

## RESEARCH ARTICLE

# The effect of layout and pacing on learning from diagrams with unnecessary text

Gertjan Rop<sup>1</sup>  | Anne Schüler<sup>2</sup> | Peter P.J.L. Verhoeven<sup>1,3</sup> | Katharina Scheiter<sup>2,4</sup> | Tamara Van Gog<sup>5,4</sup>

<sup>1</sup>Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands

<sup>2</sup>Leibniz-Institut für Wissensmedien, Tübingen, Germany

<sup>3</sup>Learning and Innovation Center, Avans University of Applied Sciences, Breda, The Netherlands

<sup>4</sup>University of Tübingen, Tübingen, Germany

<sup>5</sup>Department of Education, Utrecht University, Utrecht, The Netherlands

## Correspondence

Gertjan Rop, Department of Psychology, Education, and Child Studies, Erasmus University, Rotterdam, The Netherlands.  
Email: rop@essb.eur.nl

## Summary

Although the presentation of extraneous (i.e., irrelevant or unnecessary) information hinders learning, it is unclear whether and how layout and pacing influence this effect. In two experiments, participants learned how the heart functions using four different layouts: a diagram presented without unnecessary text (diagram only), with unnecessary text separated from the diagram (separated) or integrated into the diagram (integrated), or with separated unnecessary text and the instruction to integrate (integration instruction). In Experiment 1, study time was self-paced for half of the participants and system paced for the other half. There were no effects of layout and of pacing on learning, although system pacing was more effortful than self-pacing. In Experiment 2, which was system paced and employed eye tracking, the integrated condition showed worse learning outcomes than the separated condition. Moreover, in the integrated condition, participants made more integration attempts between the unnecessary text and the diagram than in the separated condition.

## KEYWORDS

cognitive load, coherence effect, eye tracking, multimedia learning, redundancy effect

## 1 | INTRODUCTION

The majority of contemporary instructional materials are multimedia materials, comprising a combination of text (written or spoken) and pictures (e.g., diagrams, illustrations, graphs, or animations). Learning from multimedia has been widely investigated in research inspired by the cognitive theory of multimedia learning (Mayer, 2014) and cognitive load theory (Sweller, Ayres, & Kalyuga, 2011). This research has shown that when a mutually referring text and picture are unintelligible in isolation, the combination of text and picture tends to improve learning (the multimedia effect: Butcher, 2014; Mayer, Bove, Bryman, Mars, & Tapangco, 1996). However, when either text or picture is intelligible in isolation, their combination might not help and could even hamper learning (the redundancy effect: Kalyuga & Sweller,

2014; the coherence effect: Mayer & Fiorella, 2014). Accordingly, the redundancy and coherence principles of multimedia learning state that presentation of extraneous information should be avoided, because it hampers learning compared with instructional materials in which this information is eliminated.

### 1.1 | Why does extraneous information hamper learning?

The negative effect of extraneous information on learning can be explained in terms of working memory load. Working memory is limited in both capacity and duration (e.g., Baddeley, 2000; Barrouillet & Camos, 2007; Cowan, 2001). Regarding the limited capacity, Cowan (2001) suggests that our memory span is limited to around four chunks, where a

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2018 The Authors Applied Cognitive Psychology Published by John Wiley & Sons Ltd

chunk is a collection of items that are remembered together. Regarding the limited duration, Barrouillet and Camos (2007) propose that working memory resources have to be shared between maintenance of “old” information (prior knowledge from long-term memory or previously processed information during task performance) and processing of new, incoming information. The higher the number of old information elements that have to be maintained and the faster new elements need to be processed, the higher the working memory load. Therefore, a task that is cognitively undemanding when time is unlimited and can become very demanding when time is limited.

Learning (i.e., schema construction/elaboration/automation in long-term memory; Sweller, 1994) requires that old information is maintained in working memory and successfully integrated with the new essential information that the learner selects (by attending to it) from the learning materials. The addition of extraneous information negatively affects the selection and organization of essential information. When extraneous information is present, learners have to spend cognitive capacity on selecting the most relevant information, a process that is absent when no extraneous information is present. In turn, learners have more trouble organizing the essential information into a coherent mental model or schema. A disruption of these processes early in the learning process can cascade, also hampering the integration of the essential information, as a greater portion of learners' time and cognitive resources have to be expended on these early processes (cf. McCray & Brunfaut, 2018). When (due to hampered selection and organization) learners have less time available to integrate, the essential information schema construction is hampered.

The negative effects of extraneous information processing on learning have been shown with a wide variety of materials (for reviews, see Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). A seminal study regarding the negative effect of extraneous information on learning was performed by Chandler and Sweller (1991). They studied the relationship between extraneous text added to diagrams and learning outcomes, in a series of six experiments. In all experiments, study time was self-paced. When the text contained essential information not conveyed in the diagram (Experiment 1 and 6), participants performed better when the text was integrated in the diagram (integrated condition) than when it was presented spatially separated (separated condition; this is evidence of the “split-attention effect”; see Ayres & Sweller, 2014). In Experiment 2, the text was unnecessary for learning, as the information it described could also be inferred from the diagram. In this case, participants showed better learning outcomes in the separated condition than in the integrated condition, presumably because the unnecessary information was harder to ignore in the integrated condition. In Experiment 3, all participants learned with the separated learning materials used in Experiment 2, but half of the participants were instructed to study the diagram, whereas the other half was instructed to read and integrate the textual information with the diagram (integration-instruction condition). Learning outcomes were lower in the latter group, presumably because participants in the integration-instruction condition spent cognitive resources on processing the unnecessary text. In Experiment 4 and 5, participants learned in one of three conditions: diagram only, integrated, and integration instruction. Learning outcomes were significantly higher in the diagram-only condition than in the integrated and

integration-instruction conditions. Although the integrated condition performed numerically better than the integration-instruction condition, this difference was not significant.

## 1.2 | The effect of layout and pacing on extraneous processing

These results show that the presentation of unnecessary information alongside a diagram hampers learning in a self-paced learning environment. This effect seems to be stronger when the unnecessary information comes integrated with the essential information or when learners are instructed to integrate it, than when the information is presented spatially separated. An interesting question, therefore, is whether students in the separated condition would be able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared with a diagram-only condition. This question cannot be answered on the basis of the Chandler and Sweller study because they did not make a direct comparison of all four (i.e., diagram only, separated, integrated, integration instruction) conditions: Experiments 2 and 3 did not include a diagram-only condition, and Experiments 4 and 5 did not include a separated condition without the integration instruction.

This question is relevant, because recent research suggests that students may learn to ignore extraneous information once they have gained some task experience (Rop, Van Wermeskerken, De Nooijer, Verkoeijen, & Van Gog, 2018; Rop, Verkoeijen, & Van Gog, 2017). However, in these studies by Rop and colleagues, the extraneous information was pictorial and mismatched the verbal information that participants had to remember. In the materials from Chandler and Sweller's (1991) study, the extraneous information is textual, which may be harder to ignore as learners often focus more quickly and more strongly on text than on the associated pictures (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010). Moreover, in the Chandler and Sweller study, the extraneous information was unnecessary, in the sense that the information provided by the text could also be inferred from the diagram, but it was not irrelevant or mismatching (i.e., it did have a relation with the picture). As such, it may be harder for learners to come to realize that they can ignore the text. Therefore, the present study addresses the question of whether learners would be able to ignore unnecessary textual information when it is presented spatially separated from the diagram, by building on the Chandler and Sweller study and comparing a diagram only, separated layout, integrated layout, and separated layout with integration-instruction condition. When the layout of the instructional materials makes it harder for learners to ignore unnecessary information, it is more likely that learners will attend to that information, affecting the selection and organization of essential information more than when it is easier to ignore the unnecessary information. In turn, these processes might hamper integration of the essential information (cf. McCray & Brunfaut, 2018).

In addition, we study the effects of pacing in the present study. In the Chandler and Sweller (1991) study, learning was self-paced. Consequently, the time that participants spent on processing the unnecessary text did not have to go at the expense of time spent on processing the essential information from the diagram. When learning is user paced, participants can invest more time to select and organize both the

unnecessary and essential information. In contrast, when learning is system paced, the selection and organization of unnecessary information takes up some of the limited time, leaving less room for the selection and organization of essential information (cf. Barrouillet & Camos, 2007). Thus, it is more likely that learning is hampered by unnecessary information when learning is system paced than when it is user paced. Kalyuga, Chandler, and Sweller (2004; see also Mayer & Jackson, 2005) provide some empirical evidence for this hypothesis. In two experiments, one system paced and one self-paced, participants had to learn how to use a cutting speed nomogram with or without unnecessary text. Although both experiments showed a negative effect of the unnecessary text, the effect sizes were larger in the system-paced experiment. However, this study did not directly manipulate learning in a self-paced environment versus learning in a system-paced environment within one experiment, so a causal relationship between pacing and the negative effect of unnecessary text on learning could not be established.

### 1.3 | The present study

The present study aimed to replicate and extend the results from the seminal study of Chandler and Sweller (1991) by making a direct comparison of all four conditions (i.e., diagram only, separated, integrated, and integration instruction) and by investigating the effects of self-pacing versus system pacing, using (an adapted version of) their materials about the human circulatory system (Experiment 5). In two experiments, participants learned using a self-contained diagram of the heart, which was accompanied by unnecessary textual explanations. The text was unnecessary for learning because it described processes that could also be inferred from the diagram, which, according to the findings by Chandler and Sweller (1991), could be understood without references to the text. Therefore, we concluded that the addition of this unnecessary text was the crucial difference between conditions that should lead to the observed effect (i.e., the primary information focus; Schmidt, 2009). In both experiments, four conditions were compared: The unnecessary text was not presented at all (diagram-only condition), or was presented separated from the diagram (separated condition), integrated in the diagram (integrated condition), or separated from the diagram with the instruction to integrate it (integration-instruction condition).

In Experiment 1, half of the participants learned at their own pace (self-paced conditions), whereas learning was time constrained for the other half of the participants (system-paced conditions). Learning outcomes were assessed using three retention tests (out of the six tests used by Chandler & Sweller, 1991). With regard to learning outcomes, we expected to replicate Chandler and Sweller's overall findings (i.e., 1: diagram only > integrated and integration instruction; 2: separated > integrated and integration instruction) and sought to answer the open question of whether students are able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared with a diagram-only condition (i.e., diagram only = separated > integrated and integration instruction), or whether the presentation of unnecessary text should be avoided entirely as it does hamper learning even when presented spatially separated (i.e., diagram only > separated > integrated and integration instruction). In addition, we hypothesized that system pacing would aggravate the negative effects of unnecessary text on learning particularly

in the integrated condition and the integration-instruction condition because these conditions are assumed to impose the highest cognitive load on the learner. Experienced cognitive load was measured by having participants rate their invested mental effort (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). With regard to cognitive load experienced while studying, we expected the reverse of the learning outcomes pattern (i.e., diagram only  $\leq$  separated < integrated and integration instruction) and pacing (i.e., integrated and integration-instruction conditions: self-paced < system-paced).

In Experiment 2, we employed eye tracking to investigate how the presentation of unnecessary text affects attention allocation under system-paced conditions. This allowed us to directly measure whether or not learners paid less attention to the unnecessary text when it was spatially separated from the diagram compared with integrated in the diagram and compared with when they had been given integration instructions.

## 2 | EXPERIMENT 1

### 2.1 | Method

#### 2.1.1 | Participants

Participants ( $N = 302$ ) were recruited via Amazon's Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010) and were paid 0.60 dollar for their participation (which took about 6 to 7 min). A priori we decided that participants would be excluded when they used a device with a small screen and could not see the diagram and explanations sufficiently (i.e., could not see at least six lines of explanations) without scrolling ( $n = 61$ ), when they participated in the experiment twice ( $n = 5$ ), or when they participated in a noisy environment (i.e., self-reported noise of seven or higher on a scale of one to nine;  $n = 5$ ). Furthermore, we decided post hoc to exclude participants who experienced technical problems (in the system-paced condition, for some participants, the diagram was presented for longer than the predefined presentation time;  $n = 3$ ), who did not follow orders ( $n = 2$ ), and who reported using memory influencing drugs ( $n = 1$ ). Thus, our final sample comprised 225 participants ( $M_{\text{age}} = 37.40$  years,  $SD = 12.97$ , range = 18–87; 108 females).

#### 2.1.2 | Design

Participants were randomly assigned to one of the eight conditions resulting from a  $2 \times 4$  factorial design with between-subjects factors "pacing" (self-paced vs. system paced) and "layout" (diagram only, separated, integrated, and integration instruction): system-paced diagram only ( $n = 35$ ), system-paced separated ( $n = 21$ ), system-paced integrated ( $n = 28$ ), system-paced integration-instruction condition ( $n = 26$ ), and self-paced diagram only ( $n = 34$ ), self-paced separated ( $n = 37$ ), self-paced integrated ( $n = 24$ ), and self-paced integration instruction ( $n = 30$ ).<sup>1</sup>

#### 2.1.3 | Materials

The materials were programmed and presented in Qualtrics software (Qualtrics, Provo, UT).

<sup>1</sup>This is the  $n$  after exclusion (hence the unequal sample sizes); see Section 2.1.1.

### Subjective and objective prior knowledge measures

In order to measure participants' prior knowledge on this topic, a subjective and an objective measure were used. The subjective measure asked participants to indicate how much they knew about the blood flow in our body, ranging from 1 (*nothing at all*) to 5 (*a great deal*). The objective measure asked participants to name the four major components of the heart and the two arteries through which blood exits the heart.

### Learning material

The learning material consisted of a simplified diagram of the circulatory system, which was adapted from Chandler and Sweller (1991; Experiment 5) with slight modifications. We vectored the image, added color (blue for the deoxygenated blood and red for the oxygenated blood), changed the font size and type, and increased the resolution of the image. The diagram was either the sole source of information on the screen (diagram-only condition; see Figure 1), or it was accompanied by unnecessary textual explanations separated from the diagram (separated and integration-instruction conditions; see Figure 1), or it was accompanied by the same explanations integrated in the diagram (integrated condition; see Figure 1). The unnecessary text consisted of 96 words. In the diagram-only, integrated, and separated conditions, participants were instructed as follows: "Please do your best studying the diagram, because afterwards you will be tested on your knowledge of the blood flow in the heart and body." The instruction for the participants in the integration-instruction condition was as follows: "Please try to integrate the diagram and text as much as possible, because afterwards you will be tested on your knowledge of the blood flow in the heart and body."

Participants in the self-paced conditions could spend as much time studying the learning material as they wished, although their time on task was measured. Participants in the system-paced conditions had 80 s to study the learning material, which was ~10 s more than

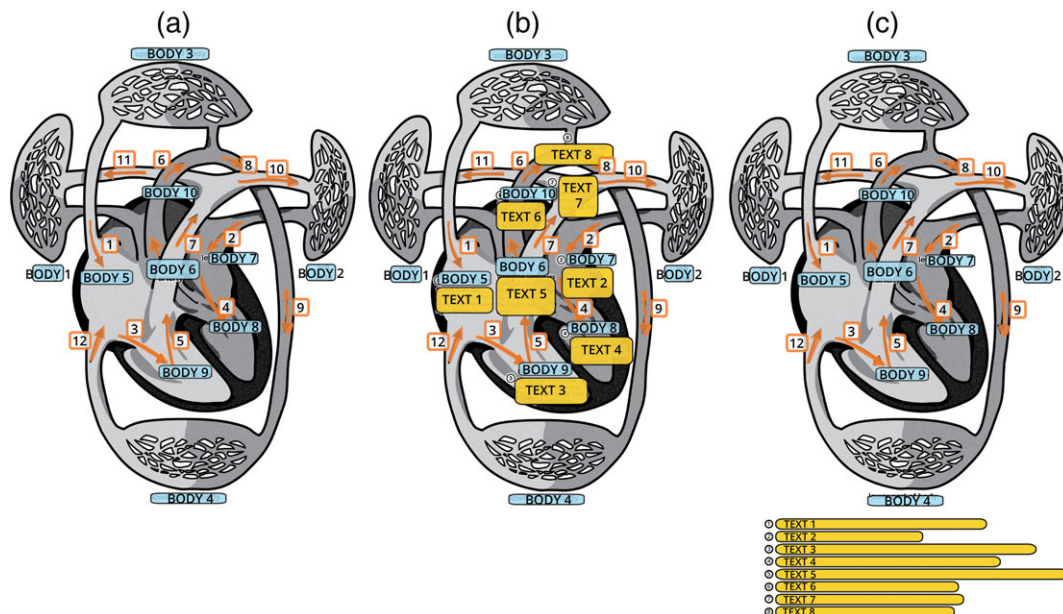
the average time spent by participants in the diagram-only condition in the Chandler and Sweller study (1991; i.e., 69.1 s). Participants were informed about this time limit.

### Posttest

The posttest assessed retention of the learning materials using three of the six outcome measures used by Chandler and Sweller (1991). These were the measures most suitable for online, computer-based testing, as the other three outcome measures required physical manipulation of the diagram (e.g., placing arrowheads to indicate the correct flow of the blood). In the Chandler and Sweller study, the first two outcome measures showed a significant advantage of diagram-only presentation, whereas this advantage was marginally significant for the third. First, participants were asked to name the four major components of the heart and the two arteries through which blood exits the heart (components test). This was a single question for which participants had to give up to six answers. Then they were asked to complete two blood flow chains presented on separate slides (blood chains test): Participants were told that the second component was either the left atrium or the right atrium (e.g., \_\_, left atrium, \_\_, \_\_, \_\_) and had to complete the first, third, fourth, and fifth component. Finally, participants were given five fill-in-the-gap questions, all presented on separate slides, in the form of "blood in the left ventricle flows to the \_\_" (blood flow test). The five questions regarded the left atrium, right ventricle, left ventricle, pulmonary artery, and aorta. The three outcome measures were always presented in the same order, although the order of the questions within each measure was randomized. The posttest was self-paced, and participants could not go back to previous questions to adjust their answers.

### Mental effort

Participants were asked to indicate how much effort they invested in studying the learning material and in answering the posttest questions



**FIGURE 1** Learning materials in Experiment 1 for the diagram-only condition (a), the integrated condition (b), and the separated and integration-instruction conditions (c) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

on a 9-point rating scale (Paas, 1992), ranging from 1 (*extremely low effort*) to 9 (*extremely high effort*). Mental effort is an indicator of experienced cognitive load (see Paas et al., 2003).

### Control questions and demographic questionnaire

To obtain information on the circumstances under which the experiment was completed, participants were asked some control questions after the posttest. Participants in the separated and integration-instruction conditions were asked whether they could see the entire diagram and all explanations without scrolling on a scale of 1 (*yes*) to 9 (*no, I could not see any of the explanations*). Participants in the diagram-only and integrated conditions were asked whether they could see the entire diagram without scrolling with a yes or no question. All participants had to self-report on the noise in their environment on a scale ranging from 1 (*quiet and no distractions*) to 9 (*noise and many distractions*). Finally, participants were asked to provide their age, gender, and highest achieved education level.

### 2.1.4 | Procedure

Participants first received a short introduction about the experiment and instructions on what they had to do. Then they filled out the subjective prior knowledge questions and completed the objective prior knowledge test, after which they studied the learning material. After the learning phase, participants were asked to indicate how much mental effort they had invested in studying the learning material. Then participants completed the posttest and reported how much mental effort they invested in answering the posttest questions. Finally, they answered the control and demographic questions

### 2.1.5 | Scoring

For all measures, participants were awarded one point per correct response. When they gave a partially correct answer (e.g., pulmonary instead of pulmonary artery), they received 0.5 points. If they did not provide an answer, or if it was completely wrong, zero points were awarded. Thus, participants could get a maximum of six points for the component test, eight points for the blood chains test (four for the left and four for the right chain), and five points for the blood flow test. A random subset of the data (10.67%) was scored by a second rater, and interrater reliability was high ( $\kappa = .90$ ).

## 2.2 | Results

In this study, we maintained an alpha level of 0.05. For the parametric tests, effect size measures used were partial eta-squared ( $\eta_p^2$ ) and Cohen's *d*. 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).

### 2.2.1 | Prior knowledge

To check whether conditions did not differ on prior knowledge (data presented in Table 1), we performed a 2 (pacing)  $\times$  4 (layout) analysis of variance (ANOVA) on the self-reported prior knowledge, which showed a main effect of pacing,  $F(1, 217) = 7.92$ ,  $p = .005$ ,  $\eta_p^2 = .04$ , indicating that participants in the self-paced conditions ( $M = 2.46$ ,  $SD = 0.83$ ) rated their domain knowledge higher than participants in the system-paced conditions ( $M = 2.16$ ,  $SD = 0.75$ ). There was no effect of layout and no interaction ( $F_s < 1$ ). The objective prior knowledge test did not confirm higher prior knowledge in the self-paced conditions though. Because the assumptions of normality were violated for the objective domain knowledge test, we performed non-parametric tests on these data. A Mann-Whitney test indicated no significant differences between the self-paced and system-paced conditions,  $U = 5552.50$ ,  $p = .111$ . Kruskal-Wallis tests showed no significant differences between layout conditions when the presentation was self-paced,  $\chi^2(3) = 2.52$ ,  $p = .472$ , or system paced,  $\chi^2(3) = 1.18$ ,  $p = .759$ .

### 2.2.2 | Time on task

We checked how much time participants in the self-paced conditions spent on the learning materials: On average, participants in the self-paced conditions studied the materials for 100.77 s ( $SD = 106.77$ ,  $Mdn = 67.83$ ). Mean time on task was highest in the diagram-only condition (129.25 s,  $SD = 147.69$ ,  $Mdn = 71.25$ ), followed by the integrated condition (104.06 s,  $SD = 107.62$ ,  $Mdn = 68.24$ ), the integration-instruction condition (98.49 s,  $SD = 69.87$ ,  $Mdn = 84.43$ ), and the separated condition (64.42 s,  $SD = 64.42$ ,  $Mdn = 42.27$ ). Thus, except for the separated condition, the average time on task was higher in the self-paced than in the system-paced (i.e., 80 s) conditions.

**TABLE 1** Mean (*SD*) self-estimated prior knowledge (range = 1–5) and mean (*SD*) and median (range) performance on the objective domain prior knowledge test (max = 6) as a function of pacing and layout in Experiment 1

|              |                         | Self-estimated<br>Mean ( <i>SD</i> ) | Objective<br>Mean ( <i>SD</i> ) | Median (range) |
|--------------|-------------------------|--------------------------------------|---------------------------------|----------------|
| Self-paced   | Diagram only            | 2.56 (0.93)                          | 2.59 (2.16)                     | 2.00 (6.00)    |
|              | Separated               | 2.59 (0.89)                          | 2.09 (2.14)                     | 1.50 (6.00)    |
|              | Integrated              | 2.42 (0.65)                          | 2.67 (2.07)                     | 2.50 (6.00)    |
|              | Integration instruction | 2.27 (0.78)                          | 2.70 (1.61)                     | 3.00 (6.00)    |
|              | Total                   | 2.46 (0.83)                          | 2.52 (1.99)                     | 2.00 (6.00)    |
| System paced | Diagram only            | 2.20 (0.80)                          | 2.44 (2.20)                     | 1.50 (6.00)    |
|              | Separated               | 2.10 (0.77)                          | 2.02 (1.41)                     | 2.00 (4.00)    |
|              | Integrated              | 2.21 (0.74)                          | 1.82 (1.61)                     | 1.50 (5.00)    |
|              | Integration instruction | 2.12 (0.71)                          | 1.90 (1.79)                     | 1.00 (6.00)    |
|              | Total                   | 2.16 (0.75)                          | 2.07 (1.81)                     | 1.50 (6.00)    |



### 2.2.3 | Learning outcomes

For all learning outcome measures (see Table 2), the assumptions of normality were violated. Therefore, we performed nonparametric tests on these data. On the components test, a Mann-Whitney test showed no significant differences between the self-paced and system-paced conditions,  $U = 6069.00$ ,  $p = .574$ . Kruskal-Wallis tests showed no significant differences between layout conditions when the presentation was either self-paced,  $\chi^2(3) = 0.47$ ,  $p = .926$ , or system-paced,  $\chi^2(3) = 4.40$ ,  $p = .221$ .

On the blood chains test, the self-paced conditions seemed to outperform the system-paced conditions, but a Mann-Whitney test showed that this difference was not significant,  $U = 5471.50$ ,  $p = .075$ . Kruskal-Wallis tests indicated no performance differences between layout conditions when the presentation was either self-paced,  $\chi^2(3) = 3.65$ ,  $p = .302$ , or system-paced,  $\chi^2(3) = 4.13$ ,  $p = .248$ .

Finally, on the blood flow test, the Mann-Whitney test showed a marginally significant advantage for the self-paced conditions compared with the system-paced conditions,  $U = 5393.00$ ,  $p = .054$ . Again, however, Kruskal-Wallis tests indicated no performance differences between layout conditions when the presentation was either self-paced,  $\chi^2(3) = 1.42$ ,  $p = .702$ , or system-paced,  $\chi^2(3) = 2.27$ ,  $p = .519$ .

### 2.2.4 | Invested mental effort

The data regarding invested mental effort can be found in Table 3. We performed a 2 (pacing)  $\times$  4 (layout) ANOVA on the invested mental effort during learning, which showed a main effect of pacing,  $F(1, 217) = 5.37$ ,  $p = .021$ ,  $\eta_p^2 = .02$ , indicating that participants in the system-paced conditions invested more mental effort during learning than participants in the self-paced conditions. We found no effect of layout,  $F(3, 217) = 2.04$ ,  $p = .109$ ,  $\eta_p^2 = .03$ , and no interaction,  $F(3, 217) = 1.11$ ,  $p = .347$ ,  $\eta_p^2 = .02$ . The same ANOVA on mental effort invested during the posttest indicated no effect of pacing,  $F(1, 217) = 1.20$ ,  $p = .275$ ,  $\eta_p^2 = .01$ , layout,  $F < 1$ , and no interaction  $F < 1$ .

## 2.3 | Discussion

In contrast to our hypotheses regarding layout, we did not replicate Chandler and Sweller's (1991) findings that learning outcomes would

**TABLE 3** Mean (SD) invested mental effort (max. = 9) during the learning phase and posttest as a function of pacing and layout in Experiment 1

|              |                         | Learning phase | Posttest    |
|--------------|-------------------------|----------------|-------------|
| Self-paced   | Diagram only            | 7.50 (1.31)    | 7.76 (1.28) |
|              | Separated               | 6.67 (1.82)    | 7.41 (1.37) |
|              | Integrated              | 6.71 (1.85)    | 7.25 (1.92) |
|              | Integration instruction | 7.27 (1.08)    | 7.30 (1.37) |
|              | Total                   | 7.08 (1.54)    | 7.45 (1.47) |
| System paced | Diagram only            | 7.49 (0.95)    | 7.63 (1.09) |
|              | Separated               | 7.19 (1.12)    | 7.43 (1.29) |
|              | Integrated              | 7.57 (1.14)    | 7.75 (1.29) |
|              | Integration instruction | 7.58 (1.36)    | 7.73 (1.46) |
|              | Total                   | 7.47 (1.13)    | 7.65 (1.26) |

be significantly lower in the integrated and integration-instruction conditions than in the diagram-only condition or the separated condition. With regard to our question of whether the separated condition would perform as well or worse than the diagram-only condition, we found no difference among those conditions. However, given the failure to replicate their respective superiority over the integrated and integration-instruction conditions, that finding cannot be meaningfully interpreted.

Regarding pacing, we found that self-paced learning was less effortful than system-paced learning. In other words, in line with our hypothesis, system pacing did indeed increase perceived cognitive load. Moreover, self-pacing led to a numerical (but not statistically significant,  $p = .075$ ) advantage in performance on the blood chains test and a marginally significant ( $p = .054$ ) advantage in performance on the blood flow test compared with system pacing. However, given the lack of differences among layout conditions in either self-paced or system-paced learning, we found no evidence that system pacing would aggravate the negative effects of presenting unnecessary text on learning.

We see two potential explanations for our failure to replicate Chandler and Sweller's (1991) findings. First, to control for possible differences in prior knowledge, we added an objective prior knowledge test, which Chandler and Sweller (1991) did not have in their experiment. This test might have guided participants' learning, which could have neutralized any effects of the unnecessary text. Indeed,

**TABLE 2** Mean (SD) and median (range) performance on the components (max. = 6), blood chains (max. = 8), and blood flow (max. = 5) tests as a function of pacing and layout in Experiment 1

|              |                         | Components  |                | Blood chains |                | Blood flow  |                |
|--------------|-------------------------|-------------|----------------|--------------|----------------|-------------|----------------|
|              |                         | Mean (SD)   | Median (Range) | Mean (SD)    | Median (Range) | Mean (SD)   | Median (Range) |
| Self-paced   | Diagram only            | 4.81 (1.70) | 5.75 (6.00)    | 2.32 (2.40)  | 2.00 (8.00)    | 2.65 (1.58) | 3.00 (5.00)    |
|              | Separated               | 5.00 (1.59) | 6.00 (5.00)    | 2.41 (2.66)  | 1.50 (8.00)    | 3.02 (1.74) | 3.50 (5.00)    |
|              | Integrated              | 5.02 (1.28) | 5.75 (4.00)    | 3.54 (2.68)  | 3.00 (8.00)    | 2.92 (1.74) | 3.00 (5.00)    |
|              | Integration instruction | 4.87 (1.68) | 6.00 (5.00)    | 3.07 (2.98)  | 2.00 (8.00)    | 2.58 (1.81) | 2.75 (5.00)    |
|              | Total                   | 4.91 (1.57) | 6.00 (6.00)    | 2.79 (2.69)  | 2.00 (8.00)    | 2.77 (1.70) | 3.00 (5.00)    |
| System paced | Diagram only            | 5.01 (1.29) | 5.50 (5.50)    | 1.56 (1.92)  | 1.00 (6.50)    | 2.47 (1.70) | 2.00 (5.00)    |
|              | Separated               | 4.67 (1.64) | 5.00 (6.00)    | 1.86 (1.80)  | 2.00 (7.00)    | 1.90 (1.68) | 1.50 (5.00)    |
|              | Integrated              | 4.95 (1.13) | 5.00 (3.50)    | 2.34 (2.26)  | 1.50 (7.00)    | 2.36 (1.31) | 2.00 (5.00)    |
|              | Integration instruction | 5.35 (1.21) | 6.00 (4.00)    | 2.60 (2.44)  | 2.00 (8.00)    | 2.50 (1.57) | 2.75 (5.00)    |
|              | Total                   | 5.01 (1.31) | 5.50 (6.00)    | 2.06 (2.13)  | 1.00 (8.00)    | 2.34 (1.57) | 2.00 (5.00)    |

research into test-potentiated learning shows that attempting to retrieve information may improve later encoding, even when the initial retrieval attempt was unsuccessful (Arnold & McDermott, 2013; Little & Bjork, 2016). We cannot rule out that this could have led to higher learning outcomes in the conditions with unnecessary text, in which case we should see negative effects of unnecessary text on learning when not giving a pretest. Second, we ran our experiment on Mechanical Turk, and although Mechanical Turk can yield results similar to lab studies (Paolacci et al., 2010), this type of online testing reduces experimental control somewhat. For instance, the time on task data from the self-paced conditions showed substantial variance, and we cannot be sure that participants spent all this time studying the learning materials (i.e., they might have been interrupted or attending to another task). However, although this may at least partly explain the large variance in learning outcomes, it does not explain the absence of mean/median condition differences.

### 3 | EXPERIMENT 2

In Experiment 2, we compared the four system-paced conditions on learning outcomes and perceived cognitive load in a more controlled lab setting and excluded the objective prior knowledge test for the reason mentioned in Section 2.3. We employed eye tracking to directly measure to what extent students processed the unnecessary text and to address our assumption that participants in the separated condition would attend less to the unnecessary text than participants in the integrated and the integration-instruction conditions. This should result in shorter fixation duration on the unnecessary text and less transitions between unnecessary text and the relevant parts of the diagram. With regard to processing the relevant information, that is, the arrows and labels in the diagram, we would expect that participants in the diagram-only condition would pay most attention to this information, followed by the separated condition and the integrated and integration-instruction conditions, with the question being whether the diagram-only and separated condition would differ from each other.

## 3.1 | Method

### 3.1.1 | Participants and design

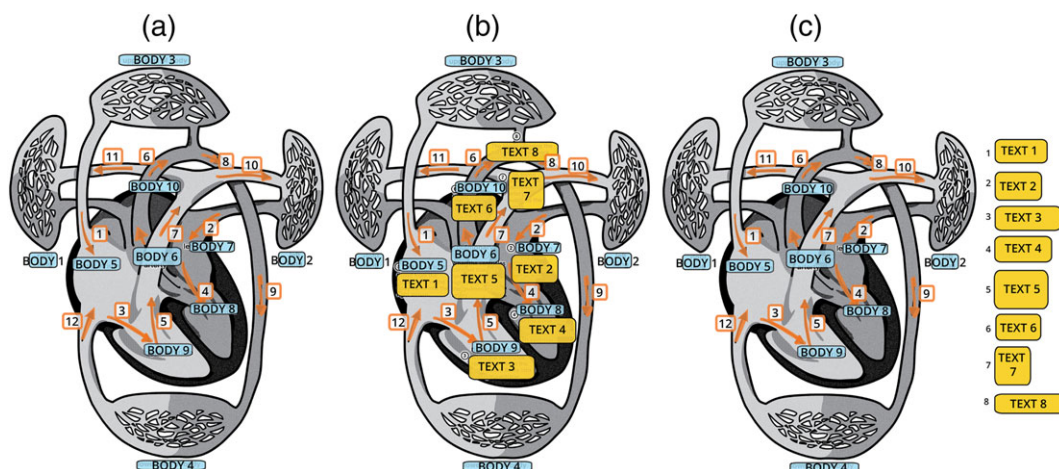
Participants were 132 German University students ( $M_{\text{age}} = 21.15$  years,  $SD = 2.34$ ; 106 female) who participated for course credit or 3 Euro. All participants had normal or corrected-to-normal vision. One participant indicated after completing the experiment that he/she wanted to retract his/her data. Moreover, nine participants were removed from the analyses because they had accidentally ended the learning phase by pressing the spacebar, which was not allowed. The remaining 122 participants were distributed across the four conditions (to which they had been randomly assigned) as follows: diagram only ( $n = 29$ ), separated ( $n = 33$ ), integrated ( $n = 32$ ), and integration instruction ( $n = 28$ ).

### 3.1.2 | Materials

The materials were programmed and presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments). They were largely similar to Experiment 1, with two exceptions: (a) The textual explanations in the separated and integration-instruction conditions were presented next to the diagram rather than underneath it (see Figures 1 and 2) and (b) the materials were presented system paced at 100 s (mean time on task in the self-paced conditions in Experiment 1) to ensure that participants had sufficient time. Furthermore, due to a programming error, the answers for the components test were not saved in Experiment 2.

### 3.1.3 | Apparatus

The materials were presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments) on a monitor with a resolution of  $1920 \times 1080$  pixels. Participants' eye movements were recorded using SMI RED 250 Mobile eye trackers (SensoMotoric Instruments) that recorded binocularly at 250 Hz using SMI iView software (Version 4.2; SensoMotoric Instruments). SMI BeGaze software (Version 3.7; SensoMotoric Instruments) was used for data analysis.



**FIGURE 2** Learning materials in Experiment 2 for the diagram-only condition (a), the integrated condition (b), and the separated and integration-instruction conditions (c). The different areas of interest are presented in a box with a continuous line (relevant text), a box with a dotted line (unnecessary text), and in black (arrows) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3.1.4 | Procedure

Participants were tested individually, or in groups of up to three participants simultaneously. At the start of the experiment, participants were seated in front of a laptop with the mobile eye tracker placed underneath, with their head approximately 60 cm from the screen. First, participants filled in their age and gender. After a short introduction about the experimental procedure, participants provided the subjective prior knowledge rating. Subsequently, 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, after which they rated how much mental effort they had invested in studying the material. Finally, participants completed the posttest and were asked to indicate how much mental effort they invested in answering the posttest questions. The experiment took 10–15 min to complete.

### 3.1.5 | Data analysis

Posttest performance on the blood chain and blood flow tests was scored compared with Experiment 1. For the eye tracking analyses, we first checked the calibration accuracy and tracking ratio (i.e., the percentage of time for which the eye tracker actually measured the eye movements). We excluded two participants from the eye-movement data analyses (one from the integrated and one from the integration-instruction condition) because of inaccurate calibration (i.e., deviation from the four validation points exceeded 1°), and 28 additional participants (diagram only:  $n = 8$ ; separated:  $n = 6$ ; integrated:  $n = 10$ ; integration instruction:  $n = 4$ ) because their tracking ratio was below 70%. The final sample for the eye-movement data analyses ( $n = 92$ ) had an average tracking ratio of 88.52% ( $SD = 6.51\%$ ), with a mean calibration accuracy of 0.27° ( $SD = 0.13^\circ$ ), and was distributed across the conditions as follows: diagram only ( $n = 21$ ), separated ( $n = 27$ ), integrated ( $n = 21$ ), and integration instruction ( $n = 23$ ).

Fixations were defined using a 40°/s velocity threshold and a minimal duration of 100 ms (cf. Holmqvist et al., 2011). Areas of interest (Aols) were created on the diagram, covering the relevant text labels and the arrows indicating the blood flow (see Figure 2). In the unnecessary text conditions, Aols were also created on the unnecessary text lines (see Figure 2). This allowed us to compute the fixation duration on the unnecessary and essential Aols; however, to control for differences in tracking ratio between participants, we calculated a relative measure of fixation duration by dividing the fixation duration on each Aol by the total fixation duration on the learning material (i.e., the sum of fixation duration on the different Aols and white space). With the

fixation duration measure, we can investigate whether the addition of unnecessary information affects the selection of essential information (i.e., when essential information is selected less, it should be reflected in lower fixation duration), and whether integrated unnecessary information is selected more than separated unnecessary information. Moreover, fixation duration can be seen as an indicator of organization processes, as longer fixation time on essential information seems to be associated with deeper processing of information (see Scheiter & Eitel, 2017).

To measure integration of the different sources of information (i.e., relevant text labels, arrows, and unnecessary text), we used transitions between the different Aols. We defined two types of transitions: unnecessary–relevant and relevant–relevant transitions. Unnecessary–relevant transitions are transitions between the unnecessary text and the relevant parts of the diagram (i.e., the relevant text labels and the arrows) or vice versa. Relevant–relevant transitions are transitions between and within two different relevant parts of the diagram, that is, between the relevant text labels and the arrows (or vice versa), within the arrows and within the relevant text labels. To control for differences in tracking ratio, we again divided the number of transitions by the total fixation duration on the learning material. Fewer transitions are indicative of less integration of the text and picture (cf. Mason, Pluchino, & Tornatora, 2015, 2016). Therefore, with the transition measure, we can investigate whether the unnecessary information is merely read, or if it is actively integrated in the essential information, disrupting the integration of the essential information. Moreover, we can examine whether the layout and pacing of the instructional materials affect these integration processes.

## 3.2 | Results

### 3.2.1 | Prior knowledge

A one-way ANOVA with layout as between-factor showed no differences between the diagram only ( $M = 2.03$ ,  $SD = 0.73$ ), integrated ( $M = 2.16$ ,  $SD = 0.63$ ), separated ( $M = 2.03$ ,  $SD = 0.64$ ), and integration-instruction conditions ( $M = 1.96$ ,  $SD = 0.43$ ) on the subjective ratings of prior knowledge,  $F < 1$ .

### 3.2.2 | Learning outcomes

For the two learning outcome measures, the assumptions of normality were violated, and therefore, nonparametric tests were conducted (see Table 4). On the blood chains test, a Kruskal–Wallis test showed a significant difference between conditions,  $\chi^2(3) = 14.97$ ,  $p = .002$ . Follow-up Mann–Whitney tests with a Bonferroni corrected  $p$  value

**TABLE 4** Mean ( $SD$ ) and median (range) performance on the blood chains (max. = 8) and blood flow (max. = 5) tests as a function of layout in Experiment 2

|                         | Blood chains  |                | Blood flow    |                |
|-------------------------|---------------|----------------|---------------|----------------|
|                         | Mean ( $SD$ ) | Median (range) | Mean ( $SD$ ) | Median (range) |
| Diagram only            | 2.00 (2.36)   | 1.00 (7.50)    | 1.28 (1.61)   | 0.50 (5.00)    |
| Separated               | 2.80 (2.36)   | 2.50 (7.50)    | 2.06 (1.36)   | 2.00 (4.50)    |
| Integrated              | 1.17 (1.88)   | 0.25 (8.00)    | 1.14 (1.12)   | 1.00 (5.00)    |
| Integration instruction | 2.64 (2.02)   | 2.25 (7.50)    | 1.82 (1.27)   | 2.00 (5.00)    |
| Total                   | 2.15 (2.24)   | 1.50 (8.00)    | 1.58 (1.38)   | 1.00 (5.00)    |



(i.e., multiplying the  $p$  value by 6, the number of tests that were performed) showed that participants in the integrated condition had lower performance than participants in the separated condition,  $U = 292.00$ ,  $p = .012$  and participants in the integration-instruction condition  $U = 222.00$ ,  $p = .006$ . No other comparisons were significant,  $p > 0.396$ .

On the blood flow test, a Kruskal–Wallis test showed significant differences between the conditions,  $\chi^2(3) = 12.91$ ,  $p = .005$ . Follow-up Mann–Whitney tests (with Bonferroni corrected  $p$  values) showed that participants in the integrated condition had lower performance than participants in the separated condition,  $U = 311.00$ ,  $p = .024$ . None of the other comparisons yielded significant results,  $p > .060$ .

### 3.2.3 | Invested mental effort

Data on the self-reported mental effort investment during the learning phase and on the posttest are reported in Table 5. A one-way ANOVA showed no significant differences among conditions in self-reported mental effort invested during learning,  $F(3, 118) = 2.23$ ,  $p = .089$ ,  $\eta_p^2 = .05$ , or during the posttest,  $F < 1$ .

### 3.2.4 | Fixation duration

See Table 6 for the fixation duration on the relevant information and the unnecessary text. A one-way ANOVA on the fixation duration on the unnecessary text in the separated, integrated, and integration-instruction conditions showed no significant differences among conditions,  $F(2, 71) = 2.46$ ,  $p = .093$ ,  $\eta_p^2 = .07$ . For the fixation duration on the relevant information, a one-way ANOVA revealed a significant main effect of layout,  $F(3, 92) = 30.61$ ,  $p < .001$ ,  $\eta_p^2 = .51$ . Bonferroni-corrected follow-up tests showed that, as expected, participants in the diagram-only condition spent more time looking at the relevant information than participants in the separated condition,  $p < .001$ ,  $d = 1.91$ ; integrated condition,  $p < .001$ ,  $d = 1.72$ ; and

integration-instruction condition,  $p < 0.001$ ,  $d = 2.35$ . All other comparisons were not significant, smallest  $p = .219$ .

### 3.2.5 | Transitions

See Table 6 for the unnecessary–relevant and relevant–relevant transitions. A one-way ANOVA on the unnecessary–relevant transitions revealed a significant main effect of layout,  $F(2, 71) = 39.92$ ,  $p < .001$ ,  $\eta_p^2 = .54$ . Participants in the integrated condition made more transitions between the unnecessary text and the relevant parts of the diagram than participants in the separated condition,  $p < .001$ ,  $d = 1.89$ , and the integration-instruction conditions,  $p < .001$ ,  $d = 1.77$ . The analysis revealed no differences between the separated and integration-instruction condition,  $p > .999$ .

Analysis of the relevant–relevant transitions again revealed a significant main effect of layout,  $F(3, 92) = 8.83$ ,  $p < .001$ ,  $\eta_p^2 = .23$ . Bonferroni-corrected follow-up tests showed that participants in the diagram-only condition made more relevant–relevant transitions than participants in the integrated,  $p = .002$ ,  $d = 0.94$ ; separated,  $p < .001$ ,  $d = 1.21$ ; and integration-instruction conditions,  $p < .001$ ,  $d = 1.39$ . All other comparisons were not significant,  $p > .999$ .

## 3.3 | Discussion

In Experiment 2, we again found no evidence for an overall negative effect of unnecessary text on learning. In contrast to Experiment 1, we did find some effects of layout: On both outcome measures, performance was higher in the separated condition than in the integrated condition, replicating Chandler and Sweller (1991). The integration instruction did not seem to influence learning as no differences emerged between this condition and the separated condition without integration instruction.

The eye-tracking measures indicate that the unnecessary text attracted attention, at the expense of attention to the relevant text labels and arrows, as participants in the unnecessary-text conditions looked less at the relevant parts of the diagram (and transitioned less between relevant parts of the diagram) than participants in the diagram-only condition. Participants in the separated condition spent as much time studying the unnecessary text as participants in the integrated and integration-instruction conditions, whereas participants in the integrated condition made more transitions between the unnecessary text and relevant parts of the diagram than participants in the separated and integration-instruction conditions. This suggests that participants' gaze behavior in the integration-instruction condition was comparable with that of the participants in the separated condition. Possibly, this could explain why we did not find the expected

**TABLE 5** Mean (SD) invested mental effort (max. = 9) during the learning phase and posttest as a function of layout in Experiment 2

|                         | Learning phase | Posttest    |
|-------------------------|----------------|-------------|
| Diagram only            | 6.38 (1.08)    | 6.24 (1.35) |
| Separated               | 7.09 (1.13)    | 6.09 (1.72) |
| Integrated              | 6.59 (1.36)    | 5.91 (1.63) |
| Integration instruction | 6.89 (1.07)    | 6.29 (1.61) |
| Total                   | 6.75 (1.19)    | 6.12 (1.58) |

**TABLE 6** Mean (and SD) fixation duration on and transitions between the relevant information and unnecessary text as a function of layout

|                         | Fixation Duration |             | Transitions       |                      |
|-------------------------|-------------------|-------------|-------------------|----------------------|
|                         | Relevant          | Unnecessary | Relevant–relevant | Unnecessary–relevant |
| Diagram only            | 0.24 (0.08)       |             | 0.12 (0.07)       |                      |
| Separated               | 0.10 (0.06)       | 0.36 (0.10) | 0.05 (0.05)       | 0.03 (0.02)          |
| Integrated              | 0.12 (0.04)       | 0.43 (0.13) | 0.05 (0.07)       | 0.20 (0.13)          |
| Integration instruction | 0.09 (0.04)       | 0.41 (0.13) | 0.04 (0.04)       | 0.04 (0.03)          |
| Total                   | 0.13 (0.08)       | 0.40 (0.12) | 0.06 (0.06)       | 0.09 (0.10)          |

superiority of the separated condition over the integration-instruction condition.

This finding provides direct evidence that spatially integrating text and diagram affects participants' study strategy, inducing them to integrate the two sources of information more. This is in line with explanations (Ayres & Sweller, 2014) for why an integrated layout is preferable over a separated layout when both the text and the diagram are necessary for learning (see also Holsanova, Holmberg, & Holmqvist, 2009). However, when the information is unnecessary for learning, the increased integration seems to hinder learning compared with a separated layout (though not compared with a diagram only).

## 4 | GENERAL DISCUSSION

The present study aimed to address the effects of layout and study time pacing on the negative influence of extraneous information on learning. We set out to replicate and extend the results from Chandler and Sweller (1991). They showed that it is better to leave out unnecessary text (diagram only) than to present it integrated into the diagram or separated from the diagram with the instruction to integrate it. They also demonstrated that a separated presentation (without integration instruction) led to better learning outcomes than an integrated presentation of unnecessary text. Their findings raised the interesting question of whether students in the separated condition would be able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared with a diagram-only condition. Another open question was whether system pacing would aggravate the effects of unnecessary information on learning, mostly when the unnecessary text is spatially integrated or when students are instructed to integrate spatially separated text. Therefore, we directly compared all four layout conditions under either system-paced or self-paced conditions: diagram only, separated, integrated, and integration instruction.

We did not replicate most of Chandler and Sweller's (1991) findings. In Experiment 1, which was conducted online, we found no significant differences in learning outcomes among layout conditions. Thus, the presentation of unnecessary text did not hamper learning compared with studying only the diagram of the heart, regardless of whether participants learned at their own pace or under system pacing. However, self-paced learning was less effortful than system-paced learning and seemed to increase posttest performance (numerically; this was not statistically significant), which is in line with theoretical arguments favoring self-pacing over system pacing (Barrouillet & Camos, 2007) and empirical studies (Kalyuga et al., 2004; Mayer & Jackson, 2005). However, we found no evidence for the hypothesis that system-paced learning would aggravate the negative effect of presenting unnecessary text on learning, as we found no differences among layout conditions when learning was self-paced or system paced.

In Experiment 2, which was system paced, conducted in a lab, and involved eye tracking to study participants' processing of the unnecessary information, we did replicate the finding that a separated presentation (without integration instruction) led to better learning outcomes

than an integrated presentation of unnecessary text. However, none of the unnecessary text conditions showed lower learning outcomes than the diagram-only condition. Moreover, although participants in the integrated condition spent about as much time looking at the unnecessary text as participants in the separated and integration-instruction conditions, they did make more transitions between the unnecessary text and the relevant parts of the diagram. A possible, but speculative, explanation is that the unnecessary text was not harmful for learning when it is merely read, without deep processing or integration attempts, but will negatively influence learning when learners attempt to actively integrate the unnecessary information with the relevant information. Possibly, a learner first has to read (or select) the unnecessary information, to decide whether it is essential for their learning and starts integrating it with the relevant information if it is. When the information seems to be unnecessary for their learning, the learner will not process it further.

One potential explanation for the failure to replicate Chandler and Sweller's (1991) findings on the effects of unnecessary text on learning is that their finding represents a false positive. Yet their study was not the only one showing that extraneous information can hamper learning (for reviews, see Mayer & Fiorella, 2014, on the coherence principle and Kalyuga & Sweller, 2014, on the redundancy principle). However, what qualified as "extraneous information" differed among studies. Because both the relation between the extraneous information and the essential information (i.e., whether the extraneous information is irrelevant or unnecessary) and the modality in which the extraneous information is presented might influence the occurrence and strength of the negative effect on learning, further studies should be performed taking into account these factors. That is, it is possible that the negative effects of irrelevant information (e.g., Harp & Mayer, 1998; Moreno & Mayer, 2000; Sanchez & Wiley, 2006) are larger than the effects of unnecessary information (e.g., Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991; Mayer et al., 1996). Moreover, the strength of the effect may depend on whether it is text (Bobis et al., 1993; Chandler & Sweller, 1991; Mayer et al., 1996), picture (Sanchez & Wiley, 2006), or sound that is extraneous (Moreno & Mayer, 2000).

Such studies should also employ eye tracking to study the attention allocation processes during learning. Indeed, theory predicts that extraneous information hampers learning because learners select this information and try to integrate it with the essential information. This presumably hampers learning by depleting valuable working memory resources that cannot be used for processes relevant for learning (Kalyuga & Sweller, 2014; Sweller et al., 2011). The eye-tracking data collected in the present study show that participants in the unnecessary-text conditions (a) selected the unnecessary information (as evidenced by the fixation duration on this information) and (b) integrated the unnecessary information with the essential information (as evidenced by the transitions between the unnecessary and relevant information). Moreover, the present study suggests that the layout of the learning materials also influences the integration process. Although participants in all unnecessary-text conditions attended to the text, participants in the integrated layout tried to integrate the unnecessary text with the relevant information more.

Note though that if such selection and integration of unnecessary information negatively affects learning, one would also expect performance to be lower in the unnecessary-text conditions compared with a diagram-only condition, which was not the case. This might be associated with a potential limitation of Experiment 2: The fact that it was system paced with a 100-s presentation time (i.e., 20 s longer than in Experiment 1) may have meant that some participants had more time available than they needed or would otherwise have taken to study the learning materials (e.g., the self-paced separated condition only took 64 s on average in Experiment 1), whereas others may not have had sufficient time (e.g., the self-paced diagram-only condition took 129 s on average in Experiment 1, though with large variability between participants). Future research should examine the influence of extra time pressure on learning outcomes and cognitive processes such as integration and organization, with or without unnecessary information present.

In sum, the present study shows no evidence for a negative effect of unnecessary information on learning compared with a diagram-only condition. However, a separated presentation led to better learning outcomes than an integrated presentation of unnecessary text, and participants made more transitions between the unnecessary text and the relevant parts of the diagram when it was integrated. This suggests that when text is presented that describes a picture, it is better not to integrate it into the picture, but to present it separately.

## ACKNOWLEDGEMENTS

This research was funded by a Research Excellence Initiative grant from the Erasmus University Rotterdam awarded to the Educational Psychology section. For Experiment 2, travel support was provided by the LEAD Graduate School (GSC1028), a project of the Excellence Initiative of the German federal and state governments, and by the Erasmus Trustfonds. The authors thank Charly Eielts for her help with designing the learning materials.

## ORCID

Gertjan Rop  <http://orcid.org/0000-0001-6204-1607>

## REFERENCES

- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 940–945. <https://doi.org/10.1037/a0029199>
- 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: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.011>
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- 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: Oxford University Press.
- Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction*, *3*, 1–21. [https://doi.org/10.1016/S0959-4752\(09\)80002-9](https://doi.org/10.1016/S0959-4752(09)80002-9)
- 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: Cambridge University Press. <https://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. [https://doi.org/10.1207/s1532690xcio804\\_2](https://doi.org/10.1207/s1532690xcio804_2)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114. <https://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. <https://doi.org/10.1016/j.cedpsych.2009.10.002>
- 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.
- Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*, 414–434. <https://doi.org/10.1037/0022-0663.90.3.414>
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphic: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology*, *23*, 1215–1226. <https://doi.org/10.1002/acp.1525>
- Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors*, *46*, 567–581. <https://doi.org/10.1518/hfes.46.3.567.1640>
- Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd rev. ed.) (pp. 247–262). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.10>
- Little, J. L., & Bjork, E. L. (2016). Multiple-choice pretesting potentiates learning of related information. *Memory and Cognition*, *44*, 1085–1101. <https://doi.org/10.3758/s13421-016-0621-z>
- 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. <https://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. <https://doi.org/10.1111/bjet.12271>
- Mayer, R. E. (Ed.) (2014). *The Cambridge handbook of multimedia learning* (2nd rev. ed.). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, *88*, 64–73. <https://doi.org/10.1037/0022-0663.88.1.64>
- 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: Cambridge University.
- Mayer, R. E., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of Experimental Psychology: Applied*, *11*, 13–18. <https://doi.org/10.1037/1076-898X.11.1.13>

- McCray, G., & Brunfaut, T. (2018). Investigating the construct measured by banked gap-fill items: Evidence from eye-tracking. *Language Testing*, 35, 51–73. <https://doi.org/10.1177/0265532216677105>
- Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, 92, 117–125. <https://doi.org/10.1037//0022-0663.92.1.117>
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, 84, 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F., 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. [https://doi.org/10.1207/S15326985EP3801\\_8](https://doi.org/10.1207/S15326985EP3801_8)
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision making*, 5, 411–419.
- Rop, G., Van Wermeskerken, M., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (2018). Task experience as a boundary condition to the negative effects of irrelevant information on learning. *Educational Psychology Review*, 30, 229–253. <https://doi.org/10.1007/s10648-016-9388-9>
- Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (2017). With task experience students learn to ignore the content, not just the location of irrelevant information. *Journal of Cognitive Psychology*, 29, 599–606. <https://doi.org/10.1080/20445911.2017.1299154>
- Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory and Cognition*, 34, 344–355. <https://doi.org/10.3758/BF03193412>
- Scheiter, K., & Eitel, A. (2017). The use of eye tracking as a research and instructional tool in multimedia learning. In C. Was, F. Sansosti, & B. Morris (Eds.), *Eye-tracking technology applications in educational research* (pp. 143–164). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-5225-1005-5.ch008>
- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13, 90–100. <https://doi.org/10.1037/a0015108>
- 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(2), 100–110.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312. [https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer. <https://doi.org/10.1007/978-1-4419-8126-4>

**How to cite this article:** Rop G, Schüler A, Verkoeijen PPJL, Scheiter K, Gog TV. The effect of layout and pacing on learning from diagrams with unnecessary text. *Appl Cognit Psychol*. 2018;1–12. <https://doi.org/10.1002/acp.3445>