

Research Article

Cognitive Load of Learner Control: Extraneous or Germane Load?

Mieke Vandewaetere^{1,2} and Geraldine Clarebout³

¹ Centre for Instructional Psychology and Technology (CIP&T), KU Leuven, Dekenstraat 2, 3000 Leuven, Belgium

² ITEC-IBBT-KU Leuven Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium

³ Center for Medical Education, Faculty of Medicine, Herestraat 49, Box 400, 3000 Leuven, Belgium

Correspondence should be addressed to Mieke Vandewaetere; mieke.vandewaetere@kuleuven-kortrijk.be

Received 11 January 2013; Accepted 31 March 2013

Academic Editor: Gwo-Jen Hwang

Copyright © 2013 M. Vandewaetere and G. Clarebout. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Computer-based learning environments become more tailored when learners can exert control over one or more parts of the learning process. Learner control (LC) demands additional efforts of learners because, in addition to learning, they also have to monitor that learning. As a consequence, LC may cause additional cognitive load and even cognitive overload. The central question in this study is what type of cognitive load is induced by LC and whether the experienced load is related to learning outcomes. For this study, half of the students had control over task selection, while the other half had not. Within each condition, students were assigned to a single treatment, with the primary task to solely focus on the learning content, and a dual treatment, comprising a primary task and a secondary task. The results indicate that LC did not impose higher cognitive load as measured by secondary task scores and mental effort ratings.

1. Introduction

In a learner-controlled environment, learners can have full or partial control on pace, sequencing, content, or presentation [1]. Irrespective of the type of LC, all types have in common that they allow the learners to (partly) define their own learning process according to their own needs, abilities, and interests.

LC is suggested to promote learning and to enhance motivation [2–4]. However, empirical research on LC has revealed mixed results on the effectiveness of LC (for overviews, see [4–7]). The inconsistent results are largely due to how LC has been defined and implemented in different studies on one hand [1] and to the characteristics of learners on the other hand.

Corbalan et al. [2] defined three threats to the effectiveness of LC as an instructional strategy. First, learners should perceive that they are in control. If learners do not perceive control or consider the available control as irrelevant, then LC will be ineffective. A second threat to the effectiveness of LC is defined by the learner's characteristics.

Several characteristics have been put forward to be related with learner control effectiveness: a learner's ability and age [8, 9], learning style [5], positive attitudes and motivation [10]; and self-efficacy and self-regulation skills [11]. Eom and Reiser [12] demonstrated that learners with low self-regulation skills performed better when they did not receive any control compared to a situation in which they received LC. Learners with high self-regulation skills are better able to effectively monitor their learning process and are thus capable of selecting the learning materials that are most beneficial for learning. A last threat to effective LC [2] is that LC can be associated with additional cognitive processing demands. Giving control to learners may result in cognitive overload, as has been demonstrated in hypermedia learning research [1]. Cognitive overload is caused when cognitive resources are no longer available for performing the main tasks because the resources are used for metacognitive processing, for instance, when dealing with LC [1]. In the next paragraph, the cognitive load theory (CLT) is presented and applied to the cognitive load of LC as an instructional strategy.

1.1. Cognitive Load of Learner Control. In learner-controlled environments, selecting and sequencing information, and monitoring one's own rate of progress imposes additional cognitive load onto learners [13]. LC demands free cognitive resources and posits an additional load on cognitive processing [1, 14, 15]. Moreover, Scheiter and Gerjets [1] caution for the risk of cognitive overload: if learning a particular content or executing a certain task puts too much load in working memory, then learning or task execution will be hampered [16]. In line with this, Granger and Levine [17, page 193] argued that *"high learner-controlled environments place a higher burden for learning [...] and cognitive load onto trainees."* As a consequence, *"if trainees want to flourish in high learner-controlled environments, they require a higher working memory capacity"* [17, page 193].

Cognitive load of instructional conditions such as LC is not to be considered as a by-product of learning but should rather be considered as the *"major factor determining the success of an instructional intervention"* [18]. This requires the measurement of cognitive load, not only from a cognitive processing perspective, but also from a motivational perspective [18].

1.2. Types of Cognitive Load and Their Measurement. CLT originally considered cognitive load as being threefold with three types of cognitive processing that are involved in cognitive load [19, 20]. First, extraneous processing can contribute to extraneous load in that this type of processing does not contribute to learning as such. Extraneous load can be attributed, for instance, to poor layout or to a surplus of information on a screen. Second, there is intrinsic load, related to intrinsic or essential processing. Intrinsic load is determined by the element interactivity or the complexity of the materials to be learned. The more interacting elements that need to be processed simultaneously, the higher the intrinsic load [21]. A last form of cognitive load is germane load, stemming from germane or generative processing [19]. This type of load is associated with processes that are directly relevant to learning such as scheme construction and automation [22]. However, in a recent review of CLT, Sweller [23] invited to define germane load in terms of intrinsic load and to associate germane load also with the element interactivity. Consequently, germane load is then assumed to represent the working memory resources that are required to deal with intrinsic cognitive load, caused by the complexity of the learning materials [24].

Summarized, when a learner experiences cognitive load, this can thus be caused by the intrinsic nature of the task (intrinsic load) or by the way the task in the information related to the task is presented. When cognitive load is caused by activities or information that do not directly foster learning, this is called extraneous load; germane load is then the consequence of processing information that contributes to learning [25]. Paas et al. [24] argued that extraneous cognitive load and germane cognitive load should be considered as communicating vessels. Reducing extraneous load, this frees cognitive resources that can be used for processing more germane load.

Generally, three measurements of cognitive load can be distinguished: self-reporting of invested mental effort, physiological measurements, and dual-task measurements [26]. First, there is the self-reporting through rating scales as a method to look at the intensity of effort that was expended by learners. Typically, a Likert scale is applied by which learners need to rate the amount of mental effort they have invested in completing the task [27]. The scale range then goes from "(very) very low" to "(very) very high." Paas et al. [26] stated that mental effort rating scales have proved to be valid, reliable, and unobtrusive. However, many authors do not ask participants to rate mental effort but to rate the perceived difficulty of the tasks [28, 29]. To some extent, the two measurements are related but important differences should be noted. van Gog and Paas [28] state that the mental effort rating focuses on the process of solving or completing the task, while the perceived difficulty is related to the task as such. As a consequence, tasks that are considered as extremely difficult by learners might be accompanied by low mental effort ratings, because learners did not want to invest too much effort. When learners perceive tasks as too difficult, they will not be motivated to invest sufficient mental effort in order to complete the tasks [18].

A second measurement method comprises physiological measurements. Despite the promising practices of physiological measurements, only a limited number of studies have been using them, partly due to the intrusive nature of such measurements [16, 19].

A last measurement method originated from the dual-task or secondary-task approach, borrowed from cognitive psychology research [16, 26, 30]. In this methodology, the learning task is considered as the main or primary task, and an additional, secondary task is offered to the learner. Often, secondary tasks require a fast reaction from the learner such as to press a spacebar as soon as the background color of the environment changes [19], to react as fast as possible to a letter that appears on the screen [31], or to press the spacebar when the color of a black letter changes into red [32]. The performance on the secondary task then indicates the cognitive load that is induced by the primary task [31]. The more resources are needed for processing the primary task, the less resources are available for secondary task processing and the slower the reaction times are on the secondary task. If the outcomes on the secondary task are low, this means that more cognitive resources were spent on processing the primary task. A major advantage of the secondary-task approach, contrary to a questionnaire, is that it offers concurrent measures of cognitive load, as it measures cognitive load on working memory while performing the learning task [16]. An important drawback of this approach however, is that it may interfere with the primary task or the learning as such [26].

In a study of DeLeeuw and Mayer [19], it was demonstrated that the different types of cognitive load (i.e., intrinsic, extraneous, and germane) may be represented by the outcomes on three different measurements of cognitive load: mental effort ratings, difficulty ratings, and secondary task outcomes. The results indicated that intrinsic load was more reflected by the learners' mental effort ratings, extraneous load was reflected by measurements of response time (on

the secondary task), and germane load could be reflected by learners' difficulty ratings. As a conclusion, the three measurements of cognitive load reflect different types of cognitive load hereby supporting a triarchic nature of cognitive load. This was also supported by the results of the DeLeeuw and Mayer study, in which no correlations between the three measurements were found.

1.3. Aim of This Study. LC induces additional cognitive load that, in turn, may cause cognitive overload due to the learner's limited working memory capacity (WMC). The central question in this study is then what type of cognitive load is induced by LC and whether the experienced load is related to learning outcomes. According to DeLeeuw and Mayer [19], if LC processing results in deep cognitive processing such as integrating knowledge and mentally organizing learning materials, then LC is related to germane or generative processing and induces germane load. If, on the other hand, LC leads to cognitive processes that do not directly support learning, then LC is likely inducing extraneous load. In line with this, Kirschner et al. [33] argued that situations of discovery learning (with high levels of LC) may result in high extraneous cognitive load as learners search for possible problem solutions, suggesting a link between LC and extraneous load.

As DeLeeuw and Mayer [19] demonstrated, if learner-controlled interaction engages the learner in deep processing (germane load), then this should affect the learners' difficulty ratings. If, however, learner-controlled interaction does not lead to deep processing but rather imposes extraneous load, this should be reflected in varying outcomes on the secondary task.

In this study, we aim to answer the following research questions. (1) Does LC impose additional cognitive load? And if so, is this additional load extraneous or germane load? (2) How is the cognitive load of LC related with learning outcomes?

2. Method

2.1. Participants. One hundred forty-five first-year bachelor students educational sciences took part in this study. All participants were recruited from the first year of educational sciences and participated in this study as part of the course "Learning & Instruction." The age range was between 18 and 20 years old, with a mean age of 19 years. Of all participants, 95% were females.

2.2. Design and Procedure. Prior to the experiment, prior knowledge on English tenses, WMC, self-regulation skills, and preexperimental motivation were measured. Afterwards, participants were assigned to the between-subjects conditions in such a way that each condition had an equal distribution of learners with low, intermediate, and high WMC and prior knowledge. The two conditions in this design are (1) no control (NC), in which participants had to complete preselected exercises on English tenses, and (2) learner control (LC), in which participants had full control over the exercises they wanted to complete. Within each

condition there were two learning phases: single-task and dual-task treatments which were randomly administered as learning phases, so that each participant completed both treatments. Single and dual treatments were administered in order to make within-subject measurements of mental effort ratings and difficulty ratings possible. After each learning phase, the postlearning questionnaire was administered, comprising a knowledge test, overall perceived difficulty, and a postexperimental motivation measurement.

2.2.1. Preexperimental Phase. The pretest measurements were organized five weeks before the actual experiment took place. During the pretest, participants first received a short introduction on English tense conjugation and then completed the prior knowledge test of 15 items. After completing the pretest, self-efficacy for learning was measured. Also, WMC was measured. Based on the results of the prior knowledge test and the WMC test, participants were assigned to the two conditions in such a way that all conditions had an equal representation of prior knowledge and WMC.

2.2.2. Learning Phase. In the learning phase, each participant received the two treatments in random order. Half of the participants received the single treatment first, followed by the dual treatment. The other half received the treatments in a reversed order. In the single treatment, the primary and single task was to complete 35 (no control condition) or at least 35 (LC condition) exercises. In the secondary treatment, participants received the same primary task, with an additional secondary task. As a secondary task, participants had to remember meaningful words (e.g., castle, pizza, and hotel) that were randomly presented after completing an exercise (e.g., "castle" after the second exercise, "pizza" after the eighth exercise, etc.). All the words were extracted from the GOSPAN WMC test [34]. Participants were not allowed to write down the words and were asked afterwards to remember as many of the words as possible. Hence, the secondary task focused on processing and rehearsal of information, rather than on the learners' attention to react as fast as possible to words. In total, 11 words were presented during the secondary treatment, both in the no control and LC condition. Table 1 presents the design schematically. Instructions for the conditions and treatments are presented in appendix.

2.2.3. Postexperimental Phase: Questionnaires and Posttest. After the learning phase, participants completed the posttest and the postexperimental motivation questionnaire (IMI) and answered additional questions. All participants were able to complete the whole study in 90 minutes.

2.3. Materials

2.3.1. Electronic Learning and Testing Environment. The open source software Moodle was adjusted to an item-based electronic language learning environment, focusing on English tenses (verb conjugation). Both the pre-test, learning phase

TABLE 1: Design of the study.

Pretest	Learning phase		Posttest N = 145		
	Condition	Treatment			
Pretest and questionnaires	No Control	Single	Secondary	Posttest n = 37	
		Secondary	Single		n = 36
	Learner Control	Single	Secondary		n = 37
		Secondary	Single		n = 35

and post-test were administered in this learning environment. Tracking and logging data on learner behavior (e.g., number of exercises completed, number of attempts, and duration of learning) were collected and stored on an external server.

2.3.2. Preexperimental Phase: Introduction and Pretest. All participants had some prior knowledge, originating from English classes in compulsory education. To activate this prior knowledge [35], participants individually consulted a brief refresher course on the function and use of the most prominent English tenses and verb conjugation (future, past, and present, simple and progressive). Subsequently, all participants completed a school-like pre-test, consisting of 15 multiple-choice items on the use and function of English tenses. This test was administered to check for equality of prior knowledge between conditions. The test was not time-limited.

After completing the pre-test, participants' WMC was measured. Therefore, a computerized group administrable operation span task was offered to each participant individually (GOSPAN, [34]). On the computer screen, participants saw simple mathematical operations (e.g., "IS(4/2) - 1 = 5?") and were instructed to indicate (by pressing a key) whether the operations were correct or wrong. After each response to a mathematical operation, a word appeared on the screen (e.g., "hotel", "car") and participants were instructed to recall as many of the unrelated words as possible. Each stimulus thus consisted of deciding on the correctness of a mathematical operation and remembering a word that was provided afterwards. The number of operation-word stimuli in one set could vary from two to six. After each set, participants were asked to write the words down. For each operation-word stimulus, responses and response latencies were recorded. A participants' score on the GOSPAN test was then the sum of recalled words for all sets that were recalled completely and in correct order. In line with the approach of De Neys and Dieussaert [36], participants were excluded from further analyses when they had made more than 15% math errors (i.e., deciding whether a mathematical operation is correct or not, 1 participant) or when their mean response latencies deviated more than 2.5 standard deviations from the sample mean (i.e., waiting too long to respond, 2 participants). The mean response latency was then $\bar{X} = 4031$ ms (min = 2680, max = 6024, SD = 67.86), and the average number of words recalled from the test was $\bar{X} = 36.27$ (min = 13, max = 60, and SD = 0.82).

To measure self-regulation and prior motivation for learning English tenses, participants were asked to complete

one subscale of the Motivated Strategies for Learning Questionnaire (MSLQ, [37]) and four subscales of the Intrinsic Motivation Inventory (IMI, [38, 39]). The MSLQ subscale comprised six-point scaled items related to self-efficacy for learning and performance (8 items, e.g., "I'm certain I can understand the most difficult material presented in the readings for this course", $\alpha = 0.92$). Prior motivation was measured by the IMI subscales interest/enjoyment (7 items, e.g., "This activity was fun to do", $\alpha = 0.88$), perceived competence (6 items, e.g., "I think I am pretty good at this activity", $\alpha = 0.86$), effort/importance (5 items, e.g., "I put a lot of effort into this", $\alpha = 0.74$), and perceived choice (7 items, e.g., "I believe I had some choice about doing this activity", $\alpha = 0.86$). All items were measured on a six-point rating scale and altogether measured a learner's prior motivation. The reliability of the motivation questionnaire (as measured by the four subscales) was $\alpha = 0.89$.

2.3.3. Learning Phase. In total, 106 exercises on English tenses were developed, focusing on recognition and reproduction. All exercises could be categorized along two dimensions. The first dimension represented the theme on which the exercise focused. There were eight themes (e.g., future tenses, present perfect simple versus present perfect progressive, and mix of tenses), and exercises within a theme were grouped by item type (second dimension). Per theme, three groups of exercises were available: multiple-choice questions, fill-in questions, and a mix of both types.

In the no control condition (NC), participants received a fixed set of 35 exercises to solve. The control condition gave participants LC over selection of exercises in terms of theme (tense) and type of exercise. In the LC condition, a counter was shown at the top of the screen to indicate how many exercises learners already had completed. After completing 35 exercises, participants could start the posttest.

All exercises were binary scored (zero for wrong answer, one for correct answer). Participants also had free access to a set of resources, such as one-page screens that contained information on the use and functionality of a specific tense.

In addition to the exercises, participants were requested to complete mental effort ratings after each task or set of exercises ("How much effort did it cost you to complete this task?"). Participants indicated the mental effort they had invested on a 7-point rating scale, ranging from 1 = very little effort to 7 = very much effort.

2.3.4. Postexperimental Phase: Posttest. After the learning phase of each of the two treatments (i.e., single task and dual

task), all participants received a new set of ten multiple-choice items, with equal content for the conditions (NC and LC) but with different content for treatment. In the posttests, items focusing on all tenses were included. The post-tests were not time-limited.

2.3.5. Motivation and Learners' Perceptions of Difficulty. Post-experimental motivation was measured after each treatment again by the IMI questionnaire. This resulted in two measurements per subscale, for each participant: one for the single treatment and one for the dual treatment. The five subscales are interest/enjoyment (7 items, $\alpha = 0.92$), perceived competence (6 items, $\alpha = 0.93$), effort/importance (5 items, $\alpha = 0.89$), value/usefulness (6 items, $\alpha = 0.90$), and perceived choice (7 items, $\alpha = 0.95$). All items were measured on a 6-point rating scale. Reliability of the total post-experimental motivation questionnaire (as measured by the five subscales) is $\alpha = 0.95$. All subscales correlated significantly positive ($P < 0.01$).

After every treatment, participants were asked to rate the difficulty of the exercises they had completed (on a 7-point rating scale with 1 = very easy to 7 = very difficult). Other measurements are learning outcomes as defined by course score and outcomes on the secondary task. Secondary task outcome was determined by the number of correctly remembered words on all words that had been presented during the dual-task treatment. In total, eleven words were presented, and the mean of correctly remembered words was 5.74 (SD = 2.29), with a minimum of zero and a maximum of eight correct words that were written down after the dual-task treatment in the learning phase.

3. Results

For all analyses, the significance level was set to $\alpha = 0.05$. Inspection of box plots and Q-Q plots for prior knowledge and preexperimental motivation revealed three outliers with extremely low prior knowledge or prior motivation. These participants were discarded from further analyses.

The equality of conditions (no control and LC) was ascertained for prior knowledge, as measured by the total score on the pre-test ($F(1, 140) = 1.35, P = 0.25$), prior motivation ($F(1, 126) = 2.98, P = 0.09$), self-efficacy for learning ($F(1, 137) = 2.44, P = 0.12$), and WMC ($F(1, 138) = 0.10, P = 0.75$).

Table 2 summarizes the mean scores per condition (standard deviations between brackets) of all variables measured in this study.

3.1. Are Ratings Confounded by the Presence of a Secondary Task? All participants completed a single-task and dual-task treatment, randomized between participants. In the single-task treatment, participants completed tasks and rated the invested mental effort and difficulty. In the dual-task treatment, participants did the same ratings but had an additional secondary task that could possibly interfere with the ratings of mental effort and difficulty. To ascertain that learners' ratings were not affected by the presence of

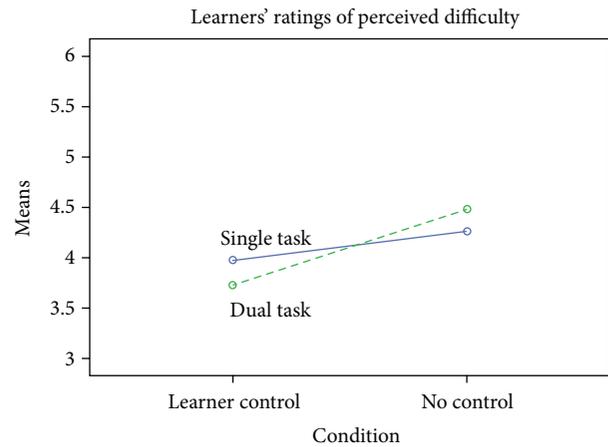


FIGURE 1: Contamination of secondary task on difficulty ratings.

the secondary task, two MANOVAs were executed. In the first, the dependent variables were the mental effort ratings of the single-task and dual-task treatment. The second analysis treated the difficulty ratings of the single- and dual-task treatment as dependent variables. In both analyses, condition (LC and NC) was entered as a factor to control for differential interferences dependent on condition.

With respect to mental effort ratings, no differences were found between treatments (Wilks' $\lambda, F(1, 139) = 0.43, ns$). This indicates that a learner's ratings were equal across the single- and dual-task treatment. Equality of ratings for the treatments was the same in both conditions, since the interaction between treatment and condition was not significant (Wilks' $\lambda, F(1, 139) = 2.06, ns$).

In Figure 1, the learners' difficulty ratings are presented. Learners' ratings were equal across treatments (Wilks' $\lambda, F(1, 132) = 0.002, ns$); however, the interaction between treatment and condition was significant (Wilks' $\lambda, F(1, 132) = 10.421, P = 0.002$, and $\eta_p^2 = 0.07$). For the single and dual treatment, the ratings of difficulty were lower in the LC condition as compared with the NC condition. This difference was however more pronounced in the dual-task treatment. Difficulty ratings may thus have been confounded by the presence of the dual task. In the single-task treatment, the differences between ratings in the LC and NC condition are not significant, while they become significant in the dual-task treatment. The means and standard deviations of the average difficulty ratings per condition and per treatment are presented in Table 3.

3.2. LC and Extraneous Cognitive Load. Secondary task outcomes were measured by the number of words that learners remembered correctly after completing the tasks. Learners' secondary task outcome was entered as dependent variable in an ANCOVA with condition (NC versus LC) as factor and prior knowledge score, prior motivation, WMC, and self-efficacy as covariates. None of the covariates was significantly related to the difficulty ratings (all ns). Having control or not was not related to the learners' secondary task outcome ($F(1, 119) = 0.08, ns$).

TABLE 2: Means (standard deviations) of all measured variables, represented per condition.

	No control	Learner control
<i>N</i>	69	73
Preexperimental		
Prior knowledge test (max. 15)	12.02 (1.89)	11.67 (1.87)
Self-efficacy (max. 48)	32.87 (5.15)	31.34 (6.12)
Prior motivation (total)	98.73 (12.62)	94.63 (13.17)
Working memory capacity (max. 60)	36.35 (9.45)	36.42 (9.35)
Learning outcomes		
Course score (in %)	46.94 (13.80)	64.12 (10.762)
Post-test score (in %)	34.92 (11.50)	31.58 (11.73)
Mean mental effort rating (max. 7)	4.19 (.99)	3.80 (.99)
Score on secondary task (max. 11)	5.74 (2.21)	5.75 (2.39)
Difficulty rating (max. 7)	4.49 (1.04)	3.84 (1.11)
Postexperimental motivation (IMI)		
Interest/enjoyment (max. 42)	23.78 (6.51)	25.90 (6.18)
Perceived competence (max. 36)	18.68 (4.71)	21.49 (5.25)
Effort/importance (max. 30)	20.49 (4.58)	19.99 (4.94)
Perceived choice (max. 42)	23.34 (6.37)	23.97 (6.87)
Value/usefulness (max. 36)	25.85 (5.07)	26.27 (5.33)

TABLE 3: Means (standard deviations) of difficulty ratings.

Treatment	Condition	Mean (SD)
Singletask	Learner control	4.07 (1.14)
	No control	4.27 (1.06)
Dual task	Learner control	3.82 (1.12)
	No control	4.52 (1.03)

3.3. *LC and Germane Cognitive Load.* Learners' difficulty ratings were entered as dependent variable in an ANCOVA with condition as factor. Learners' prior knowledge score, prior motivation, WMC, and self-efficacy served as covariates. From the covariates, only a learner's self-regulation was significantly related to the difficulty ratings ($F(1, 115) = 4.55$, $P = 0.04$, and $\eta_p^2 = 0.04$). The higher a learner's self-regulation, the lower the difficulty ratings after completing the tasks ($b = -0.04$).

The analysis also revealed a significant difference between the conditions for difficulty ratings ($F(1, 113) = 13.63$, $P < 0.001$, and $\eta_p^2 = 0.11$). Learners in the LC condition had lower difficulty ratings of tasks or reported higher germane load ($b = -0.70$) as compared to learners in the NC condition.

3.4. *LC and Intrinsic Cognitive Load.* Intrinsic cognitive load was measured by the learners' mental effort ratings. After completing a task, learners indicated the mental effort they had invested. The mean of all ratings was entered as a dependent variable in the analysis. Condition was entered as a factor, with prior knowledge, prior motivation, WMC, and self-efficacy as covariates.

Prior knowledge score ($F(1, 119) = 8.02$, $P = 0.002$, and $\eta_p^2 = 0.08$) and self-regulation ($F(1, 119) = 8.10$, $P = 0.002$,

and $\eta_p^2 = 0.08$) were significantly related to the reported mental effort. The lower the prior knowledge ($b = -0.14$) and the lower the self-regulation ($b = -0.5$), the higher the mental effort ratings of a learner.

Having LC or not was also related to the mental effort ratings ($F(1, 119) = 5.52$, $P = 0.009$, and $\eta_p^2 = 0.06$). Learners in the LC condition reported lower invested effort, or lower intrinsic load ($b = -0.43$) as compared to learners in the NC condition.

3.5. *Correlations between Measurements of Cognitive Load.* In the research of DeLeeuw and Mayer [19], the three measurements of cognitive load were not correlated, indicating that cognitive load is not a unitary construct but has a triarchic nature. Their results could not be fully corroborated in this study. We found a significant correlation between the mental effort ratings and difficulty ratings ($r = 0.66$, $P < 0.001$), indicating that, according to the arguments formulated by DeLeeuw and Mayer [19], intrinsic load and germane load were highly correlated. No correlations were found between learners' mental effort ratings and secondary task outcomes ($r = 0.04$, ns) and between difficulty ratings and secondary task outcomes ($r = 0.13$, ns). This correlational pattern was equal for the LC condition and the NC condition, as it is presented in Table 4.

3.6. *LC and Learning Outcomes*

3.6.1. *Course Score.* The three cognitive load measurements were entered as covariates in an ANCOVA to sketch their effect on the learners' course score. Condition was entered as a factor. A learner's course score was computed as the mean score (in percent) on all tasks completed during the learning phase. Of the cognitive load measurements, only

TABLE 4: Correlations between cognitive load measurements for LC and NC condition.

	Learner control			No control		
	Secondary task outcome	Mental effort rating	Difficulty rating	Secondary task outcome	Mental effort rating	Difficulty rating
Secondary task outcome (extraneous load)	1.00	0.09	0.17	1.00	-0.01	0.09
Mental effort rating (intrinsic load)		1.00	0.68*		1.00	0.59*
Difficulty rating (germane load)			1.00			1.00

*Significant with $P < 0.01$.

mental effort ratings were significantly related with learners' course scores ($F(1, 130) = 6.32, P = 0.01$, and $\eta_p^2 = 0.05$). The lower a learner's mental effort ratings, the higher the course score ($b = -3.42$). Also condition affected course score significantly and largely ($F(1, 130) = 50.23, P < 0.001$, and $\eta_p^2 = 0.28$), with learners in the LC condition scoring significantly higher ($b_{LC} = 15.32, b_{NC} = 0$).

3.6.2. Posttest Score. The same analysis was redone but with post-test scores as dependent variable. This score was computed as the mean score on the ten post-test items and is presented in percent. None of the cognitive load measurements was related to the learners' post-test score. Only condition was related with $F(1, 115) = 5.17, P = 0.03$, $\eta_p^2 = 0.04$. Learners in the LC condition had somewhat lower post-test scores ($b_{LC} = -5.05$) as compared to learners in the NC condition ($b_{NC} = 0$).

3.6.3. Postexperimental Motivation. The five subscales of the post-experimental motivation questionnaire were entered as dependents in a MANCOVA, with condition as factor and the three cognitive load measurements as covariates. Of the cognitive load measurements, mental effort ratings and difficulty ratings were significantly related with post-experimental motivation (Wilks' $\lambda, F(5, 123) = 3.84, P = 0.003, \eta_p^2 = 0.14$; Wilks' $\lambda, F(5, 123) = 7.67, P < 0.001$, and $\eta_p^2 = 0.24$, resp.). More specifically, learners' difficulty ratings were related to post-experimental perceived competence ($P = 0.002, \eta_p^2 = 0.07$) and to post-experimental perceived value and usefulness ($P = 0.006, \eta_p^2 = 0.06$). The higher the learners rated the difficulty of tasks, the lower their perceived competence ($b = -1.56$) and the higher the perceived value and usefulness of the tasks ($b = 1.56$). Although the learners' mental effort ratings revealed a multivariate significant relation with post-experimental motivation, a univariate test did not reveal significances. It can be argued that univariate tests ignore the correlations between the subscales of post-experimental motivation and that the multivariate significant effect is likely to be due to small effects of some of the variables that jointly indicate significance [40].

With respect to the factor condition, no significant differences between the LC and NC condition were found for post-experimental motivation (Wilks' $\lambda, F(5, 123) = 1.74, ns$).

4. Discussion

Giving control to learners is suggested to put additional load on learners' cognitive processing. When too much load is induced by LC this may hamper learning. Measuring the cognitive load of instructional interventions in order to align interventions with learners' processing capacities is thus of great importance in instructional research. Typically, cognitive load is considered as comprising three types of load: intrinsic, extraneous, and germane load. While intrinsic load and germane load are directly related to learning, extraneous load is not. DeLeeuw and Mayer [19] provided support for the three types of cognitive load and demonstrated that intrinsic load is more reflected by mental effort ratings, germane load is more reflected by difficulty ratings, and extraneous load is more reflected by secondary task outcomes. This study used the three measurements of cognitive load to sketch the effect of LC on cognitive load experienced by learners. It was hypothesized that differences in cognitive load, as imposed by learner control, would be related to differences in the cognitive load measurements. A second hypothesis was that learner control and experienced cognitive load would be related to learning outcomes such as course scores, posttest scores, and post-experimental motivation.

All learners completed two treatments: a single-task and a dual-task treatment. This was done to control for possible confounding effects of the secondary task on learners' ratings of mental effort and difficulty. As Paas et al. [26] stated, learners' performance on the secondary task can interfere considerably with the performance on the primary task. As a consequence, mental effort ratings and difficulty ratings of the primary task might be confounded by the presence of the secondary task. In this study, mental effort ratings, indicating intrinsic load, were not confounded by the presence of the secondary task. This was however the case with learners' difficulty ratings of the primary task. Differences on difficulty ratings between LC and NC were more pronounced when learners were in the dual-task treatment and processed both the primary and secondary tasks. Learners' ratings of difficulty and experienced germane load were probably confounded by the presence of the secondary task making the difference in germane cognitive load more pronounced. In the single-task treatment, learners' germane cognitive load was not related to having control or not. This might indicate that in the single-task treatment, a sufficient amount of cognitive resources was left to deal with the additional load imposed by LC. In the

dual-task treatment however, learners in the LC condition reported lower difficulty ratings as compared to learners in the NC condition. According to DeLeeuw and Mayer [19], this suggests that germane cognitive load in the LC condition was higher as compared to the NC condition. In line with the suggestions of Sweller [23], the increase in germane load reflects less working memory resources being spent on processing related to extraneous load and more on processing related with intrinsic load. The presence of the secondary task and the availability of learner control in this study were not related to more extraneous processing, but rather to more germane or intrinsic processing. One possible explanation for this is that the availability of control led to higher motivation, and that learners were more willing to invest mental effort in completing the tasks [18]. This suggestion could however not be demonstrated in this study, as learners in the LC condition did not report higher motivation.

The analyses in this study further revealed that LC was not related to extraneous load (as measured by secondary task performance), indicating that either learners had sufficient cognitive resources left to deal with LC, or that LC as instructional strategy was genuinely processed and led to higher germane load. This was confirmed by inspection of the difficulty ratings. Learners in the LC condition reported lower difficulty ratings, which is related to higher germane load. As germane load is stated to be associated with processes that directly contribute to learning [21], LC in this study contributed to learning without causing cognitive overload. However, LC was also related to lower invested effort or lower intrinsic load as compared to learners in the NC condition. It was expected that LC would enhance motivation, and that motivated learners are more willing to invest mental effort in processing the instructions [18]. Paas et al. [18] suggested that learners will be less motivated to invest mental effort when they perceive the tasks as not relevant for success, or when they perceive the learning task as not requiring much effort. Two explanations are possible for the lower mental effort ratings of learners in the LC condition. First, LC as such was not more motivating than NC. This might be due to the fact that learners did not perceive the control (e.g., by a lack of choice options) or did not perceive that they were in control over task selection [2], for example, because the choice options were not perceived as meaningful. Future research could address the issue of perceived relevance of choice options (e.g., choice over theme versus choice over difficulty level) and its impact on perceived control. A second explanation is that learners interpreted the rating scale differently as was intended by the research design. This was also mentioned by de Jong [16] in his overview of cognitive load measurements. Self-reporting measurements have been used in various ways and consequently, results cannot always be interpreted unequivocally. The question used in this study was "*How much effort did it cost you to complete this task?*", and learners may have interpreted this as the actual effort that was spent to the task, rather than the effort that was required to complete the task.

Summarizing, learners in LC reported higher germane load (reflected by lower difficulty ratings) and lower intrinsic load (reflected by lower mental effort ratings). Mental effort ratings and difficulty ratings were also found to be highly

correlated. This is in line with the suggestion of Sweller [23], who defined germane cognitive load in terms of intrinsic cognitive load and associated germane load also with element interactivity or the number of elements that learners need to process simultaneously. As such, working memory resources that are used to deal with intrinsic cognitive load "*are germane to the task at hand and so are referred to as germane cognitive load*" [23, page 127]. As such, creating measurements that are able to distinguish between germane and intrinsic load may be irrelevant. Future research could focus on psychometrical validations of several measurements, obtained after careful experimentation with various instructional designs that contain several degrees of intrinsic, extraneous, and germane load.

Based on the framework of CLT, it can be suggested that LC in this study did not impose additional load to the learners' cognitive processing. Reasons for this can be found in the nature of the learning materials. All participants in this study already had prior knowledge of English tenses. In fact, this knowledge was mainly acquired during high school education, where the majority of the participants followed English class at least one hour a week. Following the reasoning of Hasler et al. [41], the information on English tenses was already previously organized in long-term memory, hence putting no load on working memory. As learners acquire more expertise in the learning domain, they are better able to deal with higher element interactivity [21]. Paas et al. [24] further stated that the limitations of learners' working memory only apply to new, to be learned information. It is thus very likely that learning in this study was not affected by the additional load as imposed by LC or by the secondary task. If there was any increased load due to secondary task processing or LC processing, it did not touch upon or exceeded the limits of WMC.

Another remark to make is on the nature of the secondary task used in this study. Previous research on LC and the dual-task paradigm used time-based measurements to measure cognitive load. Participants had to react as soon as possible to a stimulus, hereby disturbing their current cognitive processes. The approach used in this study focused on placing a burden on working memory and on increasing this burden systematically by increasing the number of words to remember at the end of the learning phase. This approach is similar to the one used in the GOSPAN to measure WMC [34]. Although one would expect that this procedure would interfere much more with current cognitive processes, the results of this study invite for other interpretations. It is likely that participants rehearsed the words after completing a set of exercises, thereby temporarily lowering the cognitive load while choosing and completing the exercises. The total number of words to be remembered was eleven, and overall, six words with a standard deviation of two were remembered. This is closely related to the description of Miller's magical number seven, plus or minus two [42]. However, this description is related to short-term memory, and several theories have already acknowledged that working memory and short-term memory are two distinct, although highly correlated constructs [43]. Working memory tasks focus on maintaining and activating information that is relevant for

the current task to perform, and therefore primarily concern the central executive (e.g., the span tasks). Whereas working memory is seen as a storage and attention component, short-term memory is seen as a simple storage component [44]. Presumably, the secondary task in this study did not put any significant load on working memory because the task involved simple storage of the words and less attention was required. Future research should search for secondary tasks that involve both attention and storage processes and include these tasks to measure the cognitive load of instructional strategies such as LC.

Some limitations of this study need to be discussed. A first limitation is that learners should perceive the control they have [2]. In this study, it was not verified whether the control as offered to the learners was actually perceived as such. It might have been the case that learners were not satisfied with the degree and type of control, that they experienced a mismatch between their expectations and the actual control [45], or that they did not recognize the opportunity [46].

A second limitation is that the participants chosen for this experiment were all university students. It is very likely that all participants had a rather high WMC and high self-efficacy scores. For the WMC, this was not confirmed in post hoc analyses, as WMC was normally distributed. This was however not the case for the self-efficacy scores which were highly right skewed. This might explain why LC did not affect motivation, learning, and cognitive load. It is highly plausible that even the combined load of learning English tenses, having control, and rehearsal of words for the secondary task did not touch the limitations of a learner's WMC. Moreover, it can be suggested that the intrinsic load of the materials was rather low, because participants were familiar with the domain of English tenses [47]. When measuring cognitive load in a group of advanced learners, secondary tasks are needed which place a greater burden on WMC.

To conclude, this study focused on measuring the cognitive load imposed by LC. The triarchic nature of CL could not be corroborated in this study as intrinsic and germane load were highly correlated. This provides support for the conceptualization as proposed by Sweller [23], no longer considering germane load as an independent source of cognitive load. Whether LC as instructional strategy imposes intrinsic or extraneous load could not be answered in this study. However, the answer to this question will contribute to the research on effectiveness of LC and may help to explain the mixed results in LC research.

Appendix

Instructions per Condition and Treatment

General Instruction (All Four Courses)

In this course you can find exercises grouped according to the tenses they focus on.

There are eight themes. For each theme you can open a "refresher course," a pop-up window with some basic information on the use of the tenses.

For a general overview of the tenses (one graph) you can select the "refresher course" in this topic.

No control: a single task

You have to complete all 35 exercises.

No control: a dual task

You have to complete all 35 exercises.

After some exercises a word will appear on your screen. Please try to remember these words.

You are not allowed to write down these words.

Learner control: a single task

You can select the quizzes that interest you most and complete the exercises you want. This implies that you have full control over this learning environment, and hence that you have the freedom to select the exercises you want to solve.

You should complete (at least) 35 exercises.

Learner control: a dual task

You can select the quizzes that interest you most and complete the exercises you want. This implies that you have full control over this learning environment, and hence that you have the freedom to select the exercises you want to solve.

You should complete (at least) 35 exercises.

After some exercises a word will appear on your screen. Please try to remember these words.

You are not allowed to write down these words.

Acknowledgments

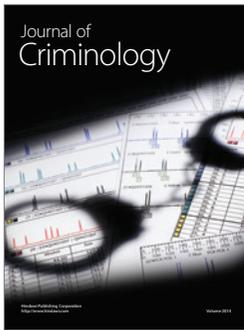
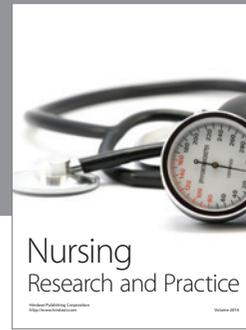
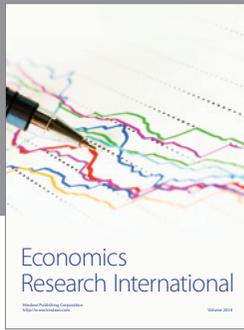
This study is based on research funded by the KU Leuven Industrial Research Fund (IOF) and conducted within the IOF Knowledge platform "Harnessing collective intelligence in order to make e-learning environments adaptive" (IOF KP/07/006). The authors would like to express their great appreciation to Martin Vanbrabant for the technical support.

References

- [1] K. Scheiter and P. Gerjets, "Learner control in hypermedia environments," *Educational Psychology Review*, vol. 19, no. 3, pp. 285–307, 2007.
- [2] G. Corbalan, L. Kester, and J. J. G. van Merriënboer, "Combining shared control with variability over surface features: effects on transfer test performance and task involvement," *Computers in Human Behavior*, vol. 25, no. 2, pp. 290–298, 2009.
- [3] M. B. Kinzie and H. J. Sullivan, "Continuing motivation, learner control, and CAI," *Educational Technology Research and Development*, vol. 37, no. 2, pp. 5–14, 1989.

- [4] M. D. Williams, "Learner-control and instructional technologies," in *Handbook of Research on Educational Communications and Technology*, D. H. Jonassen, Ed., pp. 957–982, MacMillan, New York, NY, USA, 1996.
- [5] C. L. Friend and C. L. Cole, "Learner control in computer-based instruction: a current literature review," *Educational Technology*, vol. 20, pp. 47–49, 1990.
- [6] D. Goforth, "Learner control = Decision making + Information: a model and meta-analysis," *Journal of Educational Computing Research*, vol. 11, pp. 1–26, 1994.
- [7] A. Large, "Hypertext instructional programs and learner control: a research review," *Education for Information*, vol. 14, no. 2, pp. 95–107, 1996.
- [8] M. J. Hannafin, "Guidelines for using locus of instructional control in the design of computer-assisted instruction," *Journal of Instructional Development*, vol. 7, no. 3, pp. 6–10, 1984.
- [9] W. D. Milheim and J. W. Azbell, "How past research on learner control can aid in the design of interactive video materials," in *Proceedings of the Annual Convention of the Association for Educational Communications and Technology*, pp. 459–472, New Orleans, Fla, USA, 1988.
- [10] M. B. Kinzie, "Requirements and benefits of effective interactive instruction: learner control, self-regulation, and continuing motivation," *Educational Technology Research and Development*, vol. 38, no. 1, pp. 5–21, 1990.
- [11] F. I. Winters, J. A. Greene, and C. M. Costich, "Self-regulation of learning within computer-based learning environments: a critical analysis," *Educational Psychology Review*, vol. 20, no. 4, pp. 429–444, 2008.
- [12] W. Eom and R. A. Reiser, "The effects of self-regulation and instructional control on performance and motivation in computer-based instruction," *International Journal of Instructional Media*, vol. 27, pp. 247–260, 2000.
- [13] D. S. Niederhauser, R. E. Reynolds, D. J. Salmen, and P. Skolmoski, "The influence of cognitive load on learning from hypertext," *Journal of Educational Computing Research*, vol. 23, no. 3, pp. 237–255, 2000.
- [14] I. E. Dror, "Technology enhanced learning: the good, the bad, and the ugly," *Pragmatics & Cognition*, vol. 16, pp. 215–223, 2008.
- [15] B. Schwartz, *The Paradox of Choice: Why More is Less*, Ecco, New York, NY, USA, 2004.
- [16] T. de Jong, "Cognitive load theory, educational research, and instructional design: some food for thought," *Instructional Science*, vol. 38, no. 2, pp. 105–134, 2010.
- [17] B. P. Granger and E. L. Levine, "The perplexing role of learner control in e-learning: will learning and transfer benefit or suffer?" *International Journal of Training and Development*, vol. 14, no. 3, pp. 180–197, 2010.
- [18] F. Paas, J. E. Tuovinen, J. J. G. van Merriënboer, and A. A. Darabi, "A motivational perspective on the relation between mental effort and performance: optimizing learner involvement in instruction," *Educational Technology Research and Development*, vol. 53, no. 3, pp. 25–34, 2005.
- [19] K. E. DeLeeuw and R. E. Mayer, "A comparison of three measures of cognitive load: evidence for separable measures of intrinsic, extraneous, and germane load," *Journal of Educational Psychology*, vol. 100, no. 1, pp. 223–234, 2008.
- [20] J. Sweller, J. J. G. van Merriënboer, and F. G. W. C. Paas, "Cognitive architecture and instructional design," *Educational Psychology Review*, vol. 10, no. 3, pp. 251–296, 1998.
- [21] J. J. G. van Merriënboer and D. M. A. Sluijsmans, "Toward a synthesis of cognitive load theory, four-component instructional design, and self-directed learning," *Educational Psychology Review*, vol. 21, no. 1, pp. 55–66, 2009.
- [22] J. J. G. van Merriënboer and P. Ayres, "Research on cognitive load theory and its design implications for e-learning," *Educational Technology Research and Development*, vol. 53, no. 3, pp. 5–13, 2005.
- [23] J. Sweller, "Element interactivity and intrinsic, extraneous, and germane cognitive load," *Educational Psychology Review*, vol. 22, no. 2, pp. 123–138, 2010.
- [24] F. Paas, T. van Gog, and J. Sweller, "Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives," *Educational Psychology Review*, vol. 22, no. 2, pp. 115–121, 2010.
- [25] F. Kirschner, L. Kester, and G. Corbalan, "Cognitive load theory and multimedia learning, task characteristics and learning engagement: the Current State of the Art," *Computers in Human Behavior*, vol. 27, no. 1, pp. 1–4, 2011.
- [26] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Educational Psychologist*, vol. 38, no. 1, pp. 63–71, 2003.
- [27] F. G. W. C. Paas, "Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach," *Journal of Educational Psychology*, vol. 84, no. 4, pp. 429–434, 1992.
- [28] T. van Gog and F. Paas, "Instructional efficiency: revisiting the original construct in educational research," *Educational Psychologist*, vol. 43, no. 1, pp. 16–26, 2008.
- [29] A. S. Yeung, "Cognitive load and learner expertise: split-attention and redundancy effects in reading comprehension tasks with vocabulary definitions," *Journal of Experimental Education*, vol. 67, no. 3, pp. 197–217, 1999.
- [30] P. Barrouillet, S. Bernardin, and V. Camos, "Time constraints and resource sharing in Adults' working memory spans," *Journal of Experimental Psychology*, vol. 133, no. 1, pp. 83–100, 2004.
- [31] G. Cierniak, K. Scheiter, and P. Gerjets, "Explaining the split-attention effect: is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?" *Computers in Human Behavior*, vol. 25, no. 2, pp. 315–324, 2009.
- [32] R. Brünken, S. Steinbacher, J. L. Plass, and D. Leutner, "Assessment of cognitive load in multimedia learning using dual-task methodology," *Experimental Psychology*, vol. 49, no. 2, pp. 109–119, 2002.
- [33] P. A. Kirschner, J. Sweller, and R. E. Clark, "Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," *Educational Psychologist*, vol. 41, no. 2, pp. 75–86, 2006.
- [34] W. De Neys, G. D'Ydewalle, W. Schaeken, and G. Vos, "A Dutch, computerized, and group administrable adaptation of the operation Span test," *Psychologica Belgica*, vol. 42, no. 3, pp. 177–190, 2002.
- [35] M. D. Merrill, "First principles of instruction," *Educational Technology Research and Development*, vol. 50, no. 3, pp. 43–59, 2002.
- [36] W. De Neys and K. Dieussaert, "Individual differences in rational thinking time," in *Proceedings of the Annual Conference of the Cognitive Science Society*, vol. 27, pp. 577–582, 2005.

- [37] P. R. Pintrich, D. A. F. Smith, T. Garcia, and W. J. McKeachie, *A Manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*, National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan, Ann Arbor, Mich, USA, 1991.
- [38] E. McAuley, T. Duncan, and V. V. Tammen, "Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis," *Research Quarterly for Exercise and Sport*, vol. 60, no. 1, pp. 48–58, 1989.
- [39] R. W. Plant and R. M. Ryan, "Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: an investigation of internally-controlling styles," *Journal of Personality*, vol. 53, pp. 435–449, 1985.
- [40] A. C. Rencher, *Methods of Multivariate Analysis*, John Wiley & Sons, 2nd edition, 2002.
- [41] B. S. Hasler, B. Kersten, and J. Sweller, "Learner control, cognitive load and instructional animation," *Applied Cognitive Psychology*, vol. 21, no. 6, pp. 713–729, 2007.
- [42] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956.
- [43] R. W. Engle, S. W. Tuholski, J. E. Laughlin, and A. R. A. Conway, "Working memory, short-term memory, and general fluid intelligence: a latent-variable approach," *Journal of Experimental Psychology*, vol. 128, no. 3, pp. 309–331, 1999.
- [44] R. W. Engle, M. J. Kane, and S. W. Tuholski, "Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex," in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, A. Miyake and P. Shah, Eds., Cambridge Press, London, UK, 1999.
- [45] P. Wouters, F. Paas, and J. J. G. van Merriënboer, "Observational learning from animated models: effects of studying-practicing alternation and illusion of control on transfer," *Instructional Science*, vol. 38, no. 1, pp. 89–104, 2010.
- [46] D. N. Perkins, "The fingertip effect: how information-processing technology shapes thinking," *Educational Researcher*, vol. 14, pp. 11–17, 1985.
- [47] M. McCrudden, G. Schraw, K. Hartley, and K. A. Kiewra, "The influence of presentation, organization, and example context on text learning," *Journal of Experimental Education*, vol. 72, no. 4, pp. 289–306, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

