

Measurement of cognitive load in multimedia learning: a comparison of different objective measures

Andreas Korbach¹ · Roland Brünken¹ · Babette Park¹

Received: 6 May 2015 / Accepted: 21 April 2017 / Published online: 28 April 2017
© Springer Science+Business Media Dordrecht 2017

Abstract Different indicators are interesting for analyzing human learning processes. Recent studies analyze learning performance in combination with cognitive load, as an indicator for learners' invested mental effort. In order to compare different measures of cognitive load research, the present study uses three different objective methods and one subjective method, reviewing the seductive details effect in a computer-based multimedia learning instruction. An experimental two-group design ($N = 50$) was used, with exposure to seductive details during learning as the between group factor. Eye movements were analyzed concerning the indicated cognitive activity and cognitive load was measured by the rhythm method (Park 2010; Park and Brünken 2015), the index of cognitive activity (ICA) (Marshall 2007), as well as by subjective ratings of mental effort and task difficulty (Paas 1992). Results confirm the seductive details effect for learning success with a decrease in retention and comprehension performance, an increase in total cognitive activity indicated by eye movements and significant higher cognitive load, indicated by rhythm method. The ICA values and the subjective ratings on mental effort and task difficulty show no difference in cognitive load between the groups. The results provide evidence of the suitability of different objective measures for a direct and continuous cognitive load assessment in multimedia learning. Further, the results show the benefit of combining different methods to gain detailed insight into information processing while learning with multimedia as well as a differentiated access to the single cognitive load factors.

Keywords Cognitive load measurement · Rhythm-method · Eye-tracking · Index of cognitive activity

✉ Andreas Korbach
a.korbach@mx.uni-saarland.de

¹ Department of Education, Saarland University, Campus A4 2, 66123 Saarbrücken, Germany

Introduction

In order to efficiently analyze human learning processes, many different factors must be considered. Within cognitive theories of learning and instruction, cognitive load has been stressed as an important factor for successful and efficient learning. This is specifically true when learning from multimedia (for an overview on the historical development of cognitive load theory see Moreno and Park 2010; Sweller et al. 2011). Methods of objective cognitive load measurement are of great importance for research on learning and instruction, especially methods that are objective, direct, reliable, and measure cognitive load while it is occurring (Brünken et al. 2010). These objective measures contrast subjective measures that assess cognitive load after engagement with materials. Objective measures like the rhythm method (Park and Brünken 2015), eye tracking or pupillometric analysis provide continuous information throughout the entire learning process. There is no interruption of the learning process to assess cognitive load and cognitive activity, as is the case for the frequent use of subjective ratings. In addition, eye-tracking analysis is a highly detailed measure, which provides a deep insight into human information processing regarding the allocation of visual attention and the cognitive activity on processing and integrating the related text and picture information of a multimedia-learning instruction. That leads to another big issue of cognitive-load research, the differentiating measurement of the different types of cognitive load and the identification of the related cognitive processes. Recent studies on subjective cognitive-load ratings already demonstrated the possibility to differentiate between different cognitive-load factors (Leppink et al. 2013, 2014). In the present study, the data of the selected objective methods should in combination also allow conclusions not only about the total amount of cognitive load but also about the related cognitive activity and the cognitive processes that cause the corresponding cognitive load. To this end, the present study compares three methods of objectively assessing cognitive load as well as cognitive activity and one subjective method.

Theoretical framework and predictions

Cognitive Load Theory (Choi et al. 2014; Plass et al. 2010; Sweller et al. 2011) is a commonly used theoretical framework in empirical research on learning and instruction. One basic assumption of cognitive load theory is that the available cognitive capacity is limited by working memory capacity and that knowledge acquisition is fostered by an efficient use of available resources. Another basic assumption is that there are three components (1) intrinsic, (2) extraneous, and (3) germane cognitive load (Sweller et al. 1998) that add up to the total amount of cognitive load (Brünken et al. 2010; Moreno and Park 2010; Park 2010). Intrinsic cognitive load is determined by the given complexity of the learning task and results from element interactivity. Element interactivity is defined by the number of interacting information elements that belong to the learning task. The more complex the learning task, the higher the intrinsic cognitive load. Extraneous cognitive load is caused by the format of learning instruction and hinders the learning process. An increase in extraneous cognitive load is indicated by an increase in additional information processing that does not foster mental model construction. The better the format of instruction, the lower the amount of extraneous cognitive load (Brünken et al. 2003; Paas

et al. 2003b). Germane cognitive load is the amount of load dedicated to relevant information processing and mental model construction resulting in higher learning performance.

An updated model of cognitive load theory (Choi et al. 2014; Kalyuga 2011) considers only two of the three components: intrinsic and extraneous load. The deletion of germane load was due to the close relationship between intrinsic and germane cognitive load, which manifested in the inability to separate a unique contribution of each factor to the overall cognitive load. Germane cognitive load is now considered as germane resources that reflect the amount of working memory capacity dedicated to learning. Furthermore, the cognitive load induced by relevant information processing and learning activities is incorporated into the intrinsic cognitive load factor. Due to the ongoing discussion within the literature, the present paper considers both cognitive load concepts that will be discussed according to the results.

Methods of cognitive load measurement

One goal of cognitive-load research is to examine learners' cognitive load while working with different forms of instructions and to derive instructions that efficiently support complex knowledge acquisition. Measurement of cognitive load plays a crucial role in cognitive-load research and therefore the development of practical implications for efficient instructional design. One option to measure cognitive load is the analysis of learning performance and knowledge acquisition as an indirect indicator of cognitive load. However, the use of direct measurement methods, especially of methods for continuous measurement during the learning activities, provides a more robust measurement, as increasingly discussed within the literature (e.g. Brünken et al. 2003, 2010a, b; Paas et al. 2003a, b). There are two kinds of methods that can be used to directly measure cognitive load: (1) objective and (2) subjective.

Recent objective methods for measuring cognitive load

Objective methods of cognitive-load measurement include the analysis of secondary-task performance and behavioral data such as pupil dilation, as well as the analysis of cognitive activity indicated by eye-tracking data (for a detailed classification see Brünken et al. 2010a, b). Each of these mentioned objective methods is essential due to the continuous nature of the measurement. Objective measures allow producing highly detailed data and measure cognitive load continuously during the learning process. The objective methods examined within the present study are: (1) the rhythm method, (2) eye-tracking analyses and (3) index of cognitive activity (ICA).

The rhythm method is a dual-task analysis with a rhythmic foot-tapping task as secondary task (Park and Brünken 2015). It measures cognitive load in a direct and continuous way using an intra-individual behavioral measure. Dual task means that the participants have to perform two concurrent tasks and performance of both tasks is measured. In research on learning and instruction the first task is the learning task. The established secondary tasks are mostly fulfilled by auditory or visual cues in the learning instruction and use reaction time on these up-coming cues within the learning material as a performance measure for cognitive load. The dual-task method thereby allows direct measurement of cognitive load. A series of studies provided evidence that secondary-task performance produces reliable and valid results for cognitive load measurement (e.g. Brünken et al. 2004; DeLeeuw and Mayer 2008). The rhythm method is a special kind of dual-task that uses no external cues and therefore avoids sensory interferences between the

learning instruction and the secondary task (Park and Brünken 2015). This new method was validated in a study where the participants' primary task was to work with a multimedia-learning program and the secondary task was to tap a previously presented and practiced rhythm with their foot. Because both tasks require cognitive capacity, the performance of the secondary task provides information about the amount of available cognitive capacities. For example, better performances on the secondary task indicate that less cognitive capacities are consumed by the primary task.

In more detail, the rhythm method introduced by Park and Brünken (2015) is a foot-tapping task that uses a four–four time rhythm in the form of: tap–tap–pause–pause/tap–tap–pause–pause/... and so forth. The tapping precision was calculated as the individual rhythm deviation and has been proven to be a valid indicator for cognitive load during learning in a multimedia environment. The study showed significant differences in cognitive load concerning the seductive details effect (with vs. without seductive details) and the modality effect (visual-only text vs. narration). Less precision in rhythm tapping was shown for groups with cognitively high loading versions (with seductive details or visual-only text material, respectively) of the multimedia learning instructions when compared to groups with low loading versions (material without seductive details or including narration, respectively).

Two areas of concern for the use of dual-task methodology have to be considered in research that are the additional cognitive load and the sensitivity of the measure. As it is true for all secondary-task measures, the secondary task induces additional cognitive load and therefore could impair learning. For the rhythm method, so far, there are no studies, which examine the induced cognitive load from the foot-tapping task and whether the task impairs learning. Concerning the sensitivity of the measure, dual-task methodology is not appropriate to measure the load of a single cognitive load aspect. However, given the additivity hypothesis of cognitive load theory and the theoretical explanation that rhythm production is specifically dealing with inhibition processes associated with executive control (Park and Brünken 2015), the sensitivity of this method should be associated with a general sensitivity for total cognitive load.

The established eye-tracking indicators for cognitive activity related to cognitive load are based on fixations (Haider and Frensch 1999; Jarodzka et al. 2010; Laeng et al. 2011). Eye-tracking analysis offers many different measures that provide detailed information about information processing including the allocation of attention and the cognitive activity spent for information processing. Measured alone without using additional indicators, eye-tracking analysis provides only information about the visual processing while learning. However, in combination with measures of learning performance it also provides information about the focus of cognitive activity and mental effort (Folker et al. 2005; Mayer 2010). Measurements of total fixation time, the total number of fixations, the time to the first fixation as well as the transitions between corresponding text and picture information, provide information about learners' focus of attention during visual processing. As is indicated by several studies, there is evidence of a strong positive relationship between eye-movement measures like total fixation time and cognitive load, with a long fixation time indicating high cognitive load (Just and Carpenter 1976; Holmqvist et al. 2011; Rayner 1998). For example, during learning with text and graphics, total fixation time on the relevant graphic is hypothesized to induce cognitive processing and serve as a measure of cognitive performance and load (Korbach et al. 2016; Mayer 2010; Park et al. 2015b, 2015c; Rayner et al. 2007; Reichle et al. 2003).

Another measure of perceptual processing is the total number of fixations. Just as the total fixation time, the total number of fixations can be hypothesized as an indicator of

cognitive activity on information processing. For both measures, it is assumed that long fixation times and a large number of fixations indicate high cognitive activity that is related to high cognitive load (Canham and Hegarty 2010; De Koning et al. 2010). Another measure, which is also related to cognitive activity on information processing, is the number of transitions between corresponding text and picture information. In more detail, transitions are assumed to represent integrative cognitive processes and to be directly related to schema construction out of textual and pictorial information. Thereby a large number of transitions are assumed to be associated with high cognitive engagement in integrating verbal and pictorial information (Johnson and Mayer 2012; Schmidt-Weigand et al. 2010).

The ICA introduced by Marshall (2007) is based on the short and large reactions in pupil dilation due to changes in cognitive activity that are identified by wavelet analysis and automatically calculated by designated analytics software (EyeWorks™, EyeTracking Inc.). The advantage of the ICA is that the large dilations in pupil size due to effects of illumination are automatically identified and excluded from analysis. Marshall, Pleydell-Pearce, and Dickson (Marshall et al. 2002) demonstrated that the ICA is not influenced by illumination and that the ICA reliably indicates cognitive load under high and low illumination conditions. Some recent studies support the usability of the ICA for driving tasks (Demberg et al. 2013; Schwalm et al. 2008) or mathematical tasks (Schwalm 2009), but not for learning within a multimedia instruction. Debue and van de Leemput (2014) used the ICA for cognitive load measurement concerning information processing with different types of online newspapers. However, the ICA values were not conform to the results of subjective cognitive-load ratings or performance measures.

Subjective methods for measuring cognitive load

Subjective methods commonly are ratings of perceived task difficulty, engagement or effort, which are completed by research participants. Two examples are the widely used subjective rating scale introduced by Paas (1992) and the NASA Task Load Index (NASA TLX (Hart and Staveland 1988)). The advantage of subjective methods is that ratings provide valid information about the individual experience. In addition, subjective rating scales are very easy to implement and can be used in different learning contexts with diverse learning contents and groups of participants. Several studies show the suitability of rating scales for cognitive-load measurement (Gopher and Braune 1984; Paas and Van Merriënboer 1994). However, rating scales are criticized because of methodological problems concerning the quality criteria of objectivity, validity, and reliability (Brünken et al. 2003, 2010a, b; Clark and Clark 2010; Moreno 2006). In particular, evidence of content validity is critical, as it is difficult to distinguish between different types of cognitive load with a universal subjective rating scale. Nevertheless, subjective rating scales are assumed to be the only way to differentiate the single cognitive load aspects (Leppink et al. 2013, 2014). Another disadvantage is that ratings are in general requested subsequent to the learning activities, which have to be evaluated by the subjects. Therefore, rating scales provide no continuous information about the actual cognitive load during the learning process. Given these drawbacks, subjective ratings of cognitive load are often used in research examining learning and instruction. Specifically, the ratings for task difficulty seem to provide valid information about the intrinsic cognitive load based on element interactivity (Ayres 2006). In the present study, Paas's (1992) subjective rating scale is used, as it is the most commonly used subjective method for measuring cognitive load.

Variations of cognitive load caused by instructional design

To analyze the suitability of different cognitive load measures, it is necessary to vary cognitive load in an experimentally controlled way. In order to manipulate cognitive load in the present study, seductive details were used to induce higher cognitive load during the learning process. Seductive details consist of additional information, which is highly interesting, but irrelevant for the learning goal and have been shown to have a detrimental effect on learning performance (e.g. Garner et al. 1989). The theoretical assumptions underlying this effect rely on cognitive theory of multimedia learning (Mayer 2001, 2005, 2009) and cognitive load theory (Plass et al. 2010; Sweller 1999; Sweller et al. 2011). The central assumption is that differences in performance are caused by different amounts of resource consumption when learning with or without seductive details, with higher extraneous cognitive load induced by the additional processing of the irrelevant information. In the present study, differences in cognitive load will be measured with both, the objective and subjective, load measures.

More specifically, as the extraneous cognitive load factor is assumed to be crucial for several multimedia effects on learning the present study was designed to manipulate extraneous load in order to compare the different methods of cognitive load measurement. Intrinsic cognitive load is assumed to be constant as the additional information is irrelevant for the learning goal. The seductive details are easy to understand, can be processed independent from the relevant information and therefore should not affect the element interactivity. Germane cognitive load imposed by processing relevant information is assumed to decrease in the seductive details version due to a synergetic effect. An increase in extraneous cognitive load induced by processing additional irrelevant information is expected to be accompanied by a decrease for relevant information processing, as a study by Park et al. (2015) already showed. The assumptions concerning germane and intrinsic cognitive load are true only for the former model of cognitive load theory (Plass et al. 2010; Sweller et al. 2011). With regards to the updated model (Choi et al. 2014), which only considers intrinsic and extraneous cognitive load, extraneous cognitive load is assumed to increase and therefore intrinsic cognitive load is assumed to decrease for the seductive details version. And in the case of the recent model of cognitive load theory, the revised intrinsic cognitive load factor now includes the cognitive load caused by processing relevant information and learning.

Additional variations of cognitive load can be derived from individual learner characteristics such as prior knowledge, spatial ability or working memory capacity (Korbach et al. 2016; Magner et al. 2014; Park et al. 2014, 2015c; Sanchez and Wiley 2006; Smith and Ayres 2013) that are assumed to affect the processing of relevant as well as irrelevant information or the individual task complexity. These are therefore controlled in the present study.

Goal of the present study

The present study will compare results of four methods concerning the explanatory power about cognitive activity and corresponding cognitive load (1) the rhythm method, (2) eye-tracking analysis, (3) ICA, and (4) subjective cognitive load ratings. Each method will be examined in its suitability to capture information about cognitive load within a multimedia learning content. This will be done by measuring cognitive load and cognitive activity during information processing that is induced by seductive details within multimedia

instruction. It is assumed that the presence of seductive details will increase total cognitive activity and that all measures indicate higher cognitive load for the seductive details group. Furthermore, it is assumed that the increase in total cognitive load is due to an increase in extraneous cognitive load. Eye-tracking analysis therefore is assumed to indicate a decrease for the processing of relevant information (Park et al. 2015) in combination with a large amount of additional information processing dedicated to the seductive details information which will in turn increase the cognitive activity for total information processing and result in a decrease in learning performance.

Method and data sources

Participants and design

In order to compare the different methods of cognitive load measurement, a two-group design was used with seductive details as between factor (with vs. without seductive details). All participants ($N = 50$) were university students (mean age = 22.24 years, $SD = 2.45$, 15 male, 35 female) and were randomly assigned to one of the two groups.

Materials

Both groups worked with a multimedia-learning program concerning the biological content of ATP Synthase with a pre-set, paced learning time. The program consists of 11 slides presenting relevant learning information in the format of illustrations and related text on the left side of each slide (see Fig. 1). The multimedia learning instruction was about the structure and function of the ATP Synthase, a cellular molecule responsible for synthesis of ATP. The relevant information was presented as a combination of static pictures and corresponding textual explanations. The objective of the learning task was to achieve a deep understanding of the molecule structure and the single steps associated with the process of ATP synthesis by integrating the verbal and pictorial representations. All participants were introduced to the learning objective at the beginning of the learning task. The material was successfully used in several other studies on multimedia learning and in validation studies on the rhythm method (Park and Brünken 2015; Park et al. 2011, 2015). On the one hand, the learning instruction was chosen due to the fact that the comparison of

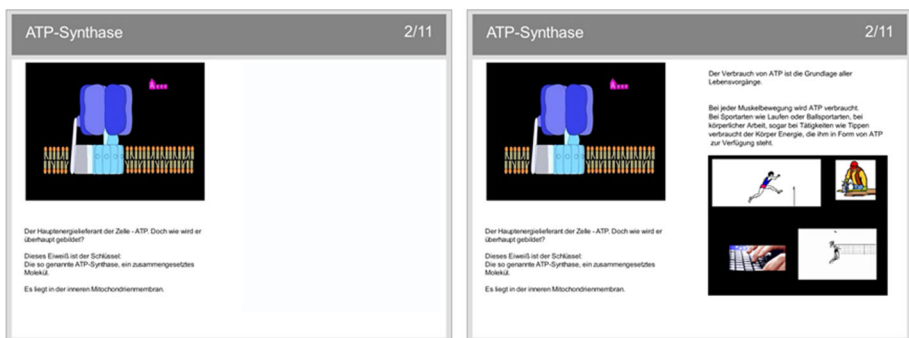


Fig. 1 Example slide of the learning instruction with and without seductive details

different cognitive load measures and the analysis of eye movements may be especially interesting for a learning instruction that showed controversial results concerning the measured cognitive load by the rhythm method and subjective ratings. On the other hand, the rhythm method can at least be assumed as a valid method of cognitive load measurement for this chosen learning instruction.

The experimental group worked with the seductive details version of the learning program that was already used in previous studies (e.g. Park et al. 2015). In this version additional, highly interesting, but irrelevant information is presented on 4 of the 11 slides in form of animated illustrations and related text on the right side of the slides (see Fig. 1). In contrast to the relevant information, seductive details provided information about the usefulness of ATP that was not part of the learning objective. According to former studies seductive details were chosen by the following aspects: interestingness, irrelevance, concreteness, conciseness, emotionality and reference to the relevant topic (Garner et al. 1989, 1992; Goetz and Sadoski 1995; Park et al. 2015; Sanchez and Wiley 2006).

Time-on-task was controlled in the way that all learners had the same amount of time to process the presented information on each slide. The time for each slide was an empirically tested required mean reading time for the seductive details version of the slide, so that the learners in the seductive details group had enough time to process all information. The overall time for the learning instruction was set to 11 min and 40 s.

Measures

Control measures

Working memory capacity, spatial ability, learning motivation, and prior knowledge served as control measures. Working memory capacity was measured by the numerical memory updating subtest of (Oberauer et al. 2000). Spatial ability was controlled by a standardized paper-folding and card-rotation test (Ekstrom et al. 1976). A revised short version of the 100-item Inventory of School Motivation (ISM; McInerney and Sinclair 1991) was used, *Cronbachs' α* = .83, to test participants' learning motivation. Finally, prior knowledge was measured by a questionnaire that included four multiple-choice and seven open-ended questions, *Cronbachs' α* = .72.

Learning performance test

Learning success was assessed with a learning performance test consisting of 12 items. All items met satisfactory item parameters with a difficulty index of $.20 < p_i < .80$. The differentiation between two levels of required cognitive processing was considered by using the two subscales retention and comprehension. The retention subscale included 5 items, 3 in multiple-choice format and 2 open-ended responses, with a *Cronbachs' α* of .73 (item examples: (1) “*The matrix is ...*”—*the inside of the mitochondrium*; the intermembrane space; an united cell structure in tissues; the space outside the mitochondrium; (2) *Describe the term “proton-motive force”*). The comprehension subscale included 7 items, 4 in multiple-choice format and 3 open-ended responses, with a *Cronbachs' α* = .75 (item examples: (1) “*What's the function of the ATP synthase's F0 complex?*”—transport of protons into the matrix; transport of protons into the intermembrane space; *the generation of proton-motive force*; the formation of the proton gradient; (2) “*Explain three requirements for the operational capability of the ATP synthase*”).

Rhythm method

The rhythm method, introduced by Park and Brünken (2015), was recorded with E-Prime psychology software in combination with a response box and the corresponding foot pedal. Prior to learning, participants were introduced to the rhythm method. First, they had to listen to the rhythm (Three example rhythm units: Tap–Tap–Pause–Pause, Tap–Tap–Pause–Pause, Tap–Tap–Pause–Pause, and so forth), then they had to tap the rhythm by accompanying the auditory example and afterwards they had to practice the rhythm alone. In order to pass the practice section, participants had to tap the rhythm for 60 s with less than 20% of deviation from the given rhythm. Recordings of these 60 s served as an individual rhythm-baseline for the calculation of precision during the learning task. The analysis of rhythm performance is based on the inter-tap interval that is the time between two taps. There is a short and a long rhythm component or in other words inter-tap interval. The short inter-tap interval is the “Tap–Tap” interval between the first two taps of each rhythm unit. The long inter-tap interval is the “Tap–Pause–Pause–Tap” interval from the second tap of the rhythm to the first tap of the next rhythm unit. For the further analysis all inter-tap intervals were accepted that were higher than 250 ms, which is the low cut-point for response time typical found in response time studies (e.g. Thorpe et al. 1996). Therefore, this is the shortest time interval participants could tap consciously. On the upper end of the measure, all inter-tap intervals were accepted if these were equal or less than the time interval for the whole rhythm unit of 2000 ms. (short inter-tap interval = 500 ms. “Tap–Tap” interval + long inter-tap interval = 1500 ms. “Tap–Pause–Pause–Tap” interval). Values longer than 1000 ms. were assigned to the long rhythm-component; values shorter than 1000 ms were assigned to the short rhythm-component. Thus, the valid inter-tap interval for the short rhythm component lies between 250 and 1000 ms and the valid inter-tap interval for the long rhythm component lies between 1000 and 2000 ms. The inter-tap intervals and the split point of 1000 ms were chosen according to several pilot studies, which demonstrated these criteria to be reliable indicators for rhythm performance (for a detailed description of the rhythm method see Park and Brünken 2015). For the rhythm method analysis, precision was calculated as an individual’s deviation from the mean rhythm values during learning. Separate analysis were conducted for both the short and long rhythm-component with a Guttman’s split-half coefficients of $r = .938$ and $r = .929$ respectively, providing evidence of an excellent internal consistency for both components (Huck 2012).

Eye-movement analysis

The participants’ eye movements were recorded with a remote eye-tracking system (Tobii TX300) while they worked on the learning program. The eye-tracking system is integrated in a 23-inch-TFT (1929 × 1080 pixel) monitor and operates with a sample rate of 300 Hz. Participants’ eye movements were analyzed with EyeWorks™ software and the standard fixation filter with a minimum duration of 100 ms and a maximum range of 30 pixels. The participants had to put their heads on a chin rest that was mounted on a table in front of the eye-tracking system in order to increase the precision of the recordings. Only participants with high calibration accuracy on all calibration points were accepted for the study. All participants were tested under the same lighting conditions in a laboratory without windows and influence of natural light. The Areas Of Interest (AOI) were defined for relevant text and relevant pictures on each slide of the learning instruction, as well as for the

seductive details text and pictures. The analysis of eye-tracking data focused on the total fixation duration and the total number of fixations on text and picture AOIs as well as on the transitions between text and picture AOIs. For the analysis of the learners' transitions from text to picture AOIs the number of transitions was counted automatically by analysis software. Eye movements on relevant information AOIs, including transitions between relevant text and picture AOIs, were assigned to processing of relevant information. Eye movements on seductive details information AOIs, including transitions between seductive details text and picture AOIs, as well as transitions between non-related relevant and seductive details AOIs, were assigned to processing of irrelevant information. For the analysis of the total information processing the amount of irrelevant information processing was added to the relevant information processing for the seductive details group.

ICA

The index of cognitive activity (ICA) introduced by Marshall (2007) was automatically calculated by EyeWorksTM-analysis software in accordance to the analysis for fixations.

Subjective cognitive load ratings

Total cognitive load was measured by subjective ratings of mental effort and task difficulty (Paas 1992). Participants were asked to rate their perceived cognitive load and task difficulty on a seven-point Likert scale. Cognitive load was assessed once immediately after the presentation of 3 slides containing seductive details (after slide 4 of 11) and again upon completion of the lesson (after slide 11 of 11). The first point in time for the subjective rating of cognitive load was chosen according to the structure of the lesson to minimize an interruption of the learning process. At each time point, learners were asked to estimate their cognitive load and task difficulty by clicking on the rating that best completed the following statement "While working on the learning material my mental effort was..." / "How easy or difficult was it to understand the learning material?" with the ratings ranging from "very low/easy", "low/easy", "rather low/easy", "neither low/easy nor high/difficult", "rather high/difficult", "high/difficult" to "very high/difficult". As several studies show that the timing of cognitive load ratings can affect the results (Schmeck et al. 2015; Van Gog et al. 2012), the subjective cognitive load ratings are analyzed and reported separately.

Procedure

Participants were randomly assigned to the experimental or control group and tested in individual sessions. The session began with the test for learning motivation, followed by the working memory capacity test, the test of spatial ability and the prior knowledge assessment. Each participant was then instructed in the rhythm method and had time to practice the rhythm. Participants who passed the practice section went on to the ATP learning program, the experimental group worked with a seductive details version and the control group worked with a no-seductive details version of the multimedia learning instruction. While they were learning, their performance in the rhythm method was recorded by E-Prime software and the participants' eye movements were automatically recorded by an eye tracker. The self-report scale for cognitive load was presented once during the learning process after slide four and again at the end of the learning session. Finally, participants completed the learning performance test. The entire session lasted for approximately 90 min.

Results and conclusions

All data analyses were conducted with the between subject factor seductive details (with vs. without). Separate analyses were used for learning performance and the different cognitive load measures, which were grouped together with respect to their inter-correlations. The first MANOVA examined learning performance, which included retention and comprehension. A second MANOVA was conducted for all eye movements together with subjective cognitive load ratings because of a high correlation between the ratings of task difficulty after slide four and the number of relevant transitions, $r = -.46$, $p = .001$. A final analysis, ANOVA, was conducted on data associated with the rhythm method. Only participants with less than 25% missing values in the eye-tracking recordings and throughout secondary task performance were included for the analysis of eye movements and rhythm performance.

Control measures

The two groups did not differ significantly concerning the control measures spatial ability, $F(1, 48) = 1.20$, *n.s.*, prior knowledge, $F < 1$, working memory capacity, $F(1, 48) = 2.71$, *n.s.*, or learning motivation, $F < 1$. In addition, the first slide of the learning program that shows only text and that is the same for all participants was used to control the variables of eye movement and the rhythm method. There were no significant differences between the groups concerning gaze duration, ICA, the number of fixations and the total fixation duration, $F_s < 1$. The deviation in rhythm method showed a significant difference between the groups for the short rhythm-component, $F(1, 48) = 6.06$, $p = .017$, $\eta^2 = .12$, but not for the long rhythm-component, $F < 1$, so the short rhythm-component was excluded from further analysis. To control illumination effects for ICA, a *t* test for dependent variables was conducted between seductive details and relevant information AOF's. Results show no significant effect concerning the ICA values, $t(22) = .59$, *n.s.*, between relevant and seductive details AOF's.

Learning performance

The first MANOVA for learning performance indicated a seductive details effect, $F(3, 46) = 4.42$, $p = .008$, $\eta^2 = .22$. Univariate testing showed a significant decrease in comprehension, $F(1, 48) = 6.01$, $p < .05$, $\eta^2 = .11$, and retention, $F(1, 48) = 8.82$, $p = .005$, $\eta^2 = .16$, for the seductive details group (see Table 1).

Cognitive load measures

The second MANOVA was conducted for all measures of eye movement including ICA and the subjective ratings of cognitive load. The analysis for the variables of eye

Table 1 Means and standard deviations of learning performance for both conditions

	Without seductive details $n = 26$		With seductive details $n = 24$	
	M	(SD)	M	(SD)
Comprehension (%)	62.86	(15.76)	49.28	(23.29)
Retention (%)	62.36	(20.28)	43.9	(23.65)

Note. M mean, SD standard deviation

movement was conducted for all AOIs over all seductive details slides. All variables were compared for differences concerning the eye movements on the relevant AOI's only and on the relevant AOI's including the seductive details AOI's for the experimental group. The MANOVA revealed an effect of seductive details, $F(23, 20) = 43.81, p < .001, \eta^2 = .98$.

Gaze

Univariate testing for the total gaze duration over all seductive details slides show a seductive details effect on the total gaze duration of the relevant picture AOI's, $F(1, 42) = 26.93, p < .001, \eta^2 = .39$, and the relevant text AOI's, $F(1, 42) = 10.22, p = .003, \eta^2 = .20$, with significantly shorter gaze durations for the seductive details group. Adding the total gaze duration for the seductive details AOI's eliminates the seductive details effect for picture AOI's, $F(1, 42) = 2.67, n.s.$, and turns the results concerning the text AOI's towards the seductive details group, with significantly longer total gaze duration across all AOI's, $F(1, 42) = 6.62, p = .014, \eta^2 = .14$. However, there was no significant difference between the groups concerning the total gaze duration across all AOI's with text and pictures, $F(1, 42) = 2.16, n.s.$, showing that both groups in total gaze spent a comparable amount of time on processing information (see Table 2).

Fixation duration

Univariate testing for the total fixation duration over all seductive details slides show a seductive details effect on the total fixation duration of the relevant picture AOIs, $F(1, 42) = 16.41, p < .001, \eta^2 = .28$, with significantly shorter fixation times for the seductive details group, but no effect on the relevant text AOI's, $F < 1$. Adding the total fixation duration for the seductive details AOI's eliminates the seductive details effect for picture AOI's, $F(1, 42) = 2.77, n.s.$, and turns the results concerning the text AOI's towards the seductive details group, with significantly longer total fixation duration across all AOI's, $F(1, 42) = 6.37, p = .015, \eta^2 = .13$. There was no significant difference between the groups concerning the total fixation duration across all AOI's with text and pictures, $F(1, 42) = 1.29, n.s$ (see Table 3).

Table 2 Means and standard deviations of total gaze duration in seconds for all seductive details slides

	Without seductive details <i>n</i> = 26		With seductive details <i>n</i> = 24	
	M	(SD)	M	(SD)
Total gaze duration on relevant picture AOIs	91.59	(36.74)	44.80	(20.95)
Total gaze duration on relevant text AOIs	229.29	(47.93)	188.37	(36.15)
Total gaze duration on all picture AOIs	91.59	(36.74)	75.83	(26.43)
Total gaze duration on all text AOIs	229.29	(47.93)	262.35	(36.49)
Total gaze duration on all AOIs	320.88	(46.32)	338.17	(29.96)

Note. *M* mean, *SD* standard deviation

Table 3 Means and standard deviations of total fixation duration in seconds for all seductive details slides

	Without seductive details <i>n</i> = 26		With seductive details <i>n</i> = 24	
	M	(SD)	M	(SD)
Total fixation duration on relevant picture AOIs	78.26	(43.26)	36.00	(22.86)
Total fixation duration on relevant text AOIs	164.35	(62.17)	151.84	(53.92)
Total fixation duration on all picture AOIs	78.26	(43.26)	59.29	(31.34)
Total fixation duration on all text AOIs	164.35	(62.17)	211.99	(62.98)
Total fixation duration on all AOIs	242.61	(84.84)	271.28	(82.30)

Note. *M* mean, *SD* standard deviation

Number of fixations

Univariate testing for the number of fixations over all seductive details slides show a seductive details effect on the total number of fixations on the relevant picture AOIs, $F(1, 42) = 18.27, p < .001, \eta^2 = .30$, with significantly fewer fixations in the seductive details group, but no effect for the relevant text AOI's, $F(1, 42) = 1.62, n.s.$ Adding the total number of fixations on the seductive details AOI's eliminates the seductive details effect for picture AOI's, $F < 1$, and there is no effect for the total number of fixations on all text AOI's, $F(1, 42) = 1.26, n.s.$, or across all AOI's with text and pictures, $F < 1$ (see Table 4).

Transitions

Univariate testing for transitions was conducted for the number of integrative transitions from text to the related relevant picture AOI's, and the number of all transitions between text and picture AOI's. Including integrative transitions between seductive details AOI's and non-integrative transitions between non-related text and picture AOI's. Furthermore, the total duration of the fixations on the picture AOI's subsequent to a relevant integrative transition was part of the analyses. Results show a seductive details effect on the number of transitions between the relevant text and picture AOI's, $F(1, 42) = 10.57, p = .002, \eta^2 = .20$, and for the total fixation duration on the picture AOI's subsequent to integrative transitions, $F(1, 42) = 32.16, p < .001, \eta^2 = .43$, with significantly fewer transitions and

Table 4 Means and standard deviations of total fixation number for all seductive details slides

	Without seductive details <i>n</i> = 26		With seductive details <i>n</i> = 24	
	M	(SD)	M	(SD)
Total fixation number on relevant picture AOIs	150.41	(58.55)	87.14	(37.32)
Total fixation number on relevant text AOIs	414.23	(209.37)	348.82	(118.92)
Total fixation number on all picture AOIs	150.4	(58.55)	146.95	(59.96)
Total fixation number on all text AOIs	414.2	(209.37)	473.00	(129.21)
Total fixation number on all AOIs	564.64	(222.66)	619.95	(174.09)

Note. *M* Mean, *SD* standard deviation

Table 5 Means and standard deviations for the transitions between text and picture AOI's and the total fixation duration subsequent to a transition for all seductive details slides

	Without seductive details $n = 26$		With seductive details $n = 24$	
	M	(SD)	M	(SD)
Transitions between relevant AOI's (N)	30.73	(11.76)	19.45	(11.24)
Transitions between all AOI's (N)	30.73	(11.76)	41.45	(18.11)
Fixation Duration after Transitions across Picture AOI's (s)	75.32	(33.26)	30.41	(16.51)

Note. *M* mean, *SD* standard deviation

shorter fixation durations for the seductive details group. Adding the total number of transitions between the seductive details AOI's and between relevant and seductive details AOI's turns the seductive details effect around for the total number of transitions, $F(1, 42) = 5.45$, $p = .024$, $\eta^2 = .12$. With the seductive details group having significantly more transitions (see Table 5).

ICA

Univariate testing for the mean ICA values across all seductive details slides show no seductive details effect for relevant picture AOIs, relevant text AOI's and all picture AOI's, all $F_s < 1$, all text AOI's, $F(1, 42) = 2.70$, *n.s.*, or the total ICA across all AOI's, $F < 1$ (see Table 6).

Subjective ratings

Univariate testing for the subjective cognitive load ratings indicate no effect of seductive details for the rating of mental effort after slide four, $F(1, 42) = 2.40$, *n.s.*, the rating of task difficulty after slide four, or the ratings of mental effort and task difficulty after the final slide of the learning instruction, all $F_s < 1$ (see Table 7).

Rhythm method

A final ANOVA was conducted for the rhythm method. The results show a seductive details effect for all seductive details slides with a significant decrease in dual-task

Table 6 Means and standard deviations of ICA for both conditions

	Without seductive details $n = 26$		With seductive details $n = 24$	
	M	(SD)	M	(SD)
Mean ICA on relevant picture AOIs	.36	(.22)	.38	(.19)
Mean ICA on relevant text AOIs	.43	(.22)	.37	(.22)
Mean ICA on all picture AOIs	.36	(.22)	.40	(.19)
Mean ICA on all text AOIs	.43	(.22)	.33	(.17)
Mean ICA on all AOIs	.39	(.20)	.37	(.18)

Note. *M* mean, *SD* standard deviation

Table 7 Means and standard deviations of cognitive load ratings for both conditions

	Without seductive details <i>n</i> = 26		With seductive details <i>n</i> = 24	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)
Mental effort rating slide 4 (max = 7)	5.18	(1.00)	4.68	(1.13)
Task difficulty rating slide 4 (max = 7)	3.82	(.85)	4.09	(1.27)
Mental effort rating slide 11 (max = 7)	4.86	(.99)	4.95	(1.00)
Task difficulty rating slide 11 (max = 7)	4.18	(1.22)	4.45	(1.18)

Note. *M* mean, *SD* standard deviation

performance for the seductive details group, $F(1, 44) = 4.10$, $p < .049$, $\eta^2 = .09$. The decrease in dual-task performance was indicated by a larger increase of the mean deviation from the given long rhythm component for the seductive details group ($M = 141.68$ ms, $SD = 50.78$ ms) in contrast to the group without seductive details ($M = 113.49$ ms, $SD = 43.32$ ms).

Correlations for learning performance

The correlations between cognitive load measures and learning success show a significant relation between the results of the rhythm method and comprehension, $r = -.48$, $p = .001$, as well as retention, $r = -.35$, $p = .02$, with a decrease in learning performance and an increase in rhythm deviation. The subjective ratings for task difficulty show significant relationships with comprehension, $r = -.48$, $p = .000$, after slide four and $r = -.32$, $p = .025$, after the last slide of the learning instruction. A similar pattern is seen for retention after slide four, $r = -.46$, $p = .001$, and after the last slide, $r = -.39$, $p = .005$, with a decrease in learning performance and an increase in the rating of task difficulty. There were no significant relations between the mental effort ratings and learning success.

The eye-movement measures of processing relevant information show significant relationship with comprehension and the number of fixations on the relevant text AOIs, $r = .31$, $p = .040$, the number of fixations on the relevant picture AOIs, $r = .37$, $p = .010$, and the fixation duration subsequent to relevant transitions, $r = .36$, $p = .010$. For retention the results show significant relationship with the number of fixations on the relevant picture AOIs, $r = .29$, $p = .046$, the fixation duration on relevant picture AOIs, $r = .35$, $p = .016$, the number of relevant transitions, $r = .29$, $p = .047$, and finally, the fixation duration subsequent to relevant transitions, $r = .42$, $p = .003$. All correlations show an increase in information processing correlated with an increase in learning performance.

Correlations for cognitive load measures

The correlations between the different cognitive load measures show no relation with the results of the rhythm method. However, a significant relationship is seen for the ratings of task difficulty after slide four and the number of relevant transitions, $r = -.46$, $p = .001$, with a decrease in processing relevant information and an increase in the rating of task difficulty. Furthermore, there are significant positive correlations for the first ratings of task

difficulty and the second ratings of task difficulty, $r = .56, p = .000$, the second ratings of task difficulty and the second mental effort ratings, $r = .51, p = .000$, the first rating of mental effort and the second ratings of mental effort, $r = .69, p = .000$, as well as for the first ratings of mental effort and the second ratings of task difficulty, $r = .38, p = .006$.

Overall, the results confirm the detrimental effect of seductive details on learning performance as well as on perceptual processing. The results of the rhythm method clearly support the assumption of an increase in cognitive load by seductive details. Furthermore, the different cognitive load measures show different results, the decrease in dual-task performance indicates higher cognitive load for the seductive details group, whereas the index of cognitive activity, as well as the ratings of mental effort and task difficulty show no significant differences between the groups. The results for the eye-movement measures in sum indicate lower cognitive activity for the seductive details group concerning the processing of the relevant pictorial information, but an increase in cognitive activity concerning the total processing of the relevant and the additional seductive details text information. Adding the eye movements dedicated to the processing of the seductive details information raises the mean values for all eye-movement measures and equals the seductive details effect for relevant picture information. The mean values for cognitive activity dedicated to the processing of relevant and seductive details information indicate a higher total cognitive activity for the seductive details group with significant differences concerning the total fixation duration on text AOI's and the total number of transitions. Therefore, these results also support the assumption of an increase in cognitive load by seductive details.

Discussion

The overall results of rhythm method and eye-movement analysis confirm our theoretical assumptions and support the cognitive load explanation of the seductive details effect with an increase in extraneous cognitive load due to the processing of the additional information. The study also demonstrates the suitability and limits of different objective measures for a direct and continuous cognitive load measurement in multimedia learning.

The first objective cognitive load measure to be discussed is the rhythm method. As the results of the rhythm method show negative correlations with both comprehension and retention, the method can be assumed to measure learning relevant cognitive load and to be sensitive for extraneous cognitive load manipulations. Furthermore, the results are in line with both models of cognitive load theory (Choi et al. 2014; Plass et al. 2010; Sweller et al. 2011) as a high extraneous cognitive load is assumed to decrease learning performance. However, the results of the rhythm method cannot differentiate between different factors of cognitive load. The only viable conclusion is that participants in the seductive details group had less available cognitive capacity to handle the dual-task and in general that can be due to an increase in total cognitive load.

The second objective measure to be discussed is the analysis of eye movements. At first, eye movements are assumed to represent the focus of attention and cognitive activity concerning information processing. However, the positive correlations for learning performance and processing of relevant information in the present study support the assumption that eye movements also provide information about cognitive load dedicated to learning (Folker et al. 2005; Mayer 2010). The eye-tracking data show a perfunctory processing of the relevant information and a large amount of irrelevant information

processing for the seductive details group with an increase in total information processing. Again the results are in line with both models of cognitive load theory (Choi et al. 2014; Plass et al. 2010; Sweller et al. 2011) as the decrease in processing of relevant information can be assumed to reflect a decrease in cognitive processes, explained by germane cognitive load in the former cognitive load theory and by intrinsic cognitive load in the updated model. The additional processing of irrelevant information leads to an increase in total information processing in the seductive details group that can further be assumed to reflect cognitive processes dedicated to the extraneous cognitive load factor. However, the meaningful negative correlations between irrelevant information processing and learning performance were missing in the present study. A limitation of this eye-movement analysis probably lies within the presentation format. In the present study, the analysis focused on learning slides that presented relevant and irrelevant information on the same slides and in the same modality. Therefore, a differentiation between relevant and irrelevant information processing was possible that allowed an interpretation respective the cognitive activity dedicated to the single cognitive load factors. Further studies should investigate contrasting presentation formats, by investigating for example the modality effect with narrated text presentation and check for similar possibilities to use eye-movement analysis to make cognitive activity observable and to get a differentiating access to the corresponding cognitive load factors.

The third objective cognitive load measure to be discussed is the ICA. The results from the ICA indicated no differences between the groups and there were no significant correlations to learning success, the results of the rhythm method or the subjective ratings of task difficulty or mental effort. One possible explanation is that the results of the ICA were caused by the task demands according to visual information processing and that the ICA was not sensitive to higher order cognitive processing like information integration and mental model construction. This assumption is supported by the results of the eye-movement analysis as significant higher cognitive activity for the seductive details group is only indicated by the total number of transitions. Another possible explanation is that the ICA is sensitive to cognitive load induced by multiple tasks or task complexity. In the present study, both groups had a primary learning task and a secondary tapping task with the same task complexity for the learning task, as defined by the element interactivity of the learning relevant information. In contrast, the studies by Marshall (2007) and Marshall et al. (2002) that were conducted to show the functionality of the ICA used an experimental setup, which compares different states of cognitive activation or manipulates cognitive load by adding another task in order to increase task complexity. Further research should investigate comparable variations in task complexity for learning tasks and manipulations of cognitive activation. However, the ICA shows no sensitivity for extraneous cognitive load manipulations in the present study and may not be suitable for this kind of multimedia research.

The results of the subjective measures within the present study provide theoretical and practical implications for the utility of subjective ratings. The ratings for task difficulty show significant negative correlations to learning success and eye-movement measures, however there is no difference between the groups. These results are in line with the results of Ayres (2006) and support the assumption that participants can reliably rate the individual task difficulty. In the present study, the results further support the assumption that the rating of task difficulty was not influenced by the between subjects factor and that intrinsic cognitive load resulting from element interactivity and task complexity was not influenced by the presentation of seductive details. The results support the suggested differentiation with ratings of task difficulty measuring intrinsic cognitive load only for the

former model of cognitive load theory (Plass et al. 2010; Sweller et al. 2011). According to the updated model (Choi et al. 2014) the rating of task difficulty seems to be insufficient to measure intrinsic cognitive load because the cognitive activity to handle the given task complexity is included in this factor and the participants should have rated a lower cognitive load for the seductive details version, as eye movements show a decrease of learning relevant information processing. Concerning the ratings of mental effort the results show no significant group differences and no correlations. One plausible explanation is that extraneous and germane cognitive load factors or the extraneous and the intrinsic cognitive load factors depending on the model of cognitive load theory were confounded in this item and the results of these ratings show the synergetic effect between relevant and irrelevant cognitive activity. The analysis of the eye movements show no significant difference between the groups in total information processing as a function of lower cognitive activity on relevant information processing in combination with a large amount of irrelevant information processing for the seductive details group. These results support the assumption that the ratings for mental effort probably measured as well learning relevant and irrelevant cognitive activity that might be due to the German wording of mental effort that incorporates as well successful and unsuccessful activity. Therefore, one item might not be a valid measure of a specific cognitive load factor, whereas self-reports in general could be suitable not only to rate the total cognitive load but also to differentiate between the cognitive load factors (Leppink et al. 2013, 2014). However, it seems difficult to formulate universally valid questions for intrinsic, germane, or extraneous cognitive load without an adaption to the respective learning content. In the present study, the subjective rating scales showed no sensitivity for the extraneous cognitive load manipulation or the differences in total cognitive activity. However, the ratings of task difficulty proved to be a valid measure for a learning relevant cognitive load facet and the increase in extraneous cognitive load might have been measured by repeated ratings instead of only two times of measurement (Schmeck et al. 2015; Van Gog et al. 2012). With respect to the redefinition of intrinsic cognitive load, the item about task difficulty is probably no longer sufficient to measure the complete intrinsic cognitive load aspect.

The final topic to be discussed is about the theoretical implications of the present study for cognitive load theory and whether the used methods measure different aspects of cognitive load. The results demonstrate that seductive details increase extraneous cognitive load by producing additional irrelevant information processing. The rhythm method in this case is sensitive to the extraneous load factor. This assumption is supported by the significant negative correlation between rhythm method and learning success and is in line with cognitive load theory that associates high extraneous cognitive load with low learning success. Furthermore, there are no significant correlations between the rhythm method and the ratings of task difficulty or the measures of eye movements, indicating that the rhythm method is sensitive to a cognitive load factor that is not related to task difficulty or relevant information processing. In combination, these three measures support the assumption of three differentiated cognitive load factors, as there was extraneous load measured by the rhythm method, intrinsic load measured by the ratings of task difficulty and germane load measured by the eye movements on relevant information processing.

Taken in combination, the results support the former model of cognitive load theory (Plass et al. 2010; Sweller et al. 2011). As the relevant information processing was clearly observable and can be assumed to represent the cognitive processes dedicated to the germane load factor. Combining different methods in this case allows differentiating the single cognitive load aspects. However, the relevant information processing that was measurable with eye tracking should not be assumed to represent all cognitive activity

that causes germane cognitive load, but at least the part that is observable. Concerning the updated model of cognitive load theory (Choi et al. 2014) there is no need to measure intrinsic cognitive load as a simple task difficulty and to differentiate germane cognitive load because the cognitive processing of the relevant information is part of the construct (Kalyuga 2011). The results of the objective measures are also in line with this assumption, showing an increase in cognitive load by the rhythm method and a decrease of relevant information processing with an increase in total information processing by the eye-movement analysis. Again, only when combining the two methods, the differentiation between single cognitive load aspects is possible. The increase in cognitive load indicated by the rhythm method can also be total or extraneous cognitive load; only in combination with eye-tracking analysis, the origin of the additional cognitive load can be identified.

In sum, the corresponding results are in line with the cognitive load theory assumptions that high extraneous cognitive load and low germane, respectively intrinsic load are associated with low learning success, whereas low extraneous cognitive load and high germane, respectively intrinsic load are associated with high learning success. The present study demonstrates that it is possible to subdivide total cognitive load by using multiple methods of cognitive load measurement. Self-reports turned out to be suitable for the measurement of task difficulty, the rhythm method was appropriate to assess extraneous cognitive load, and eye-tracking data provided information about cognitive activity dedicated to germane, respectively intrinsic cognitive load.

Further research is nevertheless needed with a systematic manipulation of the germane, respectively the intrinsic load factor and variations of modality and presentation format. Future research manipulating these aspects is needed because the interpretation of the eye-movement analysis may be limited as a function of design and presentation format. Moreover, additional evidence is needed to clarify the relation of eye movements and cognitive processes dedicated to the different cognitive load factors. One limiting factor of the present study is a quite small amount of post-test items and future research should therefore use extensive tests for learning success and take a follow-up investigation into account. Another point to pay attention for is the possible differences concerning the validity and reliability of the different methods due to the experimental design. As the present study used a between-subjects design to compare different methods of cognitive-load measurement, further studies should compare different measures in a within-subject design. However, the combination of multiple cognitive load measures should also be used to investigate other cognitive-load and multimedia-learning effects to review the suitability of this kind of analysis for cognitive-load research.

References

- Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, 16, 389–400.
- Brünken, R., Moreno, R., & Plass, J. (2010a). Current issues and open questions in cognitive load research. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 253–272). New York: Cambridge University Press.
- Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38, 53–61.
- Brünken, R., Plass, J., & Leutner, D. (2004). How instruction guides attention in multimedia learning. *Proceedings of the EARLI SIG 6 Biannual Workshop 2002*. In H. Niegemann, R. Brünken, & D. Leutner (Eds.), Erfurt: Waxmann.

- Brünken, R., Seufert, T., & Paas, F. (2010b). Measuring cognitive load. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 181–202). New York: Cambridge University Press.
- Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155–166.
- Choi, H. H., Van Merriënboer, J. J. G., & Paas, F. (2014). Effects of the physical environment on cognitive load and learning: Towards a new model of cognitive load. *Educational Psychology Review, 26*, 225–244.
- Clark, R. E., & Clark, V. P. (2010). From neo-behaviorism to neuroscience: Perspectives on the origins and future contributions of cognitive load research. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 203–229). New York: Cambridge University Press.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*, 111–122.
- Debue, N., & Van De Leemput, C. (2014). What does germane cognitive load mean? An empirical contribution to the cognitive load theory. *Frontiers in Psychology, 5*, 1099.
- DeLeeuw, K., & Mayer, R. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology, 100*, 223–234.
- Demberg, V., Sayeed, A., Mahr, A., & Müller, C. (2013). Measuring linguistically-induced cognitive load during driving using the ConTRe task. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI), Oct 2013*, Eindhoven, The Netherlands.
- Ekstrom, R. B., French, J. W., Harmann, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton: Educational Testing Service.
- Folker, S., Ritter, H. & Sichelschmidt, L. (2005). Processing and integrating multimodal material – The influence of color coding. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of 27th annual conference of the cognitive science society* (pp. 690–695). Mahwah, NJ: Lawrence Erlbaum.
- Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). “Seductive details” and learning from text. In K. A. Renninger (Ed.), *The role of interest in learning and development* (pp. 239–254). Hillsdale: Erlbaum.
- Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of “seductive details” on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6*, 41–57.
- Goetz, E., & Sadoski, M. (1995). ‘Commentary: The perils of seduction: Distracting details or incomprehensible abstractions?’: Reply. *Reading Research Quarterly, 30*, 518–519.
- Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors, 26*, 519–532.
- Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology, 25*, 172–190.
- 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 (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking – A comprehensive guide to methods and measures*. New York: Oxford University Press.
- Huck, S. W. (2012). *Reading statistics and research*. Boston: Pearson Education.
- Jarodzka, H., Scheiter, S., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*, 146–154.
- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied, 18*, 178–191.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.
- Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23*, 1–19.
- Korbach, A., Brünken, R., & Park, B. (2016). Learner characteristics and information processing in multimedia learning: A moderated mediation of the seductive details effect. *Learning and Individual Differences, 51*, 59–68.
- Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop effects. *Cognitive Processing, 12*, 13–21.
- Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavioral Research, 45*, 1058–1072.

- Leppink, J., Paas, F., Van Gog, T., Van der Vleuten, C. P. M., & Van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30*, 32–42.
- Magner, U. I. E., Schwonke, R., Alevan, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*, 141–152.
- Marshall, S. (2007). Identifying cognitive state from eye metrics. *Aviation, Space and Environmental Medicine, 78*, B165–B175.
- Marshall, S., Pleydell-Pearce, C., & Dickson, B., (2002). Integrating psychophysiological measures of cognitive workload and eye movements to detect strategy shifts. *HICSS '03 Proceedings of the 36th Annual Hawaii International Conference on System Sciences*.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning*. Cambridge: University Press.
- Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction, 20*, 167–171.
- McInerney, D. M., & Sinclair, K. E. (1991). Cross cultural model testing: Inventory of school motivation. *Educational and Psychological Measurement, 51*, 123–133.
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning, 22*, 149–158.
- Moreno, R., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 9–28). New York: Cambridge University Press.
- Oberauer, K., Süß, H. M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity – Facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017–1045.
- 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.
- Paas, F., Renkl, A., & Sweller, J. (2003a). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1–4.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. M. (2003b). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63–71.
- Paas, F., & Van Merriënboer, J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 351.
- Park, B. (2010). *Testing the additivity hypothesis of cognitive load theory*. Dissertation, Saarbrücken: Universität des Saarlandes. <http://scidok.sulb.uni-saarland.de/volltexte/2010/3478/>.
- Park, B., & Brünken, R. (2015). The rhythm method: A new method for measuring cognitive load – An experimental dual task study. *Applied Cognitive Psychology, 29*, 232–243.
- Park, B., Flowerday, T., & Brünken, R. (2015a). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior, 44*, 267–278.
- Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015b). Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of antropomorphisms. *Computers & Education, 86*, 30–42.
- Park, B., Korbach, A., & Brünken, R. (2015c). Do learner characteristics moderate the seductive-details-effect? A cognitive-load-study using eye-tracking. *Journal of Educational Technology & Society, 18*, 24–36.
- Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior, 27*, 5–10.
- Park, B., Plass, J. L., & Brünken, R. (2014). Cognitive and affective processes in multimedia learning. *Learning and Instruction, 29*, 125–127.
- Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. New York: Cambridge University Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
- Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research, 47*, 2714–2726.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–526.
- Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34*, 344–355.

- Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: Differences between immediate and delayed ratings. *Instructional Science*, *43*, 93–114.
- Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*, 100–110.
- Schwalm, M. (2009). *Pupillometry as a method for measuring mental workload within an automotive context*. Dissertation, Saarbrücken: Universität des Saarlandes. <http://scidok.sulb.uni-saarland.de/volltexte/2009/2082/>.
- Schwalm, M., Keinath, A., & Zimmer, H. D. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. In D. de Waard, F. Flemisch, B. Lorenz, H. Oberheid, & K. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 75–88). Maastricht: Shaker.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell: ACER Press.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., van Merriënboer, J. G., & Paas, F. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.
- Van Gog, T., Kirschner, F., Kester, L., & Paas, F. (2012). Timing and frequency of mental effort measurement: Evidence in favour of repeated measures. *Applied Cognitive Psychology*, *26*, 833–839.