Applied Cognitive Psychology, *Appl. Cognit. Psychol.* **29**: 232–243 (2015)
Published online 15 December 2014 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.3100

The Rhythm Method: A New Method for Measuring Cognitive Load—An Experimental Dual-Task Study

BABETTE PARK* and ROLAND BRÜNKEN

Department of Education, Saarland University, Saarbrücken, Germany

Summary: The present study joins a series of studies that used the dual-task paradigm to measure cognitive load while learning with multimedia instruction. The goal of the current work was to develop a secondary task, to measure cognitive load in a direct and continuous way using intra-individual, behavioral measures. The new task is achieved by utilizing internalized cues. More specifically, a previously practiced rhythm is executed continuously by foot tapping (secondary task) while learning (primary task). Precision of the executed rhythm was used as indicator for cognitive load—the higher the precision, the lower cognitive load. The suitability of this method was examined by two multimedia experiments ($n_1 = 30$; $n_2 = 50$). Cognitive load was manipulated by seductive details (Experiment 1: with vs. without) and modality (Experiment 2: on-screen text vs. narration). Learners who learned under low cognitive load conditions (Experiment 1: without seductive details; Experiment 2: narration) showed significantly higher rhythm precision. These results provide evidence that rhythm precision allows for a precise and continuous measurement of cognitive load during learning. Copyright © 2014 John Wiley & Sons, Ltd.

OBJECTIVES

From the perspective of cognitive load theory (CLT), the design of learning material to efficiently support complex knowledge acquisition requires information about the amount of cognitive load imposed by different forms of instruction. This requirement establishes a need for well-defined measures of cognitive load. To address this need, the present study utilizes the dual-task paradigm, which requires participants to perform two tasks simultaneously, as a behavioral method of cognitive load measurement. In order to effectively measure cognitive load, interference between primary and secondary task must be reduced, requiring the cognitive load measurement to be conducted in a way that learners do not have to consciously interrupt the primary task (learning) and that the secondary task is using a different sensory channel than the learning material. Therefore, the objective of the present work was the development and evaluation of a secondary task to measure cognitive load. This measurement is characterized by sensory independence of the presentation mode of learning instruction (primary task). Additionally, beyond sensory independence, the secondary task also has to rely on the same working memory resources as the primary task to be able to measure differences in resource consumption caused by the variation of the learning task. To achieve these two characteristics, the secondary task should be associated with executive control processes (Baddeley, 1992), by developing a secondary task that includes inhibition processes, which have been cited as good indicators for executive control and cognitive load. In the development of this new secondary task, inhibition processes are explicitly evoked by participants performing a task called the rhythm method. To provide evidence for the effectiveness of this secondary task to measure cognitive load, the current study presents a series of studies that demonstrate the usability of the dual-task paradigm to measure cognitive load (Brünken, Plass, & Leutner, 2004; Brünken, Steinbacher, Plass, & Leutner, 2002; Chandler & Sweller, 1996; DeLeeuw & Mayer, 2008; Marcus, Cooper, & Sweller, 1996; Renkl, Gruber, Weber, Lerche, & Schweizer, 2003).

*Correspondence to: Juniorprof. Dr. Babette Park, Department of Education, Campus A 4.2, Saarland University, 66 123 Saarbrücken, Germany.
E-mail: b.park@mx.uni-saarland.de

THEORETICAL FRAMEWORK

Cognitive load theory

Cognitive load theory (Paas, Renkl, & Sweller, 2003; Plass, Moreno, & Brünken, 2010; Sweller, Ayres, & Kalyuga, 2011) is currently one of the most used frameworks in empirical research on learning and instruction; this is evidenced by numerous publications in high-ranking international journals within the last decade (e.g., special issues in Applied Cognitive Psychology: Paas & Kester, 2006, and in Computers in Human Behavior: Kirschner, Kester, & Corbalan, 2011). One goal of cognitive load research is to formulate practical implications for the design of learning environments. These implications should be derived from empirical studies on the relationship between information presentation and characteristics of the cognitive system (Paas et al., 2003; Plass Moreno, & Brünken, 2010). CLT is characterized by the central assumption that the amount of knowledge acquisition depends on the efficiency of the use of available but limited cognitive resources in working memory.

In addition, the extent of cognitive load is determined by the following three components (Sweller, Van Merrienboer, & Paas, 1998). First, *intrinsic cognitive load* is associated with the given complexity of the learning task. The more complex the learning task, the higher the intrinsic cognitive load. The complexity of a task can be defined by the amount of information elements and their interactions (element interactivity) in relation to the learner's prior domain specific knowledge. Second, *extraneous cognitive load* is a 'negative' load, which is caused by ineffective instruction and does not lead to efficient schema acquisition (Brünken, Plass, & Leutner, 2003; Paas, Tuovinen, Tabbers, & van Gerven,

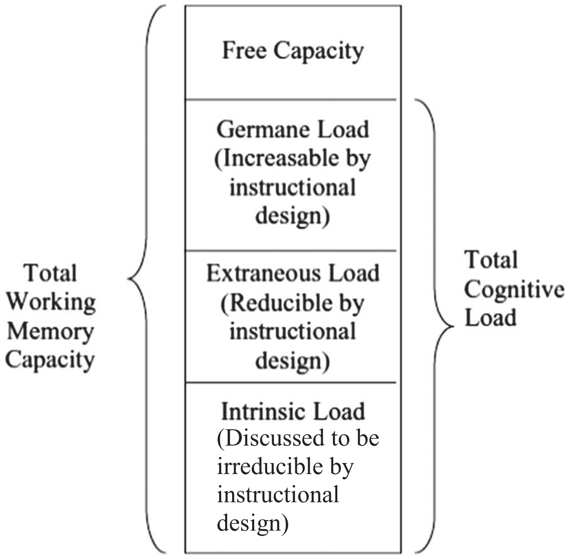

Figure 1. Triarchic model of cognitive load theory, adapted from Moreno and Park (2010), ©Cambridge University Press, reprinted with permission

2003). Information should therefore be presented in a manner that allows for essential information to be extracted easily without unnecessary add-ons. Finally, *germane cognitive load* evolves from learning activities, which foster understanding. Germane load is a 'positive' cognitive load attributed to the processing, construction, and automation of schemata. Although extraneous sources of load hinder learning, intrinsic sources of load reflect the complexity of the given learning task in relation to the learners' level of expertise and germane sources of load promote learning by helping students to engage in the process of schema formation and automation. A basic assumption of CLT is that the total cognitive load experienced during learning is the summation of these three load types, the so-called additivity hypothesis (Brünken, Plass, & Moreno, 2010; Moreno & Park, 2010; Park, 2010). The triarchic model of CLT is shown in Figure 1, which is adapted from a summary on the historical development of CLT by Moreno and Park (2010).

In sum, from the current state of cognitive load research, it can be concluded that the usefulness of CLT as a framework is impressively confirmed. A broad fundament of empirical studies demonstrates the instructional application of cognitive load principles to design learning material in an effective and efficient way. However, a systematic investigation to test and compare different methods for measuring cognitive load is still missing (Brünken, et al., 2010). Such an investigation is the indispensable prerequisite for the valid testing of main assumptions of CLT. Moreover, additional empirical confirmations of cognitive load research for questions of educational psychology can only be considered as valid and substantiated, if the used measurement instruments have been systematically tested (Gerjets, Scheiter, & Cierniak, 2009).

Measurement of cognitive load

In addition to the more traditional analysis of knowledge acquisition parameters as an indirect indicator of cognitive load, the possibility to measure cognitive load in a direct behavioral way during learning has recently been discussed (e.g., Brünken et al., 2003; Brünken, et al., 2010; Paas et al., 2003). Indirect subjective measures of cognitive load are mostly composed of rating scales (e.g., Paas, 1992), although direct objective measures are grounded within the dual-task paradigm (Brünken, Plass, & leutner, 2004; Brünken Steinbacher, Plass, & Leutner, 2002; Chandler & Sweller, 1996; DeLeeuw & Mayer, 2008; Marcus, Cooper, & Sweller, 1996; Münzer & Holmer, 2009; Renkl, Gruber, Weber, Lerche, & Schweizer, 2003). The following sections discuss the advantages and disadvantages of both subjective and objective measures.

Subjective measures

Within the last decades, subjective rating scales have been predominantly used as measures in cognitive load research. Examples include the NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988), the estimation of task difficulty (Kalyuga, Chandler, & Sweller, 1998; Paas & van Merriënboer, 1993), and the subjective rating scale of Paas (1992), which is the most often and widespread used measure. The NASA-TLX was originally constructed as a multidimensional subjective rating scale to test mental and physical demands of aircraft pilots. Six dimensions are considered (mental demand, physical demand, temporal demand, performance, effort, and frustration level), which are each rated on an 18-point Likert scale. In contrast, the instrument developed by Paas (1992) is more convenient to administer as it consists of only one item for task difficulty and one item for mental effort. Both items are answered immediately after the learning phase or the learning performance test, and learners are asked to retrospectively estimate the task difficulty or their invested mental effort during learning usually on a 9-point Likert scale.

One advantage of subjective measures is that the person is asked, who is the only one who can estimate the individually experienced difficulty of a task and the personally invested mental effort. This face validity is confirmed by findings of a study on workload published by Gopher and Braune (1984), showing that individuals are able to introspect their cognitive processes. Paas and Van Merrienboer (1994) referred to this study and additionally assumed that humans have no difficulty in assigning numerical values to the imposed mental load. Numerous studies of cognitive load research that have used subjective rating scales are based on this evidence. Additionally, subjective rating scales are advantageous as they can be used in a very easy, fast, and economic way without much effort in different learning contexts.

Disadvantages of subjective measures are measurement-technical problems (Brünken et al., 2003, 2010; Clark & Clark, 2010; Moreno, 2006), which concern the quality criteria objectivity, validity, and reliability. Two critical reviews already summarize these measurement-technical problems in detail, which are published by De Jong (2010) and Brünken, et al. (2010) and can be recommended for a good overview. As highlighted in these publications, one major problem with subjective rating scales is related to the validity of the measurement: Learners tend to translate their ratings into operationalizations as 'how difficult' or 'how complex', but not 'how much effort'. Thus, the face validity assumption that a person always knows how to correctly rate the own cognitive capacity can be put into question. This also leads to the conclusion that the comparability of individual ratings is questionable. Further reliability studies are still needed to prove the sensitivity of subjective ratings in comparison with other methods for measuring cognitive load. Moreover, the recently initiated work to develop rating scales that clearly differentiate between the three cognitive load types should be enhanced (Gerjets et al., 2009; Koch, Seufert, & Brünken, 2008; Park, 2010).

Objective measures

The present work originated owing to recent publications in cognitive load research that discussed widespread and most frequently used methods for measuring cognitive load and recommend using more object and direct methods (Brünken et al., 2003; De Jong, 2010; DeLeeuw & Mayer, 2008; Plass et al., 2010). Direct measurement of cognitive load can be measured by brain activity or dual-task performance (Brünken et al., 2003). Brain activity measures seem to be promising but so far have not used very often in the frame of learning studies for economic and practical reasons (Antonenko, Paas, Grabner, & van Gog, 2010; Whelan, 2007). In contrast, a series of studies have already begun to show that dual-task performance is a 2reliable and valid method to measure cognitive load (e.g., Brünken et al., 2004; DeLeeuw & Mayer, 2008). Within the dual-task paradigm, cognitive load is measured by the performance of a secondary task that is executed in parallel to the primary learning task. In detail, the dual-task method measures cognitive load at different times of measurement during learning (primary task) with the help of the secondary task performance of the learner (e.g., reaction time to a signal), which reflects the amount of cognitive load in the primary task. In other words, differences in a learner's resource consumption caused by different presentations of the learning material, for instance, can be measured by differences in performance on the secondary task. The established secondary tasks usually include either an auditory or visual cue in the learning instruction. For example, in a study of Brünken et al. (2004), participants had to monitor a letter in the upper part of the computer screen and had to react by pressing the space bar, when a color change was observed.

In cognitive load research, different secondary tasks were used that are appropriate in general or for specific primary tasks. The advantage of secondary tasks is that cognitive load can be measured online, when the secondary task has to be executed at different points of time over the learning phase. In addition, secondary tasks are objective measures, as the measurement, analysis, and interpretation of data are measured independently of a self-report rating. Different secondary tasks have been noted to be useful with different learning materials in empirical studies. Examples from the literature include the following: the reaction to an auditory signal (Brünken et al., 2004), to a color-changing stimulus in form of a letter within the learning material (Brünken et al., 2002), or to a color-changing background screen of the multimedia learning instruction (DeLeeuw & Mayer, 2008). Another, established secondary task was used in a study by Renkl et al. (2003), where participants had to press every 10 seconds one of the 12 possible buttons that was associated with the randomized presented letter and the instruction 'before' or 'after' (e.g., button A, if the screen shows 'B' + 'before'). In addition, Chandler and Sweller (1996) developed a secondary task that had high demands of the working memory because learners had to memorize an additional letter in order to correctly retrieve it again within the learning phase.

The major disadvantage of all these dual-task methods revolves around the frequently discussed problem of a possible interference between the primary and secondary tasks (Brünken et al., 2003). This interference is especially of concern when the primary and secondary tasks are both presented within the same modality (e.g., acoustically). A certain presentation format of the primary task (e.g., acoustic in the format of narrated information) is possibly more prone to interference (e.g., with the acoustic secondary task) in contrast to another presentation format (e.g., visual-only material). Confirming results were found by the analysis of the modality effect, which used visual (Brünken et al., 2002) and acoustic (Brünken et al., 2004) secondary tasks. This reactivity of the dual-task method is one methodological problem. A limitation of using a secondary task to measure cognitive load is that it can only be recommended for certain formats of the learning material. For instance, auditory or visual secondary tasks do not measure total cognitive load but do measure modality-specific aspects of visual or phonological processes in working memory. Moreover, these mentioned secondary tasks are dependent on the learning instruction (primary task), as they require a conscious disruption in the learning process to respond to the auditory or visual cue. The present study aims to overcome these limitations by developing a secondary task that measures cognitive load in a direct and continuous way using intra-individual, behavioral measures with a method that is characterized by its sensory independence from the presentation mode of the learning instruction (primary task).

A more general disadvantage of using dual-task methods is that the implementation is complex in comparison with the use of subjective rating scales. This perhaps is one reason that dual-task methods have rarely been used. In sum, comparatively few studies have been investigated, and the different developed methods are not comparable with each other. With further development of appropriate secondary tasks and systematic comparisons, the methodological problems of dual-task methods will be analyzed and overcome.

STATEMENT OF PROBLEM, GOAL OF THE PRESENT WORK, AND HYPOTHESES

The first objective of the present work is to develop a secondary task that can be characterized by minimal reactivity and sensory independence of the presentation mode of learning instructions (primary task). In other words, the study primary objective is to develop a secondary task that does not interrupt the primary task (learning) consciously and is associated with a different sensory channel in contrast to learning material that is presented in either visual or auditory formats. The second objective of the present work is to develop a secondary task that uses the same working memory resources as the primary task. This is necessary in order to be able to measure differences in resource consumption caused by the variation of the learning task. Both of the stated objectives can be achieved by the use of a secondary task that has a direct reference to executive control processes (Baddeley, 1992).

Inhibition processes are a good indicator for executive control processes (Cohen et al., 1997) and can be interpreted as a modality-unspecific total cognitive load indicator. The idea of the present work is to employ a rhythm method, which includes inhibition processes for cognitive load measurement, as an easy rhythm with pauses (e.g., tap–tap–pause–pause/tap–tap–pause–pause/….) already requires inhibition, when learners have to consciously inhibit to tap frequently by foot. Moreover, a foot-tapping rhythm method is characterized by sensory independence from the learning instruction (primary task) and does not require conscious disruption of the primary task, as it can be realized by internalization (without external prompts). Thus, a rhythm method can be a possible solution for the earlier described methodological problems of modality specificity and conscious learning disruption. Moreover, this new method is also characterized by intra-individual measures such as other secondary task measures: Independent of individual differences in feeling for rhythm and expertise, the own rhythm baseline of the learner serves for intra-individual comparisons with the produced rhythms under diverse experimental conditions. This fact creates the opportunity for cognitive load to be measured in experimental between-group designs based on individual measures and comparisons of their own rhythm scores, because the individual scores are always related to the individual cognitive load baseline.

Foot tapping as well as manual spatial tapping was only used with frequent tapping tasks, which do not require inhibition processes and are therefore very easy to automatize. This automation is possibly the reason that some authors showed that an additional motor task such as frequent tapping even seems to stimulate cognitive processing of and the performance in motor or visual primary tasks (Brown & Marsden, 1991; Emerson & Miyake, 2003). This phenomenon was observed in a study of Hegarty, Shah, and Miyake (2000), where performance in a paper-folding task (Ekstrom, French, & Harmann, 1976) increased while participants engaged in frequent foot tapping when compared with decrements shown by other secondary tasks such as random number generation, an *n*-back task, or spatial tapping. Another study shows that frequent foot tapping seems also to correlate with the performance in primary motor tasks with a study revealing that the higher the performance in the primary motor task of pressing an up button, the higher was the rate of frequent foot tapping (Brown & Marsden, 1991). All in all, foot-tapping tasks in dual-task research require frequent tapping without conscious interruptions and therefore lead to automation. In contrast, a rhythmic tapping task has to be monitored continuously and cannot be automatized because it requires inhibitory processes to stop the tapping within the different pauses of the rhythm. Rhythm tapping includes both short and long foot taps that are associated with inhibition processes and should require intense executive control processing. Therefore, the secondary task in the present experiments, which required participants to tap the rhythm 'tap–tap–pause–pause–tap–tap–pause–pause–….', should be a good indicator of cognitive load during learning.

In the present work, two experiments are conducted to investigate if the rhythm method allows a precise and continuous measurement of cognitive load from a dual-task paradigm. For the primary task, multimedia learning material was used, which was varied by two well-investigated instructional design effects, seductive details effect and the modality effect. Both of these instructional design effects are assumed to affect the amount of induced cognitive load during learning (e.g., Garner, Gillingham, & White, 1989; Ginns, 2005; Harp & Mayer, 1998; Lehman, Schraw, McCrudden, & Hartley, 2007). In the first experiment, seductive details material of the multimedia instruction was varied (with vs. without seductive details); in the second experiment, the modality of the multimedia instruction with pictures and accompanying text was varied (narration vs. on-screen text). Both experiments are designed to induce the hypothesized cognitive load effect and allow for the examination of the invented secondary task performance in its ability to measure cognitive load. Moreover, learning outcomes with respect to the primary task were measured. In both experiments, the higher learning performance is expected in the lower cognitive load condition (Experiment 1: without seductive details; Experiment 2: narration).

More specifically, it is hypothesized for Experiment 1 that learners who learn without seductive details will show a significantly higher learning performance (retention and comprehension) than learners who learn with seductive details. In addition, these learners will show a significantly lower level of cognitive load as measured by rhythm precision and self-report ratings than learners who are confronted with seductive details. In Experiment 2, it is hypothesized that learners who learn with narration will show a significantly higher learning performance (retention and comprehension) than learners who learn with on-screen text. Moreover, these learners will show a significantly lower level of cognitive load as measured by rhythm precision and self-report ratings than learners who learn with on-screen text.

METHOD

The suitability of the new secondary task of rhythm precision for continuous measurement of cognitive load while learning was examined within two separate experiments ($n_1 = 30$; $n_2 = 50$). To examine the rhythm method, it was necessary to use well-known instructional design effects on learning, which have been associated with cognitive load effects. Typical learning material characteristics that are associated with extraneous load include the modality of the learning material as well as the presence of seductive details. Both characteristics can be varied by teachers as well as instructional designers and are often found in educational learning materials at schools and universities (Mayer, 2005b).

According to CLT, when visual representations (e.g., pictures, diagrams, and animations) are presented with an accompanying text, they force students to invest significantly more mental effort owing to the presentation format during learning. This creates a detrimental effect on learning performance owing to the higher cognitive load that is imposed upon the learner (Sweller, 2005; Sweller & Chandler, 1994; Sweller et al., 1998). The empirical base of the learning benefits caused by replacing the accompanying text with narration (the modality effect) is quite robust (Brünken & Leutner, 2001; Brünken et al., 2004; Ginns, 2005; Low & Sweller, 2005; Mayer, 2005a, 2005b, 2009; Mayer & Moreno, 1998; Moreno & Mayer, 2002; Seufert, Schütze, & Brünken, 2009); however, the theoretical explanation of this effect is still under discussion (Rummer, Schweppe, Fürstenberg, Seufert, & Brünken, 2010; Schnotz, 2011). One explanation for the modality effect is that on-screen text material requires splitting the limited available visual working memory capacity, the so-called visuospatial sketchpad (Baddeley, 1992). This is an explanation that is analogous to the explanation of the split attention effect (Sweller, 2005; Sweller et al., 1998). This perspective offers the argument that when the text to be learned is distributed, processing occurs through the visuospatial sketchpad and the phonological loop (Baddeley, 1992) with the overall consequence of reduced cognitive load.

In a similar fashion to the modality effect, seductive details also interfere with learning owing to an associated extraneous load. The term 'seductive details' was first introduced by Garner et al. (1989) to refer to the addition of interesting but unnecessary information to text, which reduce the recall or learning of 'non-seductive' relevant text ideas. Research on the effect of seductive details has focused on seductive text passages or seductive illustrations in text comprehension studies. Several studies have shown a detrimental effect of seductive details (e.g., Garner et al., 1989; Harp & Mayer, 1998; Lehman et al., 2007; Mayer, 2009). Three different explanations for the seductive details effect have been proposed and examined within the literature, and these include diversion, and disruption of or distraction from the relevant learning material (Harp & Mayer, 1998; Lehman et al., 2007; McCrudden & Corkill, 2010; Rowland, Skinner, Davis-Richards, Saudargas, & Robinson, 2008; Sanchez & Wiley, 2006). Consequently, seductive details impose high cognitive load during learning by forcing students to spend their limited resources in processing materials that divert, disrupt, or distract from the construction of a coherent mental model during the learning process.

Participants and materials

In the first experiment, the seductive details effect was employed. Thirty high-school students (53.3% women) with an average age of 16.9 years ($SD = 1.3$) participated in the study. Participants were randomly assigned to one of the two experimental groups (with vs. without seductive details). Two participants had to be excluded owing to technical problems. This exclusion created two groups: one group with 15 participants that learned without seductive details and another group with 13 participants who learned with seductive details. In the second experiment, the modality of the multimedia learning instruction with pictures and accompanying texts was varied (narration vs. on-screen text). Fifty university students (74.0% women) with an average age of 22.2 years ($SD = 2.6$) took part and were randomly assigned to the two experimental groups (narration vs. on-screen text). Five participants had to be excluded because of technical problems, so 21 participants learned with narration and 24 participants learned with on-screen text.

The learning material in both experiments consisted of a learner-paced multimedia instruction about the structure and function of a cellular molecule, the ATP-Synthase. It included 11 screens, each one with static pictures and accompanying verbal explanations. The objective of the learning task was to understand the complex structure and function of the molecule by integrating the verbal representations (Figure 2, bottom left) with the corresponding pictorial representations (Figure 2, top left). The learning objective was explicitly stated during the introductory portion of the program that was common to all treatment conditions of both experiments. In contrast, understanding the usefulness of ATP was not part of the objectives of the learning task, which is the reason that examples of the usefulness of ATP in different areas (e.g., sports and work) were chosen as presentation for the participants in the seductive detail conditions of experiment 1 (Figure 2, right side). In Experiment 2, there were no seductive details included, but the relevant verbal representations (Figure 2, bottom left) were varied by replacing them with identical narrated explanations in the form of a male voice.

Learners had to execute the foot-tapping rhythm method simultaneously to learning and continuously throughout the learning session. This was measured by using a foot pedal. The rhythm that learners had to tap on the foot pedal was written in four–four time that is the easiest meter for playing music: tap–tap–pause–pause/tap–tap–pause–pause/…. and so forth (Figure 3). As shown in Figure 3, the rhythm includes one short rhythm component of 500 milliseconds, the short inter-tap interval, and one long rhythm component of 1500 milliseconds including the two pauses, which is simply a long pause or long inter-tap interval.

Learners were introduced to this rhythm method before the learning session began. First participants had to listen to an example of the rhythm and then had to tap the rhythm by accompanying the example; after this introduction, participants had to tap the rhythm alone without the help of the example for 1 minute. This last unit was recorded and used as individual rhythm baseline. This baseline was measured so that a comparison of the baseline with the produced rhythm during the learning phase in dual-task situation (learning = primary task; rhythm = secondary task) could be performed. The tapping activity was automatically measured by auditory software that is normally used to record music (Freeware: Audacity 1.3.5 Beta). The tapping apparatus was measuring only on versus off signals and could therefore accurately record the tapping of the participants. For the analysis of the rhythm method, all produced inter-tap intervals were accepted that were higher than 250 milliseconds. This is a low cut-point for a response time typical found in response time studies (e.g., Thorpe, Fize, & Marlot, 1996). This value was chosen because participants can only tap consciously

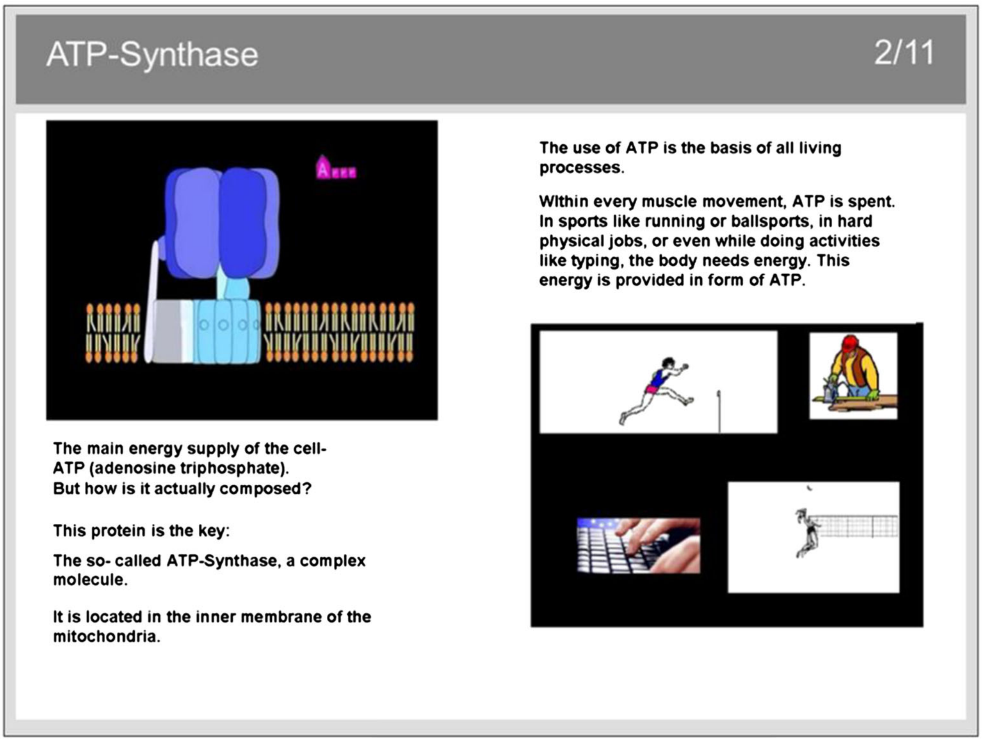

Figure 2. A screenshot of the learning environment used in the seductive detail condition showing the verbal representations (lower left corner), corresponding pictorial representations (top left corner), and seductive details (additional text and pictures on the right side); original version in German, translated by the authors

Rhythm Method

tap tap pause pause tap tap pause pause tap …

| 500 msec | 1500msec | 500 msec | 1500msec | …

Figure 3. The rhythm method that has to be executed by foot tapping; written in four–four time; including two rhythm components defined by two inter-tap intervals

with more than 250 milliseconds distance from one tap to the next tap within an inter-tap interval. Therefore, all data were higher than 250 milliseconds. On the upper end of the measure, all inter-tap intervals were accepted if they were equal to or less than the whole rhythm unit of 2000 milliseconds (short rhythm component = 500 milliseconds + long rhythm component = 1500 milliseconds). Data that were greater than 1000 milliseconds were assigned to the long rhythm component, whereas all data that were lower than 1000 milliseconds were assigned to the short rhythm component. Every rhythm unit including one short and one long rhythm component was included and analyzed, but rhythm units including two long or two short rhythm components were counted as an error and not included in the analysis.

Measures

Learning success was assessed with a learning performance test including retention versus comprehension items in open and multiple-choice formats, and mapping items where learners needed to identify corresponding verbal and pictorial elements. The scales used for Experiments 1 and 2 differed owing to item analyses. In Experiment 1, the subscale retention included eight items yielding a Cronbach α of 0.71 (item examples: (i) 'The matrix is…'—'…the area out of the mitochondrion'/'…the inter-membrane area'/'…*the inner of the mitochondrion*'/'…a cellular compound within the web'; (ii) Please describe the difference between ATP-Synthase and ATP-Synthese:), with the subscale for comprehension composed of five items yielding a Cronbach α of 0.70 (item examples: (i) 'Which cells do feature the highest number of mitochondrions?'—neurons; skin cells; *myocardial muscle cells*; intestinal epithelium cells; (ii) 'Please explain how it comes to the release of ATP (steps of process):'). In Experiment 2, the subscale of retention included eight items yielding a Cronbach α of 0.71 (item examples: (i) 'Please describe the term "proton motoric power"'; (ii) learners have to demonstrate the composition of the molecule by identifying corresponding verbal and pictorial elements), with the subscale of comprehension composed of six items yielding a Cronbach α of 0.70 (item examples: (i) 'What happens with the sub elements Alpha and Beta during the rotation of the axis?'; (ii) 'Imagine that no protons are transported and thus no rotation takes place —To which deformity of the ATP-Synthase could this be attributed? What is defective?'). These items were not identical to the 13-item prior knowledge test described later, which was only used to control for prior knowledge.

Cognitive load was measured by subjective ratings of mental effort (Paas, 1992), in which learners rated their perceived cognitive load during learning in the middle of the lesson (after screen 4 of 11) and immediately after the lesson on a 7-point Likert scale. At each one of these points in time, learners were asked to estimate their cognitive load by clicking on the rating that best completed the following statement 'While working on the learning material my mental effort was…' with the ratings ranging from 'very low', 'low', 'rather low', 'neither low nor high', 'rather high', 'high' to 'very high'. Moreover, cognitive load was measured by the rhythm method, which allowed analyzing the precision of rhythm execution. The precision is defined by the mean rhythm in milliseconds from the learning phase measurement minus the individual's rhythm baseline score in milliseconds. Therefore, participants with perfect precision received a score of zero. The higher the absolute value of the deviation from zero, the lower the rhythm precision. The precision can be calculated for both inter-tap intervals, the short rhythm component (digitally played: 500 milliseconds), and the long rhythm component (digitally played: 1500 milliseconds; Figure 3). For example, learner A with a baseline of 510 milliseconds that is the mean of the short rhythm component in the baseline condition (instead of the 500 milliseconds digitally played short rhythm component) showing a mean of 530 milliseconds within the learning phase executes the short rhythm component with a precision of 530 MINUS 510 = 20 milliseconds. In contrast, learner B with a baseline of 510 milliseconds showing a mean of 590 milliseconds within the learning phase executes the short rhythm component with a precision of 80 milliseconds that is not as precise as learner A executed this rhythm component. The reliability of the rhythm method was confirmed when using Guttman's split-half method by comparing the mean precisions of the learning screens in the first half of the learning material with the last half of the learning material in both experiments of the present work. In Experiment 1, the reliability analysis resulted in Guttman's split-half coefficient of $r = .96$ for the short rhythm component and $r = .78$ for the long rhythm component; the analysis in Experiment 2 shows Guttman's split-half coefficients of $r = .72$ for the short rhythm component and $r = .61$ for the long rhythm component. In addition, the used data of the two rhythm components were checked for outliers, but no outliers were found. Moreover, when calculating the final precision indicator by subtracting the individual baseline value of the rhythm component from the produced mean rhythm component within the learning phase, the precision indicators were again checked for outliers. Here, one to three of the participants in both experiments were checked as outliers. Results and conclusions are the same, when excluding these participants.

In addition, four control measures were used: prior knowledge, measured with a questionnaire including five multiple-choice and eight open-ended questions about the content domain (Cronbach's $\alpha = .79$); spatial ability, measured by a standardized paper-folding and card-rotation test (Ekstrom et al., 1976); time-on-task, which was automatically recorded by the computer; and working memory capacity, measured by the computer-based 'Numerical Memory Updating Subtest' from Oberauer, Süß, Schulze, Wilhelm, and Wittmann (2000).

Procedure

The learning lesson took approximately 60 minutes and started for all participants with the three control measure tests for spatial ability, working memory capacity, and prior knowledge. Thereafter, the foot-tapping rhythm method was taught before the learning phase started. In the learning phase, the multimedia instruction was presented for participants to learn about the ATP-Synthase. In Experiment 1, participants received a version with seductive details or the control version without seductive details; in Experiment 2, participants received a version with narration or with on-screen text. Participants were asked to execute the foot-tapping rhythm method simultaneously to learning and continuously throughout the entire learning phase in all conditions of both experiments. In the middle of the learning phase (after screen 4 of 11), there was a little pause where learners did not have to tap the rhythm but had to rate their perceived cognitive load for the first time. The learning lesson ended with the second cognitive load rating and the learning performance test.

RESULTS AND SUBSTANTIATED CONCLUSIONS

Experiment 1

In Experiment 1, no statistical significant between-group differences were detected in any of the control variables working memory capacity, $F < 1$, spatial ability, $F(1, 27) = 2.29$, *n.s.*, and prior knowledge, $F(1, 27) = 2.05$, *n.s.* (Table 1). Time-on-task differed significantly over the two conditions, $F(1, 27) = 6.50$, $p < .05$, $d = 1.41$. Learners in the seductive details condition took more time to learn ($M = 10.98$, $SD = 3.97$) than learners who learned without seductive details ($M = 8.00$, $SD = 1.48$; Table 1). Therefore, time-on-task was used as a covariate in the following analyses. Separate analyses of covariance (ANCOVAs) were conducted using learning performance scores (retention and comprehension) and cognitive load ratings, respectively, as dependent variables. However, results and conclusions are the same when running only analyses of variance.

A seductive details effect could be shown in the learning performance, as the learning success was significantly higher for learners who learned without seductive details ($M = 3.15$, $SD = 1.49$) in contrast to learners who learned with seductive details ($M = 1.83$, $SD = 1.75$), when testing the directed hypothesis in the comprehension subscale, $F(1, 26) = 5.48$, $p < .05$, $d = 1.15$ (Table 1). In the subscale retention, no significant difference between the learning groups was found,

Table 1. Experiment 1: Means and standard deviations of all dependent and control variables for both conditions

	Control without seductive details $n=13$		Experimental with seductive details $n=15$	
	M	(*SD*)	*M*	(*SD*)
Working memory capacity (Max = 9)	3.50	(1.24)	3.27	(1.03)
Spatial ability (%)	73.08	(15.86)	64.50	(14.12)
Prior knowledge (Max = 13)	5.31	(3.15)	3.83	(2.28)
Time-on-task (min.)	8.00	(1.48)	10.98	(3.97)
Retention (Max = 8)	4.77	(1.86)	4.73	(2.82)
Comprehension (Max = 5)	3.15	(1.49)	1.83	(1.75)
Cognitive load—subjective rating scale (Max = 7)	4.85	(1.07)	4.80	(1.42)
Cognitive load—dual-task method—long rhythm component[a] (milliseconds; 0 = absolute precision)	59.17	(33.98)	75.77	(43.09)
Cognitive load—dual-task method—short rhythm component[b] (milliseconds; 0 = absolute precision)	36.70	(25.75)	80.06	(64.65)

Note:
M, mean; *SD*, standard deviation; Max, maximum.
[a]Description of the long rhythm component: rhythm precision = absolute value of mean rhythm component within learning phase MINUS individual baseline of long rhythm component, the 1500-millisecond inter-tap interval.
[b]Description of the short rhythm component: rhythm precision = absolute value of mean rhythm component within learning phase MINUS individual baseline of short rhythm component, the 500-millisecond inter-tap interval.

$F < 1$. The seductive details effect could also not be found in the subjective cognitive load rating scale, $F < 1$.

However, the secondary task performance confirmed that learners of the seductive details condition executed the rhythm not as precise ($M = 80.06$, $SD = 64.65$) as learners who learned without seductive details ($M = 36.70$, $SD = 25.75$), $t(27) = -2.39$, $p < .05$, $d = 1.25$, when analyzing the short rhythm component, the 500 milliseconds short tap-to-tap interval (Table 1). For the long rhythm component, no significant difference was found, $t(27) = -1.12$, *n.s.*

In addition, the correlation between the different cognitive load methods could not be shown, when comparing the subjective cognitive load rating scale with each of the two cognitive load indicators of the rhythm method. Neither the correlation between the subjective ratings and the short rhythm component was significant ($r = .02$, *n.s.*), nor the correlation between the subjective ratings and the long rhythm component reached significance ($r = -.01$, *n.s.*).

In sum, the significantly lower comprehension performance of learners under the seductive details condition is accompanied by significantly lower rhythm precision. Thus, learners undergo higher cognitive load under this condition. This difference in cognitive load could not be measured by subjective ratings, which argues for the sensitivity of the rhythm method for measuring cognitive load. In sum, results of Experiment 1 demonstrate first evidence for the suitability of the new method. Moreover, the result that there were no significant correlations found between the different cognitive load measures will be further examined in Experiment 2.

Experiment 2

In Experiment 2, no statistical significant between-group differences was detected in any of the control variables working memory capacity, $F(1, 44) = 3.53$, *n.s.*, spatial ability, $F < 1$, and prior knowledge, $F < 1$ (Table 2).

Time-on-task differed significantly over the two conditions, $F(1, 44) = 4.56$, $p < .05$, $d = 0.89$. Learners in the narration condition took more time to learn ($M = 10.52$, $SD = 5.02$) than learners who learned with on-screen text ($M = 8.03$, $SD = 2.54$; Table 2). Thus, time-on-task was used as a covariate in the following analyses. Separate ANCOVAs were conducted using learning performance scores (retention and comprehension) and cognitive load ratings, respectively, as dependent variables. Again, results and conclusions are the same when running only analyses of variance. In retention performance, no difference was found between the two learning groups, $F < 1$. A modality effect could be shown in comprehension performance in the expected direction, as the learning success was significantly higher for learners who learned with narration ($M = 3.81$, $SD = 1.36$) in contrast to learners who learned with on-screen text ($M = 2.75$, $SD = 1.61$), $F(1, 43) = 6.12$, $p < .05$, $d = 1.01$ (Table 2). However, the modality effect could not be found in the subjective cognitive load ratings, $F(1, 43) = 1.31$, *n.s.*

As in Experiment 1, the secondary task performance confirmed that learners of the on-screen text condition were not as precise in the execution of the rhythm ($M = 59.58$, $SD = 42.35$) when compared with learners who learned with narration ($M = 43.33$, $SD = 19.24$), $t(44) = 1.70$, $p = .05$, $d = 0.70$, when comparing the 500-millisecond short tap-to-tap interval of the rhythm (Table 2). Again, no significant differences were found, $t(44) = 0.57$, *n.s.*, when analyzing the 1500-millisecond long tap-to-tap interval.

As already shown in Experiment 1, the correlation between the different cognitive load methods could not be shown, when comparing the subjective cognitive load rating scale with each of the two cognitive load indicators of the rhythm method. Neither the correlation between the subjective ratings and the short rhythm component was significant ($r = .13$, *n.s.*), nor the correlation between the subjective ratings and the long rhythm component reached significance ($r = -.03$, *n.s.*).

Table 2. Experiment 2: Means and standard deviations of all dependent and control variables for both conditions

	Control on-screen text $n = 24$		Experimental narration $n = 21$	
	M	(*SD*)	*M*	(*SD*)
Working memory capacity (Max = 9)	3.33	(1.24)	3.95	(0.92)
Spatial ability (%)	63.15	(19.99)	68.21	(17.69)
Prior knowledge (Max = 13)	3.63	(3.17)	3.60	(3.08)
Time-on-task (min)	8.03	(2.54)	10.52	(5.02)
Retention (Max = 8)	5.58	(2.72)	6.31	(2.09)
Comprehension (Max = 6)	2.75	(1.61)	3.81	(1.36)
Cognitive load—subjective rating scale (Max = 7)	4.71	(1.23)	5.24	(0.89)
Cognitive load—dual-task method—long rhythm component[a] (milliseconds; 0 = absolute precision)	135.23	(102.93)	118.07	(78.72)
Cognitive load—dual-task method—short rhythm component[b] (milliseconds; 0 = absolute precision)	59.58	(42.35)	43.33	(19.24)

Note:
M, mean; *SD*, standard deviation; Max, maximum.
[a]Description of the long rhythm component: rhythm precision = absolute value of mean rhythm component within learning phase MINUS individual baseline of long rhythm component, the 1500-millisecond inter-tap interval.
[b]Description of the short rhythm component: rhythm precision = absolute value of mean rhythm component within learning phase MINUS individual baseline of short rhythm component, the 500-millisecond inter-tap interval.

These results confirm that significantly lower learning success of learners under the on-screen text condition is accompanied by significantly lower rhythm precision, indicating that learners undergo higher cognitive load under this condition. This again argues for the sensitivity of the rhythm method for measuring cognitive load. In sum, results of Experiment 2 are in line with those of Experiment 1 and show additional evidence for the suitability of the new method. The confirmation by Experiment 2 that there is no significant correlation between the subjective ratings and the rhythm method is discussed in the following section.

DISCUSSION AND FUTURE DIRECTIONS

The present work offers a new secondary task, which measures inhibition processes in an objective, direct, and continuous way as an indicator for executive control and cognitive load. This method is characterized by intra-individual measures, continuity, and sensory independence of the presentation mode of learning instructions (primary task) by using an internalized task. In addition and in contrast to other secondary tasks, the rhythm method includes inhibition processes, as learners have to stop tapping in the rhythm pauses, which can be used as an indicator for executive control and total cognitive load. Rhythm precision therefore allows a precise measurement of cognitive load in the learning process.

Conclusions for research on cognitive load measures

Both experiments of the present work provide evidence for the suitability and sensitivity of the rhythm method. In Experiment 1, learners reached a significantly higher learning performance without seductive details and showed a higher rhythm precision; in Experiment 2, learners also reached a significantly higher learning performance with narration and showed a higher rhythm precision.

One point that is interesting within the results of the present study is that only the short rhythm component was a good indicator for differences in cognitive load. This could be due to the expected inhibition process when learners have to stop frequent tapping, which seems only to influence the short inter-tap intervals directly after having to inhibit tapping. This phenomenon of late influence of inhibition processes on further processes is consistent with that of studies that are using event-related potentials, which are very often only found after the event and not directly in time with the event (e.g., Berti, 2008). These results demonstrate that inhibitory processes seem not to have a direct influence on the inhibition task (in the case of the present rhythm method: to inhibit tapping within the long tap-to-tap interval) but not on the following action (in the rhythm method: to stop inhibition by tapping the next short tap-to-tap interval in the right frequency).

Concerning the subjective cognitive load ratings, the present study shows results that have to be discussed. The fact that both experiments could not confirm any cognitive load effect by the subjective rating scale and that no correlation between the two indicators of the rhythm method and the subjective rating scale was found is a hint to limitations of using subjective ratings. The first conclusion might be to state that self-estimated mental effort is not a valid method to measure cognitive load. However, in cognitive load research, there are many studies showing the usability and sensitivity of subjective ratings. These contrasting results could be due to the material and learning effects used within the present experiments. In cognitive load research, the cognitive load effects are often found with problem-solving tasks, when, for example, having a mathematical problem to solve step by step. These tasks are different to the continuous learning task in the present ATP-Synthase material, where learners do have to cognitively integrate a complex learning content in a continuous learning session. This may be one reason that the subjective rating scale was not appropriate to measure such a subtle process of mental effort in contrast to tasks with high peaks of cognitive load due to problem-solving processes. This explanation however needs to be empirically tested within future studies. In addition, an advantage of continuous measurement should be mentioned here: Using the subjective rating scale frequently within the learning task is one solution to gain process information; however, it cannot detect all peaks and troughs of cognitive load. The rhythm method is in addition a high-resolution measure that is more precise and therefore advantageous when it is the goal to analyze fluctuations in cognitive load within the learning process.

Another plausible explanation to why the present study did not find a correlation between the subjective ratings and the rhythm method could be due to different concepts that are measured by these indicators. The rhythm method measures inhibition and executive control processes, whereas the self-experienced invested mental effort could always be different to the indeed found cognitive processing. Perhaps learners are not able to estimate their mental effort because they are distracted by surface aspects of the task at hand, which lead to estimate the difficulty or complexity of the task, but not their invested mental effort. This could explain that there are no differences measured especially when varying the modality or seductive details that are associated with extraneous load and do not explicitly vary the complexity of the learning task itself in a way that is obvious for learners. This would be in line with the results of a study by DeLeeuw and Mayer (2008), where it was concluded that subjective ratings are valid and reliable for measuring intrinsic cognitive load and that dual-task measures are the method of choice for measuring differences in extraneous cognitive load. This conclusion for the present study would therefore lead to the theoretical implication that extraneous cognitive load is conceptually associated or even based on inhibition and executive control processes, which can therefore be most appropriately measured by high-resolution measures such as the rhythm method.

Conclusions for research on the modality and the seductive details effect

The modality effect and the seductive details effect are both learning effects that are recently discussed. Some boundary conditions of the modality effect were recently summarized (Schnotz, 2011), and the negative learning effect of seductive details was also discussed to reverse owing to the motivational factor of such interesting information when learners have enough cognitive resources free for integrating such additional information (Park, Moreno, Seufert, & Brünken, 2011; Park, Plass, & Brünken, 2014). However, confirming the meta-on the seductive details effect of Rey (2012) and meta-analyses on the modality effect (e.g., Ginns, 2005 analysis), the present study shows the seductive details effect and the modality effect again. Perhaps the overall higher cognitive load due to the dual-task situation can explain these clear results of the present study in contrast to the contradicting studies mentioned earlier, which challenge both well-known multimedia effects.

Limitations

In sum, validity, reliability, and sensitivity of the rhythm method should be confirmed again with other learning effects. The new dual-task method should, for example, be examined in other multimedia studies, which use learning-conducive tools such as mental animation tasks that are associated with cognitive load or the variation of expertise to vary cognitive load. Another future direction is to compare the present method with other methods for measuring cognitive load in an empirical study to gain more insight into differential effects of operationalized cognitive load methods. These validation studies should also invest in refining the instruments on how to measure the rhythm precision with the rhythm method by using an e-prime version (software for psychological studies) that also includes a rhythm training session with systematic feedback before the learning phase. This rhythm method tool should be prepared in the way that it can be distributed also to other researchers who want to apply the rhythm method in further research.

Especially studies that compare different objective methods for measuring cognitive load could show the different advantages and disadvantages of the diverse objective methods. These studies could lead to the practical implication which of the objective methods should be chosen for the different situations where it is interesting to measure cognitive load. With the results of the present study, it can already be concluded that the rhythm method has some advantages in contrast to other already mentioned objective methods. With the rhythm method, a low reactive dual-task method is available that can be recommended for different presentation formats (auditory or visual) of the learning material. Moreover, the rhythm method does not require conscious disruption of the learning process because no auditory or visual cues are necessary. Disadvantages in contrast to other objective methods should also be mentioned. First, the rhythm method requires a short training sequence before the learning phase. In addition, it is highly recommended to use the rhythm method only within laboratory studies, where participants take part in an individual learning lesson. For group setting, foot tapping might not be the method of choice owing to the emerging noise level. However, in order to use the rhythm method within classroom settings, for example, the tapping could be realized by finger tapping that should lead to a lower noise level.

Future directions

One question still remains unanswered. With the present work, it is shown that the rhythm method is suitable to measure cognitive load. Further research is now needed to focus on the modality specificity of this new secondary task. The rhythm method is assumed not to be modality specific in comparison with visual or auditory secondary tasks, which were introduced in cognitive load research so far. However, an interesting phenomenon was observed during the study and when speaking with professionals: Some professional musicians as well as some of the participants reported that for them, the rhythm method was a visual task; others reported it to be an auditory task. Thus, further studies should investigate in the research question if the rhythm method is an auditory or a visual task and should therefore be associated with one of the subsystems of Baddeley's (1992) working memory model, the phonological loop or the visuospatial sketchpad. On the other hand, it is also possible that the rhythm method is a modality-unspecific task that is clearly associated with the central executive. An examination of these research questions could be realized by the variation of cognitive load by clear and short tasks that can be varied in their complexity and demand on cognitive processing such as an *n*-back task. The participant is presented with a sequence of stimuli. The goal of the participant is to press on the button when the current stimulus matches the one from *n* steps earlier in the sequence. The variable *n* can be adjusted to make the task more or less complex ($n=2$, low complexity vs. $n=3$ vs. $n=4$…, high complexity) and thereafter varying cognitive load. The use of different visual versus auditory material should not result in any differences in the variation of cognitive load owing to the task complexity if the rhythm method is modality unspecific. In detail, an *n*-back task with $n=2$ should result in low cognitive load with auditory as well as with visual material, and an *n*-back task with $n=4$ should result in high cognitive load also independently of the given material.

One last idea that evolved during the implementation of the innovative secondary task is that probably music can have a training function in cognitive processing as the present task shows that active musicians always have to regulate executive control processes, which could foster cognitive working memory processes in general. Therefore, studies that use the present rhythm method not only for measuring cognitive load but also as a training component to foster working memory processes especially by children and older adults could be an interesting future direction.

ACKNOWLEDGEMENTS

This research was supported by the German Federal Ministry of Education and Research. The authors wish to thank Andreas Korbach, Laura Kielinger, and Patrick Beyer, who were very engaged to realize the present study by programming the learning material and managing the recruitment and instruction of participants. We also thank Kira Carbonneau for proofreading as well as the editor Tracy Packiam Alloway and all anonymous reviewers for their very helpful comments.

REFERENCES

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438.

Baddeley, A. D. (1992). *Working memory*. New York: Oxford University Press.

Berti, S. (2008). Cognitive control after distraction: Event-related brain potentials (ERPs) dissociate between different processes of attentional allocation. *Psychophysiology, 45*(4), 608–620.

Brown, R. G., & Marsden, C. G. (1991). Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. *Brain*, *114*(1), 215–231.

Brünken, R., & Leutner, D. (2001). Aufmerksamkeitsverteilung oder Aufmerksamkeitsfokussierung? Empirische Ergebnisse zur "Split-Attention-Hypothese" beim Lernen mit Multimedia [Attention splitting or attention focusing? Empirical results concerning the split attention hypothesis in learning with multimedia]. *Unterrichtswissenschaft Zeitschrift für Lernforschung*, *29*(4), 357–366.

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 53–61.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science*, *32*(1–2), 115–132.

Brünken, R., Plass, J. L., & Moreno, R. (2010). Current issues and open questions in cognitive load research. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 253–272). Cambridge: University Press.

Brünken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load. In J. L. Plass, R. Moreno, & R. Brünken (Eds), *Cognitive load theory* (pp. 181–202). Cambridge: University Press.

Brünken, R., Steinbacher, S., Plass, J., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, *49*(2), 109–119.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology*, *10*(2), 151–170.

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. Plass, R. Moreno, & R. Brünken (Eds), *Cognitive load theory* (pp. 203–208). New York, NY US: Cambridge University Press.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*(2), 105–134.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223–234.

Ekstrom, R. B., French, J. W., & Harmann, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Emerson, M., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*(1), 148–168.

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*(1), 41–57.

Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Psychology Review*, *21*(1), 43–54.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*(4), 313–331.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *26*(5), 519–532.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414–434.

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.

Hegarty, M., Shah, P., & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition*, *28*(3), 376–385.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(1), 1–17.

Kirschner, F., Kester, L., & Corbalan, G. (2011). Cognitive load theory and multimedia learning, task characteristics and learning engagement: The current state of the art [special issue]. *Computers In Human Behavior*, *27*(1), 1–4.

Koch, B., Seufert, T. & Brünken, R. (2008). One more expertise reversal effect in an instructional design to foster coherence formation. In J. Zumbach, N. Schwartz, T. Seufert & L. Kester (Eds), *Beyond knowledge: The legacy of competence. Meaningful computer-based learning environments* (pp. 207–215). Berlin: Springer.

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*(4), 569–587.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147–158). Cambridge: University Press.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*(1), 49–63.

Mayer, R. E. (2005a). Principles for managing essential processing in multimedia learning. Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 169–182). Cambridge: University Press.

Mayer, R. E. (Ed.), (2005b). *The Cambridge handbook of multimedia learning*. Cambridge: University Press.

Mayer, R. E. (2009). *Multimedia learning*. Cambridge: University Press.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312–320.

McCrudden, M. T., & Corkill, A. J. (2010). Verbal ability and the processing of scientific text with seductive detail sentences. *Reading Psychology*, *31*(3), 282–300.

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*(3), 149–158.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*(3), 598–610.

Moreno, R., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. L. Plass, R. Moreno, & R. Brünken (Eds), *Cognitive load theory* (pp. 9–28). New York, NY US: Cambridge University Press.

Münzer, S., & Holmer, T. (2009). Bridging the gap between media synchronicity and task performance: Effects of media characteristics on process variables and task performance indicators in an information pooling task. *Communication Research*, *36*(1), 76–103.

Oberauer, K. K., Süß, H. M., Schulze, R. R., Wilhelm, O. O., & Wittmann, W. W. (2000). Working memory capacity: Facets of a cognitive ability construct. *Personality and Individual Differences*, *29*(6), 1017–1045.

Paas, F. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434.

Paas, F., & Kester, L. (2006). Emerging topics in cognitive load research: Using learner and information characteristics in the design of powerful learning environments [special issue]. *Applied Cognitive Psychology*, *20*.

Paas, F. C., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4.

Paas, F. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71.

Paas, F. C., & Van Merriënboer, J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(4), 737–743.

Paas, F. C., & Van Merrienboer, J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*(4), 351–371.

Park, B. (2010). Testing the additivity hypothesis of cognitive load theory (Doctoral dissertation, Saarland University, Saarbrücken, Germany). Retrieved from http://scidok.sulb.uni-saarland.de/volltexte/2010/3478/ [accessed November 21, 2014].

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*(1), 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.

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive Load beim Lernen aus Lösungsbeispielen [Cognitive load during learning from worked-out examples]. *Zeitschrift fur Pädagogische Psychologie*, *17*(2), 93–101.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, *7*, 216–237.

Rowland, E., Skinner, C. H., Davis-Richards, K., Saudargas, R., & Robinson, D. H. (2008). An investigation of placement and type of seductive details: The primacy effect of seductive details on text recall. *Research in the Schools*, *15*(2), 80–90.

Rummer, R., Schweppe, J., Fürstenberg, A., Seufert, T., & Brünken, R. (2010). Working memory interference during processing texts and pictures: Implications for the explanation of the modality effect. *Applied Cognitive Psychology*, *24*(2), 164–176.

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, *34*(2), 344–355.

Schnotz, W. (2011). Colorful bouquets in multimedia research: A closer look at the modality effect. *Zeitschrift für Padagogische Psychologie*, *25*(4), 269–276.

Seufert, T. Schütze, M. & Brünken, R. (2009). Memory characteristics and modality in multimedia learning: An aptitude–treatment-interaction study. *Learning and Instruction*, *19*(1), 28–42.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge: University Press.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*(3), 185–233.

Sweller, J., Van Merrienboer, J. G., & Paas, F. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.

Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review*, *2*(1), 1–12.

