American Psychological Association

Journal of Experimental Psychology: General

0096-3445/19/$12.00

2019, Vol. 148, No. 11, 2058–2075
http://dx.doi.org/10.1037/xge0000578

The Cognitive Basis for the Split-Attention Effect

Wim Pouw
Erasmus University Rotterdam and University of Connecticut

Gertjan Rop and Bjorn de Koning
Erasmus University Rotterdam

Fred Paas
Erasmus University Rotterdam and University of Wollongong

The split-attention effect entails that learning from spatially separated, but mutually referring information sources (e.g., text and picture), is less effective than learning from the equivalent spatially integrated sources. According to cognitive load theory, impaired learning is caused by the working memory load imposed by the need to distribute attention between the information sources and mentally integrate them. In this study, we directly tested whether the split-attention effect is caused by spatial separation per se. Spatial distance was varied in basic cognitive tasks involving pictures (Experiment 1) and text–picture combinations (Experiment 2; preregistered study), and in more ecologically valid learning materials (Experiment 3). Experiment 1 showed that having to integrate two pictorial stimuli at greater distances diminished performance on a secondary visual working memory task, but did not lead to slower integration. When participants had to integrate a picture and written text in Experiment 2, a greater distance led to slower integration of the stimuli, but not to diminished performance on the secondary task. Experiment 3 showed that presenting spatially separated (compared with integrated) textual and pictorial information yielded fewer integrative eye movements, but this was not further exacerbated when increasing spatial distance even further. This effect on learning processes did not lead to differences in learning outcomes between conditions. In conclusion, we provide evidence that larger distances between spatially separated information sources influence learning *processes*, but that spatial separation on its own is not likely to be the only, nor a sufficient, condition for impacting learning *outcomes*.

Keywords: split-attention effect, cognitive load theory, working memory, educational psychology, cognitive psychology

The combination of instructional text (written or spoken) and pictorial information (static or dynamic) is ubiquitous nowadays in textbooks and e-learning resources. Research on this so-called multimedia learning, which is typically based on cognitive load theory (CLT; Sweller, Ayres, & Kalyuga, 2011) and the cognitive theory of multimedia learning (CTML; Mayer, 2014), has shown

This article was published Online First April 11, 2019.

Wim Pouw, Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, and Department of Psychological Sciences, University of Connecticut; Gertjan Rop and Bjorn de Koning, Department of Psychology, Education & Child Studies, Erasmus University Rotterdam; Fred Paas, Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, and School of Education/Early Start, University of Wollongong.

Wim Pouw and Gertjan Rop have contributed equally to this article and share first authorship. We would like to express our gratitude to Alex van Straaten, Sven Cammeraat, and Tudor Cristea for their help during the data collection. We would also like to thank Rolf Plötzner for providing the materials used in Experiment 3. This research was funded by the Excellence Initiative grant from the Erasmus University Rotterdam awarded to the Educational Psychology section. We also thank the Erasmus Behavioral Lab, especially Marcel Boom, for their continued help in the lab. Raw data, analyses scripts, and Python experiment code and stimuli supporting this research report can be retrieved from The Open Science Framework (https://osf.io/ruqfk/). Wim Pouw and Gertjan Rop were main contributors of writing the Introduction and Discussion, Bjorn de Koning cowrote the Introduction and Discussion, and Fred Paas provided critical revisions. Gertjan Rop and Wim Pouw designed Experiments 1, 2, and 3, with critical guidance by Fred Paas. Wim Pouw programmed and analyzed Experiments 1 and 2, and wrote the Results sections. Wim Pouw wrote the Method section of Experiment 1 and Gertjan Rop wrote the Method section for Experiment 2. Gertjan Rop programmed and analyzed Experiment 3, and wrote the Method and Results sections with Wim Pouw contributing to the JASP analyses. Wim Pouw and Gertjan Rop supervised data collection of Experiment 1, Gertjan Rop supervised data collection of Experiment 2 and 3. This research was funded by the Excellence Initiative grant from the Erasmus University Rotterdam awarded to the Educational Psychology section. The authors have no conflicts of interest to report. This experiment was designed and conducted in accordance with the guidelines of the ethical committee of the Department of Psychology, Education, and Child Studies, at the Erasmus University Rotterdam. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Correspondence concerning this article should be addressed to Gertjan Rop, Rotterdam School of Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. E-mail: grop@rsm.nl

that learning generally improves when instructional materials combine pictures and text (i.e., the multimedia principle; Butcher, 2014). However, simply combining text and pictures without further deliberation about how to present them to learners likely leads to suboptimal learning (Ayres & Sweller, 2014; Mayer & Fiorella, 2014). A well-known finding in this respect is the split-attention effect (e.g., Ayres & Sweller, 2014; Chandler & Sweller, 1991, 1992; Florax & Ploetzner, 2010; Ginns, 2006; Mayer & Fiorella, 2014).[1] The effect entails that when to-be-integrated textual and pictorial learning sources are spatially separated, learning is hindered as compared to spatially integrated sources. This general finding has led instructional designers to promote spatial integration of multimedia sources (Ayres & Cierniak, 2012).

Often when the split-attention effect is obtained it is merely assumed that the effect is produced by "splitting attention" over spatially separated information; however, the underlying mechanism of the split-attention effect is rarely directly tested. The general explanation for the split-attention effect is provided by CLT (Sweller et al., 2011), and dictates that diminished learning is caused by increased cognitive load imposed by spatial separation (Paas & Sweller, 2014). The need to search for related elements in the textual and pictorial information sources, while keeping relevant information active in working memory to mentally connect corresponding information has been argued to impose an extraneous cognitive load (e.g., Sweller, Van Merrienboer, & Paas, 1998). Given that working memory is limited in capacity and duration (Baddeley, 2000; Barrouillet & Camos, 2007; Cowan, 2001; Miller, 1956; Puma, Matton, Paubel, & Tricot, 2018), this reduces working memory resources available for processes that are relevant for learning, such as schema construction, elaboration, and automation in long-term memory (Sweller, 1994). Consequently, learning is hampered. For integrated sources though, the burden on working memory is limited as information sources can be directly visually compared. A myriad of studies have shown that, in line with this explanation, spatially integrated learning materials impose a lower cognitive load and lead to higher learning outcomes than spatially separated learning materials (Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Chandler & Sweller, 1991, 1992; Mayer, Steinhoff, Bower, & Mars, 1995; Tarmizi & Sweller, 1988; see Ginns, 2006 for a meta-analysis).

However, the split-attention effect may not be caused by spatial separation of related information per se. Several techniques exist that can be used to resolve the split-attention effect while preserving spatial separation. For example, the direction of learners' attention by signaling the corresponding parts of the text and picture is frequently employed (De Koning & Jarodzka, 2017; Jamet, 2014; Van Gog, 2014). Research has shown that mental integration of textual and pictorial information is improved when corresponding text and parts of the picture are presented in the same color (De Koning, Tabbers, Rikers, & Paas, 2009; Ozcelik, Arslan-Arib, & Cagiltay, 2010; Ozcelik, Karakus, Kursun, & Cagiltay, 2009). Importantly, when a student actively integrates instructional materials consisting of two mutually referring information sources, comprehension is improved (Mayer, 2014; Schnotz, 2014). Moreover, using cross-referring labels, such as when dividing spatially separated information into smaller segments and labeling the corresponding text-picture parts with numbers, also reduces the split-attention effect (Florax & Ploetzner, 2010). It might therefore be that the searching costs that are imposed by nonintegrated information sources are not caused by spatial distance, but because the learner does not know which pieces of information belong together and has to perform an effortful search to find semantically related segments. In integrated learning materials, spatial contiguity automatically signals which information sources belong together. Thus, it is currently unclear whether split attention is caused by spatial separation, or due to obscurities about which informational sources should be related.

In the present study, we therefore investigated to what extent spatial distance of information sources can account for the split-attention effect. That spatial distance plays a role in the split-attention effect, as originally conceived, is consistent with classic cognitive psychological research on *embedded cognition* (Ballard, Hayhoe, & Pelz, 1995; Gray & Fu, 2004; for a review see Pouw, van Gog, & Paas, 2014). It has been shown that when information that needs to be integrated is spatio-temporally separated (Ballard et al., 1995; Gray & Fu, 2004), problem solvers change from a perceptually intensive strategy (less prone to mistakes; higher saccade counts) to what seems to be a memory intensive strategy (leading to more mistakes; lower saccade counts). More precisely, Ballard et al. (1995) used a task in which participants had to copy a pattern of colored blocks. When the distance between the model and the workspace in which participants had to copy the pattern of blocks was small (model and workspace were separated by 15°), and the cost of direct acquisition of information was small, participants made more saccades, implying less use of working memory resources. When this distance between information sources was increased (70°), thereby elevating the cost of direct visual comparison, participants made fewer saccades implying more use of working memory resources (see also Lagers-van Haselen, van der Steen, & Frens, 2000 for a replication). In the study by Gray and Fu (2004), it was found that participants who memorized task-relevant information before the main task were more likely to use this memorized information when it would take more time to attain this information from the computer display given a number of mouse-clicks. When the number of mouse clicks that were needed to attain the relevant information was reduced, participants were less likely to rely on their own memory and would attain this information in the digital environment. Interestingly, this had the effect that more mistakes were made when information was less easily attainable as participants were more likely to rely on imperfect "information-in-the-head," as opposed to participants that could rely on perfect "information-in-the-world" in the condition where there was low cost of retrieving the information (i.e., time and effort to get the relevant information from the display). Together, these studies (Ballard et al., 1995; Gray & Fu, 2004) indicate that there seems to be a trade-off between *spatiotemporal* separation and the use of memory resources.

In more applied settings, comparable findings have been obtained. Johnson and Mayer (2012), for example, recorded participants' eye movements while they learned how car brakes work using a single-slide multimedia lesson consisting of a diagram and text. When the text was integrated in the relevant parts of the

[1] Note that the "split-attention effect" can have a different meaning outside of educational psychology. This then concerns the degree to which humans can visually track or detect two or more (moving) objects at once (e.g., Awh & Pashler, 2000). In this study, we focus on the split-attention effect as it is referred to in educational psychology.

diagram, participants made more saccades between these two sources of information than when the text was presented separately from the diagram. Moreover, participants' understanding of how car brakes work was better in the integrated condition. In a study by Bauhoff, Huff, and Schwan (2012) participants judged whether or not two depictions of a mechanical pendulum clock were identical. The spatial distance between these two depictions was varied, and Bauhoff et al. (2012) observed that participants made fewer saccades between the two depictions of these clocks when the spatial distance between the pictures was increased, suggesting higher working memory constraints. Together, these studies suggest that nonintegrated information (Johnson & Mayer, 2012), or increased spatial distance between information sources (Bauhoff et al., 2012) leads to more memory-intensive strategies, which provides indirect evidence for the CLT-based explanation of the split-attention effect described above (Paas & Sweller, 2014; Sweller et al., 2011). So far, however, it has not yet been investigated within more basic cognitive tasks whether spatial separation affects cognitive load and task performance directly and to what extent varying the spatial distance between two unintegrated information sources influences learners' perceptual and cognitive processing in more applied settings. If spatial distance is the key factor in producing the split-attention effect, it could be argued that given a linear relationship between the spatial distance between two information sources and working memory load (Hardiess, Gillner, & Mallot, 2008; Inamdar & Pomplun, 2003), the further apart two information sources are, the more likely it will be that learners' working memory is overloaded and that learners experience the negative consequences of split-attention.

The Present Study

The aim of this study was to investigate whether the split-attention effect can be explained by the spatial distance between information sources, and whether and how this basic cognitive phenomenon affects learning processes and outcomes in more ecologically valid learning materials. We therefore conducted three experiments in which the distance between two information sources was varied. Experiment 1 intended to assess distance effects using a basic paradigm wherein participants made similarity judgments based on two pictorial information sources (cards with symbols) that were separated at different spatial distances. In Experiment 2, for half of the cards we replaced the symbols on the card with a written description of the information presented on the card. This enabled us to examine whether the results from Experiment 1 would replicate when participants had to actively integrate pictorial and textual information and laid the foundation for the next experiment in which distance effects were investigated with more educationally relevant material. In Experiment 3, participants learned about human brain processes from a multimedia presentation consisting of a picture with accompanying text. Both information sources were unintelligible in isolation, so participants had to mentally integrate the pictorial and textual information to understand the process. The picture and text were either presented in a spatially integrated way (i.e., integrated condition) or spatially separated in such a way that the picture and text were presented in closer proximity to each other (i.e., small-separation condition) or were separated at a larger spatial distance (i.e., large-separation condition). Additionally, Experiment 3 applied eye-tracking methodology to investigate participants' information gathering strategies. Across experiments, the general prediction was that with greater distance between two information sources learners would show decreased performance.

Experiment 1

Drawing on fundamental cognitive science research (e.g., Ballard et al., 1995), this first experiment aimed to establish an effect of spatial distance when processing two pictorial information sources. Participants judged the similarity of two cards each containing three symbols and the spatial distance between the cards was varied. In half of the trials participants maintained a visual pattern in working memory, leading to additional cognitive load during information integration. We predicted that greater distance of to-be-integrated information would lead to diminished performance (Hypothesis 1), and that such diminished performance would be more pronounced under a higher cognitive load condition (Hypothesis 2). Furthermore, we predicted that a larger distance between information sources would lead to more demanding working memory strategies, which would negatively affect retrieval performance of the visual pattern that was concurrently held in working memory (Hypothesis 3).

Method

Participants and design. Fifty-two (M_{age} = 21.00 years, SD = 3.57 years; 39 female) undergraduate students from Erasmus University Rotterdam participated for course credits or a 5 euros reward. 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. Note, for the mixed effects regression analyses for repeated measures reaction time (RT) data (which usually generates small effect sizes d = 0.1) Brysbaert and Stevens (2018) recommend to use at least 1,600 observations per condition for 80% power. In the current experiment we have 3,120 observations per condition (1,560 for Hypothesis 3). A within-subjects design was employed with the factors cognitive load (two levels: load absent vs. load present) and card similarity (three levels: no similarity vs. one similarity vs. two similarities), and distance as a continuous covariate (see below).

Apparatus and stimuli. The experiment was presented on an Eizo FlexScan S2243W 22-in. monitor of 47 cm × 30 cm, and resolution was set at 1,920 × 1,080. The task was programmed in Python (Toolbox Pygaze; Dalmaijer, Mathôt, & van der Stigchel, 2014).

Stimuli for the card integration task consisted of a full card deck of the Wisconsin card sorting task (WCST; retrieved from Stoet, 2016). Each card had three feature dimensions (SHAPE + NUMBER + COLOR) with four possible levels (SHAPE: star, cross, triangle, circle; NUMBER: 1, 2, 3, 4; COLOR: blue, yellow, green, red). The total card deck of 64 cards (4 × 4 × 4 levels) was randomly placed on an 8 × 8 matrix (see Figure 1; matrix = 928 × 928, pixels = 25.77 cm × 25.77 cm). For each unique card integration trial (60 trials) card selections were pseudorandomly generated for each participant, such that 20 trials consisted of two cards that were dissimilar on all dimensions (card similarity = 0), 20 trials

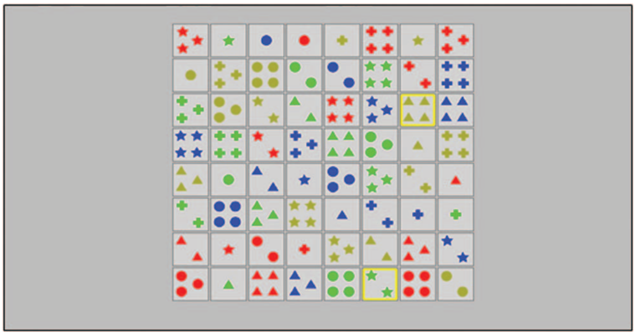

Figure 1. Example of a card integration task trial with cards selected (card similarity = 0). In the current example participants should respond with the continue button (“SPACE”) as the selected cards (signaled by yellow rectangles) were not similar on any of the dimensions SHAPE, NUMBER, COLOR. In the current example the distance was 478 pixels (12.91 cm). See the online article for the color version of this figure.

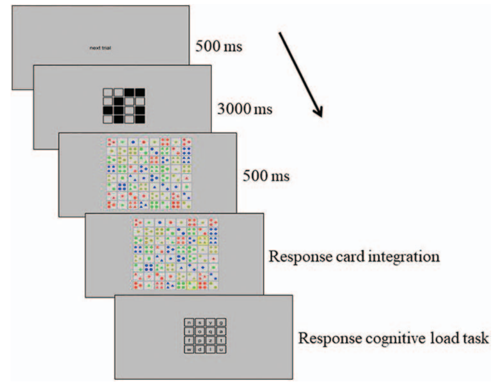

Figure 3. Example of a single trial with a secondary cognitive load task. A trial without cognitive load would not have the card integration task preceded/followed by a visual pattern presentation/response (i.e., only Slides 1, 3, and 4 as counted from above). See the online article for the color version of this figure.

contained selections of cards that were similar on one dimension (e.g., similar in COLOR; card similarity = 1), and 20 trials contained selections of cards that were similar on two dimensions (e.g., similar in COLOR and SHAPE; card similarity = 2). Note that a similarity of three dimensions was not possible because there were only unique cards in a deck. The selection of the cards to be compared for similarity was signaled by two bright yellow rectangles around the selected cards (see Figure 1). Participants responded for similarity per dimension using the response buttons “c” (SHAPE match), “v” (NUMBER match), “b” (COLOR match). If there was more than one match, participants had to push two buttons in a row (order was irrelevant). SPACE needed to be pressed to continue to the next trial. It is important to note that depending on card similarity (0, 1, or 2), more buttons needed to be pressed, as card similarity = 0 required only a SPACE press (one key press), while card similarity = 1 required a match button + a SPACE press (two keypresses), and card similarity = 2 required two match buttons + a SPACE press (three keypresses).

The unique set of 60 card integration trials was presented twice, once with and once without a secondary cognitive load task (order fully randomized; see Figure 2). As such, card integration trials were identical in nature (i.e., matched on card similarity type and distance) across cognitive load conditions. Figure 3 shows a trial flow with secondary cognitive load task. The final list of 120 trials was randomized in order of presentation. The Euclidean distance (measured in pixels) between the random selections of cards was the main variable of interest. The distances could vary between 116 pixels (3.22 cm) for directly adjacent card selections, and 1,148 pixels for card selections in opposite corners (31.88 cm).

In half of the trials (60 trials), a secondary visual cognitive load task was performed (see procedure). This task is an adapted visual patterns test (Della Sala, Gray, Baddeley, & Wilson, 1997) and has been used to measure visual working memory capacity. For each trial, a random pattern was generated of eight black squares filling an 8 × 8 matrix. This pattern was presented for memorization for 3,000 ms preceding the card integration trial. For the response phase (which was preceded by a card integration trial) participants recalled the pattern. For each trial the response buttons were randomly chosen for each matrix cell from a list from a to z (excluding the response buttons of the integration task, “c,” “v,” “b”), such that letters were not associated with particular locations across trials. Participants typed in the letters that corresponded with the pattern of black squares (order irrelevant) and could proceed to the next trial by pressing SPACE (see Figure 2).

For the instruction phase, 50 practice trials were randomly created per participant. In the first 3 × 10 trials participants learned to correctly respond on the integration task to single features, namely SHAPE, NUMBER, and COLOR. For the subsequent 10 practice trials, participants needed to respond for similarities to all features (i.e., SHAPE, NUMBER, and COLOR) at once. In the final 10 practice trials, participants learned to also perform the visual cognitive load task concurrently.

Procedure. Participants were seated in a well-lit cubicle at about 50 cm from the computer screen. To remind participants of

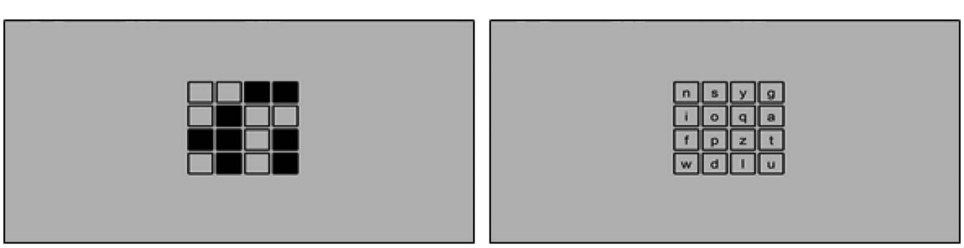

Figure 2. Visual cognitive load task: presentation phase (left) and response phase (right).

the response buttons for indicating similarity between selected cards, the experimenter had labeled the response keyboard buttons "c," "v," "b," with stickers indicating "SHAPE," "NUMBER," and "COLOR," respectively.

First, participants were instructed about the nature of the task. During this instruction phase, participants were repeatedly prompted to ask questions to the experimenter if they did not understand the task. Participants learned to press SPACE when cards were dissimilar on all dimensions (card similarity = 0), and press two buttons when there was a similarity on one dimension (card similarity = 1, e.g., pressing "c" then "SPACE" to continue), and press three buttons when there were two similarities (card similarity = 2, e.g., pressing "v" and "b," and then "SPACE" to continue). After the practice phase, participants performed the 120 experimental trials. The experiment was administered without breaks, and took about 40 min.

Performance and scoring. For the card integration task, the main measures of performance were accuracy (integration accuracy) and integration RT. Integration accuracy was a dichotomous measure of performance per dimension per trial (e.g., correct [mis]match for SHAPE, NUMBER, and COLOR; max = 3 points). Integration RT was a continuous measure of performance which entailed the time between card selection onset and participants finalizing card integration by pressing SPACE. Note, that in the analyses we only focused on integration RT as we found that accuracy was very high (>95%), leaving little meaningful variance to analyze.

For the 60 trials where a secondary cognitive load task was performed the main measure of interest was visual pattern test (VPT) score (hereinafter VPT score), which was determined by the number of correctly pressed buttons minus the number of incorrectly pressed buttons with a maximum of eight incorrect keypresses (maximum score of 8, and minimum score of −8). If participants pressed a button more than once this was only scored (in)correctly once. VPT RT was not of main interest because the measure of RT is less meaningful when the number of buttons pressed can vary.

Results

Descriptives. Table 1 shows the main descriptives of performance on the card integration task as well as the performance on the VPT.

Preanalysis and outliers. In total 6,240 trials were run (52 participants × 60 trials × 2 conditions). Following common practice, we excluded trials further than three standard deviations from the mean of the Integration RT, 64/6,240 trials (0.1%). We further restricted our main performance analyses to only correct trials for the card integration task, and excluded all incorrect trials (5.02% of the remaining trials).

Hypothesis 1 and 2. To assess Hypotheses 1 and 2 (i.e., greater distance of to-be-integrated information would lead to diminished performance, which would be more pronounced when cognitive load was present), we performed a linear mixed effects model (R Version 3.4.0, nlme Version 3.1–131). Throughout, we used maximum likelihood estimation with random intercepts for participants. See Figure 4 for a graphical overview of the integration RT data.

In building our model, we first entered cognitive load as predictor of integration RT. This added predictive value compared with a model predicting the overall mean (BIC = 101,484, chi-square change [1] = 12.19, $p < .001$). We further entered card similarity in the model, and this improved the model as compared with the model with cognitive load only (BIC = 98,694, chi-

Table 1
Mean (and Standard Deviation) Reaction Time (in Milliseconds) and Integration Accuracy With 95% Confidence Intervals Around the Mean (in Brackets) on the Card Integration Task and the Visual Pattern Task

Outcome measures	Integration RT	Integration accuracy % correct	VPT accuracy	VPT RT
Load	3,320 (1,960) [3,251, 3,388]	2.95 (.22) [2.94, 2.96] 98.33%	4.99 (1.50) [4.89, 5.09]	11,080 (5,541) [10,896, 11,264]
No Load	3,057 (2,293) [2,977, 3,138]	2.95 (.24) [2.94, 2.95] 98.33%		
Card similarity = 0 (one keypress)	2,076 (1,689) [2,004, 2,149]	3.00 (0) [3.00, 3.00] 100%	5.55 (2.57) [5.39, 5.70]	10,792 (4,995) [10,488, 11,096]
Card similarity = 1 (two keypresses)	3,405 (2,130) [3,314, 3,497]	2.95 (.23) [2.94, 2.96] 94.81%	5.01 (2.76) [4.85, 5.18]	11,135 (5,012) [10,830, 11,439]
Card similarity = 2 (three keypresses)	4,083 (2,058) [3,995, 4,172]	2.89 (.32) [2.88, 2.91] 96.33%	4.42 (3.10) [4.23, 4.60]	11,314 (5,671) [10,969, 11,659]
Overall	3,189 (2,137) [3,136, 3,241]	2.95 (.23) [2,943, 2,954] 98.33%	4.99 (1.50) [4.89, 5.09]	11,080 (5,541) [10,896, 11,264]
Distance				
r	.022	.015	−.033	.015

Note. Data before trimming. VPT = visual pattern test; RT = reaction time.

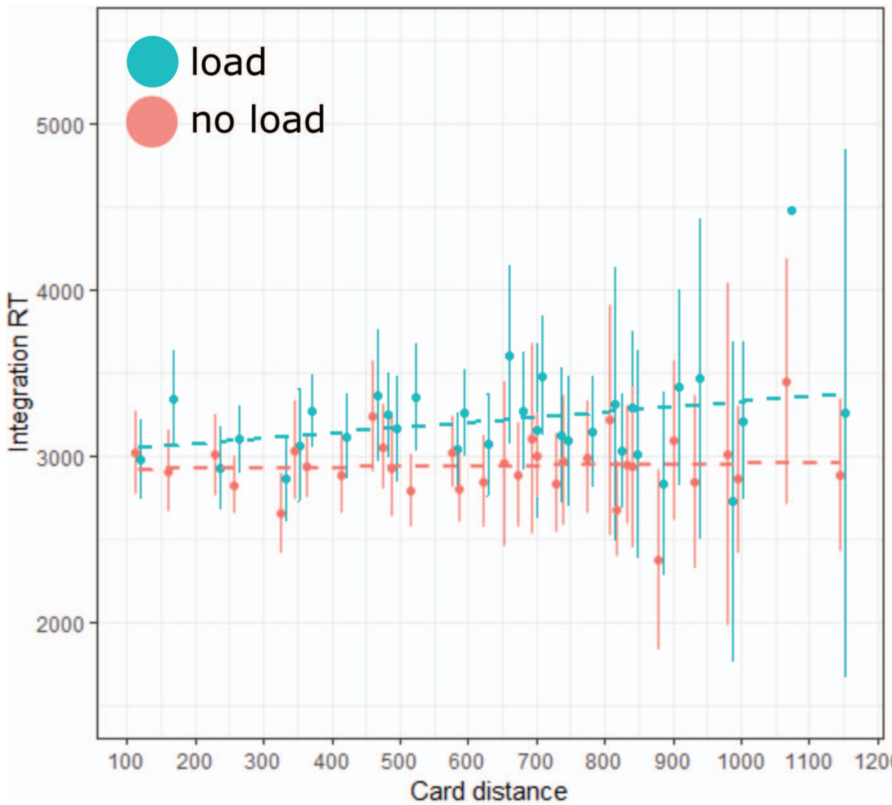

Figure 4. Effect of distance and cognitive load on subsequent integration reaction time (RT). Each point represents the mean score for all participants on that particular card distance (only successful card integrations). Error bars represent 95% CI's. Note that card positions at maximum distance concern fewer observations, and therefore CI's are wider and also less influential in the model. See the online article for the color version of this figure.

square change [2] = 2807.18, $p < .001$). Additionally, we entered distance which did not improve the model further (BIC = 98,702, chi-square change [1] = 0.478, $p = .490$). Finally, we entered the interaction between cognitive load and distance, which also did not improve the model as compared with previous models (BIC = 98,709, chi-square change [1] = 1.97, $p = .161$).

The resulting model with cognitive load and card similarity showed that there was a main effect of cognitive load on integration RT, $b = 265.55$, 95% CI [118.23, 412.85], $t(51) = 3.617$, $p < .001$. This indicates that participants were slower to successfully integrate card stimuli on trials with concurrent cognitive load. Card match type was a statistically significant predictor showing an increase in integration RT when cards were more similar (and more buttons needed to be pressed). Going from zero to one similarity increased RT by $b = 1331.70$, 95% CI [1265.87, 1396.78], $t(5760) = 39.859$, $p < .001$, from zero similarity to two, $b = 1987.51$, 95% CI [1921.07, 2053.95], $t(5760) = 58.621$, $p < .001$. Note that in the model with distance added there was a positive overall relation with RT, but this was not significant.

Hypothesis 3. Hypothesis 3 predicted that greater distance between information sources would negatively affect retrieval performance on the VPT. Figure 5 shows the relation between distance of the cards to be integrated and the subsequent performance on the VPT (thus only for cognitive load trials). We further performed a linear mixed effects model similar to the previous analyses for Hypothesis 1 and 2, with random intercepts for participants and cognitive load condition. Adding distance to the model predicting VPT score resulted in a significant increase in predictive value compared with a model predicting the overall mean, BIC = 14,044, chi-square change [1] = 5.41, $p = .020$. Adding card similarity further improved the model, BIC = 14,044, chi-square change [2] = 93.48, $p < .001$. Adding an interaction between card similarity and distance did not benefit the model, BIC = 13,981.51, chi-square change [2] = 1.99, $p = .370$.

As predicted by Hypothesis 3, the model shows that greater distance resulted in lower VPT scores, $b = -0.00049$,[2] 95% CI [−0.00091, −0.0008], $t(2767) = -2.33$, $p = .020$. This means that for every 100 pixels (ca. 2.77 cm) in distance the model predicts a decrease in performance of 0.05. Furthermore, card similarity again affected performance such that higher similarity (higher keypresses) resulted in lower VPT scores. Going from zero to one similarity decreased performance by $b = -0.546$, 95% CI

[2] Note that the b-value is so small because it expresses a relationship between 1 pixel change in distance relative to one point change in the VPT score. The effect of 100 pixels change in distance can be calculated by multiplying the current b-value by a 100 (i.e., 0.05 VPT point change per 100 pixel increased distance).

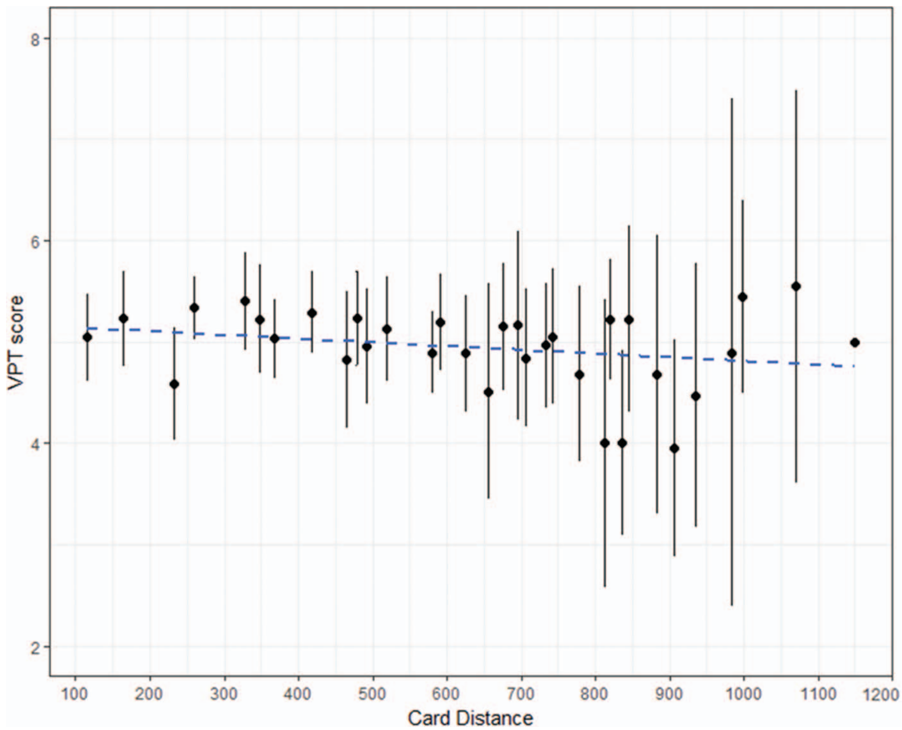

Figure 5. Effect of card integration distance on subsequent performance on the visual pattern test score (VPT score). Each point represents the mean score for all participants on that particular card distance (only successful card integrations). Error bars represent 95% CI's. Note, that card positions at maximum distance concern fewer observations, and therefore CI's are wider and data points are also less influential. See the online article for the color version of this figure.

[−0.77, −0.32], $t(102) = 3.617$, $p < .001$, from zero similarity to two, $b = -1.13$, 95% CI [−1.36, −0.90], $t(102) = -9.624$, $p < .001$.

Discussion

Based on CLT, we predicted that when information that needs to be integrated is spatially separated, problem solvers will have to mentally carry over information to integrate it with the spatially distant information source. In the current experiment, we did not find that spatial distance between information affected information integration time (Hypothesis 1 and 2). However, we did obtain that integration of information at higher spatial distances resulted in lower performance on a secondary visual working memory task (Hypothesis 3). This fits an explanation assuming that integrating information sources that are more distant from each other invites a more memory-intensive strategy that in turn leads to interference of information already maintained in working memory, leading to lower retrieval performance (Gray & Fu, 2004). That spatial distance affected performance on the visual cognitive load task and not the card integration task suggests that unintegrated information can be successfully dealt with in terms of RT losses through a more memory intensive strategy, but this comes at the cost of other processes that also make use of the working memory system. The current finding that working memory can effectively step in to solve the task in time aligns with the finding of Gray and Fu (2004) who suggested that participants adopt a strategy that allows for the quickest problem-solving solution at the cost of accuracy.

It is important to note that in Experiment 1, participants had to compare and contrast two pictorial stimuli, while the split-attention effect in multimedia is generally studied with materials consisting of a combination of text and pictures (e.g., Ayres & Sweller, 2014; Chandler & Sweller, 1991, 1992; Florax & Ploetzner, 2010; Ginns, 2006; Mayer & Fiorella, 2014). Therefore, we were interested whether the results of Experiment 1 would replicate when participants needed to integrate pictorial and textual information using the present paradigm. Experiment 2 was conducted to address this.

Experiment 2

Experiment 2 was a direct replication of Experiment 1, with one small adjustment. Half of the Wisconsin cards were substituted with a written description of the information presented on the card it replaced. For example, instead of the picture with one red star, the three dimensions (i.e., number, color, and object) were written on a card. In this experiment, participants had to compare an original card with a card containing a written description on the three dimensions, allowing us to test whether the results of Experiment 1 would replicate when participants have to integrate textual and pictorial information. This experiment, and all planned analyses were preregistered, and all analyses, data, and materials are retrievable (https://osf.io/ruqfk/).

Method

Participants and design. Fifty (M_{age} = 20.34 years, SD = 3.00 years; 46 female) undergraduate students from Erasmus University Rotterdam participated for course credits. The same within-subjects design as in Experiment 1 was used.

Apparatus and stimuli. The apparatus and stimuli were identical to Experiment 1, with two small exceptions. First, the stimuli for the card integration task were expanded with a textual variant of each WCST card. As a result, 128 cards were used, which were again randomly placed on an 8 × 8 matrix, with half of the cards pictorial, and the other half textual (see Figure 6). Second, the experiment now comprised 45 trials, as participants in Experiment 1 indicated that the experiment was quite taxing and boring to complete. Fifteen trials consisted of two cards that were dissimilar on all dimensions, 15 trials with consisted of cards that were similar on one dimension, and 15 trials that consisted of cards that were similar on two dimensions. The procedure, outlier detection, and analyses were identical to Experiment 1.

Results

Descriptives. Main descriptives for the card integration task and the VPT task are provided in Table 2.

Preanalysis and outliers. For this experiment, 4,500 trials were run (50 participants × 45 trials × 2 conditions). RT values higher or lower than three standard deviations from the mean of the integration RT, 35/4,500 trials (0.008%) were excluded from analyses. Similar to Experiment 1, our main performance analyses were executed with data for correct trials, excluding RT's for all incorrect trials (85.6% trials remaining).

Hypothesis 1 and 2. We performed a similar mixed linear analyses as in Experiment 1 (with participant as random intercept). Cognitive load was entered as predictor of integration RT, which added predictive value compared to a model predicting the overall mean (BIC = 71,756, chi-square change [1] = 17.75, $p <$ 001). Next, we entered card similarity to the model, improving the model as compared with cognitive load only (BIC = 70,642, chi-square change [2] = 1130.61, $p < .001$). Furthermore, entering distance improved the model even further (BIC = 70,593, chi-square change [1] = 7.72, p = .005). We also looked for possible interactions between distance and card similarity, but these models were not reliable.

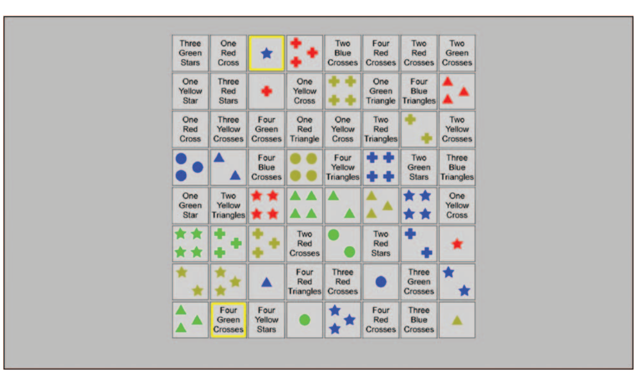

Figure 6. Example of the card integration task used in Experiment 2, in which participants had to compare a pictorial and textual version of the card. In the current trial participants should indicate a match for NUMBER (one item) and for COLOR (red). See the online article for the color version of this figure.

The resulting model with cognitive load, card similarity, and distance showed an effect of cognitive load on integration RT, b = 795.42, 95% CI [459.44, 1131.40], t(49) = 4.76, $p < .001$, indicating slowed responses on trials with concurrent cognitive load. Card match type led to increased integration RT when cards were more similar, going from zero to one similarity increased RT by b = 2404.93, 95% CI [2221.19, 2588.65], t(3719) = 25.65, $p <$ .001, from zero similarity to two, b = 3383.40, 95% CI [3191.74, 3575.06], t(3719) = 34.59, $p < .001$. Finally, and most importantly, higher distance between cards lead to higher RT's, b = 0.492, 95% CI [0.145, 0.838], t(3719) = 2.78, p = .006. In conclusion, higher distance between picture and text reliably slowed down integration RT's, confirming our main hypotheses.

Hypothesis 3. We predicted that increased distance between information sources would negatively affect retrieval performance on the secondary task (VPT). We again performed a linear mixed model, with random intercepts for participants. Adding distance to the model did however not add predictive value as compared to the model predicting the overall mean, BIC = 9283, chi-square change [1] = 0.919, p = .34. Adding card similarity to the model did improve predictive value, BIC = 9281, chi-square change [2] = 17.78, $p < .001$. Adding an interaction between card similarity and distance did not benefit the model, BIC = 9294, chi-square change [2] = 2.03, p = .363. In conclusion, card distance did not lead to reduced accuracy on the secondary VPT task.

Discussion

In the current experiment, distance between to-be-compared text-versus-picture cards led to slower responses on the main task, even after controlling for the amount of keypresses participants had to make (i.e., card similarity). This confirms our hypothesis that physical distance between information is a genuine source of interference for integrating information sources, as working memory processes are likely to be taxed. However, in contrast to Experiment 1 performance on the secondary visual working memory task did not reveal an effect of distance on visual working memory. One possible explanation for this is that the integration task in Experiment 1 was unimodal in nature (visual comparison) while in the current task it was cross-modal (text and visual comparison). Because the VPT task is a visual working memory task, it is likely to especially be affected when the concurrent primary task requires a visual comparison alone, rather than cross-modal comparison which is likely to involve more than visual working memory capacity. That the cross-modal integration is a different process than unimodal integration is further signaled by the longer integration time for the cross-modal versus unimodal task. Surprisingly though, we observed no differences across experiments in the relationship between VPT performance on cognitive load trials, while Integration performance for those trials was relatively higher for Experiment 2, r = −.11, 95% CI [−.07, −.16], t(1906) = −4.97, as compared with the same relationship for Experiment 1, r = −.07, 95% CI [−.03, −.10], t(2922) = −3.58. Thus, as the large overlap in confidence intervals of the correlation estimates indicate, we cannot draw any

Table 2
Mean (and Standard Deviation) Reaction Time (in Milliseconds) and Integration Accuracy With 95% Confidence Intervals Around the Mean (in Brackets) on the Card Integration Task and the Visual Pattern Task

Outcome measures	Integration RT	Integration accuracy % correct	VPT accuracy	VPT RT (milliseconds)
Load	5,938 (3,917) [5,777, 6,100]	2.84 (.42) [2.82, 2.85] 94.67%	5.11 (2.99) [4.99, 5.24]	13,924 (6,808) [13,643, 14,205]
No Load	5,443 (7,340) [5,140, 5,747]	2.84 (.41) [2.83, 2.86] 94.67%		
Card similarity = 0 (one keypress)	3,628 (2,261) [3,514, 3,743]	3.00 [2.99, 3.00] 100%	5.51 [5.31, 5.71]	13,430 (6,744) [12,946, 13,913]
Card similarity = 1 (two keypresses)	6,499 (8,824) [6,053, 6,947]	2.83 [2.81, 2.85] 94.33%	4.96 [4.74, 5.18]	14,074 (6,702) [13,593, 14,553]
Card similarity = 2 (three keypresses)	6,945 (3,817) [6,751, 7,138]	2.69 [2.67, 2.72] 89.67%	4.88 [4.66, 5.10]	14,270 (6,955) [13,771, 14,768]
Overall	5,691 (5,887) [5,519, 5,862]	2.84 (.411) [2.83, 2.85] 94.67%	5.11 (2.99) [4.99, 5.24]	13,924 (6,808) [13,643, 14,205]
Distance				
r	−.015	−.013	−.015	.020

Note. Data before trimming. VPT = visual pattern test; RT = reaction time.

conclusions on differences in correlation strength as to support our proposed explanation that Experiment 1 was more taxing for visual working memory capacity than Experiment 2.

All in all, Experiment 1 and 2 have confirmed that an increase in spatial distance between two stimuli has an effect on cognitive load and integration speed. A next step is to study whether these effects would scale-up, and also influence learning from more complex multimedia materials. This was investigated in Experiment 3.

Experiment 3

The aim of this experiment was to investigate whether increasing the distance between spatially separated textual and pictorial information yields a stronger split-attention effect when using a learning task. To this end, participants learned about human brain processes, with materials adapted from Florax and Ploetzner (2010) consisting of a picture with accompanying text. The text described the relevant processes also portrayed in the picture, and both sources of information were needed to fully grasp the learning material (Florax & Ploetzner, 2010). We created three conditions: the integrated condition (i.e., the text and picture are spatially integrated), the small-separation condition (i.e., the text is segmented and the picture is labeled, and they are separated by a small spatial distance), and the large-separation condition (i.e., the text is segmented and the picture is labeled, and they are separated by a large spatial distance). To enable investigating whether text segmenting and picture labeling could effectively reduce the split-attention effect (cf. Florax & Ploetzner, 2010), in the spatially separated conditions the text was segmented and the picture was labeled. Eye-tracking methodology was applied to examine whether an increase in spatial distance leads to a less visually-intensive strategy, as indicated by fewer transitions between the text and picture (e.g., Ballard et al., 1995; Gray & Fu, 2004; also see Experiment 1).

We expected that learning (i.e., retention and comprehension) and processing demands (i.e., cognitive load) in the integrated and small-separation conditions would not differ, because the segmenting and labeling would alleviate any negative effect of split-attention (Hypothesis 1). This would replicate the results of Florax and Ploetzner (2010). Based on the literature discussed above, we expected that learning would be more cognitively demanding (i.e., an increased cognitive load) and learning outcomes (i.e., retention and comprehension) would suffer in the large-separation condition compared with the small separation and integrated conditions (Hypothesis 2). To test whether an increase in spatial distance would indeed make learning more cognitively demanding, we asked participants to rate how much mental effort they invested in learning the materials (as an indicator of how much cognitive load participants experienced; Paas, 1992; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). We also asked participants to rate how much mental effort they invested during the posttest, as participants who gained more knowledge during the learning phase should be able to attain higher test performance with less investment of mental effort (Paas & Van Merriënboer, 1993; Van Gog & Paas, 2008).

An increase in spatial distance should lead to a more memory-intensive learning strategy, leading to fewer transitions between the text and picture (e.g., Ballard et al., 1995; Gray & Fu, 2004), and spatially integrating two mutually referencing information sources should lead to more transitions than spatially separated information sources (cf. Holsanova, Holmberg, & Holmqvist, 2009; Johnson & Mayer, 2012). Therefore, we expected that participants in the integrated condition would make the most transitions between the text and the picture, followed by participants in the small-separation condition, who in turn make more transitions than participants in the large-separation condition (Hypothesis 3). Fewer transitions are indicative of less integration of the text and picture, which can explain why an increase in spatial distance would hamper learning (cf. Mason, Pluchino, & Tornatora, 2015, 2016). We also measured the total fixation duration on the text and picture, as it seems that learning from text and pictures is mostly text driven, with little to no attention to the picture (cf. Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010). As an increase in distance between the text and picture should aggravate this effect, we expected that the fixation duration would be longest on the text and shortest on the picture in the large-separation condition, followed by the small separation and integrated conditions (Hypothesis 4).

Method

Participants and design. Participants were 75 undergraduate university students who participated for course credit or a 10 euros reward. We arrived at this sample size as we collected as many participants as possible before the lab facilities closed down for the summer. Given that these sample sizes are within common sample size ranges in applied educational psychology and given the research resource constraints we decided to terminate the study for these 75 participants as well as add additional Bayesian analyses to provide extra indications for the reliability of our data. All participants had normal or corrected-to-normal vision. For three participants, study times indicated that they had skipped a part of the learning phase.[3] The data of these participants were excluded for further analyses, resulting in a sample of 72 students (M_{age} = 21.68 years, SD = 2.86 years; 44 female), who were randomly assigned to one of the three conditions: integrated (n = 24), small separation (n = 25), and large separation (n = 23).

Materials. All materials were adapted from Florax and Ploetzner (2010). They were translated from German to English, and the distance manipulation was administered by moving the text closer to (small separation) or further from (large separation) the picture.

Background information. Participants were presented with a short expository text on the subject, to provide them with enough background knowledge to understand the learning materials. This background information was presented on paper, and participants could spend as much time reading the information as they wished. On average, it took around 10 min.

Prior knowledge test. Participants' prior knowledge and understanding of the background information was assessed with a paper-based multiple-choice test consisting of 12 questions about neural-chemical transmissions and communication in the human nervous system (e.g., What is a synapse?). These questions had five possible answer alternatives; four of these alternatives could possibly be correct (e.g., the correct alternative: *The connection of two nerve cells, which do not physically touch*), while the fifth alternative was always: *I do not know*. Participants were encouraged not to guess, but to pick the fifth answer alternative when they were unsure which answer alternative was correct. Participants were awarded one point when they gave the correct answer and no points when they gave an incorrect answer, or when they picked *I do not know*. Thus, they could score a maximum of 12 points on the prior knowledge test, which took approximately 5 min.

Learning materials. The learning materials consisted of one page with text and pictures concerning information transmission in the human nervous system, presented on the computer screen. The information transmission process showed how different neurotransmitters are released into the synaptic cleft, which either activate or inhibit information transfer. The text consisted of 261 words, divided over 21 numbered segments. In the integrated condition, the text segments were presented in close proximity of the relevant part of the picture (see Figure 7). In the two separated conditions, the text segments were presented above the picture, while the relevant parts of the picture were numbered in the same manner as the text segments (see Figures 8 and 9). We calculated the distance in pixels between the center of each text segment and the center of the associated picture segment (see Figures 7, 8, and 9; e.g., Text Box 1 and Picture Box 1). The average text-picture distance was 701 pixels (SD = 72.91) or 19.76 cm for the large-separation condition, 475.29 pixels (SD = 73.54) or 13.40 cm for the small-separation condition, and 150.48 pixels (SD = 49.35) or 4.24 cm for the integrated condition. The difference between the small-separation and large-separation condition was the largest difference possible on the computer screen used. As in the study of Florax and Ploetzner (2010), the learning materials were presented in a system-paced fashion and in all conditions participants had 18 min to study the materials.

Retention and comprehension tests. Knowledge was tested directly after the learning phase, by means of a paper-and-pencil multiple-choice test consisting of 30 questions. Twenty-two of these questions measured retention (i.e., What potential exists over the cell membrane of a cell that is not activated?) while eight questions required comprehension of the materials to be answered correctly (i.e., How would the potential ratio over the membrane be if a nonactivated cell would be permeable to potassium instead of sodium?). The retention questions required recall of the textual and pictorial information presented in the learning phase, while the comprehension questions required participants to make inferences based on this information. Both retention and comprehension questions had five possible answer alternatives; four of these alternatives could possibly be correct while the fifth alternative was always *I do not know*. Participants were encouraged not to guess, but to pick the fifth answer alternative when they were unsure which answer was correct. Participants were awarded one point when they gave the correct answer and no points when they gave the wrong answer, or when they picked the *I do not know* answer. Thus, they could score a maximum of 22 points on the retention questions, and eight points on the comprehension questions. Generally, participants took about 20 min to answer the retention and comprehension questions.

Invested mental effort. Participants were asked to indicate how much effort they invested in learning on a 9-point rating scale (Paas, 1992), ranging from 1 (*extremely low effort*) to 9 (*extremely high effort*). Moreover, participants were asked to

[3] These study times were logged by the eye tracker, from which it appeared that the recordings did not contain the full 18 min that the learning phase was programmed to last.

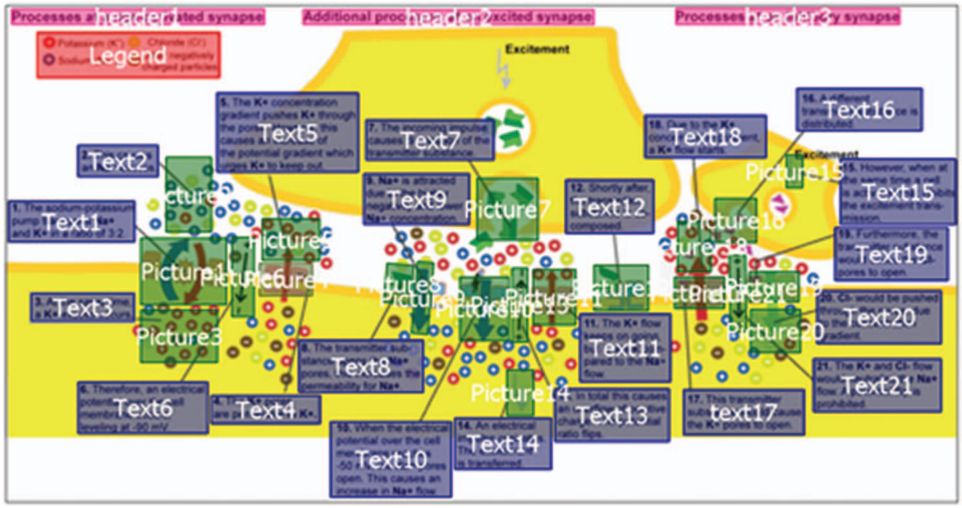

Figure 7. Learning material in the integrated condition with the area of interest's as an overlay. See the online article for the color version of this figure.

indicate how much effort they invested in answering the complete posttest (i.e., the retention and comprehension questions), using the same 9-point scale.

Apparatus. The materials were presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments), on a 22-in. monitor with a resolution of 1,680 × 1,050 pixels. Participants' eye movements were recorded using a SMI RED 250 Mobile eye tracker (SensoMotoric Instruments) that records binocularly at 250 Hz using SMI iView software (Version 2.8; SensoMotoric Instruments). The eye tracking data were analyzed using BeGaze software (Version 3.7; SensoMotoric Instruments).

Procedure. Participants were tested individually in a dedicated eye-tracking lab. First, they read the background information, after which the prior knowledge test was administered and participants were asked to provide their age and gender. Next, participants were seated in front of the computer monitor with their head positioned in a chin- and forehead rest. The distance to the monitor was approximately 60 cm. After a short introduction about the experiment, the eye tracker was calibrated using a 13-point calibration plus 4-point validation procedure, and participants were instructed to move as little as possible. Then, the learning phase started, for which participants were instructed to study the mate-

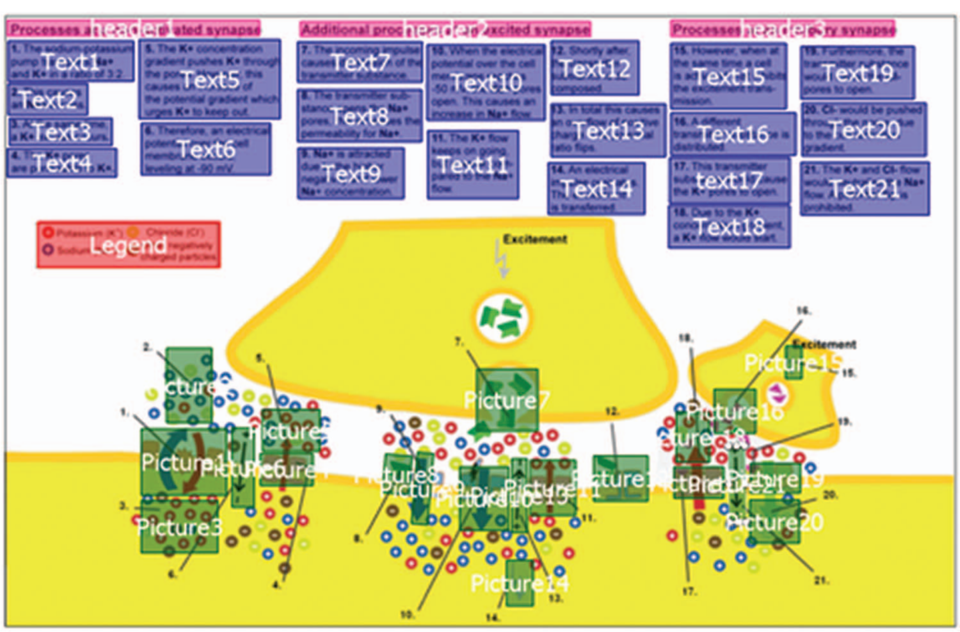

Figure 8. Learning material in the small-separation condition with the area of interest's as an overlay. See the online article for the color version of this figure.

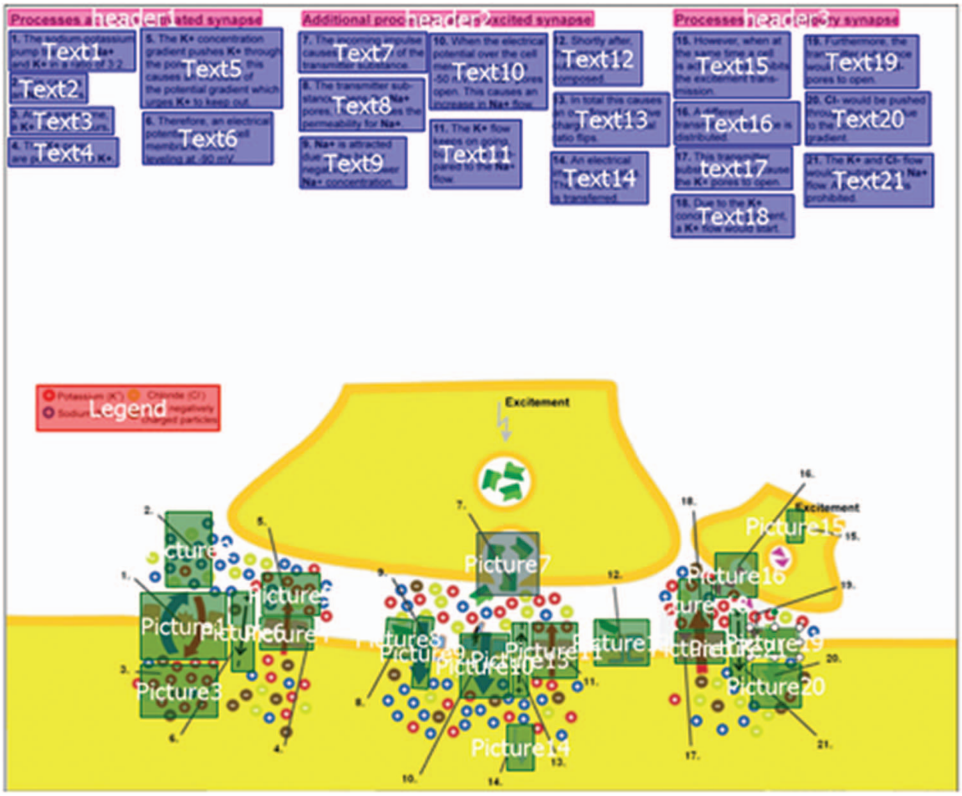

Figure 9. Learning material in the large-separation condition with the area of interest's as an overlay. See the online article for the color version of this figure.

rials to the best of their abilities, because afterward they would be tested on what they had just learnt. After the learning phase, participants indicated how much mental effort they invested during learning, and then completed the posttest. Finally, participants indicated how much mental effort they invested in answering the posttest. In total, the experiment took approximately 60 min.

Eye-tracking measures. For the eye tracking analyses, we first checked the accuracy of calibration. Based on this, five participants were excluded because of inaccurate calibration (i.e., deviation from the four validation points exceeded 1° visual angle), and three participants because the tracking ratio (i.e., the percentage of time for which the eye tracker actually measured the eye movements) was below 70%. This threshold was chosen a priori, as it leads to a high average tracking ratio, without much data loss and has been used before (e.g., Rop, Schüler, Verkoeijen, Scheiter, & Van Gog, 2018). For the remaining 64 participants, mean calibration accuracy was 0.48° visual angle (SD = 0.13°), while the average tracking ratio was 95.30% (SD = 4.62%). The participants were distributed over the conditions as follows: integrated (n = 21), small separation (n = 21), and large separation (n = 22).

For the eye tracking analyses, we defined fixations using a 40°/s velocity threshold and a minimal duration of 100 ms (cf. Holmqvist et al., 2011). We created an area of interest (AoI) for each segment of text (leading to 21 text AoIs), and for each corresponding relevant part of the picture (leading to 21 picture AoIs). The part of the screen not covered by an AoI was labeled "white space." The AoIs had exactly the same area-size across conditions; distance between text and picture AoIs was systematically varied according to our experimental conditions. To measure the amount of attention the text and picture attracted, we calculated the total fixation duration on the picture and the total fixation duration on the text by summing the fixation duration on each individual AoI (i.e., the fixation duration on the picture as reported in the results section is the grand total of the fixation duration on each of the 21 picture AoIs). To measure the text-picture integration attempts (i.e., the saccades between and within information sources), we used the number of transitions between the different AoIs. We defined three types of transitions: text–picture transitions, which are transitions between the text and the picture and vice versa; text–text transitions, which are transitions between two text blocks; and picture–picture transitions, which are transitions between two parts of the picture. We only counted the transitions between corresponding parts of the text and the picture (i.e., a transition from Text Block 1 into Picture Part 1, or vice versa), between consecutive text blocks (i.e., a transitions from Text Block 1 into Text Block 2, or vice versa), and between consecutive picture parts (i.e., a transitions from Picture Part 1 into Picture Part 2, or vice versa).

Results

All data were analyzed with one-way ANOVA's with Condition (small separation, large separation, or integrated) as between-subjects factor. We used partial eta-squared and Cohen's d as measures of effect size; both can be interpreted in terms of small ($\eta_p^2 \sim .01$, $d \sim 0.2$), medium ($\eta_p^2 \sim .06$, $d \sim 0.5$), and large ($\eta_p^2 \sim .14$, $d \sim 0.8$) effect sizes (Cohen, 1988). When post hoc follow-up tests were performed, we used a Bonferroni correction (i.e., multiplying the p value with the number of tests performed).

Prior knowledge. Performance on the prior knowledge test (see Table 3) did not differ significantly between conditions, $F(2, 69) = 0.48$, $p = .619$, $\eta_p^2 = .01$. Hence, conditions were considered similar in their knowledge about the topic before the learning phase.

Retention and comprehension performance. The means and standard deviations on the retention and comprehension questions for each of the three conditions are presented in Table 3. As can be seen in this table, for both the retention questions, $F(2, 69) = 0.39$, $p = .679$, $\eta_p^2 = .01$, and the comprehension questions, $F(2, 69) = 0.13$, $p = .876$, $\eta_p^2 < .01$, no significant difference on test performance was found between conditions. To provide an estimate of the evidence for this null-finding, we performed an additional Bayesian analysis with JASP (JASP Team, 2016, Version 0.8.4). Bayes Factors (BF) were computed while operating with default priors p(M) = 0.5 (Cauchy prior of h = .75; Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017). Jeffreys (1961) classifies Bayes factors 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. Note that Bayes values will be reported hereinafter next to standard statistical measures as an extra measure of effect strength as well as a measure of certainty to interpreting null-findings.

We obtained that there was substantial evidence for the absence of an effect for retention (BF null-model = 6.30), such that the data were 6.3 times more likely under the null-model as compared with the alternative model predicting an effect. We similarly obtained substantial evidence for the null-model for comprehension (BF null-model = 7.66), such that the observed data were 7.66 times more likely under the null-model as compared with the alternative model. To provide another estimate of how much evidence there is for a likely absence of an effect of condition, we have performed another JASP robustness analysis which provides an estimate of the Bayes factor's sensitivity to changing prior estimates. For this analysis, we compared the effect of the integrated condition versus the large-separation condition with a Bayesian t test and a concomitant robustness analysis. We contrast the integrated condition and the large-separation condition as this is the most likely contrast to detect the presence of an effect of split attention. Figure 10 shows that a prior width change will not likely render a different conclusion for the current dataset; even at a maximally constrained prior predicting an effect (H1) our data are not supportive of H1, although our evidence for the H0 does become less pronounced (from moderate evidence to anecdotal).

Invested mental effort. In Table 4, the means and standard deviations for learners' self-reported invested mental effort during the learning phase and the test phase are presented. The analyses on these scores revealed no significant differences in invested mental effort during the learning phase, $F(2, 69) = 1.75$, $p = .181$, BF null-model = 2.23 (anecdotal evidence), $\eta_p^2 = .05$, or the test phase, $F(2, 69) = 0.35$, $p = .709$, BF null-model = 6.51 (substantial evidence), $\eta_p^2 = .01$.

Transitions. The means and standard deviations for the number of transitions between the different AoIs are presented in Table 5. On the picture–text transitions (which measure the integration attempts between the text and the picture), the analysis revealed a large significant main effect of condition, $F(2, 61) = 60.55$, $p < .001$, BF alternative-model = 1.546×10^{12} (decisive evidence),[4] $\eta_p^2 = .67$. Follow-up tests showed that participants in the integrated condition made more picture–text transitions than participants in the small-separation condition, $p < .001$, $d = 2.24$, 95% CI [1.43, 2.96], and the large-separation condition, $p < .001$, $d = 3.10$, 95% CI [2.16, 3.92]. Participants in the two separated conditions did not significantly differ in their number of picture-text transitions, $p = .594$, $d = 0.54$, 95% CI [−0.08, 1.14]. On the text-text transitions (which measure the integration attempts within the text), we again found a large significant main effect of condition, $F(2, 61) = 32.09$ $p < .001$, BF alternative-model = 2.913×10^7 (decisive evidence; see Footnote 4), $\eta_p^2 = .51$. Follow-up analyses showed that participants in the integrated condition made fewer text-text transitions than participants in the small-separation condition, $p < .001$, $d = 1.57$, 95% CI [0.85, 1.25], and the large-separation condition, $p < .001$, $d = 2.94$, 95% CI [2.03, 3.74]. Moreover, participants in the small-separation condition made fewer text-text transitions than participants in the large-separation condition, $p = .040$, $d = 0.69$, 95% CI [0.06, 1.29]. Finally, on the picture-picture transitions (which measure the integration attempts within the picture), the analysis revealed no significant differences between conditions, $F(2, 64) = 0.66$, $p = .522$, BF null-model = 4.770 (substantial evidence), $\eta_p^2 = .02$.

Fixation duration. The total fixation duration on the text and picture AoIs for each of the three conditions is presented in Table 6. Please note that the fixation duration on the text and picture does not equal the 18 min that participants studied the materials. The remaining time was either spent fixating white space (which covered a considerable part of the learning materials, as we only labeled the most relevant parts of the picture and text as an AoI; see Figure 7, 8, and 9), was not fixated at one AoI long enough to be labeled a fixation, or was used to make saccades. Participants in all conditions allocated an equal amount of attention toward the text, $F(2, 61) = 0.60$, $p = .554$, BF null-model = 4.987 (substantial evidence), $\eta_p^2 = .02$. The amount of attention allocated toward the picture did not differ significantly

Table 3
Mean (and Standard Deviation) Performance With 95% Confidence Interval Around the Mean (in Brackets) on the Pretest (Max. = 8), Retention Test (Max. = 22), and Comprehension Test (Max. = 8) as a Function of Condition

Knowledge tests	Large separation	Small separation	Integrated
Pretest	4.74 (1.63)	4.20 (1.76)	4.42 (2.26)
	[4.03, 5.44]	[3.48, 4.92]	[3.46, 5.37]
Retention	11.48 (3.49)	12.48 (3.73)	12.04 (4.50)
	[9.97, 12.99]	[10.94, 14.02]	[10.14, 13.94]
Comprehension	4.00 (1.98)	3.88 (1.83)	3.71 (2.05)
	[3.14, 4.86]	[3.12, 4.64]	[2.84, 4.58]

[4] Here, we report the evidence in favor of the alternative model, which includes condition as a factor as the initial ANOVA revealed a large effect of condition.

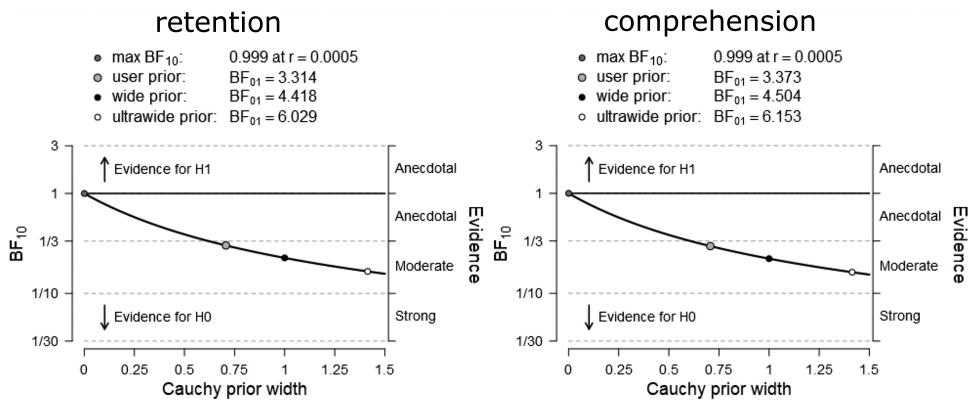

Figure 10. Prior width effects on Bayes factor (BF) estimation for retention and comprehension. Bayesian *T* tests for the integrated condition versus the large-separation condition. The figures (left panel DV = retention, right panel DV = comprehension), provide an estimate of the sensitivity of the Bayes factors as a function of Cauchy prior width changes. Higher (lower) widths indicate higher uncertainty (or higher certainty) of the effect size assuming the alternative hypothesis is true (H1). The gray dot indicates the default Cauchy prior width of .707, the red dot is the prior width where there is very high certainty that there is a presence of an effect. Figures were produced by JASP.

between the three conditions, $F(2, 64) = 2.97$, $p = .059$, BF null-model = 0.865 (no evidence), $\eta_p^2 = .09$.

Discussion

The present experiment examined whether an increase in spatial distance between text and picture leads to a stronger split-attention effect in a learning task. Moreover, we aimed to provide corroborating evidence for the finding by Florax and Ploetzner (2010) that spatial integration of a text and picture is not necessary to counteract the split-attention effect when the spatial distance is increased. The results show that spatially integrating text and picture is not a prerequisite to reduce split attention: We found no differences between the integrated and the small-separation condition on learning outcomes or cognitive load (Hypothesis 1). This replicates the result reported by Florax and Ploetzner (2010). Moreover, an increase in the spatial distance between text and picture did not seem to influence learning outcomes and cognitive load as we found no differences between the two separated (i.e., small vs. large) conditions (Hypothesis 2). Therefore, it seems that the results presented in Experiment 1 and 2 do not capitulate into cognitive benefits when learning from text and pictures.

So, we must conclude that the present results indicate that spatial distance does not influence the occurrence of the split-attention effect during multimedia learning in the present context. Importantly, one of the reviewers of the current paper suggested that it is still possible that there is an indirect effect of condition on learning outcomes via an indirect effect of condition on transitions. As such we have performed additional exploratory mediation analyses (output retrievable at https://osf.io/vx98u/) in order to ascertain whether the extent to which condition affected picture–text transitions was related to performance. We found no indication that there was a reliable indirect effect of condition on performance (retention or comprehension) via the number of picture–text transitions.

However, despite concluding an absence of an effect of condition on cognitive performance, spatially integrating the text and picture *does* promote text–picture integration at a behavioral level, as participants in the integrated condition made more text–picture transitions than participants in the two separated conditions. Unexpectedly, an increase in spatial distance between the two spatially separated information sources did not lead to fewer text–picture integration attempts (Hypothesis 3). This suggests that given a certain separation participants change information gather-

Table 4
Mean (and Standard Deviation) Invested Mental Effort (Max. = 9) With 95% Confidence Interval Around the Mean (in Brackets) During the Learning Phase and During the Test Phase as a Function of Condition

Experiment phase	Large separation	Small separation	Integrated
Learning	6.52 (.85)	6.76 (1.20)	6.17 (1.24)
	[6.16, 6.89]	[6.89, 7.00]	[5.64, 6.69]
Test	6.74 (1.42)	6.68 (1.18)	6.46 (1.06)
	[6.12, 7.35]	[6.19, 7.17]	[6.01, 6.91]

Table 5
Mean (and Standard Deviation) Number of Transitions With 95% Confidence Interval Around the Mean (in Brackets) Between the Text and Picture, Text and Text, and Picture and Picture as a Function of Condition

Transition type	Large separation	Small separation	Integrated
Text–Picture	10.27 (4.84)	15.91 (14.12)	54.29 (19.73)
	[8.13, 12.42]	[9.48, 22.33]	[45.30, 63.27]
Text–Text	153.41 (37.79)	121.90 (52.66)	56.00 (27.32)
	[136.65, 170.16]	[97.94, 145.87]	[43.56, 68.44]
Picture–Picture	34.91 (20.13)	38.19 (17.00)	31.81 (16.64)
	[25.98, 43.83]	[30.45, 45.93]	[25.98, 43.83]

Table 6
Mean (and Standard Deviation) Fixation Duration (in Seconds) With 95% Confidence Interval Around the Mean (in Brackets) on the Text and the Picture AoI's as a Function of Condition

AoI	Large separation	Small separation	Integrated
Text	397.31 (104.06)	398.90 (92.35)	428.22 (113.66)
	[351.17, 443.45]	[356.86, 440.94]	[376.48, 479.95]
Picture	118.26 (40.59)	152.52 (56.37)	125.57 (47.00)
	[100.03, 136.25]	[126.87, 178.18]	[104.17, 146.96]

Note. AoI = area of interest.

ing strategies (i.e., behavioral level), possibly indicating a nonlinear relationship between spatial separation and learning processes. That there is a drastic strategy shift is indicated by the large effect size of $d = 2.24$ for the small-separation condition and $d = 3.06$ for the large-separation condition, meaning that, on average, participants in the integrated condition undertake about four or five times as many integration attempts between the text and picture than participants in the separated conditions. Participants in the separated conditions primarily made transitions between different parts of the text, undertaking about two or three times as many integration attempts between the different parts of the text than participants in the integrated condition. These results also align with previous studies showing that learners mostly focus on the text in a split-attention format (e.g., Cromley et al., 2010; Hannus & Hyönä, 1999; Schmidt-Weigand et al., 2010). Regardless of this large effect of spatial distance on participant's processing behavior as reflected in their eye-movements, this observed increase in integrative transitions did not translate into better cognitive performance (i.e., learning outcomes) for these learning materials.

A further finding is that although participants in the separated conditions made more transitions within the text, they did not allocate more attention toward the text than participants in the integrated condition, as measured by the fixation duration (Hypothesis 4). It seems that all participants read all the text, and inspected all relevant parts of the picture. The only major difference elicited by the spatial integration of the two sources is more integration of the text and picture, and fewer integrations within the text. Possibly, this did not lead to differences in learning outcomes as participants in the separated conditions already made a reasonable amount of text–picture transitions, and further integration of text and picture was redundant for learning. Therefore, it seems that eliminating the visual search that is often required in a split format by signaling the corresponding parts of the text and picture is a robust way to avoid split-attention irrespective of the distance between the text and picture.

General Discussion

With the current experiments we probed the grounding of the split-attention effect in a more basic cognitive mechanism as predicted by cognitive load theory (CLT). We predicted that when information needs to be integrated but is spatially separated, participants will have to visually decouple for longer periods (depending on distance) from one information source so as to integrate it with the spatially distant second information source. Subsequently, an increase in spatial distance between the two sources was expected to impose higher demands on working memory as longer visual decoupling is required, which will impair learning processes. With three experiments we examined (a) whether an increase in spatial distance between two to-be-compared pictorial stimuli would increase working memory load and impair integration performance, (b) whether such an effect would be present, and perhaps be stronger with picture–text stimuli, and (c) whether these results would generalize to more complex multimedia learning materials. Results show that increasing the distance between two pictorial stimuli (i.e., a unimodal integration), hampers performance on a secondary visual-working memory task, while leaving integration speed of the visual integration task unaffected. However, when increasing the distance between picture–text stimuli (i.e., a cross-modal comparison) integration speed is reduced, but spatial distance does not affect performance on the secondary visual-working memory task. Finally, increasing the distance between text and pictures in a multimedia learning task influences learning processes as operationalized by eye-movements that reflect integration of information, but has no effect on perceived mental effort or learning outcomes.

Together, these results show that an effect of distance between two sources of information (either in a visual integration task, or in a learning task) exists, although this (a) is a small effect (Experiment 1 and 2); (b) mostly affects learning processes (Experiment 3); and (c) not always affects primary learning and problem solving outcomes (Experiment 1, 2, 3). Whether the increase in distance interferes with the learning process seems to depend on the type of integration that has to be performed, as witnessed by the differences between Experiment 1 (a pictorial–pictorial integration), and Experiment 2 (a pictorial–textual integration). The current results conceptually replicate prior research, showing that increasing the spatial distance between two information sources leads to different information gathering strategies (i.e., making more use of working memory; Ballard et al., 1995; Gray & Fu, 2004). We further extend these findings by showing that such a change in processing strategy also seems to occur in a learning context, although it does not directly influence performance on the primary learning task.

Another noteworthy finding is that in Experiment 3, fewer integration attempts did not translate into diminished learning. While previous research shows that a higher number of integrative transitions is indicative for better learning (e.g., Hannus & Hyönä, 1999; Johnson & Mayer, 2012; Mason et al., 2015, 2016), such a positive relation between transitions and learning outcomes is not always observed (Arndt, Schüler, & Scheiter, 2015; Scheiter & Eitel, 2015). Schüler (2017) surmises that, while studying a picture or reading the text, learners are able to retrieve previously seen information from memory as to mentally integrate the two sources without shifting one's gaze. As such, learners can successfully use "knowledge-in-the-head" as to replace "knowledge-in-the-world" (Gray & Fu, 2004). Of course, it is possible that with increased complexity of the learning materials this "knowledge-in-the-head" strategy will become less successful due to higher working memory demands. Future research should therefore probe whether an effect of spatial separation does translate into diminished learning outcomes when complexity of the learning task is increased. In the current study, complexity between Experiments 2 and 3 differed, but given that the nature of the tasks differed as well, it is difficult to draw any conclusions regarding the role of task complexity from these experiments.

Besides the visual search for referents in the text and picture, it has been argued that learners have to keep information active when studying spatially separated learning materials, which imposes working memory constraints. According to time-based resource sharing models of working memory, which have recently been introduced to cognitive load theory (e.g., Barrouillet, Bernardin, & Camos, 2004; Puma et al., 2018), reduced performance in split-attention materials reflects a time-related decay of the memory traces when attention is switched away from information elements. Therefore, increasing the spatial distance between information elements that need to be integrated in working memory is expected to increase the duration that the elements need to be activated in working memory and consequently to lead to more time-related decay of memory traces.

Together, the resources needed for visual search and information maintenance in working memory are assumed to lead to a high extraneous working memory load, and hamper learning (Paas & Sweller, 2014). The results of the present study show that at least spatial distance is important for the split-attention effect, but also that both spatial distance and visual search related processes are likely to underlie the occurrence of the split-attention effect, and these processes are not mutually exclusive. Indeed, an increase in distance (meaning that information has to be kept active for longer periods, while not manipulating searching processes), did not elicit a split-attention effect in learning performance in Experiment 3. However, in Experiment 1 and 2, when participants' working memory was taxed by a secondary visual-working memory task, an increase in spatial distance did lead to a split-attention effect, even though no visual search was required (because the to-be-compared stimuli were signaled by a yellow rectangle). Therefore, both searching related information in the text and picture, as well as keeping information active in working memory is likely to contribute to the split-attention effect.

Conclusion

In sum, current results indicate that increased cognitive load demands due to spatial separation of information is a viable underlying mechanism for the split-attention effect, supporting CLT (Sweller et al., 2011). As such, this study provides a more cognitively basic grounding of the split-attention effect which could help to counteract the negative effects on learning in the future. Yet, it is also clear from the results that spatial separation is likely not the only, nor a sufficient, condition for the "split-attention effect" to occur. Finally, with this study we hope to inspire further research that integrates basic cognitive research with more applied instructional design effects.

Context

This study was conceived and designed when Wim Pouw and Gertjan Rop discussed how their research backgrounds could be combined to strengthen the scientific basis for educational psychological assumptions. Wim Pouw's earlier work mostly concerns problem solving as informed by embedded/embodied approaches to cognition, while Gertjan Rop mostly works on instructional design principles based on cognitive load theory. Bjorn De Koning has an extensive background in instructional design and signaling effects, and Fred Paas is an authority on all these subjects. With this study, we wanted to combine our strengths and approach the split-attention effect from a more fundamental viewpoint, which is underrepresented in the current literature. This research fits well into all authors' respective research programs, and expands these programs by tying the fields of embedded/embodied cognition and instructional design together. At the moment, Bjorn De Koning, Fred Paas, and Gertjan Rop are continuing this line of research in their applied research, trying to shed more light on the learning detriments of spatial distance and cognitive integration on learning from text and pictures, while Wim Pouw is pursuing more fundamental topics in cognitive science.

References

Arndt, J., Schüler, A., & Scheiter, K. (2015). Text-picture integration: How delayed testing moderates recognition of pictorial information in multimedia learning. *Applied Cognitive Psychology, 29,* 702–712. http://dx.doi.org/10.1002/acp.3154

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 834–846. http://dx.doi.org/10.1037/0096-1523.26.2.834

Ayres, P., & Cierniak, G. (2012). Split-attention effect. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3172–3175). New York, NY: Springer U. S. http://dx.doi.org/10.1007/978-1-4419-1428-6_19

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

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417–423. http://dx.doi.org/10.1016/S1364-6613(00)01538-2

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66–80. http://dx.doi.org/10.1162/jocn.1995.7.1.66

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133,* 83–100. http://dx.doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., & Camos, V. (2007). The time-based resource-sharing model of working memory. In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59–80). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198570394.003.0004

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. http://dx.doi.org/10.1002/acp.2887

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualizations. *Learning and Instruction, 14,* 325–341. http://dx.doi.org/10.1016/j.learninstruc.2004.06.006

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1,* 1–20. http://dx.doi.org/10.5334/joc.10

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd rev. ed., pp. 174–206). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.010

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332. http://dx.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.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. http://dx.doi.org/10.1017/S0140525X01003922

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagram. *Contemporary Educational Psychology, 35,* 59–74. http://dx.doi.org/10.1016/j.cedpsych.2009.10.002

Dalmaijer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods, 46,* 913–921. http://dx.doi.org/10.3758/s13428-013-0422-2

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 visualizations: Innovations in research and practice* (pp. 255–278). Cham, Switzerland: Springer. http://dx.doi.org/10.1007/978-3-319-56204-9_11

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.

Della Sala, S., Gray, C., Baddeley, A., & Wilson, J. T. L. (1997). *Visual patterns test: A test of short-term visual recall*. Suffolk, UK: Thames Valley Test Company.

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. http://dx.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. http://dx.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.

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24,* 95–123. http://dx.doi.org/10.1006/ceps.1998.0987

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), 7. http://dx.doi.org/10.1167/8.1.7

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York, NY: Oxford University Press.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology, 23,* 1215–1226. http://dx.doi.org/10.1002/acp.1525

Inamdar, S., & Pomplun, M. (2003). Comparative search reveals the tradeoff between eye movements and working memory use in visual tasks. *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 25).

Jamet, E. (2014). An eye-tracking study of cueing effects in multimedia learning. *Computers in Human Behavior, 32,* 47–53. http://dx.doi.org/10.1016/j.chb.2013.11.013

JASP Team. (2018). *JASP*. (Version 0.9) [Computer software]. Retrieved from https://jasp-stats.org/

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York, NY: Oxford University 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. http://dx.doi.org/10.1037/a0026923

Lagers-van Haselen, G. C., van der Steen, J., & Frens, M. A. (2000). Copying strategies for patterns by children and adults. *Perceptual and Motor Skills, 91,* 603–615. http://dx.doi.org/10.2466/pms.2000.91.2.603

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of text and picture integration during reading: Effects on processing and learning. *Contemporary Educational Psychology, 14,* 172–187. http://dx.doi.org/10.1016/j.cedpsych.2015.01.004

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology, 47,* 1083–1095. http://dx.doi.org/10.1111/bjet.12271

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (2nd rev. ed.). New York, NY: Cambridge University Press.

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 rev. ed., pp. 279–315). New York, NY: Cambridge University Press.

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 and Development, 43,* 31–43. http://dx.doi.org/10.1007/BF02300480

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97. http://dx.doi.org/10.1037/h0043158

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. http://dx.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. http://dx.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. http://dx.doi.org/10.1037/0022-0663.84.4.429

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd rev. ed., pp. 27–425). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.004

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

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

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. http://dx.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. http://dx.doi.org/10.1007/s10648-018-9438-6

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2018). Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions. *Journal of Computer Assisted Learning, 34,* 458–470. http://dx.doi.org/10.1111/jcal.12287

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E. J. (2017). Bayesian analysis of factorial designs. *Psychological Methods, 22,* 304–321. http://dx.doi.org/10.1037/met0000057

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

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20,* 100–110. http://dx.doi.org/10.1016/j.learninstruc.2009.02.011

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd rev. ed., pp. 73–103). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.006

Schüler, A. (2017). Investigating gaze behaviour during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction, 49,* 218–231. http://dx.doi.org/10.1016/j.learninstruc.2017.03.001

Stoet, G. (2016). *Wisconsin card sorting task (WCST)*. Retrieved from http://www.psytoolkit.org/experiment-library/wcst.html

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4,* 295–312. http://dx.doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-8126-4

Sweller, J., Van Merrienboer, J. J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296. http://dx.doi.org/10.1023/A:1022193728205

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem-solving. *Journal of Educational Psychology, 80,* 424–436. http://dx.doi.org/10.1037/0022-0663.80.4.424

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd rev. ed., pp. 263–278). New York, NY: Cambridge University Press.

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

Received May 4, 2018
Revision received August 8, 2018
Accepted January 6, 2019 ■