Assessment of Cognitive Load in Multimedia Learning Using Dual-Task Methodology

Roland Brünken1, Susan Steinbacher1, Jan L. Plass2, and Detlev Leutner1

1Erfurt University, Germany, and 2New York University, USA

Abstract. In two pilot experiments, a new approach for the direct assessment of cognitive load during multimedia learning was tested that uses dual-task methodology. Using this approach, we obtained the same pattern of cognitive load as predicted by cognitive load theory when applied to multimedia learning: The audiovisual presentation of text-based and picture-based learning materials induced less cognitive load than the visual-only presentation of the same material. The findings confirm the utility of dual-task methodology as a promising approach for the assessment of cognitive load induced by complex multimedia learning systems.

Key words: cognitive load, learning with multimedia, dual task

Current research on learning with multimedia increasingly pays attention to insights on knowledge acquisition provided by cognitive theory (Brünken & Leutner, 2000; Mayer, 2001). Of central interest in this line of research is how different presentation formats used in multimedia learning systems facilitate knowledge acquisition to different degrees, and which theoretical explanations can account for these effects. While in the past few years empirical studies were primarily concerned with the extent to which different presentation modes of information, i.e., the pictorial or textual presentation of learning materials, lead to different degrees of knowledge acquisition (e.g., Brünken, Steinbacher, Schnottz, & Leutner, 2001; Mayer, 1997; 2001; Schnottz & Bannert, 1999), a few groups of researchers have begun to study whether and to what degree the sensory modality of information, i.e., auditory and/or visual presentation of verbal learning materials, influences knowledge acquisition (Brünken & Leutner, 2001; 2002; Mayer & Moreno, 1998; Moreno & Mayer 2000; Mousavi, Low, & Sweller, 1995). These studies have consistently found two effects that explain how knowledge acquisition is facilitated by multimedia materials. The multimedia dual-coding effect, focusing on the presentation mode of information, describes how learning is facilitated through the simultaneous presentation of picture-based information and closely related text-based learning materials (Mayer, 1997). The modality effect, focusing on the sensory modality of information, describes how knowledge acquisition is facilitated by an audiovisual presentation format, i.e., by simultaneously using the auditory and the visual sensory modalities instead of only the visual modality (Mayer, 2001).

In order to theoretically explain these effects, Mayer and associates (Mayer, 2001; Mayer & Moreno, 1998) have proposed a model of knowledge acquisition with multimedia learning systems that integrates various aspects of cognitive theory, such as working memory models, dual coding theory, cognitive load theory, and generative theory of learning. The fundamental assumption of multimedia dual-processing theory (Mayer & Moreno, 1998) is the existence of two separate processing systems, one for...
visual information and one for auditory information in working memory (see Miyake & Shah, 1999; Baddeley, 1986; Baddeley & Logie, 1999). Each of these systems has a limited processing capacity. Mayer makes the additional assumption that different processes of mental representation occur during the processing of pictorial and of textual information (see Paivio, 1986): Learning occurs when the learner perceives and selects relevant information, organizes it into a coherent mental representation, and builds referential connections between individual pieces of information and integrates it with prior knowledge (Mayer, 1997; Schnotz & Bannert, 1999; Schnotz, Böckheler, & Grzondziel, 1999; Wittrock, 1990). This process takes place in working memory (Mayer & Moreno, 1998). Knowledge acquisition is especially enhanced when textual and related pictorial information is presented concurrently (contiguity effect: Mayer, 1997). In line with cognitive load theory (Chandler & Sweller, 1991; Sweller, 1988), Mayer and his colleagues also assume that performance during knowledge acquisition is dependent on the cognitive resources available for information processing.

The simultaneous visual presentation of text- and picture-based learning materials (i.e., on-screen text and pictures) leads to an information processing that is, at least initially, carried out in the visual subsystem of working memory. Thus, the cognitive resources available in visual working memory have to be divided between textual and pictorial information, whereas the resources of the phonological working memory subsystem remain largely unused. By contrast, in the case of an audiovisual presentation of text- and picture-based learning materials (i.e., narrated text and on-screen pictures), the auditory information (the narration) is processed in the phonological subsystem, while the visual information (the picture) is processed in the visual subsystem. In this design, the combined resources of both subsystems can be used to process the information, so that the audiovisual presentation of information allows for the utilization of more cognitive resources for information processing than the visual-only presentation of the same learning materials. Hence, knowledge acquisition is comparatively higher with audiovisual information, as empirical studies have confirmed (modality effect: Brünken & Leutner, 2001; 2002; Mayer & Moreno, 1998; Mayer, Moreno, Boire, & Vagge, 1999; Moreno & Mayer, 1999; Mousavi et al. 1995).

Measuring cognitive load

At the heart of the model proposed by Mayer (Mayer, 2001; Mayer & Moreno, 1998) is the assumption that the processing of verbal and pictorial information presented in audiovisual modalities versus visual-only modality demands different amounts of cognitive resources. This in turn leads to different amounts of acquired knowledge given an otherwise identical content of the learning materials. Based on this difference in learning outcome, experimental studies on multimedia learning that are based on Mayer’s model use cognitive load to explain the differences in learning efficacy of different types of multimedia learning materials. Unlike other basic psychological studies on memory, these studies do not directly assess actual working memory load in the form of a manipulation check. The indirect measures used in some of these studies, using e.g., posttreatment questionnaires (Paas & van Merrienboer, 1993) or measures of time-on-task (e.g., Brünken & Leutner, 2002), do not provide conclusive evidence for the actual cognitive load experienced by the learners (Kirschner, 2001).

The goal of the present studies was therefore to develop and evaluate a measurement tool for the computer-based direct assessment of cognitive load during the learning process in multimedia learning systems. To this end, the dual-task methodology, which is well known in experimental psychology (Heuer, 1996; Verwey & Veltman, 1996; Wickens, 1984), was applied and experimentally tested for the first time in the context of multimedia learning.

The dual-task methodology is based on the assumption that the processing capacity of working memory is limited but can be flexibly allocated (see Baddeley, 1986; Miyake & Shah, 1999). If two tasks have to be processed at the same time (dual-task condition), and both require the same cognitive resources, then these resources have to be split between the two tasks. This means that fewer resources are available for processing each individual task than would be available for processing a single task (single-task condition). If the processing of a task depends on available cognitive resources, then performance in processing a secondary task will be reduced in relation to the amount of cognitive resources required by a primary task. Notably, Baddeley and colleagues (Baddeley & Hitch, 1974; Baddeley, 1986; Baddeley & Logie, 1999) have used many variations of the paradigm in testing their proposed model of working memory.

Two variations of the dual-task methodology are commonly used in the analysis of cognitive load in experimental research. In the first method, the dual-task methodology is used in order to induce cognitive load. It is assumed that the processing of a secondary task demands cognitive resources that are consequently no longer available for processing the primary task. The extent to which the addition of a secondary task causes a reduction in the performance of the primary task (compared with the single-task condition) is measured. Typical examples for such secondary tasks are articulatory suppression, tap...
ping, tracking, or verbal generation of random numbers (Baddeley, 1986; Logie, Baddeley, Mane, Donchin, & Sheptak, 1989; Klauer, Oberauer, Rossnagel, & Musch, 1996).

In the second method, the dual-task methodology is used to assess cognitive load. Here it is assumed that different variants of primary tasks require different amounts of cognitive resources. Thus, performance in processing a simultaneously presented secondary task varies according to the cognitive load induced by the primary task. Typical examples of secondary tasks used are choice reaction tasks, which comprise simple decision-making. For instance, simple equations are presented (e.g., \(4 + 3 = 7\); \(5 + 2 = 6\)), and the participant is asked to decide as quickly as possible whether the proposed solutions are correct or incorrect. In addition to numerical tasks, verbal and pictorial materials are also employed (Baddeley & Logie, 1999; Logie et al., 1989). The dependent variables in these tasks are reaction time (interval between task presentation and reaction) and error rate pertaining to the processing of the secondary task.

### Measuring Cognitive Load in Multimedia Learning Using Dual-Task Methodology

The use of secondary tasks for assessing the cognitive load of a primary task seems to be a particularly promising approach for measuring cognitive load during the processing of learning materials presented by multimedia systems. If audiovisual and visual-only presentations of the same verbal and pictorial learning materials differ in the amount of cognitive resources required for information processing in the visual subsystem of working memory, then participants working with an audiovisual or a visual-only multimedia system should also exhibit differences in performance when processing a simultaneously presented secondary visual task. Participants learning with audiovisual information in a dual-task condition should have comparatively more resources at their disposal for processing a visual secondary task than participants learning with visual-only information. Thus, participants in an audiovisual multimedia learning condition should outperform participants in a visual-only multimedia learning condition in the processing of the secondary task.

In order to test this hypothesis, two pilot experiments with different multimedia learning systems were conducted with 10 participants each. A second goal of the experiments was to determine whether and to what extent the dual-task methodology is applicable and useful within the context of research on learning with complex multimedia systems.

### Experiment 1

#### Method

**Participants and Design**

Study participants were 10 female students enrolled in the education program at Erfurt University (Germany). The average age of the participants was 22.8 years \((SD = 2.5)\). Due to expected large individual differences in reaction times, a all-within-subject design with repeated measures was applied. Dependent variables were reaction times for the secondary task and knowledge acquisition for the primary task (learning with a complex multimedia system). A pre-test was performed to allow for the assessment of prior knowledge effects for the primary task.

**Primary task.** The learning system was a multimedia computer-based training (CBT) program on how the human cardiovascular system works. The program comprised of 22 screen pages; each page contained verbal and closely related pictorial descriptions of the relevant organs (e.g., heart, lungs, and veins), the physiological processes of gas and substance transportation, as well as the electrochemical processes of heart stimulation and muscle contraction. The program was developed in Toolbook (Asymetrix, 1997) for the Windows 95 operating system (Microsoft, 1996). It was available in two variants that differed in the presentation format of information (audiovisual vs. visual-only). In the visual-only variant, the verbal information was presented visually (i.e., as on-screen text) simultaneously with the picture. In the audiovisual variant, the same verbal information was presented in auditory format (i.e., narrated by a female speaker), while the on-screen text was not visible.

The program has been successfully used in previous studies (Brünken & Leutner, 2001; 2002). In these studies, we were able to show that a modality effect for knowledge acquisition did indeed occur: Participants working with the audiovisual variant of the learning system acquired more knowledge than participants who were working with the visual-only variant of the system.

In order to implement the all-within-subject design of the present study, the CBT program was redesigned in such a way that, for each participant, the visual-only and the audiovisual presentation format of each of the 22 screen pages was alternated page by page, and the sequence was altered between participants (see Table 1). The software was programmed to automatically advance to the next screen page, and the presentation time of each page was identical for both variants, corresponding to the duration of the audio narration of the audiovisual variant (see also Brünken & Leutner, 2001).
Table 1. Research Design Used in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Sequence</th>
<th>N</th>
<th>Pretest</th>
<th>(ST)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>n-1</th>
<th>n</th>
<th>(ST)</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>X</td>
<td>X</td>
<td>V</td>
<td>AV</td>
<td>V</td>
<td>...</td>
<td>V</td>
<td>AV</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>X</td>
<td>X</td>
<td>AV</td>
<td>V</td>
<td>AV</td>
<td>...</td>
<td>AV</td>
<td>V</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>X</td>
<td>V</td>
<td>AV</td>
<td>V</td>
<td>V</td>
<td>...</td>
<td>V</td>
<td>AV</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>X</td>
<td>AV</td>
<td>V</td>
<td>AV</td>
<td>AV</td>
<td>...</td>
<td>AV</td>
<td>V</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note. V = visual-only presentation of learning materials as primary task; AV = audiovisual presentation of learning materials as primary task; ST = single-task condition (only secondary task; N = number of participants); “X” indicates that the respective test or task was completed by this group.

Two criterion referenced paper-and-pencil tests (pretest and posttest; Klauer, 1987) were used to assess knowledge acquisition concerning the information presented in the learning system. Each test consisted of 20 multiple-choice items with four answer alternatives, where each of the alternatives could be right or wrong. Right answers were presented in random order. The content of the test items referred to the organs as well as to the physiological and electrochemical processes described in the multimedia learning materials. The tests had previously been successfully implemented in assessing knowledge acquisition for these materials (Brünken & Leutner, 2001; 2002). However, in the present experiment the assessment of knowledge acquisition served as a control measure. In the dual task condition, the knowledge acquisition test should show whether the participants worked, according to the instruction they had received, on the primary task (the learning system) and whether they indeed acquired knowledge from this task. Since previous studies had shown that the modality effect could be replicated with the multimedia materials used for this study (Brünken & Leutner, 2001), we chose an within-subjects design to test the usefulness of the dual-task methodology for the measurement of cognitive load in multimedia learning, in which each participant was presented with both the visual-only and the audiovisual primary task. Thus, in this design, the knowledge acquisition test did not serve as a test for the replication of the modality effect.

Secondary Task. A simple, continuous visual observation task was selected as the secondary task: The black-colored letter “A” was displayed in a separate window on the screen (see Figure 1). After a random period of 5–10 seconds, the color of the letter changed to red. The participants were asked to press the space bar on the keyboard as soon as possible after the letter had changed to red. Once the key was pressed, the letter changed back to black, and the next countdown started. The software automatically recorded the time lapse between the appearance of the letter in red and the pressing of the space bar.

However, due to the multitasking capabilities of the Windows operating system, the measurement of reaction times in milliseconds is very problematic. Currently there are no commercially available Windows-based software tools for such a reaction time measurement that could be used in the framework of our dual-task study. PC-based measurement tools such as ERTS (Experimental Run Time System) (ERTSLab, 1999) operate on MS-DOS and cannot run simultaneously with Windows-based multimedia programs. Therefore, to be able to implement the dual-task condition, we developed the Windows-based software tool WinRT, which is able to measure reaction times in Windows-based multimedia software.

Procedure

Individual test sessions were conducted for each participant in one of the computer laboratories at Erfurt University. The apparatus consisted of an 800 MHz Pentium III PC with a 17 inch monitor that was set to a screen resolution of 1152 x 864 pixels. The experimenter was present during the entire investigation. Before starting the experiment, the experimenter instructed the participant in the operation of the software programs.

The participant first completed the paper-and-pencil pretest, which assessed prior domain-specific knowledge. Then, the study phase with the computer programs began, which involved the single processing of the secondary task (single-task condition) as well as the concurrent processing of the secondary task and of the materials in the learning system as primary task (dual-task condition). The single-task condition, which was completed either before or after the dual-task condition, was included to obtain an individual baseline measure for each participant’s reaction time. Following the last measure, the paper-
and-pencil posttest was administered to assess knowledge acquisition.

The entire procedure, including the two paper-and-pencil tests, lasted approximately 60 minutes. The duration of the work with the computer programs took 3 minutes for the single-task condition and 14 minutes 15 seconds for the dual-task condition.

The computer screen presented to the participants contained two frames: a primary task frame and a secondary task frame. The primary task frame appeared centered against a black background. The secondary task frame was placed in the center, vertically above the primary task frame, (see Figure 1). The size of the primary task frame was 22 × 16 cm (width × height), and the size of the secondary task frame was 3 × 3.5 cm. The letter displayed in the secondary task frame was 1.2 × 1.3 cm in size. Thus, the single-task and the dual-task condition represented identical technical implementations, as both programs (Toolbook-application and WinRT) were active at the same time.

In the single-task condition, the secondary-task frame was presented above the blank page of the primary task frame (black background; empty beige-colored frame). In the dual-task condition, the learning program was presented in the primary task frame. The pages of the learning program were advanced automatically; no mouse or keyboard input was required to work with the multimedia learning system.

Results

Scoring and analysis

The points scored in the pretest and in the posttest, as well as reaction times in milliseconds, were recorded. A paired sample t-test was calculated to assess knowledge acquisition. For the analysis of reaction times, the data from the secondary task were first synchronized with the learning program on the basis of time-stamped log file data. Then the measures were matched to their corresponding test condition: single task, dual-task with visual-only primary task, or dual-task with audiovisual primary task. Finally, individual outliers were eliminated. According to the statistical definition, a reaction time measure was defined as an outlier when its absolute value was more than three standard deviations above the individual mean in the corresponding test condition.

Because the secondary task was presented within a randomized time interval, and because of individual differences in reaction times and different amounts of outliers for each individual, we obtained different numbers of reaction time measures for each participant and test condition (between 23 and 47 measures per person and condition). Therefore, the reaction time measures for each of the three test conditions were averaged for each participant (see Table 2), and these three mean reaction times were used for all further analyses. Following the conventional approach for the evaluation of reaction time measures used in experimental psychology (see Donk & Sanders, 1989), we first computed a one-factor repeated measures analysis of variance (RM-ANOVA) with repeated measures on the three test conditions (single task, dual-task: visual-only, and dual-task: audiovisual). Subsequently, post hoc paired samples t-tests with α-adjustment (total α = 0.05) were calculated for two test conditions at a time.

Primary-task performance

For knowledge acquisition, the mean number of points was 6.9 (SD = 5.9) at pretest and 26.8 (SD = 6.7) at posttest. The difference is highly statistically significant (t9 = 6.97; p < 0.001), providing empirical evidence that the participants completed the primary task according to the instructions and indeed acquired knowledge from the multimedia learning system.

Secondary-task performance

Table 2 displays the individual reaction times in milliseconds (M, SD, number of measures) in the three test conditions for all 10 participants. As predicted, the reaction times were lowest for the single-task condition (M = 355.10, SD = 54.13). For the two dual-task conditions, the reaction times were lower for the audiovisual primary-task presentation (M = 631.60, SD = 183.09) than for the visual-only primary-task presentation (M = 803.60, SD = 234.92). The RM-ANOVA confirmed these results, as did the post hoc t-tests. The RM-ANOVA revealed a significant main effect of the factor “test condition” (F(2, 9) = 41.191, p < 0.001, MSE = 12429.54). The post hoc t-tests showed that the differences in the reaction times between the single-task condition and the visual-only dual-task condition (t9 = 7.318, p < 0.001, d = 2.63), the single-task condition and the audiovisual dual-task condition (t9 = 6.005, p < 0.001, d = 2.05) as well as the difference between the visual-only and the audiovisual dual-task conditions (t9 = 4.279, p = 0.002, d = 0.82) were statistically significant.

These results are in full agreement with our predictions: The reaction times in secondary task performance increase when the primary task includes visual-only learning materials compared to a primary task with audiovisual materials of the same content. To verify the results on the basis of the individual measures of reaction times, the analysis was repeated separately for each of the 10 participants. The same effect patterns were found as in the analysis based on the computed means for reaction times.
Experiment 2

To replicate the results found in Experiment 1, and to test for generalizability across different learning systems, a second experiment was conducted. The second study used the same design and the same secondary task as in Experiment 1, but included a different multimedia system as the primary task.

Method

Participants and design

Experiment 2 was conducted with 10 female students enrolled in the Psychology program at Erfurt University, Germany. The mean age of the participants was 21.2 years ($SD = 1.93$). Again, a all-within-subject design with repeated measures was used (single task; dual-task: visual-only; dual task: audiovisual). The dependent variables were knowledge acquisition in the primary task (learning with the multimedia system) and reaction time in the secondary task.

Primary task. The multimedia learning system used in the second experiment, which was also used
Table 2. Individual Measures in Experiment 1 for the Secondary Task (reaction time in milliseconds; $N$, $M$, $SD$,) Broken Down by the Three Test Conditions.

<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Single Task</th>
<th>Dual Task (audio-visual)</th>
<th>Dual Task (visual only)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$ $M$ $SD$</td>
<td>$N$ $M$ $SD$</td>
<td>$N$ $M$ $SD$</td>
</tr>
<tr>
<td>1</td>
<td>23 320 70</td>
<td>37 456 111</td>
<td>24 564 119</td>
</tr>
<tr>
<td>2</td>
<td>24 442 97</td>
<td>39 901 714</td>
<td>26 1194 920</td>
</tr>
<tr>
<td>3</td>
<td>23 392 129</td>
<td>26 951 982</td>
<td>42 992 696</td>
</tr>
<tr>
<td>4</td>
<td>23 268 47</td>
<td>43 441 309</td>
<td>44 564 381</td>
</tr>
<tr>
<td>5</td>
<td>25 410 202</td>
<td>34 585 306</td>
<td>42 678 391</td>
</tr>
<tr>
<td>6</td>
<td>23 354 128</td>
<td>43 508 163</td>
<td>40 909 639</td>
</tr>
<tr>
<td>7</td>
<td>24 335 52</td>
<td>38 612 358</td>
<td>47 736 419</td>
</tr>
<tr>
<td>8</td>
<td>23 332 68</td>
<td>37 621 264</td>
<td>31 709 342</td>
</tr>
<tr>
<td>9</td>
<td>23 397 135</td>
<td>40 768 476</td>
<td>41 1119 1039</td>
</tr>
<tr>
<td>10</td>
<td>24 301 57</td>
<td>45 476 139</td>
<td>38 571 229</td>
</tr>
</tbody>
</table>

in prior studies (Brünken et al., 2001), contained verbal and pictorial information about the historic city of Firenze (Florence), Italy. The system was programmed in Toolbook Instructor II (Asymetrix, 1997) and comprised of 16 screen pages. Each page included verbal information (on-screen text or audio narration) and closely related pictorial information (graphics and pictures) about the location and the importance of a historic building, place, or work of art (e.g., a cathedral, or Michelangelo’s sculpture David). Similar to a tourist guide, each page of the system contained at least one picture of a specific object, a verbal description with some historical facts concerning this object, and a route map of the city of Firenze that showed where the object was located. Verbal and pictorial/graphical information were presented separately in two frames.

As in Experiment 1, the presentation format of each screen page was either visual-only (on-screen text and graphic/picture) or audiovisual (audio narration and graphic/picture, with the text frame hidden), while the content of both variants was identical. The presentation format alternated from screen page to screen page, and the sequence of alternation itself was alternated among different participants (see Table 1). The screen pages were automatically advanced by the system, and the presentation time of each page was the same for both formats, determined by the duration of the corresponding narration for each screen page in the audiovisual treatment.

Knowledge acquisition of the content of the learning system was assessed using two criterion-referenced paper-and-pencil tests, which were constructed as psychometrically parallel tests and which served as pretest and posttest. Both tests have successfully been used in prior studies (Brünken et al., 2001). Each test contained 20 multiple-choice items with four answer alternatives. The test items elicited responses regarding historical facts and the location of a specific object. Again, the knowledge acquisition test served as a control measure to allow verification that the participants had indeed worked with the learning system and had acquired knowledge from the information presented.

Secondary task. The secondary task used in Experiment 2 was the same visual observation task as in Experiment 1 (see Figure 2).

Procedure

The study was conducted with individual test sessions at one of the computer labs at Erfurt University. The technical details of the computer systems, screen-sizes, and the experimental procedures were identical to Experiment 1. The mean duration of the entire experiment was approximately 50 minutes; the single-task condition (secondary task alone) lasted 3 minutes, and the dual-task condition 11 minutes 40 seconds.

Results

Scoring and analysis

As in Experiment 1, points scored in the knowledge acquisition pretest and posttest were recorded and compared using a paired samples $t$-test. The analysis of the reaction time measures was conducted — after completing the same preprocessing as in Experiment 1 — firstly on individual data, where the mean reaction times for each participant and test condition were calculated. As in Experiment 1, we observed a different number of measurements for each person and test condition (between 19 and 35; see Table 3), caused by the random generation of secondary tasks.
Figure 2. Screenshot from the materials used in Experiment 2 under the dual-task condition. In the middle is the page of the learning system (visual-only variant); the frame of the secondary task is located above center.

individual differences in reaction times, and differences in individual outliers. In a second step, the individual means were again analyzed by computing a RM-ANOVA and the appropriate post hoc tests.

Primary-task performance

Concerning the knowledge acquired from the learning system, the mean number of points was 3.2 (SD = 1.69) at pretest and 10.8 (SD = 4.24) at posttest. The difference was statistically significant ($t_{9} = 6.626, p < .001$). Again, the results indicate that the participants worked on the primary task according to the instructions they had received and acquired knowledge from the multimedia learning system.

Secondary-task performance

Table 3 shows the individual secondary-task performance (mean, SD, number of measures) for each participant for each of the three experimental conditions. As can be seen, the second experiment shows the same effect pattern as found in Experiment 1. The mean reaction times were lowest for the single-task condition ($M = 466.5; SD = 74.67$), followed by the audiovisual dual-task condition ($M = 789.9; SD = 349.04$). As in the first experiment, the reaction times were highest in the visual-only dual-task condition ($M = 1063.6; SD = 382.2$). The results were confirmed by the RM-ANOVA, which showed a statistically significant main effect of the test condition ($F(2, 9) = 20.356, p < .001, MSE = 43,887.32$). To
Table 3. Individual Measures in Experiment 2 for the Secondary-Task (reaction time in milliseconds; \(N, M, SD\),) Broken Down by the Three Test Conditions.

<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Single Task</th>
<th>Dual Task (audio-visual)</th>
<th>Dual Task (visual only)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N)</td>
<td>(M)</td>
<td>(SD)</td>
</tr>
<tr>
<td>1</td>
<td>24</td>
<td>459</td>
<td>157</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>537</td>
<td>76</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>602</td>
<td>341</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>373</td>
<td>323</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>414</td>
<td>74</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>533</td>
<td>270</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>394</td>
<td>79</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>506</td>
<td>95</td>
</tr>
<tr>
<td>9</td>
<td>22</td>
<td>437</td>
<td>156</td>
</tr>
<tr>
<td>10</td>
<td>22</td>
<td>410</td>
<td>79</td>
</tr>
</tbody>
</table>

analyze these results in more detail, post-hoc paired-sample t-tests with \(\alpha\)-adjustment (total \(\alpha = 0.05\)) were computed. The t-tests showed that the observed mean differences were statistically significant for the comparison of the single-task and audiovisual dual-task conditions \((t_9 = 3.078, p = .013, d = 1.53)\), the comparison of the single-task and the visual-only dual-task conditions \((t_9 = 5.042, p = .001, d = 2.61)\), and the comparison of the audiovisual and the visual-only dual-task conditions \((t_9 = 7.687, p < .001, d = 0.75)\). Again, as can be seen in Table 3, this effect pattern was also present for the reaction time values of each individual participant.

In summary, we found that the results of our second experiment, as well as the results of the first experiment, are in line with the underlying cognitive-load hypothesis and show a significant increase in reaction times in the secondary task when subjects work with a visual-only multimedia learning system as the primary task, compared with when they work with an audiovisual multimedia system of the same content as the primary task.

Discussion

The goal of the experimental pilot studies presented in this paper was the development and evaluation of a new approach for the direct assessment of cognitive load during learning with complex multimedia learning systems. The study was motivated by a series of recently published studies (Brünken & Leutner, 2001; 2002; Mayer & Moreno, 1998, Moreno & Mayer, 1999; Mousavi et al. 1995), which showed that the audiovisual presentation of closely related picture-based and text-based learning materials leads to enhanced knowledge acquisition when compared with the visual-only presentation of the same learning material. All of these studies explain this effect with reference to cognitive-load theory (Chandler & Sweller, 1991): It is assumed that the visual-only presentation of learning materials results in a comparatively higher visual cognitive load as compared with audiovisual presentation, and this higher cognitive load causes the observed lower amount of knowledge acquisition.

Having applied the dual-task methodology to measure cognitive load during multimedia learning, the present studies show that these hitherto only assumed differences in cognitive load of the visual subsystem of working memory can be assessed directly, based on performance in processing a simultaneously presented secondary task.

The present findings are in accordance with the theoretical predictions. On the one hand, both experiments revealed a statistically significant difference in performing the secondary task alone and in performing the same task in a dual-task condition. The processing of the learning materials (primary task) and the processing of the secondary task are shown to use the same cognitive resources, so that the secondary task is a sensitive measure of cognitive load during the processing of the primary task.

Furthermore, and of central importance for our studies, we found, again in both experiments, that the processing of the secondary task during the visual-only presentation of verbal and pictorial learning materials was significantly slower than during the audiovisual presentation of the same materials. As predicted by cognitive load theory, the simultaneous visual presentation of texts and related pictures induces a comparatively higher cognitive load than the audiovisual presentation of the same learning materials.

In addition, comparing the results of the two experiments, we found the reaction times in Experi-
ment 2 under both dual-task conditions to be longer than in Experiment 1. This could indicate that the learning scenario used in Experiment 2 was more difficult than that used in Experiment 1. If a task is more difficult in general, then it needs more cognitive resources (Kirschner, 2001) independent of the sensory modality used, which could be the reason for the overall increase in reaction times in Experiment 2. Though we do not have empirical information on this topic from the present experiment, prior studies with the learning system used in Experiment 2 (Brünken et al., 2001) indicated that these materials seem to be relatively difficult for learners.

Our findings have theoretical as well as practical implications. On the theoretical side, we found strong empirical evidence for cognitive load effects on learning outcomes by directly measuring the performance in a simultaneously processed secondary task. This direct measurement is based on the dual-task method, which is well known in the field of experimental psychology but has not been previously applied in studies on learning and instruction with complex multimedia systems. On the practical side, the dual-task approach developed in the present studies appears to be a suitable approach for measuring cognitive load in contexts of practical application, even beyond learning and instruction. For example, to evaluate variants of interface design drafts for a multimedia system that differ with respect to the format in which information is presented, it is conceivable that this approach could be used to determine the amount of cognitive load induced in the user by each design. On the basis of such an ergonomic software evaluation one could determine, for example, which interface design variant minimizes cognitive load.

These pilot studies of course leave many questions open. For instance, there are still some technical problems associated with the measurement of reaction times under the Windows operating system that have to be satisfactorily resolved in the future, and no implementation opportunities for other operating systems, such as Mac OS or UNIX, exist. However, in the present studies sufficient progress has been made so that the further development of the dual-task approach in measuring cognitive load in learning with complex multimedia systems seems worthwhile.

From a more theoretical perspective, a central issue that has to be resolved is the simultaneous test of differences in learning outcomes (e.g., modality effects) on the one hand, and of secondary task performance on the other. As a result of the experiments presented in this paper, we obtained empirical evidence for the convergent validity of the underlying cognitive load hypothesis. In former experiments (Brünken & Leutner, 2001; 2002), in which the same learning system, the same knowledge acquisition tests, and comparable participants had been used, we found modality effects on knowledge acquisition. However, using conventional analysis methods from experimental psychology in the present experiments, we had to choose a all-within-subjects design to compensate for individual differences in reaction times. This design did not allow for the testing of modality effects on knowledge acquisition within participants, because test items were not directly matched with screen pages. Implementing the assessment tool in a between-subjects design would necessitate a different and much more complex approach to analyzing reaction times. However, in such a design, our method of measuring cognitive load should be capable of explaining also other differences in knowledge acquisition, for example, between novices and experts, or differences caused by learning material redundancy (Mayer, 2001), which can also be interpreted as cognitive load effects.

Further research is also necessary with respect to the task types used for the secondary task. Until now, the secondary tasks we used involved very simple visual perception that did not require any elaborate cognitive processing. A systematic variation of the task types would not only be helpful for a validation of the present findings, but would also give a more precise indication of the type of cognitive resources that is used for processing multimedia learning materials.

Finally, in the present studies, we assessed differences in visual cognitive load produced by different presentation forms of computer-based learning materials. However, with our approach, the auditory cognitive load produced by these systems can also be assessed through the use of an auditory secondary task. Actual studies, such as by Moreno and Mayer (2000), show that the implementation of additional auditory information (such as background music) in a multimedia learning system reduces knowledge acquisition. The authors interpret this effect as an auditory cognitive load effect. We are currently working on the implementation of a auditory secondary task for our assessment tool to test for related differences in reaction times during the processing of auditory information in a dual task condition.

The direct measurement of cognitive load in multimedia learning will, for the first time, allow for more concrete conclusions to be made regarding working memory utilization in learning with multimedia, and has a high potential to provide further empirical evidence for a multimedia dual-processing theory.

References


---

Dr. Roland Brünken

Independent Research Group
Erfurt University
Postbox 900221
D-99105 Erfurt
Germany
Tel.: +49 363 7371405
Fax: +49 363 7371031
E-mail: roland.bruenken@uni-erfurt.de

© 2002 Hogrefe & Huber Publishers