilities were presented ($\alpha = 0.738$). For example, the question "According to Erikson's psychosocial development, what would a child typically do at age 4–5?" was given with the answer possibilities: "a) The child explores new objects independently.", "b) The child discovers itself.", "c) The child explores its environment.", and "d) The child increasingly differentiates himself from the environment.". Hereby, the answer options a), c), and d) were correct and had to be marked. Each correct crossing or omission of a crossing was rewarded with one point resulting in a maximum of four points per question. Overall, 20 points could be achieved in the transfer test.

Procedure. The procedure did not differ from Experiment 1. Overall, the experiment lasted 35–45 min.

6. Results

Again, MANOVAs and ANOVAs were used to explain group differences and interaction effects. Also, presentation mode and

Table 3
Mean scores of all dependent variables together with their corresponding standard deviations by treatment groups of Experiment 2.

	Experimental groups							
	Progressive drawing				Static pictures			
	With a narrative ($N = 32$)		Without a narrative ($N = 33$)		With a narrative ($N = 33$)		Without a narrative ($N = 35$)	
Type of scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mental load	2.59	0.79	2.84	0.97	2.92	0.72	3.52	0.84
Mental effort	5.09	0.95	4.50	0.83	5.20	0.84	4.51	1.07
Intrinsic motivation	5.04	1.20	4.90	1.40	4.05	1.37	4.33	1.45
External regulation	4.96	1.68	5.23	1.41	4.71	1.60	4.94	1.72
Positive activation (a-priori)	3.92	0.85	3.88	0.76	4.03	0.82	3.86	0.75
Positive activation (post-hoc)	4.66	1.25	4.08	1.31	4.77	1.10	4.09	1.19
Negative activation (a-priori)	3.15	0.69	3.10	0.66	3.04	0.75	3.26	0.99
Negative activation (post-hoc)	2.61	1.21	2.57	1.22	2.28	1.09	2.56	1.31
Retention	14.63	2.88	12.54	2.93	9.64	3.11	13.00	2.96
Transfer	15.16	3.05	12.97	2.80	13.36	3.02	9.83	2.71
Test time	9.92	1.77	9.63	1.61	9.57	1.80	9.62	1.50
Retention efficiency	1.54	0.48	1.34	0.40	1.06	0.41	1.36	0.31
Transfer efficiency	1.61	0.55	1.38	0.35	1.45	0.44	1.03	0.29

Note. The scales of mental effort and mental load each ranged from 1 to 5. The scale of intrinsic motivation and external regulation each ranged from 1 to 7. Each scale of positive and negative activation ranged from 1 to 5. The learning scores for retention and transfer ranged each from 0 to 20 points. All efficiency scores represent the corresponding learning score per invested minute in the learning test.

narrative context were used as independent variables. Prior knowledge was not included as a covariate since no participant reported any prior knowledge. The four treatment groups also did not differ in terms of age, $p = [.701; .971]$. Chi-squared tests indicated that there were no other differences with respect to gender, $p = [0.580; 0.850]$, class level, $p = [0.518; 0.615]$, and profile, $p = [0.354; 0.912]$. Consequently, these variables were not considered as covariates for the following analyses. Descriptive values of all dependent variables including mean scores and standard deviations are shown in Table 3.

Learning outcomes (see Table 4). In order to check for learning differences, a MANOVA was conducted with the two dependent variables retention and transfer. Significant main effects were found for the presentation mode, Wilk's $\Lambda = 0.82$, $F(2, 128) = 13.95$, $p < .001$, $\eta p2 = 0.18$, and the narrative context, Wilk's $\Lambda = 0.68$, $F(2, 128) = 30.21$, $p < .001$, $\eta p2 = 0.32$. The interaction also showed a significant effect, Wilk's $\Lambda = 0.70$, $F(2, 128) = 27.06$, $p < .001$, $\eta p2 = 0.30$. Based on the significant results, follow-up ANOVAs were carried out.

In terms of retention, it was found that progressive drawing led to better retention performances, $F(1, 129) = 19.33$, $p < .001$, $\eta p2 = 0.13$. However, manipulating the narrative context was not associated with significant differences, $p = .215$, $\eta p2 = 0.01$. The significant interaction, $F(1, 129) = 27.86$, $p < .001$, $\eta p2 = 0.18$, revealed that when a progressive drawing was presented, a narrative

Table 4
Summary of the results of Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
Learners exposed to a progressive build-up of a described scenario …		
H1a: achieve higher learning scores …	+	+
H1b: report less cognitive load …	+	+
H1c: report more mental effort …	✗	✗
H1d: indicate a higher motivation …	+	+
than learners not exposed to a progressive build-up in whiteboard animations.		
Learners exposed to a narrative in whiteboard animations …		
H2a: achieve higher learning scores …	+	+
H2b: report less cognitive load …	+	+
H2c: report more mental effort …	+	+
H2d: indicate a higher motivation …	✗	✗
H2e: perceived more positive, activating emotions …	+	+
than learners not exposed to a narrative in whiteboard animations.		
RQ1: Interaction for learning scores	+	+
RQ2: Interaction for cognitive load	✗	✗
RQ3: Interaction for mental effort	✗	✗
RQ4: Interaction for motivation	✗	✗
RQ5: Interaction for emotions	✗	✗

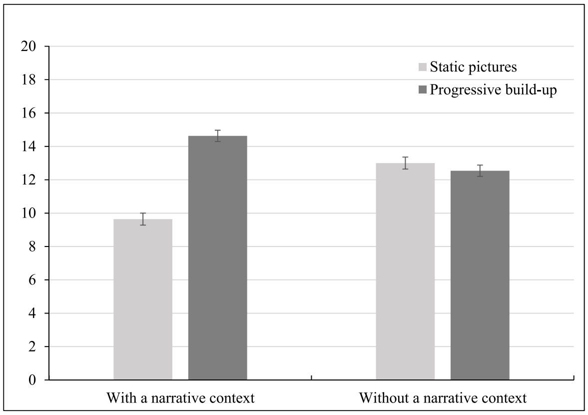

Fig. 6. Retention scores by experimental groups of Experiment 2. *Note*. Mean scores and corresponding standard errors by experimental group. The scale of retention ranged from zero to twenty points.

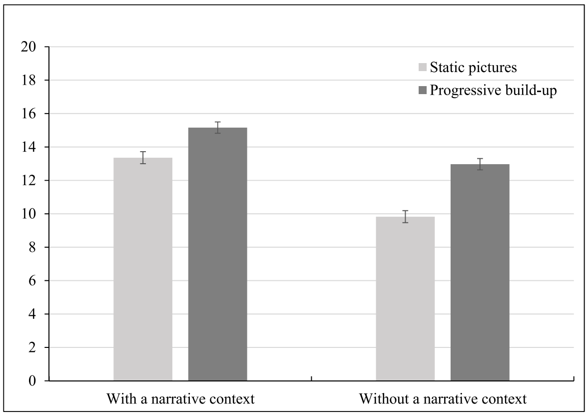

Fig. 7. Transfer scores by experimental groups of Experiment 2. *Note*. Mean scores and corresponding standard errors by experimental group. The scale of transfer ranged from zero to twenty points.

was better than an informative context, $p = .003$; η2 = 0.11. In contrast, when static pictures were presented, the informative context without narration led to better retention performances, $p < .001$; η2 = 0.24. For an overview, see Fig. 6. With respect to transfer, the significant main effect for presentation mode, F (1, 129) = 24.15, $p < .001$, $\eta_p^2 = 0.16$, indicated that progressive drawing resulted in better transfer scores. Moreover, a narrative led to better transfer scores, F (1, 129) = 32.48, $p < .001$, $\eta_p^2 = 0.20$. The interaction of both factors did not reach significance, F (1, 129) = 1.80, $p = .182$, $\eta_p^2 = 0.01$. For an overview, see Fig. 7. Overall, the hypothesis H1a can be fully confirmed while the hypothesis H2a can be confirmed for transfer only.

Cognitive processes. A MANOVA with the dependent variables mental effort and mental load showed significant main effect for the presentation mode, Wilk's Λ = 0.91, F (2, 128) = 6.18, $p = .003$, $\eta_p^2 = 0.09$, and narration, Wilk's Λ = 0.85, F (2, 128) = 11.22, $p < .001$, ηp2 = 0.15, but no significant interaction, $p = .737$, ηp2 = 0.011. Consequently, only the main effects can be analyzed for the follow-up ANOVAs.

Regarding mental load, static pictures resulted in a significant higher perceived mental load, F (1, 129) = 12.02, $p = .001$, $\eta_p^2 = 0.09$. Furthermore, the narrative context was associated with a significant lower perceived mental load, F (1, 129) = 8.79, $p = .004$, $\eta_p^2 = 0.06$. Based on these results, the hypothesis H1c and H2 can be confirmed. Concerning mental effort, only a significant main effect for narration was found, F (1, 129) = 15.68, $p < .001$, $\eta_p^2 = 0.11$, whereby a narrative increased the invested mental effort. The main effect for presentation mode, $p = .715$, $\eta_p^2 < 0.01$, failed to reach significance. Accordingly, the hypothesis H1d must be rejected while hypothesis H2d can be confirmed.

Motivational processes. Based on the proposed scales of intrinsic motivation and external regulation by Guay et al. ([97]), a MANOVA was conducted. Hereby, a significant main effect for the presentation mode occurred, Wilk's Λ = 0.90, F (2, 128) = 6.77, $p = .002$, ηp2 = 0.10. However, no significant main effect for narration, $p = .594$, ηp2 = 0.01, nor a significant interaction, $p = .672$, ηp2 = 0.01, was found. Regarding the intrinsic motivation, it could be shown that a progressive drawing resulted in higher perceptions, F (1, 129) = 11.01, $p = .001$, $\eta_p^2 = 0.08$. Regrading external regulation, no differences for presentation mode were found, $p = .335$, $\eta_p^2 = 0.01$. Overall, the hypothesis H1e can be confirmed for intrinsic motivation while hypothesis H2e must be rejected.

Affective changes. A mixed MANOVA was calculated with time as an additional within-subject factor and the positive affect and negative affect scores as dependent variables. The analysis revealed a significant main effect for time, Wilk's $\Lambda = 0.51$, $F(2, 128) = 60.61$, $p < .001$, $\eta_p^2 = 0.49$, and a significant interaction effect of time and narration, Wilk's $\Lambda = 0.90$, $F(2, 128) = 3.45$, $p = .035$, $\eta_p^2 = 0.05$. In contrast, the interaction between time and presentation mode, $p = .297$, $\eta_p^2 = 0.02$, as well as the interaction between time, presentation mode, and narrative context were not significant, $p = .984$; $\eta_p^2 < 0.01$.

Regarding the effect of time, both the positive affect, $F(1, 129) = 23.38$, $p < .001$, $\eta_p^2 = 0.15$, as well as the negative affect, $F(1, 129) = 101.37$, $p < .001$, $\eta_p^2 = 0.44$, reached significance. Watching a whiteboard animation led to an increase in the scores of positive affect and a decrease in the scores of negative affect.

In terms of positive affect, a significant interaction between time and narrative context was found, $F(1, 129) = 6.93$, $p = .009$, $\eta_p^2 = 0.05$. When the whiteboard animation included a narrative context, $p = <.001$; $\eta_p^2 = 0.07$, positive affect increased more than watching the whiteboard animation without a narrative context, $p = .018$; $\eta_p^2 = 0.02$. With respect to negative affect, no significant interaction between time and narrative context occurred, $p = .967$, $\eta_p^2 < 0.01$.

7. General discussion

Whiteboard animations are an increasingly used tool for educational and instructional purposes as more and more education providers offer them. However, experimental evidence on how to best design such animations was still missing. First studies were only able to show that whiteboard animations unfold positive effects on learning compared to other instructional formats ([15]), making a comparison that is confounded. Other studies have shown that such animations can be embedded in flipped classrooms or project-based learning settings ([16,18]) without losing the learning effect. Addressing this research gap, the goal of the two studies reported was the first experimental verification of two important factors (presentation mode and narrative context) implemented in whiteboard animations. Therefore, two experiments were conducted in order to reduce the influence of possible novelty effects. In this context, the second experiment was aimed at replicating findings from Experiment 1 with a new whiteboard animation. Results show that all effects were replicated although effect sizes slightly differ descriptively. This allows a valid and reliable generalization. In general, it can be noted that both visual and verbal components of whiteboard animations should be examined as they are first processed separately to be integrated then into a coherent model (following CTML [38]).

In both experiments, a progressive drawing (i.e., dynamic visualizations) within a whiteboard animation resulted in higher retention and transfer scores what is in line with meta-analytical findings (e.g. Ref. [11]) when not looking at the narrative context. Learners in the conditions with static pictures also reported a higher mental load indicating that more cognitive resources were needed in order to manage the inherent complexity (e.g., Klepsch & Seufert [27]). Learners might have been less cognitively burdened by a step-by-step guide resulting from the progressive build-up. The progressive drawing within the whiteboard animation also led to higher ratings of intrinsic motivation indicating that learners perceived the learning material as more interesting and pleasant leading to higher learning scores ([106]). The motivational effect of showing a drawing process may also have been enhanced by an increased social presence of the instructor in the video (for an overview on social effects on learning, see [107]).

Regarding the narrative context, both experiments provide strong support for the inclusion of story-similar elements in whiteboard animations. Thus, previous findings that a narrative context is beneficial for learning (e.g. Refs. [49,71]) are supported. In this context, a narrative context was associated with higher scores of transfer. In contrast, no effects were found for retention. A narrative context also led to an increased amount of mental effort what is basically positive as learners are supposed to be more engaged in the learning task (e.g. Ref. [108]). Moreover, learners who watched the whiteboard animations with a narrative context perceived the mental load to be lower supporting the idea of stories as automatized mental models ([46]). From an affective perspective, it can be argued that narrative elements even increase the positive affect during learning. As learning with multimedia is supposed to be affected by emotional factors (e.g. Ref. [35]), it can be assumed that implementing narrative elements into a whiteboard animation is associated with a higher engagement resulting in better transfer performances. Another explanation for learning benefit of a narrative context is based on a stronger relation with the characters through the story. This connectedness does not only help to combine past experiences with new information (see also [107]), but it also activates learners to invest more effort and brings them in a higher positive and activating affective state. By this, a connection between the emotional design hypothesis ([40]) and a narrative context is much more evident than a connection with the presentation mode, but needs to be proven with mediation hypotheses in future. Also, the voice principle (e.g. Ref. [109]) can be used to explain the results. According to this principle, a human-like but also “natural” way of presenting information verbally is helpful to increase learners' performance and motivation. This might be the case when looking at narrative story-telling in whiteboard animations. Interestingly, also positive affect was especially increased by a narrative context over time. Here, the positive interaction with learners and the animation got visible. Future studies, however, should also measure the likability of such animations in order to strengthen this hypothesis.

When considering the interdependency between presentation mode and narrative context, it becomes clear that the combination of both is not necessarily conducive to learning. When looking at students' retention performance, a narrative does only help when also a progressive build-up of the visuals is presented. To put it the other way around, a progressive build-up is only retention-enhancing when a story was used. This might be explained by the contiguity effect (e.g. Ref. [110]) stating that both visual and verbal instructions should be presented simultaneously. Especially in the case of whiteboard animations, the visual presentation of information is often presented in a comic-like and story-based character ([15]) so that too informative verbal explanation can lead to a mismatch. It has also been shown that progressive build-up decreases mental load ratings and increase learners perceived intrinsic motivation, while a story within whiteboard animations helped do not only decrease mental load but also increase learners' mental effort and positive affect. While a progressive build-up was shown to be beneficial for retention and transfer, a narrative solely enhanced learners' transfer performance. Here, stories might play an important role in embedding information into a context to be more in applied contexts ([111]).

7.1. Limitations and future directions

Although the study and its results make an important contribution to research on whiteboard animations, some limitations must be stated, which could be addressed in future studies. First, and most important is that the instructional topics covered in the two whiteboard animations were quite different. The Pygmalion effect can be understood as a process, but it is possible to learn about the eight stages of psychological development simply by recalling the factual knowledge about them. The learning materials also differed in terms of coloring. While the animation in Experiment 1 was colored, the animation in the second experiment was presented in black-and-white. Consequently, possible emotional design influences (e.g. Ref. [40]) cannot be ruled out when explaining the effects. In Experiment 2, the progressive drawing was shown by a human hand, while in the first experiment the contents appeared without a human hand. Hereby, the human hand could trigger social processes and therefore influence learning (e.g., social cues [107]).

The learning materials used in this study were system-paced indicating that the learner had no control over the progress. As learning with complex learning materials such as dynamic visualizations can cause additional demands on information processing, giving the learner more control (e.g., with the possibility to pause the animation) can be associated with learning benefits (segmentation principle [33,34]). As the success of learning dynamic visualizations depends largely on learners' spatial ability (e.g. Ref. [112]), further studies should include this variable in the experiment. Insightful findings of the role of spatial ability in learning from an instructional animation could be gained here, particularly given whether spatial ability is a compensator or enhancer when learning with whiteboard animations ([113]). Moreover, the learning materials used in this study were relatively short (about 5 min). It would be interesting to examine if the found results remain stable over a longer learning period. In terms of the study's sample, only learners from a secondary school participated. For elementary school students, for example, effects might change. Also, the selection of static pictures was done by choosing the full images of each progressive-build up in the animations. This help to reduce possible differences in the conditions with and without a progressive-buildup. However, this might also have led to an increase in cognitive load, because the full picture has to be screened for relevant information at once.

In this study, cognitive and affective variables were considered in order to explain the effects of the presentation mode and narrative context on learning. Future studies should also implement metacognitive measures, as the individual's ability to regulate cognitive resources can be beneficial for learning ([114]). Moreover, the study relied on variance analysis procedures to check group differences. A greater knowledge gain could be obtained by hypothesizing and calculating mediation analyses based on the results of this study. For instance, it might be hypothesized that the learning-beneficial effect of the narrative context is caused by the increased positive affect. Following the path, the narrative context could lead to a higher positive affect, which then leads to better learning performances.

Although it is common in multimedia learning research to conduct the learning tests immediately after the intervention (e.g. Ref. [115]), future studies should query learners' performance with a delayed testing (e.g., three weeks [116]) to capture long-term effects of a narrative context or a progressive build-up.

7.2. Implications

This study draws both theoretical and practical implications that can be used by educational designers and researchers. On the theoretical side, this experiment is the first attempt to examine components of whiteboard animations explaining their learning benefit. Without a progressive presentation of drawings and a storyline in the verbal explanation, learning results were significantly lower showing that these components might be essential to present such animations (following the CTML [38]). In addition, both a progressive build-up and narratives are not yet fully experimentally examined so that this study adds momentum in the research of these concepts, which should be also tested in other dynamic media formats such as educational videos or educational video games.

The results of this study can help instructors to more effectively design or choose whiteboard animations for educational purposes. Based on the results of this study, instructional designers and teachers are encouraged to take both the progressive build-up and the narrative context into account. Especially in the case of implementing just one of these design recommendations, one should be aware of possible negative effects for recalling information measured with a retention test in this study.

Author contribution statement

Sascha Schneider: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Felix Krieglstein: Contributed reagents, materials, analysis tools or data; Performed the experiments; Wrote the paper.

Maik Beege: Analyzed and interpreted the data; Wrote the paper.

Günter Daniel Rey: Conceived and designed the experiments; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interest's statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit author statement

Sascha Schneider: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. ***Felix Krieglstein***: Contributed reagents, materials, analysis tools or data; Performed the experiments; Wrote the paper. ***Maik Beege***: Analyzed and interpreted the data; Wrote the paper. ***Günter Daniel Rey***: Conceived and designed the experiments; Wrote the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e13229.

References

[1] S. Syafrizal, M. Muhajir, J. Pahamzah, F. Furqon, The effects of using whiteboard animation toward Indonesian students' vocabulary and grammar mastery, J. Lang. Linguist. Stud. 17 (2) (2021) 783–793.
[2] [Internet]. YouTube. YouTube; [cited 2023Jan5]. Available from: 100SekundenPhysik https://www.youtube.com/@100SekundenPhysik/featured.
[3] A. Occa, S.E. Morgan, The role of cognitive absorption in the persuasiveness of multimedia messages, Comput. Educ. 176 (2022), 104363.
[4] R. Plötzner, S. Berney, M. Bétrancourt, A review of learning demands in instructional animations: the educational effectiveness of animations unfolds if the features of change need to be learned, J. Comput. Assist. Learn. 36 (6) (2020) 838–860.
[5] S. Berney, M. Bétrancourt, Does animation enhance learning? A meta-analysis, Comput. Educ. 101 (2016) 150–167.
[6] R.K. Lowe, W. Schnotz, Animation principles in multimedia learning, in: R.E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning, second ed., Cambridge University Press, Cambridge, 2014, pp. 513–546.
[7] M. Wong, J.C. Castro-Alonso, P. Ayres, F. Paas, The effects of transient information and element interactivity on learning from instructional animations, in: S. Tindall-Ford, S. Agostinho, J. Sweller (Eds.), Advances in Cognitive Load Theory: Rethinking Teaching, Routledge, London, 2019, pp. 80–88.
[8] W. Schnotz, R.K. Lowe, A unified view of learning from animated and static graphics, in: R.K. Lowe, W. Schnotz (Eds.), Learning with Animation: Research Implications for Design, Cambridge University Press, Cambridge, 2008, pp. 304–356.
[9] S. Fischer, R.K. Lowe, S. Schwan, Effects of presentation speed of a dynamic visualization on the understanding of a mechanical system, Appl. Cognit. Psychol. 22 (8) (2008) 1126–1141.
[10] T.N. Höffler, H. Prechtl, C. Nerdel, The influence of visual cognitive style when learning from instructional animations and Static Pictures, Learn. Indiv Differ 20 (5) (2010) 479–483.
[11] J.C. Castro-Alonso, M. Wong, O.O. Adesope, P. Ayres, F. Paas, Gender imbalance in instructional dynamic versus static visualizations: a meta-analysis, Educ. Psychol. Rev. 31 (2) (2019) 361–387.
[12] T.N. Höffler, D. Leutner, Instructional animation versus static pictures: a meta-analysis, Learn. InStruct. 17 (6) (2007) 722–738.
[13] M.T. McCrudden, D.N. Rapp, How visual displays affect cognitive processing, Educ. Psychol. Rev. 29 (3) (2015) 623–639.
[14] J.C. Castro-Alonso, P. Ayres, F. Paas, Comparing apples and oranges? A critical look at research on learning from statics versus animations, Comput. Educ. 102 (2016) 234–243.
[15] S. Türkay, The effects of whiteboard animations on retention and subjective experiences when learning advanced physics topics, Comput. Educ. 98 (2016) 102–114.
[16] M. Li, C.W. Lai, W.M. Szeto, Whiteboard animations for flipped classrooms in a common core science general education course, in: J. Domenech, P. Merello, E. de la Poza, D. Blazquez, R. Peña-Ortiz (Eds.), 5th International Conference on Higher Education Advances. 2019 June 26-28, Editorial Universitat Politècnicade València, Valencia, Spain, 2019, pp. 929–938.
[17] R.K. Lowe, J. Sweller, The modality principle in multimedia learning, in: R.E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning, second ed., Cambridge University Press, Cambridge, 2014, pp. 227–246.
[18] F. Suhroh, B. Yudi Cahyono, U. Praba Astuti, Effect of using whiteboard animation in Project-Based Learning on Indonesian EFL students' English presentation skills across Creativity Levels, Arab World Engl. J. 6 (2020) 212–227.
[19] J. Sweller, Element interactivity and intrinsic, extraneous, and germane cognitive load, Educ. Psychol. Rev. 22 (2) (2010) 123–138.
[20] J. Sweller, J.J. van Merriënboer, F. Paas, Cognitive architecture and instructional design: 20 years later, Educ. Psychol. Rev. 31 (2) (2019) 261–292.
[21] J. Sweller, Working memory, long-term memory, and instructional design, J. Appl. Res. Mem. Cogn. 5 (4) (2016) 360–367.
[22] N. Cowan, The magical number 4 in short-term memory: a reconsideration of Mental Storage Capacity, Behav. Brain Sci. 24 (1) (2001) 87–114.
[23] G.A. Miller, The magical number seven, plus or minus two: some limits on our capacity for processing information, Psychol. Rev. 101 (2) (1994) 343–352.
[24] L. Peterson, M.J. Peterson, Short-term retention of individual verbal items, J. Exp. Psychol. 58 (3) (1959) 193–198.
[25] M. Krell, Evaluating an instrument to measure mental load and mental effort considering different sources of validity evidence, Cogent Education 4 (1) (2017), 1280256.
[26] F. Paas, A. Renkl, J. Sweller, Cognitive load theory and instructional design: recent developments, Educ. Psychol. 38 (1) (2003) 1–4.
[27] M. Klepsch, T. Seufert, Making an effort versus experiencing load, Frontiers in Education 6 (2021).
[28] S. Kalyuga, Effects of information transiency in multimedia learning, Procedia Soc. Behav. Sci. 30 (2011) 307–311.
[29] W. Leahy, J. Sweller, Cognitive load theory, modality of presentation and the transient information effect, Appl. Cognit. Psychol. 25 (6) (2011) 943–951.
[30] A.-M. Singh, N. Marcus, P. Ayres, The transient information effect: investigating the impact of segmentation on spoken and written text, Appl. Cognit. Psychol. 26 (6) (2012) 848–853.
[31] A. Wong, W. Leahy, N. Marcus, J. Sweller, Cognitive load theory, the transient information effect and e-learning, Learn. InStruct. 22 (6) (2012) 449–457.
[32] P. Ayres, F. Paas, Making instructional animations more effective: a cognitive load approach, Appl. Cognit. Psychol. 21 (6) (2007) 695–700.

[33] G.D. Rey, M. Beege, S. Nebel, M. Wirzberger, T.H. Schmitt, S. Schneider, A meta-analysis of the segmenting effect, Educ. Psychol. Rev. 31 (2) (2019) 389–419.
[34] I.A.E. Spanjers, T. van Gog, P. Wouters, J.J.G. van Merriënboer, Explaining the segmentation effect in learning from animations: the role of pausing and temporal cueing, Comput. Educ. 59 (2) (2012) 274–280.
[35] D. Leutner, Motivation and emotion as mediators in multimedia learning, Learn. InStruct. 29 (2014) 174–175.
[36] R.M. Ryan, E.L. Deci (Eds.), Self-determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness, Guilford Publications, New York, 2017.
[37] R. Moreno, R. Mayer, Interactive multimodal learning environments, Educ. Psychol. Rev. 19 (3) (2007) 309–326.
[38] R.E. Mayer, Cognitive theory of multimedia learning, in: Re Mayer (Ed.), The Cambridge Handbook of Multimedia Learning, second ed., Cambridge University Press, Cambridge, 2014, pp. 43–71.
[39] R.E. Mayer, Incorporating motivation into multimedia learning, Learn. InStruct. 29 (2014) 171–173.
[40] R.E. Mayer, G. Estrella, Benefits of emotional design in multimedia instruction, Learn. InStruct. 33 (2014) 12–18.
[41] R.K. Lowe, Animation and learning: selective processing of information in dynamic graphics, Learn. InStruct. 13 (2) (2003) 157–176.
[42] R.K. Lowe, W. Schnotz (Eds.), Learning with Animation: Research Implications for Design, Cambridge University Press, Cambridge, 2008.
[43] B.B. de Koning, H.K. Tabbers, R.M. Rikers, F. Paas, Towards a framework for attention cueing in instructional animations: guidelines for Research and Design, Educ. Psychol. Rev. 21 (2) (2009) 113–140.
[44] S. Schneider, M. Beege, S. Nebel, G.D. Rey, A meta-analysis of how signaling affects learning with media, Educ. Res. Rev. 23 (2018) 1–24.
[45] B.B. de Koning, H. Jarodzka, Attention guidance strategies for supporting learning from dynamic visualizations, in: R.K. Lowe, Plötzner (Eds.), Learning from Dynamic Visualization, Springer, Cham, 2017, pp. 255–278.
[46] R. Lowe, W. Schnotz, T. Rasch, Aligning affordances of Graphics with learning task requirements, Appl. Cognit. Psychol. 25 (3) (2010) 452–459.
[47] C. H`mida, O. Degrenne, N. Souissi, G. Rekik, K. Trabelsi, M. Jarraya, et al., Learning a motor skill from video and Static Pictures in physical education students—effects on technical performances, motivation and cognitive load, Int. J. Environ. Res. Publ. Health 17 (23) (2020) 9067.
[48] N.O. Ceylan, Using short stories in reading skills class, Procedia Soc. Behav. Sci. 232 (2016) 311–315.
[49] S.M. Nair, N.M. Yusof, S.C. Hong, Comparing the effects of the story telling method and the conventional method on the interest, motivation and achievement of Chinese primary school pupils, Procedia Soc. Behav. Sci. 116 (2014) 3989–3995.
[50] M.O. Riedl, R.M. Young, Narrative planning: balancing plot and character, J. Artif. Intell. Res. 39 (2010) 217–268.
[51] N. Padilla-Zea, F.L. Gutiérrez, J.R. López-Arcos, A. Abad-Arranz, P. Paderewski, Modeling storytelling to be used in educational video games, Comput. Hum. Behav. 31 (2014) 461–474.
[52] P. Ricœur, Hermeneutics and the Human Sciences: Essays on Language, Action, and Interpretation, Cambridge University Press, New York, 2016.
[53] D. Herman, Stories as a tool for thinking, in: D. Herman (Ed.), Narrative Theory and the Cognitive Sciences, Center for the Study of Language and Information, Chicago, 2003, pp. 163–192.
[54] G. Dettori, T. Giannetti, A. Paiva, Technology-mediated Narrative Environments for Learning, Sense Publishers, Rottterdam, 2006.
[55] J.V. Wertsch, Mind as Action, Oxford University Press, New York, 1998.
[56] E.M. Forster, Aspects of the Novel, Rosetta Books, New York, 2010.
[57] V. Clinton, T. Taylor, S. Bajpayee, M.L. Davison, S.E. Carlson, B. Seipel, Inferential comprehension differences between narrative and expository texts: a systematic review and meta-analysis, Read. Writ. 33 (9) (2020) 2223–2248.
[58] R.A. Mar, J. Li, A.T. Nguyen, C.P. Ta, Memory and comprehension of narrative versus expository texts: a meta-analysis, Psychon. Bull. Rev. 28 (3) (2021) 732–749.
[59] J.M. Mandler, N.S. Johnson, Remembrance of things parsed: story structure and recall, Cognit. Psychol. 9 (1) (1977) 111–151.
[60] J. Bruner, Acts of Meaning, Harvard University Press, Cambridge, 1990.
[61] M. Wu, The Effect of Story Narrative in Multimedia Learning, Arizona State University, Tempe (AZ), 2018 [doctoral dissertation].
[62] R.J. Gerrig, Experiencing Narrative Worlds: on the Psychological Activities of Reading, Yale University Press, London, 1993.
[63] F.C. Bartlett, Remembering: A Study in Experimental and Social Psychology, Cambridge University Press, Cambridge, 1932.
[64] R.C. Schank, R.P. Abelson, Scripts, Plans, Goals, and Understanding: an Inquiry into Human Knowledge Structures, Psychology Press, 1977.
[65] P.A. Tun, Age differences in processing expository and narrative text, J. Gerontol. 44 (1) (1989) 9–15.
[66] S. Klassen, The construction and analysis of A science story: a proposed methodology, Sci. Educ. 18 (3–4) (2009) 401–423.
[67] R. Isbell, J. Sobol, L. Lindauer, A. Lowrance, The effects of storytelling and story reading on the oral language complexity and story comprehension of young children, Early Child. Educ. J. 32 (3) (2004) 157–163.
[68] B.F. Ellis, Why tell stories? Storytelling Magazine 9 (1997) 21–23.
[69] L. Colon-Vila, Storytelling in an ESL classroom, Teaching Pre K- 8 (27) (1997) 58–59.
[70] L.M. Sáenz, L.S. Fuchs, Examining the reading difficulty of secondary students with learning disabilities, Remedial Special Educ. 23 (1) (2002) 31–41.
[71] A. Ogan, V. Aleven, C. Jones, Advancing development of intercultural competence through supporting predictions in narrative video, Int. J. Artif. Intell. Educ. 19 (3) (2009) 267–288.
[72] H. Qin, P.-L. Patrick Rau, G. Salvendy, Measuring player immersion in the computer game narrative, Int. J. Hum. Comput. Interact. 25 (2) (2009) 107–133.
[73] M.D. Dickey, Game design narrative for learning: appropriating adventure game design narrative devices and techniques for the design of interactive learning environments, Educ. Technol. Res. Dev. 54 (3) (2006) 245–263.
[74] P. Wouters, C. van Nimwegen, H. van Oostendorp, E.D. van der Spek, A meta-analysis of the cognitive and motivational effects of serious games, J. Educ. Psychol. 105 (2) (2013) 249–265.
[75] D.E. Alvermann, C.E. Hynd, G. Qian, Effects of interactive discussion and text type on learning counterintuitive science concepts, J. Educ. Res. 88 (3) (1995) 146–154.
[76] L.J. Cunningham, M.D. Gall, The effects of expository and narrative prose on student achievement and attitudes toward textbooks, J. Exp. Educ. 58 (3) (1990) 165–175.
[77] S. Golke, R. Hagen, J. Wittwer, Lost in narrative? the effect of informative narratives on text comprehension and metacomprehension accuracy, Learn. InStruct. 60 (2019) 1–19.
[78] J.T. Hartley, Reader and text variables as determinants of discourse memory in adulthood, Psychol. Aging 1 (2) (1986) 150–158.
[79] M.B. Wolfe, J.A. Mienko, Learning and memory of factual content from narrative and expository text, Br. J. Educ. Psychol. 77 (3) (2007) 541–564.
[80] M.B. Wolfe, J.M. Woodwyk, Processing and memory of information presented in narrative or expository texts, Br. J. Educ. Psychol. 80 (3) (2010) 341–362.
[81] H. Pashler, P.M. Bain, B.A. Bottge, A. Graesser, K. Koedinger, M. McDaniel, J. Metcalfe, Organizing Instruction and Study to Improve Student Learning. IES Practice Guide (NCER 2007-2004), National Center for Education Research, 2007.
[82] J. Szurmak, M. Thuna, Tell me a story: the use of narrative as a tool for instruction, in: D.M. Mueller (Ed.), Imagine, Innovate, Inspire. Indianapolis: the Proceedings of the ACRL 2013 Conference, 2013, pp. 546–552.
[83] E.F. Schneider, A. Lang, M. Shin, S.D. Bradley, Death with a story. How story impacts emotional, motivational, and physiological responses to first-person shooter video games, Hum. Commun. Res. 30 (3) (2004) 361–375.
[84] S.T. Murphy, L.B. Frank, J.S. Chatterjee, L. Baezconde-Garbanati, Narrative versus nonnarrative: the role of identification, transportation, and emotion in reducing health disparities, J. Commun. 63 (1) (2013) 116–137.
[85] M.C. Clark, Narrative learning: its contours and its possibilities, N. Dir. Adult Cont. Educ. 2010 (126) (2010) 3–11.
[86] A. Arguel, E. Jamet, Using video and Static Pictures to improve learning of procedural contents, Comput. Hum. Behav. 25 (2) (2009) 354–359.
[87] F. Stebner, T. Kühl, T.N. Höffler, J. Wirth, P. Ayres, The role of process information in narrations while learning with animations and Static Pictures, Comput. Educ. 104 (2017) 34–48.

[88] F. Faul, E. Erdfelder, A.-G. Lang, A.G. Buchner, *Power 3: a flexible statistical power analysis program for the social, Behavioral, and Biomedical Sciences, Behav. Res. Methods 39 (2) (2007) 175–191.
[89] Sprouts Deutschland, Der Pygmalion Effekt. [Video], 2019. Available from: https://www.youtube.com/watch?v=z9dm-9BdoRQ. (Accessed 5 January 2023). Accessed.
[90] R. Rosenthal, The Pygmalion effect and its mediating mechanisms, in: J. Aronson (Ed.), Improving Academic Achievement: Impact of Psychological Factors on Education, Academic Press, London, 2002, pp. 25–36.
[91] J.E. Tuovinen, Optimising student cognitive load in computer education, in: A.E. Ellis (Ed.), Proceedings of the Australasian Conference on Computing Education, Association for Computing Machinery, New York, 2000, pp. 235–241.
[92] R.E. Mayer, R.B. Anderson, Animations need narrations: an experimental test of a dual-coding hypothesis, J. Educ. Psychol. 83 (4) (1991) 484–490.
[93] L.J. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika 16 (3) (1951) 297–334.
[94] J.C. Nunnally, Psychometric Theory, McGraw-Hill, New York, 1978.
[95] O. Chen, S. Kalyuga, J. Sweller, The expertise reversal effect is a variant of the more general element interactivity effect, Educ. Psychol. Rev. 29 (2) (2016) 393–405.
[96] F. Krieglstein, M. Beege, G.D. Rey, P. Ginns, M. Krell, S. Schneider, A systematic meta-analysis of the reliability and validity of subjective cognitive load questionnaires in Experimental Multimedia Learning Research, Educ. Psychol. Rev. 34 (4) (2022) 2485–2541.
[97] F. Guay, R.J. Vallerand, C. Blanchard, On the assessment of situational intrinsic and ˆ extrinsic motivation: the Situational Motivation Scale (SIMS), Motiv. Emot. 24 (3) (2000) 175–213.
[98] U. Schallberger, Kurzskala zur Erfassung der positiven Aktivierung, negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS) [Scales for assessing positive/negative activation and valence in experience sampling studies], Department of Psychology, University of Zurich, Zurich, 2005. Research Report.
[99] R.E. Mayer, The Cambridge Handbook of Multimedia Learning, University of Cambridge, New York, 2014.
[100] R.E. Clark, Reconsidering research on learning from Media, Rev. Educ. Res. 53 (4) (1983) 445–459.
[101] J. Hamari, J. Koivisto, Why do people use gamification services? Int. J. Inf. Manag. 35 (4) (2015) 419–431.
[102] D.R. Sanchez, M. Langer, R. Kaur, Gamification in the classroom: examining the impact of gamified quizzes on student learning, Comput. Educ. 144 (2020), 103666.
[103] J.A. Plucker, M.C. Makel, Replication is important for educational psychology: recent developments and key issues, Educ. Psychol. 56 (2) (2021) 90–100.
[104] Sprouts Deutschland, Die 8 Stufen der Entwicklung nach Erik Erikson [Video], 2019. Available from: https://www.youtube.com/watch?v=7nCCsB6BAHk. (Accessed 5 January 2023). Accessed.
[105] E. Erikson, Childhood and Society, second ed., W. W. Norton, New York, 1950/1963.
[106] R. Moreno, Does the modality principle hold for different media? A test of the method-affects-learning hypothesis, J. Comput. Assist. Learn. 22 (3) (2006) 149–158.
[107] S. Schneider, M. Beege, S. Nebel, L. Schnaubert, G.D. Rey, The cognitive-affective-social theory of learning in Digital Environments (CASTLE), Educ. Psychol. Rev. 34 (1) (2021) 1–38.
[108] F. Paas, J.E. Tuovinen, J.J. van Merriënboer, A. Aubteen Darabi, A motivational perspective on the relation between mental effort and performance: optimizing Learner Involvement in instruction, Educ. Technol. Res. Dev. 53 (3) (2005) 25–34.
[109] R.E. Mayer, C.S. DaPra, An embodiment effect in computer-based learning with animated pedagogical agents, J. Exp. Psychol. Appl. 18 (3) (2012) 239–252.
[110] P.B. Sederberg, J.W. Miller, M.J. Howard, M. Kahana, The temporal contiguity effect predicts episodic memory performance, Mem. Cognit. 38 (6) (2010) 689–699.
[111] J.D. Bransford, R.D. Sherwood, T.S. Hasselbring, C.K. Kinzer, S.M. Williams, Anchored instruction: why we need it and how technology can help, in: D. Nix, R. J. Spiro (Eds.), Cognition, Education, and Multimedia, Routledge, London, 2012, pp. 129–156.
[112] T.N. Höffler, D. Leutner, The role of spatial ability in learning from instructional animations – evidence for an ability-as-compensator hypothesis, Comput. Hum. Behav. 27 (1) (2011) 209–216.
[113] M. Hegarty, Multimedia learning about physical systems, in: R.E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning, Cambridge University Press, Cambridge, 2005, pp. 447–466.
[114] A.O. Akturk, I. Sahin, Literature review on metacognition and its measurement, Procedia Soc. Behav. Sci. 15 (2011) (2011) 3731–3736.
[115] M. Beege, S. Schneider, S. Nebel, G.D. Rey, Look into my eyes! exploring the effect of addressing in educational videos, Learn. InStruct. 49 (2017) 113–120.
[116] M. Bannert, C. Sonnenberg, C. Mengelkamp, E. Pieger, Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance, Comput. Hum. Behav. 52 (2015) 293–306.