Computers in Human Behavior 110 (2020) 106379

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: http://www.elsevier.com/locate/comphumbeh

Effects of spatial distance on the effectiveness of mental and physical integration strategies in learning from split-attention examples

Björn B. de Koning[a,*], Gertjan Rop[b], Fred Paas[a,c]

[a] Department of Psychology, Education, and Child Studies, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands, Burgemeester Oudlaan 50, 3062, PA, Rotterdam, the Netherlands
[b] Rotterdam School of Management, Erasmus University Rotterdam, The Netherlands, Burgemeester Oudlaan 50, 3062, PA, Rotterdam, the Netherlands
[c] University of Wollongong, School of Education/Early Start, Wollongong, NSW, 2522, Australia

ARTICLE INFO

Keywords:
Self-management effect
Cognitive load theory
Split-attention effect
Spatial distance

ABSTRACT

Learning from spatially separated text and pictures is improved when learners are instructed to use a physical or mental integration strategy. This study investigated whether varying the spatial distance between text and pictures affects the effectiveness of physical and mental integration strategies. We hypothesized that a larger spatial distance would increase cognitive load and harm learning. Ninety-two university students studied the functioning of an electrical circuit from text and pictures that were presented at a small or large spatial distance from each other, while using a physical or mental integration strategy during learning. Results indicated that participants using the mental integration strategy obtained higher recall scores than participants using the physical integration strategy, but no differences were found for comprehension, transfer, and cognitive load. No effects of spatial distance were found. More research is needed to investigate to what extent spatial distance influences learning with mental and physical integration strategies.

1. Introduction

Computer-based learning environments typically contain instructional materials that consist of a combination of text and pictures. While this combination of information sources usually leads to higher learning performance than relying on text only (i.e., the multimedia effect; Mayer, 2003), this is not always the case. In a large number of situations, learners are required to mentally integrate mutually referring text and pictures that are presented in a spatially separated format. Research has shown that this format requires learners to split their attention and leads to lower learning performance than a spatially-integrated format in which the text is presented adjacent to the corresponding part in the picture. This finding has been recognized in cognitive load theory (CLT; Sweller, Van Merriënboer, & Paas, 1998, 2019) as the split-attention effect (Ayres & Sweller, 2014; Pouw, Rop, De Koning, & Paas, 2019). According to CLT, a spatially separated format is associated with high extraneous cognitive load and suboptimal learning due to the unnecessary visual search and reorienting processes that learners have to engage in to integrate the associated parts of text and pictures in working memory. Because the working memory capacity is limited, such processes use up resources for cognitive processes beneficial for learning, like schema construction and elaboration (Paas, Renkl, & Sweller, 2003).

Recently, researchers have therefore started to investigate whether teaching learners strategies to integrate spatially separated text and pictures themselves reduces their cognitive load and improves their learning outcomes (e.g., De Koning et al., 2020; Tindall-Ford, Agostinho, Bokosmaty, Paas, & Chandler, 2015). Such strategies can be either based on physical (i.e., manipulating text and/or pictures to decrease the spatial distance) or mental (i.e., instruct leaners to mentally integrate the text and pictures) integration. Together with this development towards self-managed learning (cf. self-management of cognitive load, Roodenrys, Agostinho, Roodenrys, & Chandler, 2012), the range of computer screens and digital devices available to present text and pictures has undergone considerable growth in the past years. With varying screen sizes of devices such as smartphones, tablets, laptops, and computer screens, the distance between text and picture also likely varies depending on the device or computer screen that is used. This makes it relevant to investigate whether the effectiveness of self-management strategies is related to the distance at which spatially separated text and picture are presented. Moreover, as computer-based learning materials generally leave little room for physical integration, it is

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

https://doi.org/10.1016/j.chb.2020.106379
Received 16 October 2019; Received in revised form 27 March 2020; Accepted 8 April 2020
Available online 12 April 2020

informative to compare the effectiveness of mental and physical integration strategies at varying spatial distances. In the present study, we therefore studied the extent to which teaching physical and mental integration strategies improves learning when spatially separated text and pictures are presented at a smaller and larger spatial distance from each other.

1.1. Self-managed integration

An increasing number of studies have investigated whether learning from split-attention worked examples can be improved by teaching learners a strategy to integrate spatially separated text and picture (e.g., Gordon, Tindall-Ford, Agostinho, & Paas, 2016; Sithole, Chandler, Abeysekera, & Paas, 2017). The majority of these studies have focused on teaching a physical integration strategy. That is, learners were taught to use the hands to pick up and move text segments towards the corresponding part of the picture, either by interacting directly with the learning materials on a touchscreen, by controlling a mouse, or by using cut-out text segments. The findings of several studies indicate that teaching a physical integration strategy supports learning. It has been shown that, in a university student sample, using a physical integration strategy to integrate text and pictures results in higher transfer performance than just studying a split-attention format (Roodenrys et al., 2012). Additionally, Sithole et al. (2017) showed that university students using a physical integration strategy did not only have higher recall and transfer performance than learners studying a split-attention format but also than those studying an integrated format that was presented to them. While the above studies combined the physical integration strategy with other supporting strategies that are known to improve learning, such as highlighting and drawing lines to connect textual and pictorial elements (Van Gog, 2014), a study by Tindall-Ford et al. (2015) showed that physical integration without additional supportive strategies also supports learning. In their study, secondary school students who were taught to drag-and-drop text to the corresponding part in the picture obtained higher transfer performance than learners studying the material in a split-attention format.

However, there are also studies showing less favorable results of the physical integration strategy. First, two studies that applied the physical integration strategy without additional supporting strategies (e.g., highlighting) showed that learners using the physical integration strategy obtained comparable recall, comprehension, and transfer performance as learners who studied a split-attention format (Agostinho, Tindall-Ford, & Roodenrys, 2013; De Koning, Rop, & Paas, 2020). Second, several studies have found that physical integration is less effective for learning than mental integration (without additional supporting strategies). A recent study by De Koning et al. (2020) compared the physical integration strategy with a mental integration strategy. In the mental integration strategy condition, learners were taught to imagine moving the text to the corresponding part in the picture. Results showed that learners using the mental integration strategy had higher recall and comprehension performance than learners studying the split-attention format with the physical integration strategy or without being taught a strategy. Comparable findings have been obtained by Bodemer and Faust (2006) who found that prompting learners to mentally integrate spatially separated text and pictures led to higher learning outcomes than prompting learners to physically integrate the text and pictures. These findings confirm that the engagement in active mental integration of multiple external representations, such as text and picture, are essential for creating an accurate and coherent mental representation (Mayer, 2014; Schnotz & Bannert, 2003). Specifically, mental integration supports the construction of a mental representation because text and pictures provide complementary information that learners can combine during active processing of the content (Ainsworth, 2006). Additionally, as suggested by De Koning et al. (2020), engaging in guided mental integration of text and picture has a dual function given that working towards an integrated format in mind helps to reduce unnecessary visual search processes to match corresponding textual and pictorial elements, and encourages active generative processing (e.g., making inferences) that is needed to form a coherent integrated mental representation (cf. Sweller, van Merriënboer, & Paas, 2019).

It is important to note that while the studies discussed above presented the text and pictures in a spatially separated format, the text and pictures were still presented at a relatively small spatial distance from each other. It is conceivable that this may have created a situation that was particularly beneficial for the mental integration strategy. Having the text and picture close to each other requires little cognitive effort to simultaneously process the textual and pictorial information because the information can quickly be integrated in working memory (Pouw et al., 2019; Sweller et al., 2019). Conversely, for the physical integration strategy this close proximity between text and picture might have contributed to the mixed findings given that the limited demands associated with matching and reorienting processes to integrate text and picture make it less necessary to physically integrate text and picture. We therefore investigated whether the superiority of the mental integration strategy over the physical integration strategy would persist when the distance between spatially separated text and picture is increased.

1.2. Spatial distance

Several studies have shown that increasing the distance between spatially separated information sources imposes a higher working memory load on learners and results in lower task performance. In a study by Pouw et al. (2019), for example, participants had to judge the similarity of two cards containing pictures and/or text depicting information that varied in color, number, and form. To make an accurate judgment, the information on the cards had to be mentally integrated. The cards were presented either at a small or large spatial distance from each other. Results indicated that increasing the distance between the cards was associated with higher working memory load, as indicated by reduced performance on a secondary visual working memory task, and a longer time to make a judgment. In a related study by Bauhoff, Huff, and Schwan (2012) participants had to compare two pictures of a mechanical pendulum clock to identify similarities and differences in the functioning of the depicted clocks. These two pictures varied in spatial distance and eye tracking was used to investigate how learners came to their comparison judgment. Results showed that when the spatial distance between the pictures was larger, fewer integrative saccades were made between the pictures. This suggests that learners relied more on working memory when the spatial distance between two mutually referring information sources increased. These findings align with research on embedded cognition where several studies have demonstrated that cognitive demands increase when there is a larger spatial distance between two information sources because participants switched from a perceptual-oriented strategy to a strategy that relied more heavily on working memory (Ballard, Hayhoe, & Pelz, 1995; Gray & Fu, 2004; Pouw, Van Gog, & Paas, 2014). In the study by Ballard et al. (1995), for example, participants had to copy a given block pattern. The spatial distance between the to-be-copied pattern (i.e., model) and the workspace where they could create this pattern varied. When there was a small spatial distance between the two, more saccades were made between the model and the workspace whereas fewer saccades were made when the spatial distance between model and workspace was larger. Also studies employing comparative visual search tasks provide evidence that processing two information sources that are spatially separated at a larger distance requires longer processing times and results in less gaze switches between information sources (e.g., Hardiess, Gillner, & Mallot, 2008).

Together, this indicates that there is a trade-off between working memory use and spatial distance between information sources. Particularly a larger spatial distance between information sources is associated with a processing strategy that relies more on working memory. If we extend this to the self-management strategies investigated in the present study, it could be argued that particularly the mental integration strategy might be disadvantaged when the text and picture are presented at a large spatial distance. The mental integration strategy requires learners to rely on working memory as they are not actually moving the text to the picture but only imagine doing so. This means that with text and picture spatially separated at a large distance they have to keep the textual information active in working memory for a longer time to integrate it with the pictorial information, which likely increases the amount of cognitive load they experience and reduce learning outcomes (Barrouillet & Camos, 2012; Puma, Matton, Paubel, & Tricot, 2018). In the physical integration condition, learners physically move the text to the corresponding part in the picture. While this may take somewhat longer when text and picture are initially presented at a large spatial distance compared to a small spatial distance, irrespective of spatial distance eventually the text is placed close to the picture. It is therefore less likely that increasing the spatial distance between text and picture in the initial state will negatively impact learning outcomes and cognitive load with the physical integration strategy.

1.3. The present study

The aim of this study was to investigate whether increasing the distance between spatially separated, but mutually referring text and picture differentially affects learning outcomes (i.e., recall, comprehension, transfer) and cognitive load when using the physical vs. mental integration strategy. Participants studied a picture with accompanying text about the functioning of an on/off-light-switching circuit with the text and picture separated at a small or a large spatial distance and with instructions to integrate text and pictures physically or mentally. No comparison was made to a condition in which the text and picture were presented in an integrated format given that it has already been demonstrated in previous studies using these materials that an integrated format results in higher learning outcomes than a spatially separated format (De Koning et al., 2020; Kalyuga, Chandler, & Sweller, 1998), and the primary goal of using self-managed integration strategies is to support learning from spatially separated text and pictures for which we -for the first time-investigate a potential boundary condition in this study. Based on the above theoretical and empirical findings, we hypothesized an interaction between spatial distance and type of self-managed integration strategy. Specifically, we expected that when mutually referring text and picture are presented at a large spatial distance from each other using the mental integration strategy to integrate text and picture would yield lower learning outcomes and higher cognitive load than using the physical integration strategy (e.g., Ballard et al., 1995; Pouw et al., 2019). However, following the findings of De Koning et al. (2020), we expected higher learning outcomes (recall and comprehension) and comparable cognitive load when the mental integration strategy -compared to the physical integration strategy-would be used to integrate text and picture that are presented at a small spatial distance from each other.

2. Method

2.1. Participants and design

Ninety-two psychology students (64 females) from [university name blinded for peer review] participated. Their mean age was 21.93 years ($SD = 4.36$). Participants were randomly assigned to one of four conditions that resulted from a 2 × 2 between-subjects design with the factors self-management strategy (mental vs. physical) and spatial distance (small vs. large). There were 22 participants in the condition studying with instructions to mentally integrate text and picture that were presented at a small distance from each other (mental-small condition), 24 participants studied with instructions to mentally integrate text and picture that were presented at a large distance from each other (mental-large condition), 21 participants studied with instructions to physically integrate text and picture that were presented a small distance from each other (physical-small condition), and 25 participants studied with instructions to physically integrate text and picture that were presented at a large distance from each other (physical-large condition). All participants received course credits for their participation and provided informed consent before the study.

2.2. Materials

The materials used in this study were taken from De Koning et al. (2020). The main learning task and the practice exercise were presented on a 24 inch computer screen. Tests regarding prior knowledge, learning outcomes, and cognitive load were administered on paper.

2.2.1. Prior knowledge questionnaire

A questionnaire was used to assess participants' prior knowledge of electrical circuits. The questionnaire contained one question asking participants to indicate their knowledge of electrical circuits on a 5-point scale with a score of 1 reflecting very little knowledge and 5 reflecting very much knowledge. Additionally, the questionnaire contained six checklist items about electrical circuits requiring a yes/no answer (e.g., I know what a circuit breaker is; I know what this symbol [symbol of a coil] means). Each 'yes' answer was awarded one point while 'no' answers were given zero points. The scores on the self-rating question (ranging from 1 to 5) and the checklist items (ranging from 0 to 6) were summed to yield a total prior knowledge score with a minimum score of zero and a maximum score of 11. The rather low prior knowledge scores in each of the conditions (see Table 1) indicate that participants were novices on the topic. There were no significant differences in the total prior knowledge score between the four conditions, $F(3,85) = 0.97$, $p = .410$.

2.2.2. Learning materials

The learning task contained a picture accompanied with text explaining the operation of an on/off-light-switching circuit. Comparable to previous studies (Agostinho et al., 2013; De Koning et al., 2020), this task was created and presented with SMART notebook software. In all conditions, the same picture and text were presented in a split-attention format such that the two mutually referring information sources had to be integrated during learning. In the small spatial distance conditions (see Fig. 1), the text was presented as close as possible to the picture. The distance between the top of the text and closest part of the picture was 0.6 cm. In the large spatial distance conditions (see Fig. 2), the text was presented at the largest distance from the picture that was possible on the computer screen that we used. To create maximum distance, we divided the text segments over the bottom of the screen in three columns. The distance between the top of the text and closest part of the picture was 5.4 cm. For the mental self-management conditions, the text and picture were unmovable and participants just had to imagine moving the text to the corresponding part in the picture with one text segment at the time. In the physical self-management

Table 1
Means and SDs (in Brackets) on Recall, Comprehension, and Transfer Per Condition.

	Prior Knowledge	Recall	Comprehension	Transfer
Physical-small	2.81 (2.62)	19.09 (5.77)	4.14 (2.03)	2.05 (1.16)
Physical-large	1.68 (1.75)	17.56 (6.62)	4.44 (1.81)	2.08 (0.95)
Mental-small	2.42 (2.24)	22.21 (7.89)	4.90 (2.40)	2.21 (1.58)
Mental-large	2.42 (2.60)	20.58 (5.75)	4.79 (2.11)	2.13 (1.15)

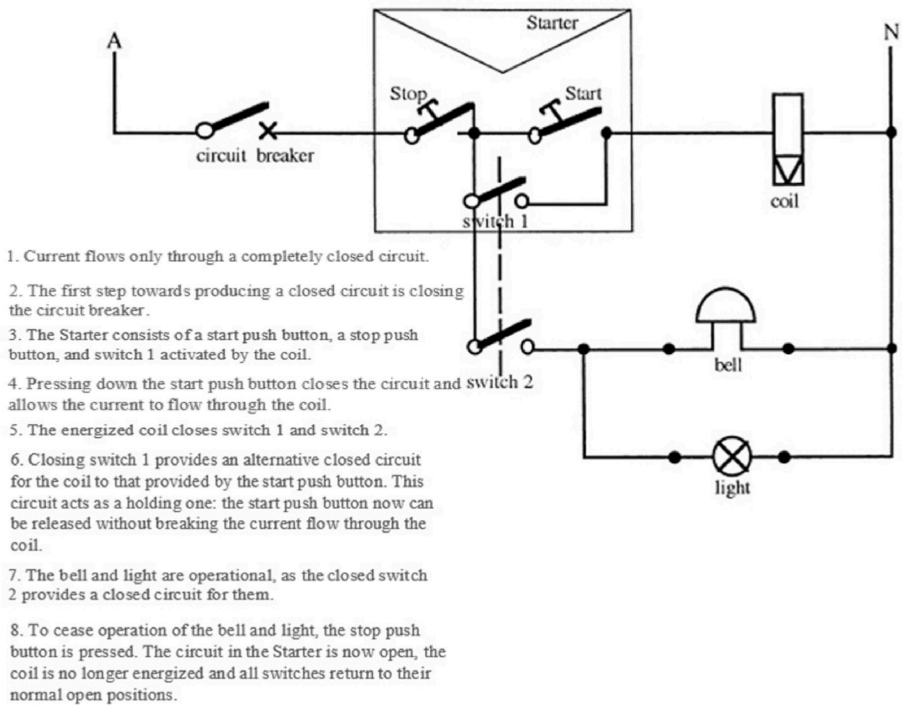

Fig. 1. Learning task in the small spatial distance conditions.

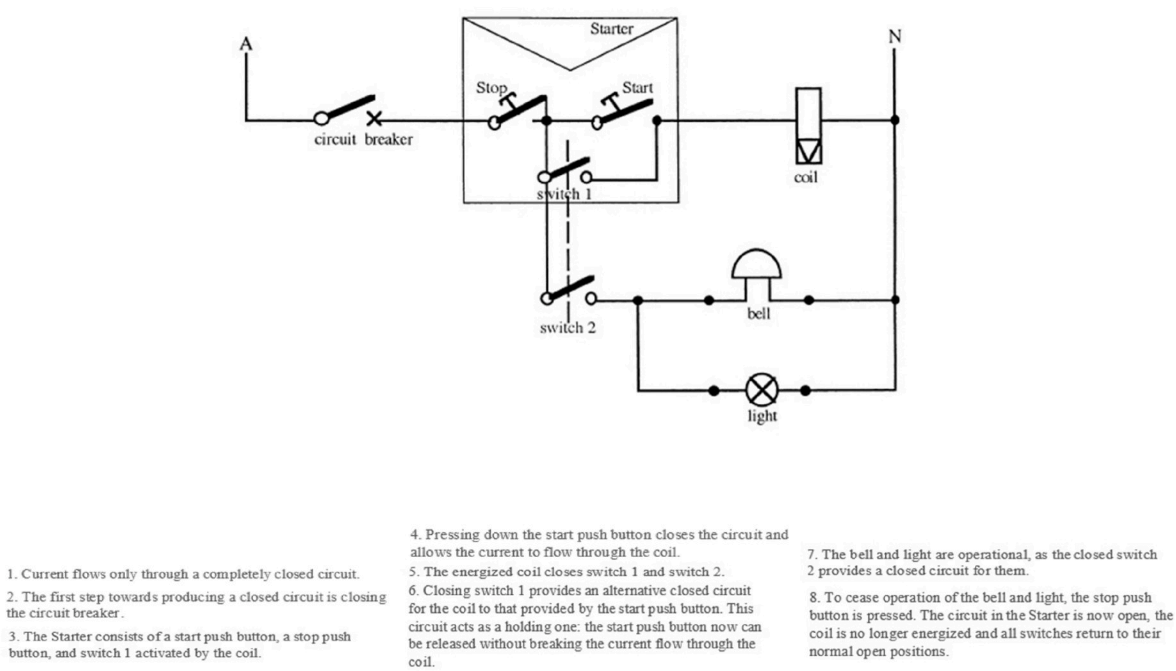

Fig. 2. Learning task in the large spatial distance conditions.

conditions, participants could move the text segments to the corresponding part of the picture (picture elements were unmovable) by dragging-and-dropping the text with the mouse. Text segments could be moved as often as participants wished but it was only possible to move one text segment at the time. For comparability between the mental and physical self-management conditions, no feedback was given in the physical conditions as to whether text segments were placed correctly. Inspection of the integrated formats that were created in the physical integration conditions showed that all learners had accurately integrated the text into the picture. In the mental integration conditions, participants were asked whether they mentally moved and imagined the text and picture information as intended. This appeared to be the case for all participants, except for three participants in the mental-small condition. These three participants were removed from the dataset for the analyses, leaving 19 participants in the mental-small condition. In all conditions, the learning task was presented for 4 min. To familiarize participants with the mental and physical integration that was required in the upcoming learning task, they engaged in a practice task. This task showed a picture of a cat and the word 'tail' in a spatially separated format (with the same spatial distance for all participants) that had to be mentally or physically integrated depending on the condition participants were in. As in the main learning task, in the mental self-management conditions the text was unmovable while in the physical self-management conditions the text could be moved with drag-and-drop functionality.

2.3. Learning outcome measures

Participants' learning outcomes were assessed with a recall test, a comprehension test, and a transfer test. As in other studies with these materials (De Koning et al., 2020; Kalyuga et al., 1998), during the comprehension and transfer tests the picture (without text) of the on/off-light-switching circuit was available to participants on paper.

The recall test contained six questions to measure participants' memory of the components and their spatial arrangement in the on/off-light circuit. One question asked participants to draw the components and relations of the electrical system from memory. One point was given for each component and/or relation that was drawn in the right location. No points were given or subtracted if participants respectively missed or incorrectly drew a component and/or relation. There were 28 components and relations that could be drawn (minimum score = 0, maximum score = 28). Five other questions required participants to recall the name of the components in the electrical system. These labeling questions showed participants a picture of a symbol from the electrical system that was studied and participants had to fill in the name of that symbol. One point was given for a correct answer and no points were awarded to incorrect answers (minimum score = 0, maximum score = 5). The scores on the drawing question and the labeling questions were summed for each participant to yield an overall recall score (minimum score = 0, maximum score = 33). Cronbach's alpha for the recall test was 0.78.

The comprehension test assessed the functioning of the on/off-light-switching circuit with 11 open-ended questions. Example question are "Which switches are pressed when the light is operating?" and "How can the operation of the light be ceased?". One point could be earned per question if the correct answer was given. Incorrect answers received zero points. The scores on all comprehension questions were summed for each participant to yield an overall comprehension score (minimum score = 0, maximum score = 11). Cronbach's alpha for the comprehension test was 0.47.

The transfer test measured participants' ability to reason about the on/off-light-switching circuit with six open-ended questions. Example questions are "After the start button is released, the bell and light stop working. What is the cause of this problem?", and "After the stop button is released, the bell and the light start working again. What is the cause of this problem?". Per question one point was given for the correct answer and no points were awarded to incorrect answers. The scores on all transfer questions were summed for each participant to yield an overall transfer score (minimum score = 0, maximum score = 6). Cronbach's alpha for the transfer test was 0.21.

The answers on the recall test, comprehension test, and transfer test were coded by one rater who was blind to experimental condition. A subset of the answers on these tests (10%) was randomly selected and scored by a second rater. Inter-rater reliability between the two coders appeared relatively high on the recall test ($r = 0.77$), comprehension test ($r = 0.81$), and the transfer test ($r = 0.80$). Therefore, we used the scores of the first coder in the analyses.

2.4. Cognitive load

To measure the cognitive load that participants experienced during the learning task and completing the tests, participants provided an assessment of their invested mental effort during the task on a 9-point self-rating scale (Paas, 1992), which is a proxy for cognitive load (Paas, Tuovinen, Tabbers, & van Gerven, 2003). A score of one indicated that participants invested very, very little mental effort during the just completed task, while a score of nine indicated that participants had invested very, very much mental effort during the task. Prior work has established that this self-rating scale provides a reliable and valid indication of the cognitive load learners experience during a task (Paas et al., 2003).

2.5. Procedure

Participants completed the experiment individually in a one-person cubicle in the university lab. Participants sat at a desk in front of a computer screen with a keyboard and mouse in front of them. The first task participants completed was the prior knowledge test. Before the main learning task was presented, all participants completed the practice task, which was monitored by the experimenter who provided additional explanations where appropriate. Then, the experimenter instructed participants to study the picture and text about the operation of the on-off-light-switching circuit. In the physical self-management conditions participants were additionally told to drag-and-drop text segments to the corresponding location in the picture using the mouse, whereas in the mental self-management conditions participants were told to imagine dragging-and dropping the text. Participants then provided a self-rating of the mental effort invested during the learning task. Subsequently, participants completed the recall test, comprehension test, and transfer test. These tests did not have a time limit, so the time to complete each test was recorded per participant by the experimenter. Participants rated the invested mental effort during each test directly after the respective test had been completed. The whole experiment lasted about 30 min.

3. Results

3.1. Learning outcomes

As there was no time limit to complete the recall test, comprehension test, and transfer test, we first checked whether the conditions spent a comparable amount of time to complete each of these tests. There were no significant differences between conditions in the time needed to compete the tests, Wilk's $\Lambda = 0.907$, $F(9, 202) = 0.924$, $p = .506$, $\eta_p^2 = 0.032$. Therefore, time to complete the tests was not considered in subsequent analyses.

Table 1 displays the mean scores and standard deviations on the recall test, comprehension test, and transfer test in each condition. Separate 2 × 2 Analyses of Variance (ANOVAs) with self-management strategy (mental vs. physical) and spatial distance (small vs. large) as between-subjects factors were conducted on the recall, comprehension, and transfer scores. Results for the recall test showed that there was a significant main effect of self-management strategy, $F(1, 85) = 4.89$, $p = .030$, $\eta_p^2 = 0.054$. Participants who used the mental integration strategy obtained higher recall scores than participants using the physical integration strategy. There was no significant main effect of spatial distance, $F(1, 85) = 1.30$, $p = .258$, $\eta_p^2 = 0.015$, and there was no significant interaction between self-management strategy and spatial distance, $F(1, 85) = 0.001$, $p = .974$, $\eta_p^2 < 0.001$.

Regarding the comprehension test, there were no significant main effects of self-management strategy, $F(1, 85) = 1.55$, $p = .216$, $\eta_p^2 = 0.018$, and spatial distance, $F(1, 85) = 0.048$, $p = .827$, $\eta_p^2 = 0.001$. The interaction between self-management strategy and spatial distance was also not significant, $F(1, 85) = 0.204$, $p = .652$, $\eta_p^2 = 0.002$.

For the transfer test, there were no significant main effects of self-management strategy, $F(1, 85) = 0.162$, $p = .688$, $\eta_p^2 = 0.002$, and spatial distance, $F(1, 85) = 0.011$, $p = .918$, $\eta_p^2 = 0.001$. Also, there was no significant interaction between self-management strategy and spatial distance, $F(1, 85) = 0.052$, $p = .820$, $\eta_p^2 < 0.001$.

3.2. Mental effort

The means and standard deviations for the mental effort ratings collected after the learning task, recall test, comprehension test, and transfer test are shown in Table 2. Separate 2 × 2 ANOVAs with self-management strategy (mental vs. physical) and spatial distance (small vs. large) were conducted on the mental effort scores.[1] Regarding the mental effort reported after the learning task, there was no significant main effect of self-management strategy, $F(1, 85) = 0.006$, $p = .939$, $\eta_p^2 < 0.001$, nor of spatial distance, $F(1, 85) = 1.72$, $p = .193$, $\eta_p^2 = 0.020$. Also, there was no significant interaction between self-management strategy and spatial distance, $F(1, 85) = 1.23$, $p = .272$, $\eta_p^2 = 0.014$.

For the mental effort scores collected after each learning task a similar pattern of findings was obtained. There was no significant main effect of self-management strategy after the recall test, $F(1, 85) = 3.77$, $p = .056$, $\eta_p^2 = 0.043$, the comprehension test, $F(1, 85) = 0.866$, $p = .355$, $\eta_p^2 = 0.010$, and the transfer test, $F(1, 85) = 0.906$, $p = .344$, $\eta_p^2 = 0.011$. Also, there was no significant main effect for spatial distance on mental effort reported after the recall test, $F(1, 85) = 0.85$, $p = .360$, $\eta_p^2 = 0.010$, comprehension test, $F(1, 85) = 0.010$, $p = .919$, $\eta_p^2 < 0.001$, and transfer, $F(1, 85) = 0.162$, $p = .689$, $\eta_p^2 = 0.002$. The interaction between self-management strategy and spatial distance was not significant regarding the mental effort reported after the recall test, $F(1, 85) = 0.010$, $p = .921$, $\eta_p^2 < 0.001$, comprehension test, $F(1, 85) = 0.614$, $p = .436$, $\eta_p^2 = 0.007$, and the transfer test, $F(1, 85) = 0.221$, $p = .640$, $\eta_p^2 = 0.003$.

4. Discussion

This study investigated whether varying the distance between spatially separated but mutually referring text and picture impacted learning when using mental and physical self-managed integration strategies. It was hypothesized that the mental integration strategy would yield higher learning outcomes when text and picture were spatially separated by a small distance, whereas the mental integration strategy was expected to result in lower learning outcomes than the physical integration strategy when text and picture were spatially separated by a large distance. In contrast to our hypothesis, we did not find support for such an interaction between the type of self-management strategy and spatial distance in our results. Rather, it appeared that irrespective of spatial distance between text and picture participants who used the mental integration strategy outperformed participants who used the physical integration strategy on recall of information. This finding is consistent with prior research showing that prompting mental integration of text and picture (Bodemer & Faust, 2006) and teaching a specific mental integration strategy (De Koning et al., 2020) support learning. While in these prior studies the text and picture were presented at a small spatial distance from each other, the present study extends this prior work by demonstrating that using a mental self-management strategy also contributes to learning when text and picture are separated by a slightly larger spatial distance.

Moreover, these findings corroborate the finding of De Koning et al. (2020) that self-managed mental integration is superior to physical self-managed integration. Our findings provide additional support for the finding that using a physical self-management strategy to integrate spatially separated text and picture does not necessarily support learning. If benefits of the physical self-managed integration strategy were found in previous studies (e.g., Sithole et al., 2017), the taught physical integration strategy in the majority of cases contained more than just the physical movement of text segments to the corresponding part in the picture. It also involved additional strategies such as highlighting which are known to support the integration of information and learning (Van Gog, 2014). The findings of the present study combined with those of other studies which used the physical integration strategy without additional supporting strategies (Agostinho et al., 2013; De Koning et al., 2020) provide increasing evidence that a physical integration strategy that simply relies on moving text to the parts of the picture it corresponds to is unlikely to support learning. Encouraging or explicitly instructing learners to engage in additional support strategies thus appears to be a critical factor in the effectiveness of the physical integration strategy (e.g., Roodenrys et al., 2012). Future research could further investigate this by making a direct comparison between physical integration strategies with or without additional supporting strategies. Furthermore, it would also be worth to investigate whether the benefits of the mental strategy remain when a comparison is made with a physical strategy that contains additional supporting strategies.

Following De Koning et al. (2020), the better performance of the mental integration strategy could be explained in terms of the cognitive activities elicited by the specific guidance provided in the mental integration strategy. Learners were taught to imagine moving text segments to the corresponding part in the picture, which encourages active integration of textual and pictorial information in a mental representation (Leahy & Sweller, 2004). When gradually building and refining a mental representation that eventually represents the text and picture information in an integrated format, with each text-picture link that is established in mind learners have to engage in less search and matching processes to find the part of the picture the next text segment corresponds to (i.e., extraneous cognitive load; Sweller et al., 2019). Hence, they can devote considerable working memory capacity to generative (Mayer, 2014) or germane (Sweller et al., 2019) cognitive processing which is reflected in higher learning outcomes. However, while learners using the physical integration strategy receive guidance in how to physically move text segments to the corresponding parts in the picture, using the strategy does not guarantee that active mental integration of the textual and pictorial information takes place. Learners could just have completed the task by moving the text segments to the corresponding parts in the picture and devoting little or no working memory capacity to generative activities that contribute to the construction of an accurate and coherent mental representation. Also, the instruction to move the text segments with the mouse to the corresponding part in the picture on the computer screen requires extra motor coordination not present in the mental integration strategy which imposes additional working memory demands that hinder processing the content of the learning task (i.e., extraneous cognitive load; cf. Skulmowski, Pradel, Kuhnert, Brunnett, & Rey, 2016). This is reflected in the present findings by the lower performance on the recall test for the physical integration conditions.

Table 2
Means and SDs (in Brackets) of Cognitive Load Scores After Learning and After Each Test Per Condition.

	Learning	Recall	Comprehension	Transfer
Physical-small	5.76 (1.98)	6.42 (1.35)	7.19 (1.40)	7.19 (1.57)
Physical-large	5.84 (1.77)	6.80 (1.53)	6.92 (1.44)	6.88 (1.54)
Mental-small	5.37 (1.69)	5.74 (1.76)	6.67 (1.53)	6.68 (1.80)
Mental-large	6.29 (1.52)	6.04 (2.10)	6.88 (1.33)	6.71 (1.78)

[1] Next to analyzing participants' raw cognitive load scores, we used the formula developed by Paas and Van Merriënboer (1993) to calculate instructional efficiency, which is based on the combination of participant's test performance with the cognitive load they experienced during learning. Separate instructional efficiency scores were calculated based on participant's performance on the recall test, comprehension test, and transfer test. An instructional strategy (e.g., self-managed mental integration at a short distance) is more efficient than another one if it produces the same test performance while this is achieved with fewer cognitive resources (Van Gog & Paas, 2008). Separate 2 × 2 ANOVAs with self-management strategy (mental vs. physical) and spatial distance (small vs. large) as between-subject factors showed that there we no significant main or interacting effects for instructional efficiency (*ps* vary from 0.126 to 0.961).

We are aware that this interpretation does not take into account differences in spatial distance between text and picture. In the present study, the learning effectiveness of the mental and physical integration strategies appeared to be comparable irrespective of whether the text and picture were presented at a relatively smaller or larger distance from each other. This finding deviates from prior research showing that increasing the distance between information sources that need to be integrated in a mental representation reduces task performance due to a heavier reliance on working memory (as inferred from less integrative eye movements between the information sources; e.g., Bauhoff et al., 2012). A potential explanation for the failure to find an influence of distance is that the learners might have processed the text and pictures in a way that is unaffected by spatial distance. According to Schüler (2017), an approach that learners use when studying mutually referring text and picture is to first mentally represent the textual and pictorial information separately. Later on, the two mental representations are integrated to form one coherent mental representation of all information together. When using this strategy, the distance between spatially separated text and picture is not relevant because learners do not attempt to step-by-step integrate each text segment with a pictorial element and then move on to the next but rather focus on the text and picture separately and later on integrate the two at once. As we did not collect process-oriented data (e.g., eye-movements, verbal protocols), we do not have insight into whether this actually was an approach learners have used. Based on the learning outcome measures it seems unlikely that this was the case. If all learners had used such a two-step strategy, this would have encouraged active mental integration in all conditions which would have equally benefitted learning outcomes across conditions.

Another more likely explanation is that in the large distance conditions the text and picture were still presented relatively close to each other. For reasons of transferability to actual practice, we used a widely available and commonly sized computer screen (24 inch). In the large distance conditions, the text and picture were presented at maximum distance from each other given the constraints of the computer screen. It is conceivable that in this situation, for the learning task used in the present study, the large distance conditions may not have increased working memory demands up to a level that challenged learners to make them change their processing behavior such that they shifted to a strategy that requires them to rely more heavily on working memory. Recent studies using a comparable screen size, and thus spatial distance between text and picture (De Koning, Rop, & Paas., 2020; Pouw et al., 2019), have reported findings that are consistent with such an interpretation. For example, Pouw et al. (2019) showed that learners studying text and pictures that were spatially separated by a small or large distance both had fewer integrative eye movements than learners studying spatially integrated text and pictures. If the larger spatial distance would have elicited a more working memory-intensive strategy by the requirement to keep information active longer in working memory, more integrative eye movements would have been expected. Together, it is thus likely that in the present study, the distance manipulation was not powerful enough to have a significant impact on learning and cognitive load. Additional research is therefore needed to further investigate the role of distance between spatially separated text and picture on learning and working memory. A promising avenue for future research in this regard is to investigate the mental and physical integration strategies with text-picture materials that are spatially separated by a larger distance, for example by presenting them on a smartboard or multitouch table which typically have larger screens. These devices also allow for physical interaction which makes them especially suitable tools for comparing physical and mental integration strategies at small and large distances. In pursuing this direction of research, complementary information about learners' processing behavior could be obtained by using process measures such as eye-tracking, which could deepen our understanding of why and how variations in spatial distance impact learning or not.

Related to this latter explanation, it should be noted that the constraints of the computer screen not only limited the distance between the text and the picture but also led to the practical decision to segment the text into three columns in the large distance conditions to maximize the distance between text and picture. This has reduced the comparability of the large distance conditions to the small distance conditions where the text was presented in a single column. Previous research has shown that segmenting textual information into smaller units improves retention and comprehension performance (e.g., Florax & Ploetzner, 2010). Therefore, the segmentation of text in the large distance conditions might have increased learning outcomes compared to the small distance conditions and thereby potentially reduced the negative effects of a larger spatial distance on learning. It is, for example, possible that segmenting has made it easier for learners to process the text and picture in the large distance conditions because they could more quickly find and process the text segment they were looking for without being distracted by the rest of the text. Future work investigating the effects of distance and segmentation separately could elucidate to what extent segmentation of the text contributes to learning from text and pictures that are presented at a small or large spatial distance and whether this differentially affects the effectiveness of mental and physical self-managed integration strategies.

A number of educational implications can be tentatively drawn from our findings. First, providing learners with specific instructions (how) to mentally integrate mutually referring text and pictures supports learning. This is particularly helpful when the learning environment offers no or only limited possibilities for interacting with textual and pictorial instructional materials. Second, using a computer-based physical integration strategy that just requires learners to move text to the corresponding part in the picture without stimulating the use of additional strategies to support deeper processing (e.g., highlighting) does not contribute to learning. A more effective alternative for supporting learning and encouraging active integration of text and picture is to use a mental integration strategy. Third, relatively small variations in the distance between spatially separated text and picture presented on a computer screen do not differentially affect learning outcomes or cognitive load when using mental and physical integration strategies. More concretely, for instructional materials comparable to the one used in this study both the smallest and largest distance between text and picture that is possible on a 24 inch computer screen yield comparable learning outcomes and cognitive load. While these implications focus on self-managed strategies to support learning from spatially separated text and pictures in a computer-based learning environment, they are also applicable to situations where instructors aim to design more optimal instructional materials containing textual and pictorial information and in non-computer-supported learning situations such as learning from paper-based materials (cf. Sithole et al., 2017).

In conclusion, the present study built on recent research investigating the effectiveness of mental and physical integration strategies to support learning from spatially separated text and pictures. In an attempt to extend this work, we investigated the distance between spatially separated text and pictures as a potential boundary condition for the effectiveness of self-managed integration of text and picture, particularly in relation to the mental integration strategy. Our findings indicate that the mental integration strategy yielded higher learning outcomes (i.e., recall) than the physical integration strategy overall. To what extent spatial distance impacts learning and cognitive load when using a mental or physical self-managed integration strategy requires further research. Additional research is also warranted to substantiate this conclusion and to more generally investigate when and why self-managed learning strategies contribute to learning.

CRediT authorship contribution statement

Björn B. de Koning: Conceptualization, Investigation, Formal analysis, Writing - original draft. **Gertjan Rop:** Writing - review &

editing. **Fred Paas:** Conceptualization, Writing - review & editing, Supervision.

References

Agostinho, S., Tindall-Ford, S., & Roodenrys, K. (2013). Adaptive diagrams: Handing control over to the learner to manage split-attention online. *Computers & Education, 64*, 52–62. https://doi.org/10.1016/j.compedu.2013.01.007.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001.

Koning, De, B B Rop, G., & Paas, F. (2020). *Learning from split-attention materials: Evidence for a mental self-managed integration effect*. Manuscript submitted for publication.

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd ed., pp. 206–226). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.011.

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.

Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science, 21*, 413–419. https://doi.org/10.1177/0963721412459513.

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.

Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior, 22*, 27–42. https://doi.org/10.1016/j.chb.2005.01.005.

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.

Gordon, C., Tindall-Ford, S., Agostinho, S., & Paas, F. (2016). Learning from instructor-managed and self-managed split-attention materials. *Applied Cognitive Psychology, 30*, 1–9. https://doi.org/10.1002/acp.3159.

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.

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.

Leahy, W., & Sweller, J. (2004). Cognitive load and the imagination effect. *Applied Cognitive Psychology, 18*, 857–875. https://doi.org/10.1002/acp.1061.

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.

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (2nd, rev. ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.

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

Paas, F., & Van Merriënboer, J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*, 737–743. https://doi.org/10.1177/001872089303500412.

Pouw, W., Rop, G., De Koning, 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.

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.

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.

Roodenrys, K., Agostinho, S., Roodenrys, S., & Chandler, P. (2012). Managing one's own cognitive load when evidence of split attention is present. *Applied Cognitive Psychology, 26*, 878–886. https://doi.org/10.1002/acp.2889.

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

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.

Sithole, S. T. M., Chandler, P., Abeysekera, I., & Paas, F. (2017). Benefits of guided self-management of attention on learning accounting. *Journal of Educational Psychology, 109*, 220–232. https://doi.org/10.1037/edu0000127.

Skulmowski, A., Pradel, S., Kuhnert, T., Brunnett, G., & Rey, G. D. (2016). Embodied learning using a tangible user interface: The effects of haptic perception and selective pointing on a spatial learning task. *Computers & Education, 92–93*, 64–75. https://doi.org/10.1016/j.compedu.2015.10.011.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–295. 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*, 261–292. https://doi.org/10.1007/s10648-019-09465-5.

Tindall-Ford, S., Agostinho, S., Bokosmaty, S., Paas, F., & Chandler, P. (2015). Computer-based learning of geometry from integrated and split-attention worked examples: The power of self-management. *Educational Technology & Society, 18*, 89–99.

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.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16–26. https://doi.org/10.1080/00461520701756248.