



The influence of spatial distance and signaling on the split-attention effect

Sven Cammeraat^a, Gertjan Rop^{a,b}, Bjorn B. de Koning^{a,*}

^a Department of Psychology, Education and Child Studies, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA, Rotterdam, the Netherlands

^b Rotterdam School of Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:

Split-attention effect
Cognitive load theory
Signaling
Spatial distance
Educational psychology

ABSTRACT

The split-attention effect seems a robust finding in the literature and shows that complex learning materials consisting of text and pictures lead to better learning when they are spatially integrated instead of spatially separated. In the current study, two experiments were conducted to investigate the effects of distance and signaling on the split-attention effect. In Experiment 1, a 2×2 design was used in which spatial distance (large vs. small) and signaling (signaling vs. no signaling) were manipulated, while in Experiment 2 it was tried to replicate the split-attention effect previously observed with the learning materials used in Experiment 1. In contrast to our hypotheses, spatial distance and signaling did not influence mental effort ratings and transfer performance in Experiment 1. Moreover, no split-attention effect was found on retention, transfer or mental effort ratings in Experiment 2. As no split-attention effect was obtained with the learning materials, it is not possible to examine the effects of distance and signaling on the split-attention effect. Concluding, the current results raise doubts about how robust the split-attention effect is, and demonstrate the importance of replication studies.

1. Introduction

In contemporary education, students are increasingly required to learn from instructional materials that combine text and pictures. For example, in a typical textbook, the functioning of the heart is explained in a picture with accompanying text. Such multimedia learning materials usually lead to a better learning performance than learning materials that solely consist of text (multimedia principle, Mayer, 2003). However, a critical aspect that has to be taken into account in this regard is the mental integration of the presented text and picture which likely has an influence on learning performance (Ayres & Sweller, 2014; Mayer & Fiorella, 2014). It has been demonstrated that when text is placed near the part of the picture that the text corresponds to (i.e., when text is spatially integrated in the picture), learning outcomes are higher than when the corresponding text and picture elements are presented as spatially separated sources of information (Ayres & Sweller, 2014; Ginns, 2006; Schroeder & Cenkci, 2018). This effect is referred to as the split-attention effect (Ayres & Sweller, 2014; Chandler & Sweller, 1991, 1992; Mayer & Fiorella, 2014) and is typically explained by Cognitive Load Theory (CLT; Paas, Renkl et al., 2003; Sweller, Ayres, & Kalyuga, 2011). According to CLT, learning materials presented in a spatially

separated format result in suboptimal learning because the spatial distance between text and picture imposes a high cognitive load on working memory. Specifically, in a spatially separated format learners need to visually search and reorientate between text and picture, while processing it in working memory, which imposes a high extraneous cognitive load (Ayres & Sweller, 2014; Kalyuga, Chandler, & Sweller, 1998; Sweller, Van Merriënboer, & Paas, 1998). Such unnecessary activities are not required when studying integrated text-picture materials, which increases the chance that learners' working memory capacity can be used for learning-related processes like schema construction (germane cognitive load, Sweller et al., 1998).

At present, the split-attention effect has been demonstrated in a myriad of studies which show that spatially integrated text and pictures leads to less extraneous load and/or higher learning outcomes than spatially separated materials (Angeli, Valanides, & Kirschner, 2009; Ayres & Sweller, 2014; Cierniak, Scheiter, & Gerjets, 2009; Ginns, 2006; Harter & Ku, 2008; Schroeder & Cenkci, 2018). However, prior studies typically compared a condition that studied spatially integrated text and pictures to a condition that studied text and pictures that were spatially separated from each other without taking into account design aspects in the spatially separated condition. That is, studies varied in how the text

* Corresponding author.

E-mail address: b.b.dekoning@essb.eur.nl (B.B. de Koning).

<https://doi.org/10.1016/j.chb.2019.106203>

Received 17 January 2019; Received in revised form 17 October 2019; Accepted 15 November 2019

Available online 16 November 2019

0747-5632/© 2019 Elsevier Ltd. All rights reserved.

and pictures in the spatially separated condition were presented. For example, there was variation in terms of the spatial distance between the textual and pictorial representations (e.g., very close to each other or at a larger distance) or the location where the text was presented (e.g., under the picture or next to the picture) (Pouw, Rop, de Koning, & Paas, 2019; Florax & Ploetzner, 2010). Recently, the precise role of varying spatial distance in learning from spatially separated text-picture materials is receiving increasing attention (Pouw et al., 2019). In the present study we build on this emerging research by empirically investigating the influence of varying spatial distance and using signaling (i.e., using labels to match text and picture elements) on learning from spatially separated text and integrated and picture materials.

1.1. The influence of spatial distance and signaling on cognitive load and learning

In a recent study by Pouw et al., 2019 three experiments were conducted to investigate whether the split-attention effect is actually caused by the spatial separation of the information sources. In the first two experiments, participants had to judge the similarity of two spatially separated cards with symbols that could vary in form, number, and colour. The results of both experiments showed that an increase in distance between the cards led to a higher load on working memory and a slower integration of the information sources. In the third experiment, they tried to replicate these findings with more complex multimedia learning materials about information transmission in the human nervous system by comparing three conditions: a spatially integrated condition, a spatially separated condition with a small distance between text and picture and a spatially separated condition with a large distance between text and picture. In contrast to their expectations, no split-attention effect was observed in this experiment on learning outcomes and experienced mental effort during learning. Moreover, no effects of spatial distance were found on learning outcomes or experienced mental effort. These results suggest that spatial distance did not lead to a split-attention effect with these text-picture instructional materials. This finding could, however, be explained by the fact that signaling was used in both spatially separated conditions. It is relatively well-established that signaling can facilitate mental integration of information sources by reducing cognitive demands associated with having to search for and link corresponding parts of the text and picture (De Koning & Jarodzka, 2017; Florax & Ploetzner, 2010; Ozcelik, Karakus, Kursun, & Cagiltay, 2009; Van Gog, 2014). Importantly, according to CLT searching and matching activities to link textual and pictorial information requires more working memory capacity if text and pictures are presented at a larger spatial distance from each other. With a larger distance between text and picture, learners have to keep information active in working memory for longer periods of time, which increases cognitive load and reduces learning outcomes (Puma, Matton, Paubel, & Tricot, 2018). It is conceivable that these negative effects of increasing spatial distance were alleviated by the use of signaling in Experiment 3 of Pouw et al., 2019. To test this hypothesis, the present study aimed to investigate whether signaling reduces the negative effect of a large (vs. small) spatial distance between spatially separated text and picture on learning and mental effort, thereby gaining more clarity about the cognitive processes underlying the split-attention effect.

The influence of distance on cognitive processing has mainly been investigated in research fields such as comparative visual search and embedded cognition. Several studies using comparative visual search tasks have shown that increasing the spatial distance between two information sources leads to longer processing times, longer fixation times, and less gaze switches between information sources (Hardiess, Gillner, & Mallot, 2008; Inamdar & Pomplun, 2003). Research in the field of embedded cognition also has shown that increasing the spatial distance between information sources increases cognitive demands because people rely more on memory and less on attention switches (Ballard, Hayhoe, & Pelz, 1995; Gray & Fu, 2004; Pouw, Van Gog, &

Paas, 2014). For example, Ballard et al. (1995) showed that people prefer to use a more perceptual strategy in case of a short distance between two information sources, but rely more on their memory when spatial distance increases. While these studies did not use typical multimedia learning tasks, a few studies exist in multimedia learning research which corroborate these findings using more educationally relevant materials (Bauhoff, Huff, & Schwan, 2012; Huff, Bauhoff, & Schwan, 2010). Together, these studies indicate the importance of distance and its effect on cognitive load, and are in accordance with the results of Experiment 1 and 2 reported by Pouw et al., 2019.

In learning from spatially separated information sources one possible strategy to support learning is to signal corresponding pieces of information in both information sources (cf. De Koning & Jarodzka, 2017; De Koning, Tabbers, Rikers, & Paas, 2009). The usefulness of signaling has been well-established in multimedia learning. Multiple studies have shown that signaling can have a positive effect on retention, transfer, and cognitive load when learning from textual and pictorial representations (Kalyuga, Chandler, & Sweller, 1999; Ozcelik et al., 2009; Schneider, Beege, Nebel, & Rey, 2018; Van Gog, 2014). For example, Kalyuga et al. (1999) showed that colour coding (i.e., a form of signaling in which colours are used to guide learners' attention) alleviated split-attention effects when learning about electrical circuits: Participants reported lower mental effort and obtained higher learning outcomes when colour coding was added to spatially separated learning materials than when no colour coding was added to spatially separated learning materials. This suggests that signaling can effectively decrease extraneous load by directing learners' attention to relevant parts of the learning materials or by facilitating organization and integration of the learning materials (cf. De Koning et al., 2009; Kalyuga et al., 1999; Ozcelik, Arslan-Arib, & Cagiltay, 2010; Ozcelik et al., 2009; Van Gog, 2014). The influence of signaling can be used to explain why Pouw et al. (2019; Experiment 3) did not find a split-attention effect. Even though spatial distance between the textual and pictorial information sources makes it more difficult to integrate textual and pictorial information (Bauhoff et al., 2012; Gray & Fu, 2004), signaling likely facilitates integration processes (De Koning et al., 2009; Van Gog, 2014). More specifically, given that a larger distance between spatially separated text and picture likely leads learners to rely more on working memory processes (cf. Ballard et al., 1995) and increases extraneous cognitive load, signaling is expected to reduce the (extraneous) working memory demands by directing learners to connections between the spatially separated information sources.

1.2. The present study

We conducted two experiments to investigate the influence of spatial distance and signaling in learning from split-attention instructional materials. The current study extends past research by specifically studying the effects of signaling and variations in spatial distance (in Experiment 1) and not just studying the effects of spatial integration of textual and pictorial information (in Experiment 2). In other words, previous research mostly compared spatially integrated learning materials to spatially separated learning materials (Ayres & Sweller, 2014; Ginns, 2006), whereas the present study also compares spatially separated learning materials with a small distance between text and picture to separated learning materials with a large distance between text and picture. Furthermore, the interaction between signaling and distance is considered to investigate whether signaling alleviates the negative effect of increased spatial distance. In both experiments, the materials covered the functioning of car brakes for which we used the materials of Johnson and Mayer (2012), who have shown a split-attention effect with these materials evidenced by higher transfer performance for the integrated condition than for the spatially separated condition.

2. Experiment 1

Experiment 1 investigated whether increasing the spatial distance between spatially separated text and pictures harms learning compared to presenting the same information at a small spatial distance and whether signaling alleviates this negative effect of a distance increase. Therefore, we employed a 2×2 design with distance (large vs. small) and signaling (yes vs. no) as between-subjects factors. We investigated the effect of distance and signaling on retention, transfer and invested mental effort (which is an indication of experienced cognitive load; Paas, 1992; Paas, Tuovinen et al., 2003). We expected that a large spatial distance would lead to worse transfer performance, worse performance on a label test where parts of the system had to be named, and higher mental effort than a small spatial distance (Bauhoff et al., 2012; Gray & Fu, 2004). Furthermore, we expected that conditions studying the materials with signaling would obtain higher transfer performance and label performance, and experience less mental effort compared to when signaling is not used (Schneider et al., 2018; Van Gog, 2014). Consistent with the findings reported by Johnson and Mayer (2012), we did not expect an effect of spatial distance or signaling on the free recall test. Finally, we expected to find an interaction between spatial distance and signaling, showing that the influence of spatial distance on mental effort and learning (transfer and labeling) is smaller when signaling is used compared to when signaling is not used (Pouw et al., 2019; Mayer, Steinhoff, Bower, & Mars, 1995).

2.1. Method

2.1.1. Participants and design

Participants were 107 ($M_{age} = 21.18$, $SD = 4.13$, range 18–48 years old; 89 females) undergraduate psychology students who participated for course credit. They were randomly assigned to one of four conditions: a condition with a small spatial distance between text and picture in which corresponding textual and graphical elements were coloured similarly (i.e., small signaling condition; $n = 27$), a condition with a small spatial distance between text and picture without colour-coding (i.e., small condition; $n = 27$), a condition with a large spatial distance between text and picture in which corresponding textual and graphical elements were coloured similarly (i.e., large signaling condition; $n = 27$), and a condition with a large spatial distance between text and picture without colour-coding (i.e., large condition $n = 26$). This study was designed and conducted in accordance with the guidelines of the ethical committee of Erasmus University Rotterdam, Department of Psychology, Education, and Child Studies. All participants gave informed consent before the start of the study.

2.1.2. Materials

2.1.2.1. Demographics. Basic demographic questions were asked to assess prior education, age, gender, and the participant's native language. Participants could indicate their gender with three possible answers: male, female, or other. Prior education was assessed by asking what the highest level of education was that they had completed. This question had three possible answers: secondary education, higher vocational education, or other. There were no significant differences between conditions in age, $F < 1$; gender, $\chi^2(6) = 2.35$, $p = .504$; level of prior education, $\chi^2(6) = 2.65$, $p = .852$; and native language, $\chi^2(6) = 4.87$, $p = .561$ (see Table 1 for descriptives).

2.1.2.2. Prior knowledge test. Prior knowledge was assessed by asking participants to rate their prior knowledge of car mechanics and repair on a scale of 1 (very high) to 5 (very low), and by asking whether or not six statements applied to them. The statements were: "I have a driver's license", "I have put air into a tire on a car", "I have changed a tire on a car", "I have changed the oil in a car", "I have changed spark plugs in a

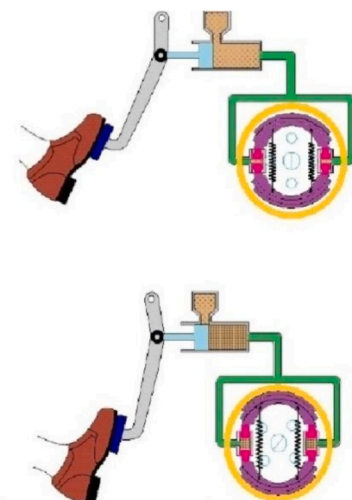
Table 1

Descriptive statistics of demographics variables and prior knowledge in Experiment 1.

| | Small | Large | Small signaling | Large signaling |
|---|--------------|--------------|-----------------|-----------------|
| Gender (%) | | | | |
| Male | 3 (11.11) | 4 (15.38) | 4 (14.81) | 7 (25.93) |
| Female | 24 (88.89) | 22 (84.62) | 23 (85.19) | 20 (74.07) |
| Level of prior education (%) | | | | |
| Secondary education | 20 (74.07) | 19 (73.08) | 22 (81.48) | 21 (77.78) |
| Higher vocational education | 5 (18.52) | 3 (11.54) | 2 (7.41) | 4 (14.81) |
| Other | 2 (7.41) | 4 (15.38) | 3 (11.11) | 2 (7.41) |
| Native language (%) | | | | |
| Dutch | 7 (25.93) | 11 (42.31) | 10 (37.04) | 14 (51.85) |
| English other | 5 (18.52) | 2 (7.69) | 3 (11.11) | 3 (11.11) |
| | 15 (55.56) | 13 (50.0) | 14 (51.85) | 10 (37.04) |
| Age M (SD) | 20.41 (1.28) | 21.58 (3.80) | 21.30 (5.68) | 21.44 (4.59) |
| Prior knowledge ratings M (SD) | 4.48 (0.58) | 4.50 (0.58) | 4.67 (0.56) | 4.63 (0.63) |
| Prior knowledge statements M (SD) | 10.63 (1.15) | 10.85 (1.08) | 10.85 (1.35) | 11 (0.92) |

car", and "I have replaced brake shoes on a car". This resulted in a prior knowledge self-rating (1–5) and a prior knowledge statement score (1–6). There were no significant differences between conditions in prior knowledge ratings ($F < 1$) and prior knowledge statements ($F < 1$). Table 1 displays the descriptives for the prior knowledge scores in each condition.

2.1.2.3. Learning material. The learning material was shown on a single page on a computer screen with a resolution of 1920×1200 pixels and consisted of a text and picture about the functioning of a car's braking system. The picture depicted the car braking system when the brake pedal is pushed down and when it is not pushed down. The text consisted of 73 words and explained how the car braking system works. The text and picture were spatially separated in all conditions. In this experiment, spatial distance was manipulated by using a large or a small distance between text and picture. The small distance was set at 10 pixels



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the shoes press against the drum, the wheel stops or slows down.

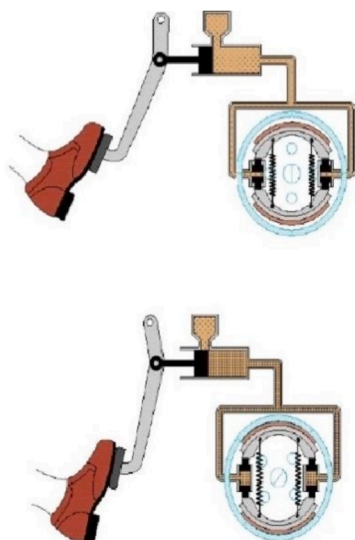
Fig. 1. Learning materials in the small signaling condition.

(see Figs. 1 and 2), while the large distance was set at 410 pixels (see Figs. 3 and 4). The large distance of 410 pixels was used, as this was the largest distance possible on a 1920×1200 pixels computer screen while still keeping the learning material clear and readable. Signaling was manipulated by using colour-coding or not. Signaling was applied by matching the colours of important parts of the car braking system and their corresponding words in the text, see Figs. 1 and 3.

2.1.2.4. Retention tests. Retention was measured with a free recall test by asking the participants to explain how a car's braking system works. The free recall test was scored by giving one point for each correct step of the braking system mentioned in the explanation which lead to a maximum of 8 points. Additionally, we used a labeling test by asking participants to write down the correct name of eight labelled parts in the picture. Six of these labelled parts were colour-coded and two were not. The label test was scored by giving one point for each correctly labelled part which lead to a maximum of 8 points in total. To explore whether differences between signaled and non-signaled items of the label test would emerge, separate scores were calculated for signaled and non-signaled parts: participants could earn 6 points for the colour-coded parts and 2 points for the parts that were not colour-coded.

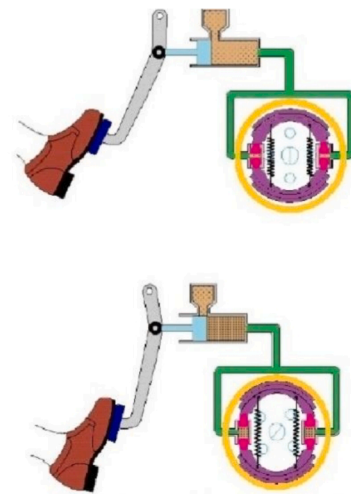
2.1.2.5. Transfer test. Transfer was measured by asking three open-ended questions about understanding of the braking system. For each transfer question, multiple correct answers were possible. The first transfer question was: "What could be done to make brakes more reliable – that is, to make sure they would not fail?" For example, a correct answer on this question could be: "make the pedal more sensitive". The second transfer question was: "What could be done to make brakes more effective – that is, to reduce the distance needed to bring a car to a stop?" A possible correct answer on this question was: "use larger pistons". The third transfer question was: "Suppose you press on the car's brake pedal in your car but the brakes don't work. What could have gone wrong?" An example of a correct answer on this question could be: "the tube had a leak". Each transfer question was scored by giving one point for each correct explanation and by summing up the points of the three questions, yielding one transfer score per participant.

2.1.2.6. Mental effort. Mental effort was measured by asking



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the shoes press against the drum, the wheel stops or slows down.

Fig. 2. Learning materials in the small condition.



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the shoes press against the drum, the wheel stops or slows down.

Fig. 3. Learning materials in the large signaling condition.

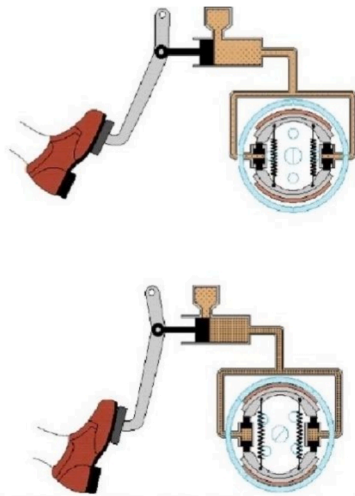
participants to indicate on a Likert scale of 1 (*very, very low effort*) to 9 (*very, very high effort*) how much effort they invested in the task (Paas, Renkl et al., 2003; Paas, Tuovinen et al., 2003). This self-rating instrument has proven to provide a valid and reliable indication of the mental effort experienced during a task (Paas, Renkl et al., 2003; Paas, Tuovinen et al., 2003).

2.1.3. Procedure

Participants were tested in groups (up to eight students) and worked individually on computers. They were first asked to fill in the demographics questions, followed by the prior knowledge questions. Then, the learning phase started in which they were presented with the learning materials for 50 s according to their assigned condition. All participants were instructed to use both the text and picture during learning. Participants in the signaling conditions were also told that the main parts of the text and picture were presented in the same colour. After the learning phase, participants were asked to rate the mental effort they invested during the learning phase. Subsequently, they completed the retention (free recall followed by labeling test) and transfer questions. Participants were given 2.5 min to answer the free recall question, after which the program automatically progressed to the label test. This test had no time limit, participants had to click the next button if they finished the question to progress to the transfer questions. Participants were given 2.5 min to answer each of the transfer questions, after which the next question was automatically presented. Lastly, participants were asked to rate their mental effort invested in answering all questions.

2.1.4. Scoring

The free recall test and the transfer test were scored with the coding



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the shoes press against the drum, the wheel stops or slows down.

Fig. 4. Learning materials in the large condition.

scheme of Johnson and Mayer (2012). For the label test we designed a coding scheme ourselves. Ten percent of all data was scored by a second rater, with both raters being unaware of the participants' conditions while scoring. We calculated the interrater-reliability, which was almost perfect, $K = 0.86$ (Landis & Koch, 1977). Therefore, the scorings of the first rater were used in the analyses.

2.2. Results

Separate two-Way ANOVAs with distance and signaling as between-subjects factors were conducted on the learning outcomes (free recall test, signaled items label test, non-signaled items label test, and transfer

Table 2

Means (and standard deviations) of the free recall test (max. = 8), signaled items of the label test (max. = 6), non-signaled items of the label test (max. = 2), and the transfer test.

| | Free recall test | Label test Signaled items | Label test non-signaled items | Transfer test |
|-----------------|------------------|---------------------------|-------------------------------|---------------|
| Small | 2.56 (1.42) | 1.83 (0.94) | 0.26 (0.53) | 1.81 (1.42) |
| Large | 2.77 (1.42) | 1.65 (1.12) | 0.29 (0.51) | 1.81 (1.81) |
| Total | 2.66 (1.41) | 1.75 (1.05) | 0.27 (0.51) | 1.81 (1.61) |
| Small signaling | 2.74 (1.85) | 2.02 (1.23) | 0.11 (0.29) | 1.56 (1.60) |
| Large signaling | 2.70 (1.73) | 2.70 (1.75) | 0.26 (0.42) | 1.89 (2.10) |
| Total | 2.72 (1.77) | 2.36 (1.54) | 0.19 (0.37) | 1.72 (1.86) |

test) and mental effort (during learning and during testing). Table 2 and Table 3 show descriptive statistics. Because we found mostly null-findings, we performed additional Bayesian analyses with JASP (JASP Team 2016, Version 0.8.4). Bayes Factors (BF) were computed using non-informative default priors $p(M) = 0.5$ (Cauchy prior of $h = 0.75$; Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2016). Bayes Factors can be classified as follows: no evidence $BF = 1$, anecdotal evidence $BF = 1-3$, substantial evidence $BF = 3-10$, strong $BF = 10-30$, very strong $BF = 30-100$, decisive $BF > 100$ (Jeffreys, 1961).

2.2.1. Retention

On the free recall test, there were no significant main effects of Distance, $F < 1$ or Signaling, $F < 1$, nor a significant interaction, $F(1, 103) = 1.60$, $p = .690$, $\eta_p^2 < 0.01$. The Bayesian analysis showed there was substantial evidence for the absence of an effect of Signaling (BF null-model = 4.80; showing that the data were 4.802 times more likely under the null-model than under the model predicting an effect of Signaling) and Distance (BF null-model = 4.72), while there was strong evidence for the absence of an interaction (BF null-model = 22.94). For the signaled items of the label test we found a significant main effect of Signaling, $F(1, 103) = 3.29$, $p = .016$, BF alternative-model = 2.657 (anecdotal evidence in favor of the model predicting an effect of Signaling),¹ $\eta_p^2 = 0.06$, showing that participants in the signaling conditions outperformed participants in the non-signaling conditions. We found no significant main effect of Distance, $F(1, 103) = 1.00$, $p = .320$, BF null-model = 3.11 (substantial evidence), $\eta_p^2 = 0.01$, nor a significant interaction between Distance and Signaling on signaled items, $F(1, 103) = 2.92$, $p = .090$, BF null-model = 1.00 (no evidence), $\eta_p^2 = 0.03$. On the non-signaled items of the label test, we found no significant main effects of Distance, $F(1, 103) = 1.10$, $p = .308$, BF null-model = 3.06 (substantial evidence), $\eta_p^2 = 0.01$ or Signaling, $F(1, 103) = 1.10$, $p = .308$, BF null-model = 3.06 (substantial evidence), $\eta_p^2 = 0.01$. Moreover, we found no significant interaction between Distance and Signaling, $F < 1$, BF null-model = 9.41 (substantial evidence).

2.2.2. Transfer

On the transfer test, there were no significant main effects of Distance, $F < 1$, BF null-model = 4.40 (substantial evidence) or Signaling, $F < 1$, BF null-model = 4.74 (substantial evidence), nor a significant interaction, $F < 1$, BF null-model = 19.91 (strong evidence).

2.2.3. Mental effort

On invested mental effort after the learning phase, there were no significant main effects of Distance, $F < 1$, BF null-model = 3.46 (substantial evidence) or Signaling, $F < 1$, BF null-model = 4.25 (substantial evidence), nor an interaction, $F < 1$, BF null-model = 14.57 (strong evidence). On invested mental effort after the testing phase, also no significant main effects of Distance, $F < 1$, BF null-model = 4.47

Table 3

Means (and standard deviations) of the mental effort ratings of the learning phase and test phase (max. = 9).

| | Mental effort learning | Mental effort test |
|-----------------|------------------------|--------------------|
| Small | 6.00 (0.92) | 5.30 (1.59) |
| Large | 6.00 (1.55) | 5.69 (1.41) |
| Total | 6.00 (1.26) | 5.49 (1.50) |
| Small signaling | 6.33 (0.83) | 6.07 (1.66) |
| Large signaling | 5.93 (1.39) | 5.41 (1.65) |
| Total | 6.13 (1.15) | 5.74 (1.67) |

¹ Here, we report the evidence in favor of the alternative model, which includes Signaling as a factor as the initial ANOVA revealed a moderate effect of signaling.

(substantial evidence) or Signaling, $F < 1$, BF null-model = 3.64 (substantial evidence) were found, nor a significant interaction between Distance and Signaling, $F(1, 103) = 3.02$, $p = .085$, BF null-model = 16.21 (strong evidence) $\eta_p^2 = 0.03$.

2.3. Discussion

With Experiment 1 we aimed to investigate whether signaling alleviates the negative effect of increased spatial distance between spatially separated text and pictures. Consistent with Johnson and Mayer (2012), no differences between conditions were found in free recall performance. Consequently, it is not surprising that we did not find an effect of spatial distance and signaling either, as a prerequisite for finding such effects is that the learning materials elicit a split-attention effect. We did find an effect of signaling on the label test, namely that colour-coded items were remembered better by participants in the signaling conditions than participants in the conditions without signaling. It seems therefore that signaling facilitated retention of important parts of the braking system, which is in line with previous research showing that signaling can lead to a higher retention performance (De Koning, Tabbers, Rikers, & Paas, 2007; Ozelik et al., 2009; Schneider et al., 2018; Van Gog, 2014). However, we found no effects of signaling on transfer performance or mental effort. In contrast to our expectations we found no effect of increased distance on any of the dependent variables. Possibly, spatial distance does not have a large influence on learning these materials. It could be that the difference in spatial distance was not large enough to elicit an effect on learning. For this experiment, we chose the maximum spatial distance possible on a 24 inch computer screen which might not have challenged learners' working memory in the large spatial distance condition considerably more than learners in the small spatial distance condition; this is corroborated by the mental effort results indicating that learners in the large spatial distance condition did not experience higher cognitive load than learners in the small spatial distance condition.

Furthermore, we failed to replicate the split-attention effect on transfer performance reported by Johnson and Mayer (2012). However, they compared a spatially separated condition with a spatially integrated condition, whereas in our experiment in all conditions text and picture were spatially separated. Together our findings suggest that either distance or signaling do not influence the split-attention effect, or the split-attention effect is very small to non-existent in the present materials. To clarify this, a second experiment was conducted in which we specifically aimed to replicate the primary findings of Johnson and Mayer (2012), that transfer-test performance was lower in the spatially separated condition than in the spatially integrated condition. If we are able to replicate the split-attention effect, it is highly likely that increasing the spatial distance between text and picture does not hinder the learning of the present materials.

3. Experiment 2

Experiment 2 aimed to replicate and extend the findings of Johnson and Mayer (2012) by including a spatially separated condition with signaling besides the spatially separated and spatially integrated conditions. The materials used in the spatially separated condition and the spatially integrated condition were identical to the first experiment of Johnson and Mayer. The dependent variables were the same as in Experiment 1 of the current study. On the free recall test, we expected to find no differences across conditions, in line with Johnson and Mayer (2012). On the signaled items of the label test, we expected that both the spatially integrated condition and signaling condition would outperform the spatially separated condition. On the non-signaled items of the label test we expected the spatially integrated condition to outperform the spatially separated conditions. On the transfer test and mental effort ratings we expected the spatially integrated condition to perform better and report lower invested mental effort than both spatially separated

conditions.

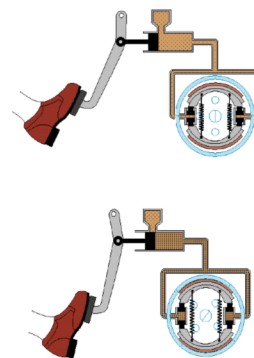
3.1. Method

3.1.1. Participants and design

Participants ($n = 180$) were recruited via Amazon's Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010), and were paid 2 dollars for their participation in this experiment which took approximately 20 min. A priori we decided that participants would be excluded from the study when they used a screen with a resolution smaller than 1024×768 and therefore could not see the picture and text on the screen ($n = 20$). This cutoff was chosen based on an earlier study by Rop, Schüler, Verkoeijen, Scheiter, and Van Gog (2018), in which 6.62 percent of Mechanical Turk participants used a screen with a resolution smaller than this. Therefore, this cutoff would ensure maximal spatial distance without much data loss. Participants were also excluded when they participated in the experiment twice ($n = 1$), when they participated in a noisy environment (i.e., self-reported noise of seven or higher on a scale of one to nine; $n = 3$), when they cheated by using google search ($n = 10$), or when they provided non-sensible answers to the questions (e.g., responding "it is great and good one"; $n = 6$). The final sample consisted of 140 participants ($M_{age} = 37.68$, $SD = 11.57$, range 19–71, 81 males, 58 females and 1 other), who were randomly assigned to one of three between-subject conditions: separated ($n = 43$), separated with signaling ($n = 49$), and integrated ($n = 48$), see Figs. 5–7.

3.1.2. Materials

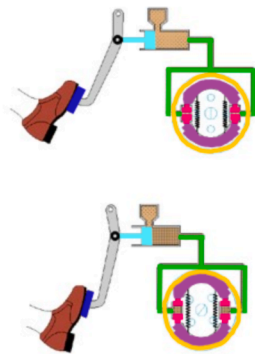
3.1.2.1. Demographics. The demographic questions were the same as in Experiment 1. Additionally, in Experiment 2 participants were asked about what type of computer monitor they used during the study and about how noisy their environment, the latter of which they indicated on a Likert scale ranging from one (*quiet and no distractions*) to nine (*noise and many distractions*). Table 4 shows the descriptives for these variables in each of the conditions. There were no significant differences between conditions in age, $F < 1$; testing environment, $F < 1$; type of monitor, $\chi^2(6) = 4.04$, $p = .672$; gender, $\chi^2(6) = 8.57$, $p = .073$; and native language, $\chi^2(6) = 1.87$, $p = .392$. There was however a significant difference between conditions in the level of prior education, $\chi^2(6) = 9.58$,



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the brake shoes press against the drum, the wheel stops or slows down.

Fig. 5. Learning materials in the separated condition.

$p = .048$; in the integrated condition, there were more participants who



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder. The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders. In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move outward. These smaller pistons activate the brake shoes. When the brake shoes press against the drum, the wheel stops or slows down.

Fig. 6. Learning materials in the separated signaling condition.

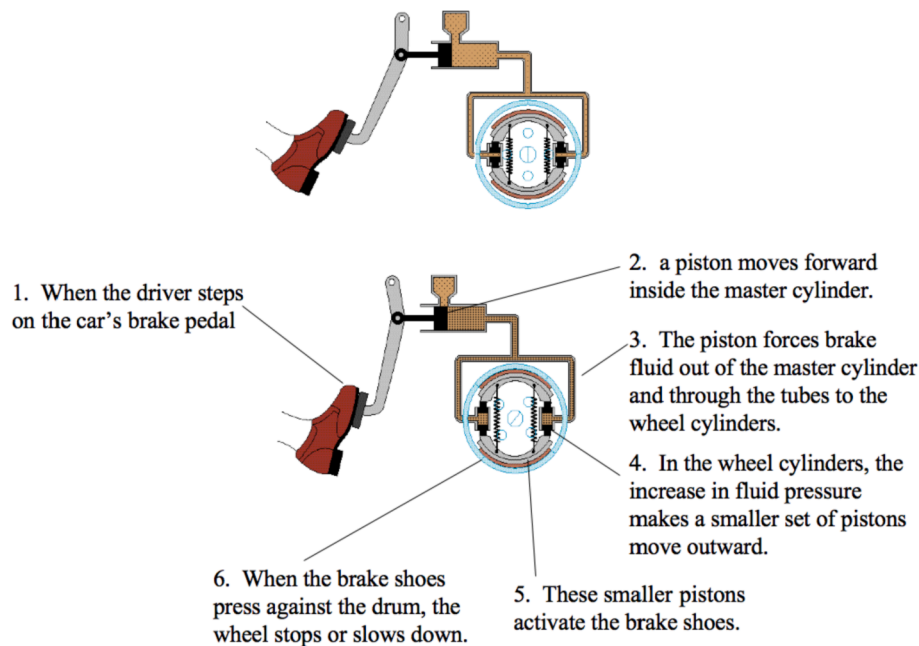


Fig. 7. Learning materials in the integrated condition.

had completed higher vocational education than in the other conditions. Given that level of prior education did not significantly interact with the factor Condition on all outcome measures, we report the results of the analyses without this factor in the Results.²

3.1.2.2. Learning material. Both spatially separated conditions showed the text and picture separated at a small spatial distance. In the separated condition with signaling, colour-coding was added like in Experiment 1 (see Figs. 5 and 6 for the two separated format conditions). The integrated format condition consisted of the same text and picture as the separated format conditions, but in the integrated condition text was placed at corresponding parts close to the picture (see Fig. 7). The separated and integrated formats were identical to those used in Johnson and Mayer (2012).

² Conducting the two-way ANOVAs on each of the outcome measures with level of prior education as additional factor yielded the same results as when conducting the analyses without this factor.

3.1.2.3. Dependent variables. The mental effort ratings, prior knowledge questions, retention questions, and transfer questions were the same as in Experiment 1. Participants in the different conditions did not differ in their prior knowledge based on their self-rating, $F < 1$, and statements score, $F(2, 137) = 1.22$, $p = .300$, $\eta^2 = 0.02$ (see Table 4).

3.1.3. Procedure

The procedure was the same as in Experiment 1, except that (1) participants sat in front of a computer screen outside the university lab when they participated in the study, (2) the demographic questions were completed at the end of the experiment instead of at the beginning.

3.2. Results

Separate one-Way ANOVAs with condition as between-subjects factor were conducted on the learning outcomes (free recall test, signaled items label test, non-signaled items label test, and transfer test) and mental effort (during learning and during testing).³ Tables 5 and 6 show descriptive statistics. Again, we complemented these ANOVAs with

Bayesian analyses.

3.2.1. Retention

On the free recall test, there were no significant differences across conditions, $F(2, 137) = 1.91$, $p = .153$, BF null-model = 5.11 (substantial evidence), $\eta^2 = 0.03$. For the signaled items of the label test, we found no significant differences across conditions, $F(2, 137) = 1.22$, $p = .300$, BF null-model = 3.45 (substantial evidence), $\eta^2 = 0.02$. On the non-signaled items of the label test, we found no significant differences across conditions, $F(2, 137) = 2.18$, $p = .117$, BF alternative-model = 1.70 (anecdotal evidence), $\eta^2 = 0.03$.

³ Based on a reviewer comment, we have directly compared the integrated condition with the separated condition (i.e., the conditions used in the Johnson and Mayer [2012] study) in separate t -tests. Consistent with our ANOVA results, these analyses yielded no significant differences between conditions.

Table 4

Descriptive statistics of demographics variables and prior knowledge in Experiment 2.

| | Integrated | Separated signaling | Separated |
|---|---------------|---------------------|---------------|
| Gender (%) | | | |
| Male | 21 (43.75) | 33 (67.35) | 27 (62.79) |
| Female | 27 (56.25) | 15 (30.61) | 26 (37.21) |
| other | 0 (0) | 1 (2.04) | 0 (0) |
| Level of prior education (%) | | | |
| Secondary education | 25 (52.08) | 32 (65.31) | 21 (48.84) |
| Higher vocational education | 91 (18.75) | 7 (14.29) | 16 (37.21) |
| Other | 14 (29.17) | 10 (20.41) | 6 (13.95) |
| Native language (%) | | | |
| English | 48 (100) | 48 (97.96) | 43 (100) |
| other | 0 (0) | 1 (2.04) | 0 (0) |
| Type of monitor (%) | | | |
| PC | 19 (39.58) | 20 (40.82) | 16 (37.21) |
| Laptop | 29 (60.42) | 27 (55.10) | 27 (62.79) |
| Tablet | 0 (0) | 1 (2.04) | 0 (0) |
| Other | 0 (0) | 1 (2.04) | 0 (0) |
| Age <i>M</i> (<i>SD</i>) | 38.19 (11.11) | 38.35 (12.26) | 36.35 (11.44) |
| Testing environment <i>M</i> (<i>SD</i>) | 1.31 (0.83) | 1.41 (1) | 1.30 (0.80) |
| Prior knowledge ratings <i>M</i> (<i>SD</i>) | 3.90 (0.93) | 3.86 (1.08) | 3.74 (1.09) |
| Prior knowledge statements <i>M</i> (<i>SD</i>) | 8.69 (1.85) | 8.20 (1.84) | 8.19 (1.61) |

Table 5

Means (and standard deviations) of the free recall test (max. = 8), signaled items of the label test (max. = 6), non-signaled items of the label test (max. = 2), and the transfer test.

| | Free recall test | Label test Signaled items | Label test non-signaled items | Transfer test |
|---------------------|------------------|---------------------------|-------------------------------|---------------|
| Integrated | 4.00 (1.86) | 2.93 (1.34) | 0.61 (0.68) | 2.10 (1.57) |
| Separated signaling | 3.33 (1.51) | 2.93 (1.55) | 0.40 (0.55) | 2.00 (1.55) |
| Separated | 3.74 (1.77) | 2.52 (1.34) | 0.64 (0.63) | 1.86 (1.70) |

Table 6

Means (and standard deviations) of the mental effort ratings of the learning phase and test phase (max. = 9).

| | Mental effort learning | Mental effort test |
|---------------------|------------------------|--------------------|
| Integrated | 7.33 (1.56) | 7.29 (1.38) |
| Separated signaling | 7.45 (1.24) | 7.29 (1.51) |
| Separated | 7.67 (1.15) | 7.47 (1.05) |

3.2.2. Transfer

On the transfer test, there were no significant differences across conditions, $F < 1$, BF null-model = 12.04 (strong evidence).

3.2.3. Mental effort

On invested mental effort after the learning phase, there were no significant differences across conditions, $F < 1$, BF null-model = 12.24 (strong evidence). On invested mental effort after the testing phase, also no significant differences across conditions, $F < 1$, BF null-model = 15.17 (strong evidence) were found.

3.3. Discussion

The aim of Experiment 2 was to replicate and extend the findings of Johnson and Mayer (2012). In contrast to our expectations, and findings by Johnson and Mayer (2012), we failed to replicate the split-attention effect. There are several possible explanations that may explain this

discrepancy. One possible explanation could be that in our experiment we recruited participants via Amazon's Mechanical Turk, whereas Johnson and Mayer recruited psychology students who were tested in a controlled lab setting. However, participants in Mechanical Turk pay attention and exhibit biases very similarly to other sources of subjects (Paolacci et al., 2010), making this unlikely. Another possible explanation for the current findings is that the split-attention effect obtained with the present materials is not as robust as previously assumed (cf. Pouw et al., 2019; Florax & Ploetzner, 2010). Furthermore, since no split-attention effect was found in Experiment 2, it is not possible to examine the influence of varying spatial distance and signaling on the split-attention effect.

4. General discussion

The split-attention effect refers to the finding that learning materials consisting of text and pictures which are spatially integrated lead to better learning than when they are spatially separated. In two experiments we attempted to investigate the effects of signaling and variations in spatial distance on the split-attention effect to gain more clarity about the cognitive processes underlying this effect. It was hypothesized that a larger spatial distance would hamper learning because it would lead to more reliance on working memory thereby imposing a higher cognitive load (Ballard et al., 1995; Bauhoff et al., 2012; Gray & Fu, 2004; Har-diess et al., 2008; Inamdar & Pomplun, 2003), whereas signaling could facilitate learning in a spatially separated format because it would direct learners to relevant information and facilitate integration of information sources (Florax & Ploetzner, 2010; Ozcelik et al., 2009, 2010; Schneider et al., 2018; Van Gog, 2014). The results of both experiments showed no differences in transfer performance, mental effort ratings, free recall or retention of non-signaled items of the label test.

4.1. Cognitive load theory and the split-attention effect

According to CLT, people need to visually search while keeping information active in working memory when learning from spatially separated learning materials, resulting in worse learning performance and a higher cognitive load compared to spatially integrated learning materials (Ayres & Sweller, 2014; Ginns, 2006). The current results showed no effects on learning or mental effort, which is not in line with CLT. Furthermore, the mental effort ratings seem to indicate that keeping information active in working memory did not lead to higher cognitive load, which raises doubts about the explanation of CLT for the split-attention effect. Although based on our study, the CLT explanation cannot be tested directly (which would require for example eye tracking methodology), there is some evidence suggesting that despite comparable learning outcomes, a spatially integrated format elicits different processing behavior than a spatially separated format. For example, Pouw et al. (2019; Experiment 3) found no differences in learning outcomes between an integrated condition and separated format conditions, but they did find that participants in the spatially integrated condition made more integration attempts than participants in the separated format conditions. Furthermore, Schüler (2017) showed that learners first make mental representations of the text and picture and subsequently integrate them mentally. Translating this to the present findings, participants in the separated format conditions might have successfully used those representations to mentally integrate the learning materials later. However, this should be reflected in higher mental effort ratings in the separated format conditions (particularly the non-signaled one) because keeping mental representations in working memory while integrating them is expected to impose a higher cognitive load.

The failed replication of the split-attention effect in Experiment 2 raises doubts about the robustness of the split-attention effect. Similar to our study, other researchers have reported to not always find the split-attention effect (e.g., Pouw et al., 2019; Florax & Ploetzner, 2010).

The fact that several studies failed to replicate this effect raises questions about the actual strength of the split-attention effect. Furthermore, the failed replications limit the practical use of the split-attention effect, because spatially separated learning materials do not always appear to lead to a worse learning performance and higher cognitive load than spatially integrated learning materials. Future research could identify possible boundary conditions under which the split-attention effect is weaker or absent, like the type or complexity of the learning task. Furthermore, eye-tracking could be used to investigate explanations of the split-attention effect, such as whether spatially separated learning materials lead to more visual search compared to spatially integrated learning materials. All in all, future research should focus on trying to replicate earlier findings of the split-attention effect, and possible boundary conditions for this effect.

4.2. Spatial distance, signaling, and the split-attention effect

The results of the present study indicate that increasing the spatial distance between two spatially separated sources of information does not influence learning. Although previous research did not always find an effect of distance on learning (Pouw et al., 2019; Bauhoff, Huff, & Schwan, 2012; Florax & Ploetzner, 2010; Mayer, Steinhoff, Bower, & Mars, 1995), many studies did show that a larger distance leads to more reliance on memory and less gaze switches between information sources (Ballard et al., 1995; Bauhoff et al., 2012; Gray & Fu, 2004; Hardiess et al., 2008; Inamdar & Pomplun, 2003). Possibly, the spatial distance in our experiments was not large enough to elicit an effect on learning outcomes. Future research could investigate this possibility by enlarging the spatial distance between two information sources by using two screens that are separated from each other at a considerable distance. Additionally, the results of Experiment 2 showed no split-attention effect, which made it impossible to examine the effects of varying spatial distance and signaling on the split-attention effect and draw conclusions about them, because the split-attention effect is a prerequisite for finding such effects.

We found a small effect of signaling in Experiment 1, indicating that participants in the signaling conditions remembered more signaled items of the label test than participants in the non-signaling conditions. The finding that signaling can lead to better retention of important parts of the learning materials is consistent with previous research (Schneider et al., 2018; Van Gog, 2014). However, it does not seem to lead to lower cognitive load or increased transfer performance, which is in contrast with previous research (Ozcelik et al., 2009; Schneider et al., 2018; Van Gog, 2014). Possibly, the text already guided participants through the picture well enough, meaning that signaling did not have much added value in directing participants' attention (Pouw et al., 2019). Future research could further investigate whether and how the type of information presented in the text has an influence on the attentional processes of the reader when learning from text-picture materials. Related to this, in the current study, we did not check whether the signaling manipulation worked as intended. Future research is warranted to further investigate this issue, for example by using eye-tracking to see if a signaling manipulation successfully guides participants' attention during learning (De Koning, Tabbers, Rikers, & Paas, 2010).

4.3. Theoretical and practical implications

The current study extends our knowledge of the robustness and boundary conditions of the split-attention effect. More specifically, the impact of varying spatial distance and signaling on the split-attention effect was investigated. A contribution of this study is that an increase in spatial distance of 400 pixels does not necessarily affect learning or cognitive load from textual and pictorial information presented in a split-attention format. Practically, this suggests that irrespective of the spatial distance between text and picture comparable learning outcomes and mental effort can be expected for learners. Another finding was that

signaling helps learners remember specific parts of the textual and pictorial information. Therefore, the implication that can be derived is that if the learning goal is to support memory for specific information, elements in the text and picture should be signaled, for example by using colour to highlight the information. Furthermore, we did not find evidence for the split-attention effect previously reported by Johnson and Mayer (2012) in a direct replication, using their instructional materials and tests. This raises questions about the robustness of the split-attention effect, and warrants more direct replications in the field of educational psychology. Like stated above, such replications would benefit greatly from adding manipulations checks and process data like eye-tracking or (retrospective) reporting (e.g., Van Gog, Paas, Van Merriënboer, & Witte, 2005). For instructional designers, a tentative practical implication emanating from this finding is that the spatial integration of text and pictures has no priority when designing instructional materials. Moreover, as spatially integrating two information sources can even hamper learning under certain circumstances (e.g., when one of the information sources is self-containing, cf. the redundancy effect, Kalyuga & Sweller, 2014), instructional designers should be very careful with integrating text and pictures. However, please note that more research on this topic is warranted before we can substantiate this claim.

4.4. Conclusion

This study shows that it is important to not take previously found instructional design effects (see Sweller, Van Merriënboer, & Paas, 2019) for granted and stresses the importance of replicating earlier findings. Specifically, this study raises questions about the robustness of the split-attention effect, and offers new possibilities for research into possible boundary conditions of the split-attention effect.

Acknowledgment

We would like to express our gratitude to Dr. Cheryl I. Johnson for sharing the materials used in this study.

References

- Angeli, C., Valanides, N., & Kirschner, P. (2009). Field dependence–independence and instructional-design effects on learners' performance with a computer-modeling tool. *Computers in Human Behavior*, 25, 1355–1366. <https://doi.org/10.1016/j.chb.2009.05.010>.
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, pp. 206–226). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.014>.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7, 66–80. <https://doi.org/10.1162/jocn.1995.7.1.66>.
- Bauhoff, V., Huff, M., & Schwan, S. (2012). Distance matters: Spatial contiguity effects as trade-off between gaze switches and memory load. *Applied Cognitive Psychology*, 26, 863–871. <https://doi.org/10.1002/acp.2887>.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293–332. <https://doi.org/10.1207/s1532690xci0804.2>.
- Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233–246. <https://doi.org/10.1111/j.2044-8279.1992.tb01017.x>.
- Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25, 315–332. <https://doi.org/10.1016/j.chb.2008.12.020>.
- De Koning, B. B., & Jarodzka, H. (2017). Attention guidance strategies for supporting learning from dynamic visualizations. In R. Lowe, & R. Ploetzner (Eds.), *Learning from dynamic visualization* (pp. 255–278). Cham: Springer. https://doi.org/10.1007/978-3-319-56204-9_11.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology*, 21, 731–746. <https://doi.org/10.1002/acp.1346>.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21, 113–140. <https://doi.org/10.1007/s10648-009-9098-7>.

- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20, 111–122. <https://doi.org/10.1016/j.learninstruc.2009.02.010>.
- 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, 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, 511–525. <https://doi.org/10.1016/j.learninstruc.2006.10.001>.
- Gray, W. D., & Fu, W. T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, 28, 359–382. <https://doi.org/10.1016/j.cogsci.2003.12.001>.
- Hardiess, G., Gillner, S., & Mallot, H. A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision*, 8, 1–13. <https://doi.org/10.1167/8.1.7>.
- Harter, C. A., & Ku, H. Y. (2008). The effects of spatial contiguity within computer-based instruction of group personalized two-step mathematics word problems. *Computers in Human Behavior*, 24, 1668–1685. <https://doi.org/10.1016/j.chb.2007.06.006>.
- Huff, M., Bauhoff, V., & Schwan, S. (2010). A closer look at the split attention effect: Integrated presentation formats for troubleshooting tasks. *Proceedings of the 9th International Conference of the Learning Sciences*, 1, 174–181.
- Inamdar, S., & Pomplun, M. (2003). Comparative search reveals the tradeoff between eye movements and working memory use in visual tasks. *Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society*, 25, 599–604.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press, Clarendon Press.
- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18, 178–191. <https://doi.org/10.1037/a0026923>.
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40, 1–17. <https://doi.org/10.1518/001872098779480587>.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351–371. [http://doi.org/10.1002/\(SICI\)1099-0720\(199908\)13:4%3c351::AID-ACP589%3e3.0.CO;2-6](http://doi.org/10.1002/(SICI)1099-0720(199908)13:4%3c351::AID-ACP589%3e3.0.CO;2-6).
- Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, pp. 247–262). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.10>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. <https://doi.org/10.2307/2529310>.
- Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, 13, 125–139. [https://doi.org/10.1016/S0959-4752\(02\)00016-6](https://doi.org/10.1016/S0959-4752(02)00016-6).
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, pp. 279–315). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.015>.
- Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research & Development*, 43, 31–41. <https://doi.org/10.1007/BF02300480>.
- Ozcelik, E., Arslan-Arib, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, 26, 110–117. <https://doi.org/10.1016/j.chb.2009.09.001>.
- Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education*, 53, 445–453. <https://doi.org/10.1016/j.compedu.2009.03.002>.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, 84, 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1–4. https://doi.org/10.1207/S15326985EP3801_1.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63–71. https://doi.org/10.1207/S15326985EP3801_8.
- Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411–419. [10/10630a/jdm10630a](https://doi.org/10.10630a/jdm10630a).
- Pouw, W. T. J. L., Van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, 26, 51–72. <https://doi.org/10.1007/s10648-014-9255-5>.
- Pouw, W. T. J. L., Rop, G., De Koning, B. B., & Paas, F. (2019). The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General*, 148, 2058–2075. <https://doi.org/10.1037/xge0000578>.
- Puma, S., Matton, N., Paubel, P. V., & Tricot, A. (2018). Cognitive load theory and time considerations: Using the time-based resource sharing model. *Educational Psychology Review*, 30, 1199–1214. <https://doi.org/10.1007/s10648-018-9438-6>.
- Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2018). The effect of layout and pacing on learning from diagrams with unnecessary text. *Applied Cognitive Psychology*, 32, 610–621. <https://doi.org/10.1002/acp.3445>.
- Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E. J. (2016). Bayesian analysis of factorial designs. *Psychological Methods*, 22, 304–321. <https://doi.org/10.1037/met0000057>.
- Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, 23, 1–24. <https://doi.org/10.1016/j.edurev.2017.11.001>.
- Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, 30, 679–701. <https://doi.org/10.1007/s10648-018-9435-9>.
- Schüler, A. (2017). Investigating gaze behaviour during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction*, 49, 218–231. <https://doi.org/10.1016/j.learninstruc.2017.03.001>.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer. <https://doi.org/10.1007/978-1-4419-8126-4>.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296. <https://doi.org/10.1023/A:1022193728205>.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 1–32. <https://doi.org/10.1007/s10648-019-09465-5>.
- Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, pp. 263–278). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.014>.
- Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, 11(4), 237–244. <https://doi.org/10.1037/1076-898X.11.4.237>.