Learning from text and animations: a study into the need for cross-representational signaling

Desiron, Juliette; Betrancourt, Mireille; de Vries, Erica

How to cite

DESIRON, Juliette, BETRANCOURT, Mireille, DE VRIES, Erica. Learning from text and animations: a study into the need for cross-representational signaling. In: L'Année psychologique, 2021, vol. 121, n° 4, p. 393–416. doi: 10.3917/anpsy1.214.0393

This publication URL: https://archive-ouverte.unige.ch/unige:157856
Publication DOI: 10.3917/anpsy1.214.0393

NUMERO SPECIAL "APPRENDRE AVEC LES TECHNOLOGIES DE L'INFORMATION"

L'ANNEE PSYCHOLOGIQUE/ TOPICS IN COGNITIVE PSYCHOLOGY

Learning from text and animations: a study into the need for cross-representational signaling

Authors

Juliette C. Désiron[a*], Mireille Bétrancourt[a], Erica de Vries[a]

[a] University of Geneva, TECFA

[b] Univ. Grenoble Alpes, LaRAC

To cite this article

Désiron, J., Bétrancourt, M. & de Vries, E. (2021). Learning from text and animations: a study into the need for cross-representational signaling. *L'Année psychologique*, 121, 393-416. https://doi.org/10.3917/anpsy1.214.0393

[*] *Corresponding author : Juliette.Desiron@unige.ch; Mireille.Betrancourt@unige.ch; erica.devries@univ-grenoble-alpes.fr*

Abstract. Current theories on multimedia learning posit that the presence of static or animated visualizations improve text comprehension but only if learners accurately integrate the multiple sources of information. This study investigated the effect of labels and colour coding to link the co-referring verbal and pictorial information in a document, i.e. cross-representational signaling (CRS). Sixteen-year-old students of secondary pre-vocational school (n =148) with low prior knowledge and reading abilities studied a document about the black kite migration on tablets, in one of three formats: multimedia (text and animations) with CRS, multimedia without CRS and text only. Learning outcome was measured with a questionnaire and a drawing task. The results show that format had a significant effect on learning, with the multimedia format outperforming the text format only when CRS was included. Animated visualizations effectively support text comprehension and learning when the necessary guidance for text-picture integration is provided.

Keywords: Text, Animation, Multimedia learning, Signaling, Inference

1. Introduction

These past twenty years, digital learning technologies were developed with the underlying goals of facilitating global access to knowledge and promoting learning through the use of advanced functionalities, such as interactivity, personalization, collaboration or multimedia. Tablets are of particular interest in education because of their multipurpose functionalities, the high level of acceptability among teachers and both beneficial effects on learning and enjoyment in the classroom (Griffith et al., 2020; Haßler et al., 2016; Xie et al., 2018). Yet, few studies provide recommendations on how to adapt existing instructional material or how to exploit the particular additional features of tablets to create effective multimedia documents. Abundant research on learning from multimedia documents indicates that providing an additional source of information in a pictorial format, such as a static or animated visualization, can support text comprehension (e.g. Hegarty & Just, 1993; Mayer, 1989; Peeck, 1974; Schnotz & Bannert, 2003). However, this multimedia effect only occurs if learners have the ability to integrate verbal and pictorial information in a coherent representation. The present research investigates the effect of signaling the links between verbal and pictorial information on learning from text and animated visualizations when considering learners with low prior knowledge and reading abilities.

1.1. Learning from multimedia documents

Following the theory of text processing developed by Kintsch and van Dijk (1978), the construction of a mental model requires learners not only to comprehend the text base, but also to generate inferences. Indeed, inference generation is the process by which learners integrate elements distributed throughout the text in order to reach full understanding. A difficulty arises when documents include information in multiple representational formats, such as text and pictures. Two prominent theories, proposed by Mayer (Mayer, 2005, 2014) and Schnotz (Schnotz, 2014; Schnotz & Bannert, 2003), currently describe the specificities of multimedia

learning. The two models build on Paivio's dual coding theory (DCT; Clark & Paivio, 1991; Paivio, 1971) which describes how the processing of text-picture combinations affects comprehension. DCT predicts higher learning outcomes after a presentation with both a picture and a verbal label of an object than with twice the information in one media. Anchoring information in two representational channels, rather than a single one, results in the construction of a stronger memory trace. Many studies (e.g. Hegarty & Just, 1993; Mayer, 1989; Peeck, 1974; Schnotz & Bannert, 2003) provide evidence of higher comprehension when learning from text and pictures compared to text alone. According to the cognitive theory of multimedia a learning (CTML) (Mayer, 2005, 2014), this multimedia effect results from the fact that complementary information provided by the pictorial representation supports the construction of a mental model of the situation, which is required for deep understanding and thus learning (see for example Schüler et al. 2013 for a more detailed explanation).

Digital technology offers the possibility to include animated or interactive visualizations and not only static illustrations. As demonstrated by several meta-analyses, animated visualizations benefit learning compared to static visualizations (Berney & Bétrancourt, 2016, Höffler & Leutner, 2007), but the effect is moderated by a large number of variables related to individual characteristics (e.g. visuospatial abilities) or to material characteristics (e.g. interactivity). In a recent meta-analysis, Ploetzner and colleagues (2020) demonstrated that animation is particularly effective for presenting complex changes over time.

1.2. Improving multimedia integration with signaling

The integration process required to successfully link multiple sources of information can be challenging for students (Ainsworth et al., 2002). Schnotz (2014), in the Integrated Text and Picture Comprehension (ITPC) model, states that creating a mental model of a text-picture document relies on both perceptual and semantic processing. Thus, learners have to integrate prior knowledge stored in long-term memory with new information handled in working

memory. This model includes a coherence principle, which predicts that the higher the overlap across verbal and pictorial representations (where all or most information overlaps so that readers can easily connect the two representations), the higher comprehension. On the contrary, with none or only few instances of information overlap across representations, lower comprehension is expected. Signaling features, such as arrows, colour coding, temporal cues, may highlight the correspondence between co-referring verbal and pictorial representations. Numerous studies investigated the effectiveness of signaling to support learning from multimedia documents (see van Gog, 2014, Richter et al., 2016 and Schneider et al., 2018). In their meta-analysis, Richter et al. (2016) found a small to medium positive effect of signaling moderated by prior knowledge, especially benefiting learners with low prior knowledge. In the framework of cognitive load theory, a positive effect of signaling by colour coding is explained by the fact that the search for co-referring information in pictorial and verbal format is facilitated when both are signaled by the same colour, thus reducing the split-attention effect (Kalyuga et al., 1999). A recent study (Cammeraat, Rop & de Koning, 2020) showed that colour coding is beneficial to learning whereas reducing spatial distance between text and picture is not. Their findings are in line with those of Florax and Ploetzner (2010) who did not find an effect of spatial integration on learning but demonstrated an effect of text segmentation and picture labelling. Therefore, colour coding and picture labelling seem more promising features than spatial integration for facilitating mental integration of co-referring verbal and pictorial information. Following theory and previous research (Boucheix et al., 2013; Jamet et al., 2008; Kalyuga et al., 1999; Mason et al., 2013; Scheiter & Eitel, 2015; Richter et al., 2017), we chose to implement cross-representational signaling (CRS) via colour coding and picture labelling to highlight links between co-referring verbal and pictorial information.

1.3. Prior knowledge and reading abilities affect the multimedia effect

Based on Schnotz' previous research (Schnotz, 2002, 2005; Schnotz & Bannert, 2003), the ITPC model (2014) includes several predictions regarding the effectiveness of using multiple sources of information to improve or facilitate learning. These predictions cover effects of design combinations and of learner characteristics, namely prior knowledge and reading abilities. Hence, Schnotz (2014) considers that there is a risk of redundancy for high prior knowledge readers. Sweller and colleagues (2003) coined the term "expertise reversal effect" and argue that learners with high prior knowledge process multimedia documents with pre-existing schemas that help them organize information in working memory. Thus, whereas novices benefit from instructional guidance, which compensates for their missing schemas, redundant information is not necessary for experts, and might even be harmful to their comprehension (Richter and colleagues, 2016, 2017; Richter & Scheiter, 2019; Seufert, 2019). Therefore, research on multimedia learning often focuses on low prior knowledge learners (e.g. Boucheix & Guignard, 2005; Lowe & Boucheix, 2008; Scheiter & Eitel, 2015). Learning with multimedia documents with written text may be influenced by other individual characteristics, such as reading ability. For Schnotz (2014), struggling readers process verbal representations with difficulty and therefore additional visual representations may enhance comprehension. The picture offers an additional route for mental model construction (p. 88). Yet, unlike prior knowledge, which tends to be controlled in multimedia research involving written text, reading ability is often ignored.

1.4. Research aim and hypotheses

Previous research in multimedia learning showed that signaling benefits learners with low prior knowledge (e.g. Scheiter & Eitel, 2015). However, few studies (e.g. Ozcelik et al., 2009) address the effect of signaling when learning from instructional material where animations are adjunct to written text. Despite evidence that spoken text is more effective than written text

when learning from animation, schools still use written text with the educational goal to improve reading skills. Moreover, whereas prior knowledge is usually controlled for, this is rarely the case for reading abilities. Thus, this study investigates the effect of animation on learning from a written text, controlling for both prior knowledge and reading abilities.

Hypothesis 1. Following the multimedia principle and animation research, we expect that learning from text with an adjunct animation will lead to higher learning outcome to inference questions than learning from text only (*multimedia effect*).

Hypothesis 2. Our second prediction is that the presence of cross-representational signaling results in higher learning outcomes compared to both text only and text with animations (*signaling principle*). Indeed, cross-representational signaling aims at guiding attention and thus should foster better text-animation integration in the mental model.

2. Method

2.1. Participants

One hundred and forty-eight students (61.5% females) in their first year of a prevocational track school took part in the experiment (mean age =16 years, *SD* = 11 months). Teachers proposed the task to their students as part of a 45-minutes class activity, and students who did not give their informed consent were given a silent reading task. Within the context of one school, ten classes could take part in the study. To avoid awareness of the other conditions, two classes were assigned to the text only condition (n = 25) and eight to the multimedia conditions. Within the multimedia classes, participants were randomly assigned to one of the two conditions either with (n = 60) or without CRS (n = 63).

This study was approved by the ethics and the “research in the school” committee of the first author’s University.

2.2. Material

The experimental material was a four-page long document on the migration of the Black kite (thus involving spatial change over time), from the school curriculum in French. It was presented using 9.7 inches tablets in landscape format with the text on the left side. In the multimedia conditions, an animation was presented on the right side. The right side was left blank in the control condition. The text was 430 words long and animations lasted between 10.56 seconds and 22.80 seconds.

The multimedia without CRS group studied the material with a text and an animation on each page (Appendix A1). The multimedia with CRS group studied the material in multimedia format including cross-representational signaling and the text only group studied the material without animations nor signaling. Following Désiron and colleagues (2018), changes to the document for including CRS thus consisted in the insertion of captions and arrows in the animation and the use of mutually referring colour coding in both text and animation (Appendix A2).

2.3. Measures

Prior knowledge. To control for prior knowledge of black kite migration, we used a questionnaire with 6 statements on a 5-item self-rating scale (maximum score: 30) ranging from ‘do not know’ to ‘know very well’ (e.g. *I […] what animal migration is*) or from ‘cannot explain’ to ‘can explain very well’ (e.g. *I […] how a thermal column is formed*). Thus, in accordance with multimedia research (Mayer, 2014; Schmidt-Weigand, et al., 2010) and recommendations from Ozuru and colleagues (2013), a self-report measure of prior knowledge was used.

Reading abilities. To measure reading ability, we used two tests found to be good predictors of multimedia comprehension in a previous study (Désiron, et al., 2021). The vocabulary test is a French version of the Hill Mill assessment developed by Deltour (1993), which asks

participants to determine 33 synonyms with 6 options each and an 8-minute time limitation (minimum score: 11, maximum score: 44). The verbal reasoning test is a translation of the test designed by Meteyard and colleagues (2015), which assesses the ability to generate inferences in answering four open-ended questions (maximum score: 12 points) after reading short texts. Each question was scored on two (0,1) to three levels (0, 0.5, 1) depending on the level of inference, and following procedure provided by Meteyard and colleagues (2015).

Comprehension questionnaire. Following research on text processing (e.g. Ozuru et al., 2009), the comprehension test covered three levels. Four text-based questions measured retention of elements explicitly stated in the text (e.g. *What do black kites do during their stay in Geneva?* – referring to page 1). Four local inference questions measured comprehension of elements requiring the generation of bridging inferences from elements no more than a sentence apart (e.g. *After crossing the Pyrenees to go south, where do black kites pass? Give, in the order of the migratory journey, three places mentioned in the document.* – referring to page 3). Finally, four global inference questions measured comprehension of elements requiring the generation of bridging inferences from elements dispersed in the text (e.g. *Do black kites need thermal ascendency for both directions of crossing the Pyrenees? Justify your answer.* – referring to pages 2 and 4). Moreover, findings from Ozuru and colleagues (2013) indicate that open-ended questions are a more sensitive measure of inference generation than multiple-choice questions. Therefore, expected answers to comprehension questions ranged from one word up to two sentences depending on the level of comprehension. Each question was scored following an analysis grid considering idea units from the text and pictures. The value of idea units ranged from .50 to 1 depending on the number of ideas expected (Appendix B). The total score for each level of comprehension ranged from 0 to 4 points. Answers were evaluated by the first author, and a second rating was done by the second author on a random subset of 10% of the comprehension questions (n = 30). Inter-rater reliability was determined by intraclass

correlation coefficients and was ICC (3, 1) = .98 for the comprehension questions. The raters jointly settled the few differences in the two ratings (α = .70).

Drawing task. A nonverbal task was designed to assess the mental model of visuospatial elements needed for an accurate internal representation of bird migration. Following Lowe and Boucheix (2008, 2011) who used a manipulation task to check the acquisition of the rotation of gears, we designed a drawing task. Learners were asked to retrace the journey of the birds on a background map displayed on the tablet. The spatial representation was assessed with an analysis grid paired with an annotated background map (Appendix C). Learner productions were matched to the description in the document regarding the General path (4 points) and the number of Key locations (5 points). Thus, points for the General path could range from 0 to 4, and for the Key locations from 0 to 5. Answers were evaluated by the first author, and a second rating was done by the second and third authors on a random subset of 20% of the productions (n = 15). Inter-rater reliability was determined by intraclass correlation coefficients and was ICC (3, 1) = .991 for the General path, and ICC (3,1) = 1.00 for the Key locations. The raters jointly settled the few differences in the two ratings (α = .85).

Procedure. This experiment was conducted in school, during classes lasting 45 minutes, with up to ten participants at a time. Participants first completed the knowledge questionnaire, before reading the experimental material in either one of the three format conditions, without time limit. Participants were then prompted to answer the comprehension questions, complete the drawing task, and complete the two reading ability tests (presented in random order). At the end of the allocated time, participants were debriefed with regard to the research hypotheses.

3. Data analysis

Data analyses were conducted using the statistical software Jamovi (v.1.2.27 for mac OS), and the significance level was set at p =.05. We used a multivariate analysis of covariance (MANCOVA) to assess between-group differences, and used η^2_p as the effect size index,

taking .01, .06 and .14 as the values for small, medium, and large effect sizes (Cohen, 1988). Planned covariates were individual characteristics (prior knowledge and reading abilities).

As group sizes were unequal, we checked the assumption of homogeneity of variances for our five outcome variables. Levene's test for equality of variances was non-significant for all the outcome variables ($p < .05$).

4. Results

4.1. Individual characteristics

Following recommendations from the ITPC model, a multimedia effect is more likely to occur when learners have either low prior knowledge or low reading abilities.

Learners in this study score between 0 and 17 at the prior knowledge questionnaire, with a mean score of 4.28 ($SD = 3.35$). They were qualified as low prior knowledge learners on the topic of the black kite migration.

With regard to reading abilities, scores to the verbal reasoning assessment were between .50 and 9.50, with a mean score of 5.61 ($SD = 1.97$). This test is not normed, but descriptive data indicate overall low to medium verbal reasoning ability. Scores to vocabulary ranged between 11 and 33, with a mean score of 23.2 ($SD = 4.01$). Following norms from the vocabulary test (Deltour, 1993b), readers scoring below 33.5 are considered to be struggling. Therefore, based on scores at the vocabulary and the verbal reasoning tests, we qualified participants in this experiment as struggling readers.

Pearson correlations indicated that prior knowledge was significantly positively associated with global inferences questions (r (146) = .221, $p = .007$). Regarding verbal abilities, vocabulary performance was correlated with text-based (r (146) = .282, $p < .001$) and local inference questions (r(146) = .344, $p < .001$), and verbal reasoning with text-based questions (r (146) = .328, $p < .001$). Thus, prior knowledge and reading abilities were entered as covariates in the analyses.

4.2. Effect of animation and CRS

A 3-document format (text only, multimedia without CRS, multimedia with CRS) MANCOVA was conducted with scores on comprehension questions and drawing task as the dependent variables. Covariates included were prior knowledge, vocabulary and verbal reasoning. Table 1 displays the mean scores and standard deviations for scores to comprehension questions and drawing task, for each condition.

The multivariate test was significant, Roy's largest root = .175, F (5, 139) = 4.87, p < .001. While descriptive data show a tendency confirming our hypotheses (Table 1), we further analysed main effects when distinguishing comprehension questions by level and drawing task by dimension, with a Bonferroni correction.

Table 1. Estimated Marginal Means (SE) of comprehension questions and drawing task (n = 148) - Moyennes marginales estimées (ES) pour les questions de compréhension et la tâche de dessin (n = 148)

	Text only	Multimedia without CRS	Multimedia with CRS
Comprehension questions			
Text-based (max. 4)	1.10 (.20)	1.14 (.13)	1.26 (.12)
Local-inference (max. 4)	0.90 (.19)	1.09 (.12)	1.19 (.12)
Global-inference (max. 4)	0.57 (.15)	0.98 (.10)	1.16 (.09)
Drawing task			
General path (max. 4)	1.52 (.26)	2.05 (.17)	2.67 (.16)
Key locations (max. 5)	2.62 (.24)	2.89 (.15)	3.65 (.15)

4.2.1. Effect on comprehension questions

There was a main effect of format for global inference questions, F (2, 142) = 5.18, p = .007, η^2_p = .068 (Figure 1). Pairwise comparisons indicated that the only significant difference was between the text only and multimedia with CRS conditions (p = .005). There were no significant differences for the other two levels of comprehension (text-based and local inference questions).

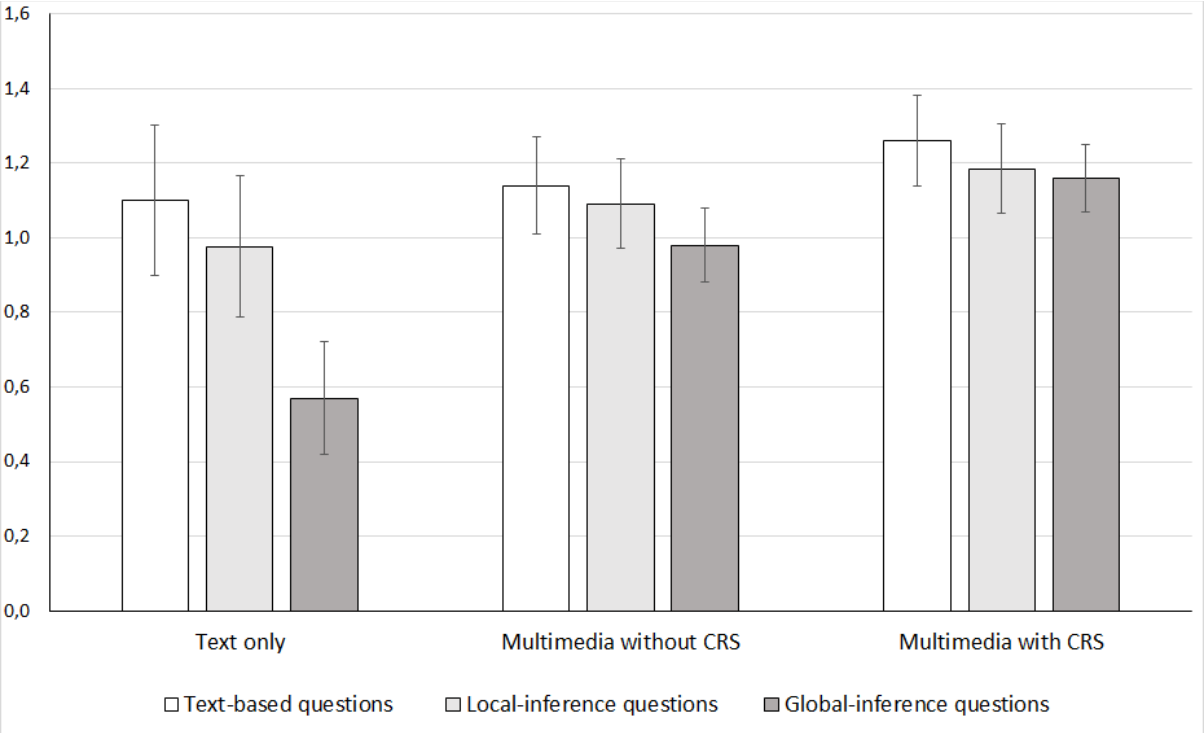


Figure 1. Estimated marginal means and standard errors scores for the three formats as a function of comprehension question level - Moyennes marginales estimées et erreurs standards pour les scores des trois formats en fonction du niveau de compréhension

4.2.2. Effect on drawing task

There was a main effect of format on scores in the drawing task for both for the General path, F (2, 142) = 7.96, p < .001, η^2_p = .101, and the Key locations, F (2, 142) = 9.63, p < .001, η^2_p = .119 (Figure 2). For the General path, pairwise comparisons revealed significant differences between text only and multimedia with CRS (p < .001), and between multimedia without CRS

and multimedia with CRS (p < .022) conditions. There were no significant differences for General path between text only and multimedia without CRS. Differences for Key locations were also significant between text only and multimedia with CRS (p < .001), and between multimedia without CRS and multimedia with CRS (p < .022). There were no significant differences for Key locations between text only and multimedia without CRS.

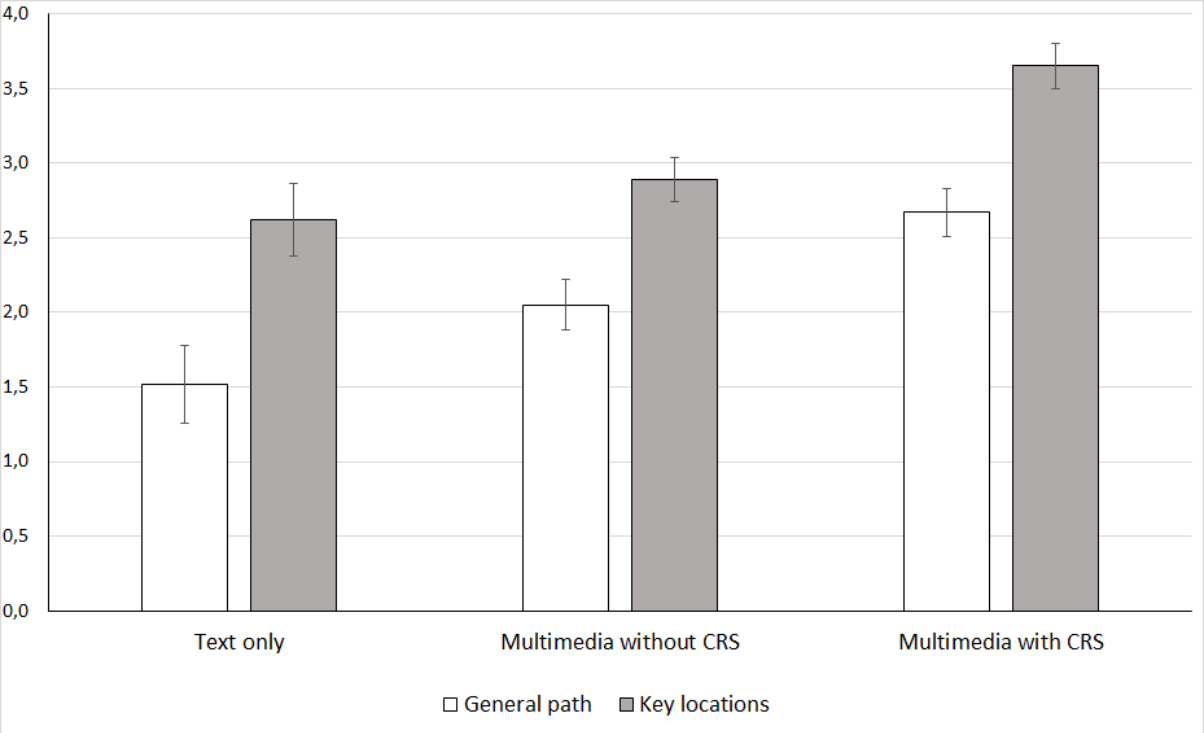


Figure 2. Estimated marginal means and standard errors scores for the three format conditions as a function of drawing task dimension - Moyennes marginales estimées et erreurs standards pour les scores des trois formats en fonction des dimensions pour la tâche de dessin

5. Discussion

This quasi-experimental study investigated whether cross-representational signaling (CRS), i.e. linking the verbal and pictorial information through colour coding and labelling, improves learning from a multimedia document composed of text and animations. Based on Schnotz' (2014) ITPC model, we expected that supporting learners in the construction of links between the verbal and the pictorial representations would facilitate the construction of a mental model of the situation described by the text (Kintsch & van Dijk, 1978). As Schnotz hypothesized that

cross-representational linking (correspondence) is important for learners with low prior knowledge and low reading abilities, we measured these two variables to ensure that the participants matched these criteria. Besides the two multimedia conditions (with and without CRS), a third text only condition was used as a control for the supportive function of the pictorial information.

Based on two meta-analyses from Richter and colleagues (2016) and Schneider and colleagues (2018), as well as on the text-processing model from Kintsch and van Dijk (1978), reading a multimedia document (text with animations) is expected to lead to higher scores than a text only document for questions requiring the generation of local or global inferences, and even more so when including cross-representational signaling. Multivariate analyses indicated a significant effect of condition on learning outcomes. For comprehension questions, the differences were significant only for global inferences, with worse performance in the text only condition than in the multimedia condition only when including CRS. Learning outcome assessed with the drawing task showed higher performance for multimedia with CRS than both text-only and multimedia without CRS. Thus, our results partially support hypothesis 1 corresponding to a beneficial effect of visualization (multimedia effect, Mayer, 2005, 2014) for text comprehension, and in the case at hand, of animated visualization when the content involves change over time (Ploetzner et al., 2020). The findings suggest that the multimedia effect in this study is dependent on cross-representational signaling, as specified in hypothesis 2. It is probable that the multimedia effect did not appear without CRS because learners had not only low prior knowledge on the topic but also low reading abilities. Therefore, they needed help to link the verbal and pictorial co-referring information, just as Schnotz hypothesized in his ITPC model. The beneficial effect of cross-representational signaling in multimedia documents is in line with previous research on signaling for text with static visualizations (Kalyuga et al., 1999; Mason et al., 2013), as well as the rare research on

signaling with text and animations (Ozcelik et al., 2009; Scheiter & Eitel, 2015). Yet, in this study, the effect appeared only for comprehension questions requiring global inferences and not for those requiring local inferences. Our assumption is that although prior knowledge and reading abilities were low, the local inferences did not require as much visual support as expected. Nonetheless, more research is needed to precisely ascertain the interaction between reading abilities and the generation of both local and global inferences from multimedia information.

5.1. Limitations and perspectives

One may argue that the lack of replication of the multimedia effect could be due to the fact that the multimedia material in this study was designed in a separated rather than in an integrated format, thus generating a split-attention effect, largely replicated in the literature (see Ginn's (2006) meta-analysis). As mentioned in the introduction, recent studies question the actual strength of the split-attention effect. These studies showed that the spatial distance between text and pictorial information is less important than the presence of features that help linking text and picture information together such as colour coding (Cammeraat et al., 2020) or picture labelling (Florax & Ploetzner, 2010). Future research should investigate the effect of integrating segments of text appearing sequentially as the animation unfolds, as Jamet, Gavota and Quaireau (2008) investigated for static visualizations.

In this study, we used a drawing task as an additional means to measure a multimedia effect. This task seemed particularly relevant to the topic presented in the instructional text, which included spatial and temporal information. A limitation is that the background map was shown in the multimedia conditions, whereas learners in the text only condition had to build their mental representation. This may have provided an advantage to learners in multimedia conditions. However, viewing the map in the learning phase (multimedia conditions) did not provide such an advantage, as no significant differences were found between text only and

multimedia without CRS. The significant differences between multimedia with CRS and the two other conditions corroborate that cross-representational signaling is needed to learn from a text and animation conveying spatio-temporal information.

As expected, correlation analyses showed that both prior knowledge and reading abilities had an impact on comprehension performances. However, students in vocational education in this study had both low prior knowledge on the topic and low reading abilities compared to their age group, so that a median split would have brought no additional information. Future research should investigate the effect of prior knowledge and reading abilities as predictors of the effectiveness of cross-representational signaling in multimedia learning with an adequate sample of participants varying on both variables. Another limitation of this study is the quasi-experimental design and as an unforeseen consequence, the unbalanced number of participants in the three conditions which could have lowered the power of the format factor.

5.2. Practical implications

The results of this study support the claim that multimedia instructional material that makes use of features afforded by tablets, such as the insertion of animations to support text comprehension, can be effective to promote learning. Moreover, this study shows that learners with low prior knowledge and low reading abilities benefit from animation only in the presence of cross-representational signaling. Publishers could consider designing material with on-demand support, where readers could activate signaling, either for the whole document, or during a rereading phase, similar to the study from Kalyuga et al. (1999). Finally, as tablets allow assessment of both verbal and pictorial productions (drawing task), possibilities for automated feedback could be exploited in the future.

Appendix A. Samples of experimental material – Extraits du matériel expérimental

A1. Sample page 3 with animation, no signaling – Extrait de la page 3 avec animation, sans signalement

Gibraltar, et l'hivernage en terres africaines

Les milans noirs traversent ensuite l'Espagne en quelques jours pour rejoindre l'Andalousie. Entre le 10 et le 20 août, la majorité de la population d'Europe centrale gagne l'Afrique en empruntant le détroit de Gibraltar.

Ils doivent encore franchir un obstacle, et pas des moindres : le désert du Sahara. Sans nécessairement longer la mer, ils le traversent au plus court et gagnent le sud du Sahara fin août.

A2. Sample page 3 with animation, with CRS - – Extrait de la page 3 avec animation, avec CRS

Gibraltar, et l'hivernage en terres africaines

Les milans noirs traversent ensuite l'**Espagne** en quelques jours pour rejoindre l'Andalousie. Entre le 10 et le 20 août, la majorité de la population d'Europe centrale gagne l'Afrique en empruntant le **détroit de Gibraltar**.

Ils doivent encore franchir un obstacle, et pas des moindres : le **désert du Sahara**. Sans nécessairement longer la mer, ils le traversent au plus court et gagnent le sud du Sahara fin août.

Appendix B. Sample of the analysis grid and participants' answers for the comprehension questions (in French) – Extrait de la grille d'analyse et réponses de participants pour les questions de compréhension

Question	0	0.5	1	Sample answers	Points awarded
Text base Que font les milans noirs lors de leur séjour à Genève?			Ils se reproduisent	"ils se reproduisent"	1
				"ils se reproduisent"	1
Local inference Après la traversée des Pyrénées en direction du sud, par où passent les milans noirs? Donnez, dans l'ordre du parcours migratoire, trois endroits cités dans le document		2 des endroits sur 3	1) L'Andalousie/ Espagne 2) le détroit de Gibraltar 3) le désert du Sahara	"espagne (andalousie) , afrique (gibraltar) et apres Genève"	0.5
				"par l'espagne puis le detroit de gibraltare pour arriver au desert du sahara"	1
Global inference Les milans noirs ont-ils besoin d'ascendance thermique pour les deux sens de traversée des Pyrénées? Justifiez votre réponse.		(0.5) de la France vers l'Espagne/ au retour/ équivalent (0.5) Parce que le côté français est en pentes verticales alors que le côté espagnol est en pente douce		"non pas pour la deuxième mais je ne saurai pas l'expliquer correctement je pense que c'est parce que c'est de l'autre côté donc nous n'avons pas besoin de ça"	0.5
				"non car les pentes sont moins raides depuis l'espagne"	1

Appendix C. Drawing task background map and sample answers – Tâche de dessin: carte de fond et exemples de réponses

C1. Annotated background map and analysis grid for Drawing task – Fond de carte annoté et grille d'analyse pour la tâche de dessin

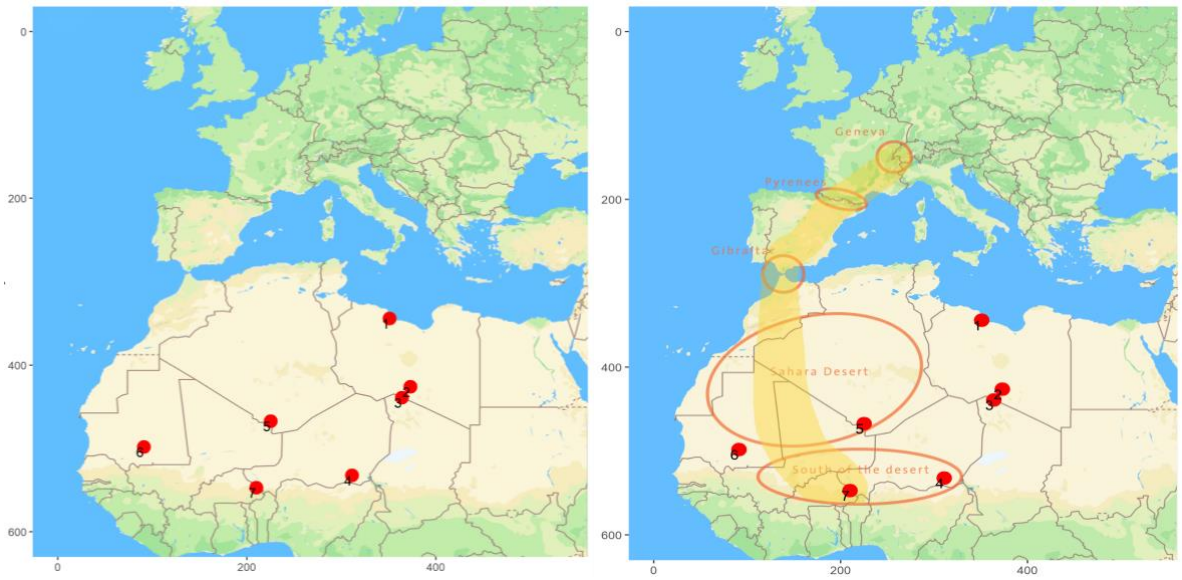
General Path

0	0.25	0.5	0.75	1
The path is much incomplete. Le trajet est très incomplet. Particularly : one of the 3 zones (France, Spain, Africa) is missing.	The 3 zones are covered but part of the path is off track. Example: goes above the sea, or another country.	Start AND finish areas are incorrect, OR numerous points are off track.	Start OR finish areas is not where it should be, but everything else is good. Nb. Start in Germany accepted.	Everything is good, including start and finish areas.

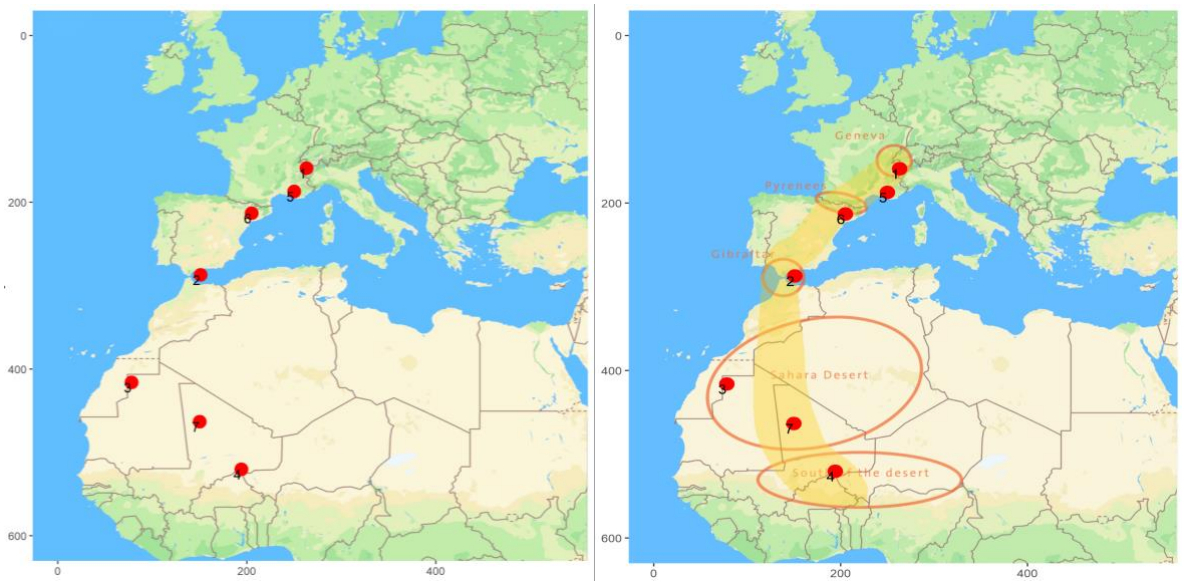
Key locations

0	1	2	3	4	5
None of the points is located at key locations.	There are points for only one key location.	There are points for only two key locations.	There are points for only three key locations.	There are points for only four key locations.	There is at least one point per key location. Nb. Ok if either in zone or touching delineation.

C2. Sample of participants maps (raw on the left, with annotated background on the right) – Exemples de cartes de participants (vue participant gauche, vue sur fond annoté à droite)



Path: 0 points Locations: 2 points



Path: 4 points Locations: 5 points

References

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, *11*(1), 25–61. https://doi.org/10.1207/S15327809JLS1101_2

Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, *101*, 150–167. https://doi.org/10.1016/j.compedu.2016.06.005.

Boucheix, J., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students ? Format , signaling and control of the presentation. *European Journal of Psychology of Education*, *20*(4), 369–388.

Boucheix, J.-M., Lowe, R. K., Putri, D. K., & Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learning and Instruction*, *25*, 71–84. https://doi.org/10.1016/j.learninstruc.2012.11.005

Cammeraat, S., Rop, G., & Koning, B. B. De. (2020). The influence of spatial distance and signaling on the split-attention effect. *Computers in Human Behavior*, *105*(October 2019), 106203. https://doi.org/10.1016/j.chb.2019.106203

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114*, 647–662.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Deltour, J. J. (1993a). Echelle de vocabulaire de Mill Hill de JC Raven: Adaptation française. In *Manuel des Raven*. Oxford Psychologists Press.

Deltour, J. J. (1993b). Echelle de vocabulaire de Mill Hill de JC Raven: Adaptation française.

In *Manuel des Raven*. Oxford: Oxford Psychologists Press.

Désiron, J. C., Bétrancourt, M., & de Vries, E. (2018). How cross-representational signaling affects learning from text and picture: An eye-tracking study. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10871 LNAI*, 725–728. https://doi.org/10.1007/978-3-319-91376-6_68

Désiron, J. C., Bétrancourt, M., & de Vries, E. (2021). Cross-Representational Signaling and Cohesion Support Inferential Comprehension of Text–Picture Documents. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.592509

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect ? The role of text segmentation , picture labelling , and spatial proximity. *Learning and Instruction*, *20*(3), 216–224. https://doi.org/10.1016/j.learninstruc.2009.02.021

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. Learning and Instruction, 16(6), 511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Griffith, S. F., Hagan, M. B., Heymann, P., Heflin, B. H., & Bagner, D. M. (2020). Apps As Learning Tools: A Systematic Review. *Pediatrics*, *145*(1). https://doi.org/10.1542/peds.2019-1579

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, *32*(2), 139–156. https://doi.org/10.1111/jcal.12123

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717–742. https://doi.org/10.1006/jmla.1993.1036

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: a meta-analysis. Learning and Instruction, 17(6), 722–738.

https://doi.org/10.1016/j.learninstruc.2007.09.013.

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*(2), 135–145. https://doi.org/10.1016/j.learninstruc.2007.01.011

Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior*, *24*(3), 852–861. https://doi.org/10.1016/j.chb.2007.02.018

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *25*(SUPPL. 1), 351–371. https://doi.org/10.1002/acp.1773

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *American Psychologist*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Larkin, J. H., & Simon, a. (1987). Why a Diagmm is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, *99*, 65–99. https://doi.org/10.1016/S0364-0213(87)80026-5

Lowe, R. K., & Boucheix, J.-M. (2008). Learning from animated diagrams: How are mental models built? In *Diagrammatic Representation and Inference* (pp. 266–281). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-87730-1_25

Lowe, R. K., & Boucheix, J.-M. (2011). Cueing complex animations: Does direction of attention foster learning processes? *Learning and Instruction*, *21*(5), 650–663. https://doi.org/10.1016/j.learninstruc.2011.02.002

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. *Reading Research Quarterly*, *48*(2), 199–214. https://doi.org/10.1002/rrq.41

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, *81*(2), 240–246. https://doi.org/10.1037/0022-0663.81.2.240

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31–46). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Meteyard, L., Bruce, C., Edmundson, A., & Oakhill, J. (2015). Profiling text comprehension impairments in aphasia. *Aphasiology*, *29*(1), 1–28. https://doi.org/10.1080/02687038.2014.955388

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers and Education*, *53*(2), 445–453. https://doi.org/10.1016/j.compedu.2009.03.002

Ozuru, Y., Briner, S., Kurby, C. A., & McNamara, D. S. (2013). Comparing comprehension measured by multiple-choice and open-ended questions. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, *67*(3), 215–227. https://doi.org/10.1037/a0032918

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Paivio, A. (1971). *Imagery and Verbal Processes*. Lawrence Erlbaum Associates, Inc.

Peeck, J. (1974). Retention of pictorial and verbal content of a text with illustrations. *Journal of Educational Psychology*, *66*(6), 880–888. https://doi.org/10.1037/h0021531

Ploetzner, R, Berney, S, Bétrancourt, M. (2020). A review of learning demands in instructional animations: The educational effectiveness of animations unfolds if the features of change need to be learned. Journal of Computer Assisted Learning, *36*, 838– 860.

https://doi.org/10.1111/jcal.12476

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Richter, J., Scheiter, K., & Eitel, A. (2017). Signaling text-picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology*, *110*(4), 544–560. https://doi.org/10.1037/edu0000220

Richter, J., & Scheiter, K. (2019). Studying the expertise reversal of the multimedia signaling effect at a process level : evidence from eye tracking. In *Instructional Science* (Vol. 47, Issue 6). Springer Netherlands. https://doi.org/10.1007/s11251-019-09492-3

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: evidence, theory, and practice. In S. B. Neuman, D. Dickinson, & K. (Eds.), *Handbook of early literacy research* (pp. 97–110). Guilford Press.

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, *23*(November 2017), 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

Schnotz, W. (2002). Towards anintegrated view of learning from text and visual displays. *Educational Psychology Review*, *14*(1), 101–120. https://doi.org/10.1023/A:1013136727916

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 49–69). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.005

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning, Second Edition* (2nd ed., pp. 72–103). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.005

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*(2), 141–156. https://doi.org/10.1016/S0959-4752(02)00017-8

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior*, *29*, 1590–1601. https://doi.org/10.1016/j.chb.2013.01.047

Spanjers, I. a. E., Wouters, P., van Gog, T., & van Merriënboer, J. J. G. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior*, *27*(1), 46–52. https://doi.org/10.1016/j.chb.2010.05.011

Sweller, J., Ayers, P. L., Kalyuga, S., & Chandler, P. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31.

van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 263–278). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.014

Xie, H., Peng, J., Qin, M., Huang, X., Tian, F., & Zhou, Z. (2018). Can touchscreen devices be used to facilitate young children's learning? A meta-analysis of touchscreen learning effect. *Frontiers in Psychology*, *9*(DEC), 1–15. https://doi.org/10.3389/fpsyg.2018.02580

Conflict of interest: none

Funding: This work was supported by the Swiss National Science Foundation [grant number P0GEP1_165256 / 1].