A Closer Look at the Split Attention Effect: Integrated Presentation Formats for Troubleshooting Tasks

Markus Huff, Vera Bauhoff, Stephan Schwan,
Knowledge Media Research Center Tübingen, Konrad-Adenauer-Str. 40, 72072 Tübingen
m.huff@iwm-kmrc.de, v.bauhoff@iwm-kmrc.de, s.schwan@iwm-kmrc.de

Abstract: According to the spatial contiguity principle of multimedia, learning interrelated information should be presented in an integrated instead of in a disconnected way. In this experimental study (80 participants), we examined performance with a troubleshooting task preventing such an integrated design. Instead, we used a display technology enabling a so-called "vexing image presentation", where there is no need for two foci of visual attention. Compared to a classic split screen setup in which participants have to shift their visual focus frequently, troubleshooting performance was superior in the "vexing image" mode. Hence, not the separation of external representations but processes of re-orientation after switching the visual focus seem to be responsible for split attention effects.

Theoretical Background

As a well-established principle in multimedia design, the "spatial contiguity principle" is concerned with the spatial arrangement of different sources of information, for example a picture and an accompanying text description. Basically, the principle assumes that inter-related information (e.g., a geometric sketch and its related mathematical formula; Tarmizi & Sweller, 1988) should be presented in an integrated format in order to help the learners to mentally connect corresponding elements of the content (Ginns, 2006; Mayer, 1997). In contrast, if inter-related information is presented in a disconnected way (at locally distinct locations), learners have to search for corresponding pieces of information. This process wastes cognitive resources that are then missing for processes of elaboration and transfer into long-term memory (Mayer & Moreno, 2002). Another relevant example of the split attention effect is troubleshooting. Here learners have to find an error in a mechanical device and – if there is one – judge if this is critical for the proper functioning of the system. This is usually done by visually comparing a real device (or its depiction) with a reference device, be it another real device or a picture in a manual. Again, one finds the typical characteristics of split attention, namely, the necessity to switch the focus of attention frequently between two representations in order to mentally connect its corresponding elements in a time-consuming and error-prone manner.

However, although for combinations of text and pictures integrated presentations that avoid split attention are abundant (e.g., Scheiter, Gerjets, & Schuh, in press; Schnotz, 2002), until recently, similar solutions for combinations of real devices and pictures, or pairs of pictures, have been scarce. Accordingly, there have been few empirical studies to date that have examined split attention effects for learning tasks which require the comparison and mental integration of two pictures (Lee, Plass, & Homer, 2006; Huff & Schwan, subm.). But current digital technologies, such as augmented reality or auto-stereoscopic displays, offer new design possibilities for the problem of split attention. In particular, auto-stereoscopic displays allow for the realization of vexing images, that is, the presentation of two pictures in a way that enables participants to switch between them just by moving their head without a need for saccades (Huff & Schwan, subm.). Hence, one major goal of the present study was to examine the suitability of digital vexing images to substantially reduce the split attention effect for troubleshooting tasks that require extensive comparison of pairs of pictures.

Further, if one takes a closer look at the troubleshooting process, one can decompose it into at least two (distinct) sub-tasks, namely, a stage that can be termed as "comparative visual search" (Hardiess, Gillner, & Mallot, 2008) that is followed by a second stage which can be termed as "decision by mental animation".

During the first stage of troubleshooting, the learners have to identify one or more elements in the depiction of the target device that deviate from the respective parts of the correct reference depiction. In order to do this, correspondences between the elements of both pictures have to be established and for each corresponding pair, its similarity or dissimilarity has to be determined. This requires extensive visual search processes that are accompanied by frequent eye saccades and therewith shifts of the visual focus (Hardiess et al., 2008). Saccading alters the projection on the retina and consequently disrupts visual input. This in turn leads to a disruption of the visual working memory because the so-called "transsaccadic memory" store has a limited capacity and maintains visual characteristics of the learning material for only a brief time after stimulus offset (Phillips, 1974). If it would be possible to switch between different external representations without saccades, knowledge acquisition strategies could be based on the iconic memory store of high capacity (Irwin, 1991, 1992; Philips, 1974). In the present study, it is assumed that presenting the picture pairs as a vexing image via an auto-stereoscopic display might substantially reduce the need for shifts of visual focus and therewith the frequency and amplitude of saccades that are necessary to process the relevant information.

Once one or more possible problematic elements have been identified, the learner has to decide for each element whether or not it causes a malfunctioning. During this second stage, a promising strategy for diagnosing a breakdown could be to consider the identified element as part of a causal chain of events, and to simulate this chain of events by mental animation (Hegarty, Kriz, & Cate, 2003). Typically, mental animation is a process that poses high demands on cognitive resources. By reducing cognitive load during the first stage, more cognitive resources should be available for the second, decisional stage. Additionally, if participants can switch between two pictorial representations without moving their visual focus, the causal chain of events should also be easier to trace. Therefore, the second stage of processing should profit from an integrated presentation via vexing image, too.

Finally, an additional consideration concerns the kind of differences between a target device and its correct reference depiction. Here a distinction can be made between differences that are critical for the proper functioning of the target device (e.g., a missing gear in a clockwork) and differences that are not critical (e.g., a differentially colored part). Compared to functioning devices, the causal chain of events should be distorted and shorter for non-functioning devices. Therefore, identifying malfunctioning devices should be more accurate and faster than identifying a different but functioning clockwork. Hence, in sum, a second major goal of the study was to develop a more differentiated account of the cognitive processes that contribute to the split-attention effect.

Hypotheses and Design

In the present study, participants were asked to solve a troubleshooting task by comparing two depictions of a mechanical pendulum clock and deciding whether one of the clocks was malfunctioning. After a learning phase in which the participants learned the basic principles of mechanical clockworks, picture pairs were presented either side by side in a classic split screen presentation mode (Split) or via an auto-stereoscopic display as a "vexing-image" in an integrated mode of presentation (Vex). While one of the pictures showed a functioning pendulum clock as the correct reference device, the second picture showed either the same (functioning) clock, or a different, but also functioning clock, or a different malfunctioning clock. The task was two-fold: First, learners had to visually search the picture pairs for differences and to state whether the two clocks were identical. If the two clocks differed, a second phase followed in which the learners had to decide whether or not the target clock was a functioning one.

We expected troubleshooting performance to be superior in the Vex condition for both the search and the decision sub-task. More specifically, visual search should be more accurate and faster in the Vex mode as there is no need to make a saccade when switching between the two external representations. This in turn should free some resources for the subsequent decision phase, again providing an advantage for the Vex mode. Further, we expected task performance to be superior for malfunctioning target clocks in the decision task, because in this case the causal chain of events is disrupted, which in turn should lead to faster and more accurate decisions.

Compared to the Split mode, where the two external representations were presented on two displays standing next to each other, in the Vex mode, the representations were presented on an auto-stereoscopic display. With such a display, participants could easily switch between two representations without shifting their visual focus. Instead, they just had to move their head. First participants, who were trained on the physical characteristics of a mechanical pendulum clock, had to accomplish a troubleshooting task. Second, they were asked to judge whether both clocks function properly. For this to work, we constructed three kinds of clockwork sets, namely, pairs of identical clockworks, and functioning and nonfunctioning pairs of non-identical clockworks. In the visual search sub-task, participants just needed to search for the respective identical and non-identical clockworks.

Method

Participants and Design

80 students of the University of Tübingen, Germany participated in the study. They received compensation for their participation and were randomly assigned to one of two cells of the between-subjects factor "presentation mode", which described whether information was presented in the Vex or in the Split mode.

Materials and Apparatus

For each participant, the materials consisted of a computer-based learning environment that also included the troubleshooting tasks and a set of questionnaires, including one questionnaire for demographic information and a post-test containing transfer questions. Further we measured self-assessed cognitive load with a German adaptation of the NASA-TLX (Hart & Staveland, 1988), participants´ spatial abilities with the Mental Rotation

Test (MRT, Vandenberg & Kuse, 1978), and the Spatial Orientation Task (Hegarty & Waller, 2004). Visual span was measured with the Visual Patterns Test (VPT, Della Sala, Gray, Baddeley, & Wilson, 1997) and verbal span was assessed via the "Zahlen Merken" (digit span) scale from the Wilde Intelligence Test (Jäger & Althoff, 1983).

The computer-based learning environment was three-tiled. It took approximately 50 minutes to complete. First, participants were asked to learn the basic physical principles of a mechanic pendulum clock that were displayed on 7 HTML-pages using static images and short descriptive texts. More specifically, participants had to learn the functionalities of the weight as energy source, the escapement as central time base, and the different gears with the hands for indicating the time. The second part of this learning environment was a comprehension test that assessed participant's knowledge of pendulum clocks with 5 multiple-choice questions. For each question, three answers were displayed of which one was correct. (e.g., "Which statement is true?" (a) The ticking sound of a mechanical pendulum clock is generated by the pendulum. (b) The escapement is causal for the movement of the gears. (c) The weight is the energy source of the mechanical pendulum clock. Finally, the third part of this learning environment was the troubleshooting task. Participants were presented with two stills of a mechanic pendulum clock that were created using blender3D (http://www.blender.org). The two pictures either depicted exactly the same or different clockworks (see Figure 1). The "same" pictures showed clockworks that were exactly the same (including colors and the indicated time). This could be either two functioning clockworks or two malfunctioning clockworks. There were two kinds of "different" picture pairs. The functioning pairs showed different clockworks that both function properly but where one surface feature was different (e.g., the clocks indicated different times). In contrast, the malfunctioning pairs showed one functioning clockwork and one malfunctioning clockwork (e.g., a clockwork with a missing gear). Additionally, we varied the color of the clockworks presented pairwise across participants. Altogether, the troubleshooting task consisted of 40 pairwise clockworks (20 identical and 20 non-identical clockworks pairs). A training phase at the beginning of the troubleshooting task consisted of 8 clockwork pairs. These trials were excluded from further analysis.

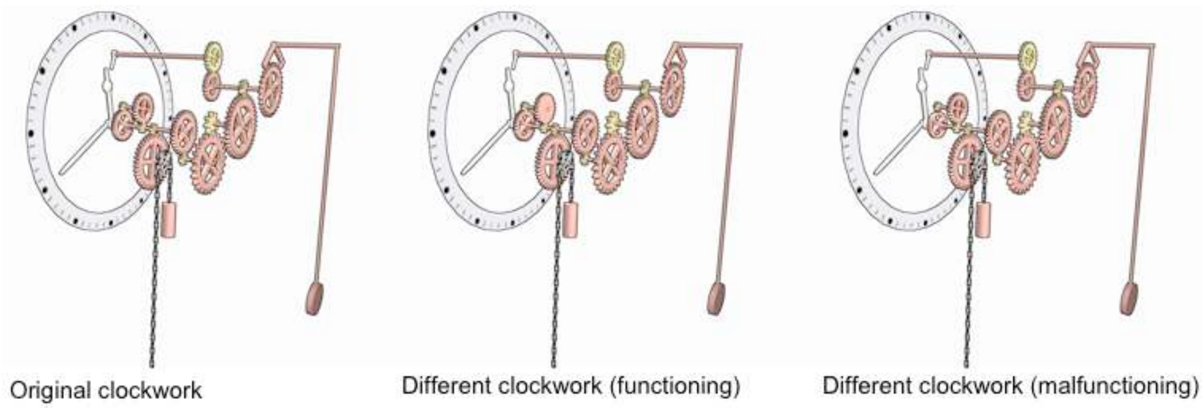

Figure 1. Examples of the stimulus material. In each trial, participants were presented with a combination of an original clockwork with a different clockwork (either functioning with a filled gear or malfunctioning with a gear without cogs).

There were two presentation modes: In the Split mode, the pictures were presented side by side on two displays standing next to each other. In contrast, in the Vex mode we used an auto-stereoscopic display that presented the two pictures such that participant were able to alternate between them simply by shifting their head. In this condition it is not necessary to shift the visual focus (see Figure 2; Huff & Schwan, subm.).

The experimental procedures were controlled by a PC and programmed using custom software programmed with python (http://www.python.org/). In both presentation modes, the first part of the learning environment was presented on a 24" display. The third part of the learning environment (troubleshooting task) was either presented on two 24" displays that were placed next to each other (Split mode) or on a single auto-stereoscopic display (Vex mode). In all presentation modes, the visual angles of the stills were comparable.

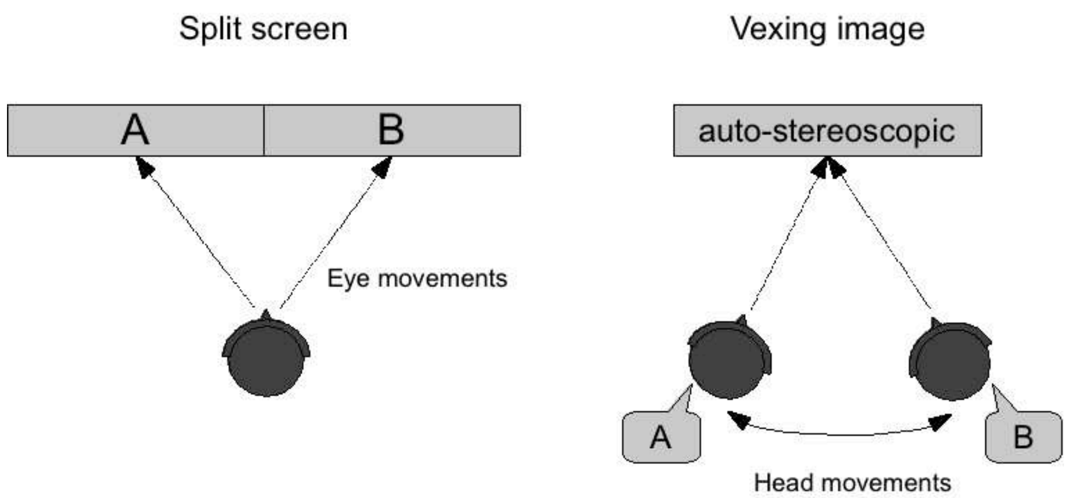

Figure 2. Split screen (left) and vexing image (right) presentation modes both presenting two pictorial representations ("A" and "B"). Whereas participants had to shift their visual focus in the Split mode, participants in the Vex mode just needed to move their head in order to see the second representation.

Procedure

All participants were tested individually and received instructions via the computer monitor. At the beginning, participants completed the first two parts of the learning environment (learning phase and comprehension test). In the third part (troubleshooting task), participants were presented with 48 sets of clockwork pairs. Learners' task was two-fold. After the stills appeared on the screen, the participant had to visually compare the two clockworks and to answer the question whether both clockworks were identical by pressing the key "J" for "Yes" and "N" for "No, the clockworks are different.". Second, if the participant correctly detected that the two clockworks were different she/he had to decide whether both clockworks function properly or if one clockwork was defective. Again, the participant had to press the corresponding key in order to answer the question "Do both clockworks function properly?" ("J" for "Yes" and "N" for "No"). Half of the 40 experimental sets contained functioning and the other half non-functioning clockworks (see Table 1); the serial order and the left/right position of the original clockwork were chosen at random. The first 8 trials were for exercise purposes and were excluded from analysis.

Finally, the post-test, the mental-rotation test, the spatial orientation test, the visual patterns test, the test for the verbal memory span, and the cognitive load assessment questionnaire were filled out by the participants.

Results

Pre-Test

After the first part of the learning environment in which the participants worked through a description of the basic principles of clockworks, they received five multiple-choice questions about pendulum clocks in order to check their knowledge level. Each question had three possible responses of which one was true. Each correctly answered question was scored with one point. A total number of 5 points was possible. Three participants who obtained less than 3 points in this test were excluded from further analysis because they did not learn the relevant principles of a mechanical pendulum clock.

Troubleshooting task

The troubleshooting task consisted of two components. In the comparison task, the participants had to decide whether the two depicted clockworks were identical or not. If the participants had detected a difference between the two clockworks, they were given a second task. In this task, they were required to decide whether the target clockwork was a functioning or a malfunctioning one.

Comparison task

Comparison task performance was measured with the sensitivity measure from the signal detection theory (Green & Swets, 1966) and search times to correctly identified sets of two original clockworks.
Mean hit rate (correct answers to target items – the clockworks were identical) and to false alarm rates (wrong answers to differing picture pairs – although the clockworks were different, participants answered "yes", both

for functioning and malfunctioning differing target clocks) were calculated for each participant in both conditions (Split vs. Vex). Next, two corresponding A' measures (Pollack & Norman, 1964) as sensitivity measure were calculated and submitted to a 2 (presentation mode; between-subjects) x 2 (target clock functionality; within-subjects) mixed factor ANOVA (see Table 1). Participants' performance was significantly influenced by both presentation mode and target clock functionality as the significant interaction of these factors showed, F (1, 75) = 15.88, $p < .001$, η_p^2 =.17. A Bonferroni correction with α = 0.008 was used to correct for the 6 multiple comparisons in this experiment. Performance in the "vexing image" mode was not affected by the kind of target clock type (functioning or malfunctioning, p = .996). In contrast, in the split screen mode, search performance was higher for malfunctioning target clocks than for functioning target clocks ($p < .001$). The main effect for "presentation mode" indicates higher performance in the Vex mode and the main effect for "target clock functionality" higher performance with malfunctioning clockworks, F (1, 75) = 22.13, $p < .001$, η_p^2 = .22 and F (1, 75) = 15.52, $p < .001$, η_p^2 = .17, respectively.

Table 1: Sensitivity measures of the comparison task (Question 1). Arithmetic means and standard deviations in parentheses.

Presentation mode	Clock functionality	
	Functioning	Malfunctioning
Vexing image	.79 (.07)	.79 (.09)
Split screen	.69 (.06)	.74 (.05)

The time from stimulus onset until the decision regarding the similarity of the clockworks was calculated for each participant in each condition (see Table 2). Data as presented in Table 2 were submitted to a mixed-factor ANOVA including the within-subject factor "clockwork type" (search times for identical, non-identical functional, and non-identical malfunctioning clockworks) and the between-subjects factor "presentation mode" (Vex, Split). Participants' search times were faster in the Vex mode as the main effect for presentation mode shows, F (1, 68) = 38.03, $p < .001$, η_p^2=.36. Further, clockwork type influenced search times too, F (2, 136) = 6.14, p = .003, η_p^2 = .08. A Bonferroni correction with α = 0.024 was used to correct for the 2 multiple comparisons in this experiment. Whereas there was no difference between the two non-identical clockwork types functioning (M = 15.71 sec., collapsed over "presentation mode") and malfunctioning (M = 15.95 sec., collapsed over "presentation mode"), $t < 1$, p = .62, search times for non-identical clockworks were faster than search times for identical clockworks (M = 17.19 sec., collapsed over "presentation mode"), ts > 2.62, ps < .011. However, the interaction of clockwork type and presentation mode did not reach the level of significance, F (2, 136) = 2.23, p = .112, η_p^2 = .03.

Table 2: Search times of the comparison task in seconds (Question 1). Arithmetic means and standard deviations in parentheses.

Presentation mode	Clockwork type		
	Identical	Functioning	Malfunctioning
Vexing image	11.52 (6.08)	10.83 (5.56)	11.27 (7.79)
Split screen	21.97 (7.01)	19.81 (6.85)	19.90 (6.55)
Mean (collapsed over "presentation mode")	17.19 (8.40)	15.71 (7.71)	15.95 (8.31)

Decision task

If participants were presented with a pair of non-identical clockworks (with the target clock either functioning or malfunctioning) and they had correctly classified them as "not identical" in the comparison task, they were asked to state whether both clocks function properly. Again, we analyzed the data with the signal detection theory measure, sensitivity A', and reaction times for both correctly identified pairs of functioning clockworks and for correctly identified malfunctioning clockworks. Note, this task uses a two alternative forced choice paradigm. That is, the sensitivity measure as plotted in Table 3 reflects the participants' ability to distinguish between functioning and malfunctioning clockworks in the different presentation modes. Hence, it is not necessary to include this factor in the analysis.

Data as displayed in Table 3 were submitted to a t-test with the independent variable presentation mode for the sensitivity and to a mixed factor ANOVA with the between-subjects factor "presentation mode" and the within-subjects-factor "functionality" for the search times. Neither the sensitivity measure A', t (54.40) = 1.52, p = .134, nor the search times were influenced by the presentation mode Fs < 2.65, ps > .11.

Taken together, both sensitivity and search time indicate that participants in the Vex mode outperformed participants in the Split mode. The superiority of the Vex condition is based on higher performances in the comparison task. In contrast, performance measures in the decision task were comparable across presentation modes.

Table 3: Results of the identification task (Question 2). Arithmetic means and standard deviations in parentheses.

Presentation mode	Sensitivity (A')	Search time (seconds)	
		Functioning clockworks	Malfunctioning clockworks
Vexing image	.78 (.07)	4.99 (4.24)	5.77 (7.82)
Split screen	.74 (.14)	4.99 (4.34)	6.64 (5.00)

Transfer test

In the final phase of the experiment, participants were asked to answer 6 multiple-choice questions that each described one malfunctioning clockwork (e.g., "The hour hand stands still, all other hands move correctly."). Participants were presented with 3 alternative solutions (e.g., "The axis for the hour hand transmission is missing.") of which one was true. Each correct answered question was coded with 1; maximum score was 6. Mean performance was 4.43 (SD = 1.14), there were no significant differences between presentation modes, t (74.90) < 1, p = .591.

Cognitive load and learner related factors

Cognitive load was assessed with a German adaptation of the NASA-TLX (Hart & Staveland, 1988). The six questions were analyzed separately with a t-test including the independent variable presentation mode. Questions 1 (mental activity), 3 (time pressure), and 4 (effort) showed no significant differences between presentation modes (ts < 1.58, ps >.12). However, compared to the Split presentation mode, participants in the Vex mode reported a higher amount of physical activity (Question 2), t(64.45) = 2.51, p = .01, were more confident in scaling their own performance (Question 5), t(65.40) = 3.39, p = .001, and were less frustrated (Question 6), t(71.92) = -3.12, p = .003.

The experimental groups were checked for a-priori differences on the two learner-related factors spatial abilities (assessed with MRT and SO), visual memory span (VPT), and verbal memory span (assessed with ZM). We conducted a t-test for each of the tests separately with the independent variable presentation mode (Vex vs. Split presentation mode). There were no significant differences between groups with regard to spatial ability. Neither MRT, t(74.68) < 1, p = .996, nor VPT, t (74.99) = 1.31, p = .19, and SO, t (66.58) < 1, p = .874 reached level of significance. The same was true for working memory span, t(74.74) = 1.08, p = .28.

Discussion

Although the spatial contiguity principle is well established in research on multimedia learning (Ginns, 2006; Mayer, 2001, 2006), few attempts have been made to extend the principle from picture-text combinations to picture-picture combinations. Hence, the present study attempted to develop an integrated format for the presentation of picture pairs in the context of a troubleshooting task and to compare it experimentally to a non-integrated split-screen presentation. Also, the troubleshooting task was decomposed into a two distinct phases, namely a "comparative visual search" phase (Hardiess et al., 2008), and a "decision by mental animation" phase.

We observed troubleshooting performance to be superior with a new display technology that allows for the integrated presentation of two pictures, compared to a traditional split-screen condition. The integrated presentation was realized by a vexing image which was presented via an auto-stereoscopic display. Here, it is not necessary to change the location of the visual focus in order to switch between the pictures. Instead, learners just had to move their head (see Figure 2). Hence, processes of reorientation were reduced to a minimum. In contrast to the split screen condition, visually comparing the two clockworks in the vexing image condition was both twice as fast and significantly more accurate. This finding suggests that error-prone processes of reorientation are one major reason for lower performance scores in learning environments with low spatial contiguity. In addition, participants were more confident in scaling their own performance and were less frustrated. If the split screen presentation mode can be characterized by frequent shifts of the visual focus and – as a consequence – frequent re-orientation processes, we can conclude that the split attention effect is reduced in the vexing image mode, which seems to be an adequate presentation mode for pairs of pictorial representations.

An interesting finding was that the functioning or malfunctioning of the non-identical clockworks did influence troubleshooting performance in the Split mode but not in the Vex mode. Because viewers in the vexing image mode could keep their visual focus at the critical position while switching between the representations, one can assume that the comparison task can be solved primarily by means of perceptual pattern matching. In contrast, in the Split mode participants cannot rely on perceptual information alone. Switching between representations requires also memory based processing. Participants have to encode parts of one representation and compare this mental representation with the corresponding part of the second representation. Such memory-based processes might involve mental animating of the static clockworks (Hegarty et al., 2003). A central part of mental animation as described by Hegarty et al. (2003) is the identification of the causal chain of events. Learners have to identify the energy source and the transmitting gears. This causal chain is disrupted in malfunctioning clockworks. As participants were instructed in the mechanical principles of pendulum clocks at the beginning of the experiment, this might have helped them during troubleshooting.

Surprisingly, we observed all these effects in the first phase when participants were asked to visually search for differences between the clockworks. Originally, we hypothesized that such processes take place in the decision phase and not – as observed in this study - during the visual search phase. One reason for this might be that participants intermixed the tasks of comparing the clocks and deciding whether the clocks work properly. If this is true, the decision about the proper functioning of the clockworks might have been finished before the decision task began and participants used the decision phase just to verify their decision. For a clearer distinction between those two phases, measures of process related variables such as gaze behavior could be employed. Further research should include eye gaze as dependent measure in order to gain deeper insights in the cognitive processing of separated information.

Finally, we measured participants transfer knowledge in a verbal-based transfer test. As there was no need for costly re-orientation processes in the Vex mode, we hypothesized that participants should be able to encode a more elaborate and more precise mental representation of the mechanical clockwork. As a consequence, compared to participants in the Split mode, transfer test performance should be higher. However, transfer test performance was comparable across the two presentation modes. Faster search times in the Vex mode indicate that learners did not use the free mental resources to encode an elaborated mental model. Instead, they used the free capacity to speed up their responses.

Conclusions

In conclusion, we were able to show that the process of troubleshooting as employed in this study is highly dependant on visual search processes. Shifting the visual focus frequently involves costly processes of re-orientation. The subsequent decision, whether or not the identified difference between two mechanical devices is critical for its proper functioning, is independent of the presentation mode. Solving this task does not require any shifts of visual attention. Instead, learners can solve it by mentally animating both clockworks in a row. They are finished if they find an error or if they successfully animated both clockworks. An integrated design not only reduces the number of shifts of visual focus but also leads to lower frustration scores and higher confidence ratings. Hence, multimedia design reducing costly re-orientation processes is beneficial for the troubleshooting task as a whole. This finding has also consequences for our understanding of multimedia learning. Whereas previous research focused mainly on the combination of pictorial and verbal information, this study showed that common principles of multimedia learning also apply to purely pictorial material. Finally, this study also demonstrated that new display technology is helpful in designing new kinds of learning environments and, in turn, in gaining new insights in cognitive processes involved in learning from multiple external representations.

References

Della Sala, S., Gray, C., Baddley, A.D., and Wilson, L. (1997). *Visual Pattern Test*. Bury St. Admunds: Thames Valley Test Company.

Ginns, P. (2006). Integrating information: Meta-analyses of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511-525.

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford: Wiley.

Hardiess, G., Gillner, S., & Mallot, H.A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision, 8,* 1-13.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Hg.), *Human mental workload*. (pp. 139-183). North-Holland.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21*(4), 325-360.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32,* 175-191.

Huff, M., & Schwan, S. (subm.). Integrating information from two pictorial animations: Complexity and cognitive prerequisites influence performance.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23*(3), 420-456.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*(2), 307-317.

Jäger, A. O., & Althoff, K. (1983). *Der WILDE-Intelligenztest (WIT)* [The WILDE intelligence test]. Göttingen, Germany: Hogrefe.

Lee, H., Plass, J. L. & Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *Journal of Educational Psychology, 98*(4), 902-913.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,*1-19.

Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press.

Mayer, R. E. (2006). *Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14,* 87–99.

Phillips, W. A. (1974) On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*, 283-290.

Pollack, L. & Norman, D. A. (1964). Non-parametric analysis of recognition experiments. *Psychonomic Science, 1,* 125–126.

Scheiter, K., Gerjets, P., & Schuh, J. (in press). The acquisition of problem-solving skills in Mathematics: How animations can aid understanding for structural problem features and solution procedure. *Instructional Science*.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101-120.

Tarmizi, R. & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424-436.

Vandenberg, S. G. & Kuse, A. R. (1978). Mental rotations, a group test of three dimensional spatial visualizations. *Perceptual and Motor Skills, 47*, 599-604.

Acknowledgments

Markus Huff is indebted to the LANDESSTIFTUNG Baden-Württemberg for the financial support of this research project by the Eliteprogramme for Postdocs. The authors thank Kathrin Stalder and Marco Dehez for their help in conducting this experiment.