Exploring Variability in Cognitive Load across Different Writing Mode

Naser Danesh Pouya[1], Masood Siyyari[1*]

[1]Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran

Email: na.daneshpouya@iau.ac.ir
*Corresponding Author's Email: m.siyyari@srbiau.ac.ir

ABSTRACT

This study examined the effect of cognitive load on L2 learners' descriptive and expository essay writings reflected in their obtained readability indices by taking a mixed-method approach. Randomly selected 58 out of 83 intermediate to upper-intermediate EFL learners in three groups underwent varying methods of Topic-Only, Topic-plus Argument/Counter-argument, and Topic-plus Mechanics, followed by retrospective interviews. The writings were analyzed in terms of density, diversity, and syntactic complexity based on the certain variables. Following the research interventions, 12 (five boys and seven girls) participants, randomly selected, attended retrospective semi-structured interviews. ANCOVA results indicated that, unlike descriptive writings, expository ones virtually underwent significant changes, and learners' attentions were highly affected. The results obtained in the quantitative writing analysis were consistent with the records in the interviews. The study revealed that participants performed better when they received either no or meta-cognitive supports. The results suggest that content support may decrease the load and facilitate linguistic encoding. The findings corroborate cognitive load theory and emphasize adopting sensitive measures to meet increasing task demands in writing tasks.

INTRODUCTION

Cognitive Load Theory (CLT) posits that working memory (WM) resources play a crucial role in task performance, as outlined by Sweller (2023). CLT aims to predict successful learning outcomes and effectively integrate cognitive architecture into instructional design, as emphasized by Kirschner et al. (2018). By considering the impact of instructional design and materials on cognitive load (CL), as Paas and Sweller (2014) discussed, we recognize CL as a multidimensional construct, which manifests as the cognitive load placed on learners' cognitive systems during task execution. This load significantly influences the learning environment and learners' experiences during instructional design integration with their cognitive structures (Paas & Sweller, 2021). CLT (Sweller, 2010b) underscores the importance of attending to the structures of human cognitive architecture. By developing efficient instructional methods (Rikers et al., 2004), we optimize information processing within our limited cognitive capacity (Sweller et al., 2011). This optimization enhances our ability to apply acquired knowledge and skills when addressing novel learning challenges (Kirschner et al., 2009a).

Sweller (2004) states that instructional design and human cognitive architecture contribute to the development of our knowledge of cognitive structures and help us to systematically organize learning environment because efficient instructional designs (Sweller et al., 2019) shed light on the human cognitive architecture, without

which it is not easy to decipher human cognitive structure. Sweller argues that the cognitive effort in means-ends analysis imposes a high CL on WM to accomplish a problem-solving task, leaving insufficient cognitive resources for schema acquisition (Plass et al., 2010). In WM, novel information is processed, which results in CL affecting learners' performance, and stored in long-term memory (LTM) as schemas (Kirschner et al., 2011; Sweller, 1999). Human cognitive architecture has a WM with a limited capacity (Sweller, 2004; Van Merriënboer & Sweller, 2005; Verhoeven et al., 2009) associated with all conscious activities and an unlimited LTM to store increasingly sophisticated schemas comprised of multiple elements treated as a single element. It helps WM capacity be freed up and allows information processing (Sweller & Chandler, 1994) and fluid performance on standard features of tasks and varying degrees of performance on unfamiliar ones (Sweller et al., 2019). Berndt et al. (2022) reported that mindful cognitive processing might increase by higher transitions between text elements fostering learners' performance.

Coordination among multiple elements in complex tasks incurs higher CL than in less demanding tasks. Research shows that instructional techniques reducing the extraneous cognitive load (ECL) enhance efficiency in complex tasks only as low complexity content will not require a high number of mental resources (Clark et al., 2006); however, in such instructional techniques proposed by Leppink and Van Den Heuvel (2015) to reduce or eliminate the source of ECL, it may be unavoidable such as interruptions or internal distraction (Young & Sewell, 2015).

Human cognitive capacity (Sweller et al., 2011), having limited storage (Cowan, 2001; Miller, 1956), and a short duration generates CL during information processing in WM (Baddeley, 1992). Unlike the function of WM, there are seemingly controversial views on its structure, ranging from unitary models (Cowan, 1999) and multi-component models (Baddeley & Hitch, 1974) to psychometric models of cognition (Chalmers & Freeman, 2019). Conscious engagement in Atkinson and Shiffrin's structural component model of cognitive processing (Atkinson & Shiffrin, 1968) is limited to some processes at a time occurring in short-term memory, which Baddeley and Hitch (1974) called working memory; its capacity to store and process information is limited (Sweller, 1988) allowing learners to consciously process only a limited amount of information at once (Kuldas et al., 2013).

CLT (Sweller, 2004) assumes a limited WM capacity of about seven elements for storing and two to four elements for processing novel information (Miller, 1956), and virtually unlimited long-term memory (LTM) encompassing cognitive schemas varied in complexity and automation are stored in and retrieved, and through which human expertise often consciously develops (Van Merriënboer & Ayres, 2005). The interaction between the two facilitates transferring acquired knowledge and skills through schema construction, mental models, and automation (Sweller, 2010a; Sweller et al., 2019), in which manifold of visual/spatial and auditory/verbal information are transformed into single elements in biologically secondary cognitive schemas or domain-specific knowledge structures (Klepsch & Seufert, 2020; Van Merriënboer & Ayres, 2005), which are consciously built by activating learners' knowledge through the performance of a task (Zavgorodniaia et al., 2020). The process helps to circumvent the limitations by bypassing WM during mental processing, which is the principal notion in instructional designs; therefore, CLT-based tasks require less time and effort in learning and transfer performance. Exceeding the available cognitive capacity, misallocating cognitive resources, or both may lead to failures in learning and performing complex tasks (Paas, Tuovinen et al., 2003).

CLT deals with cognitive information-processing limitations in WM and its association with learning before novel information is stored in long-term memory (Paas et al., 2004; Paas et al., 2011). The differential characteristics of serial processing, attending to and processing one item at a time, and parallel processing, simultaneously processing incoming stimuli of varying quality, are fundamental distinctions in cognitive psychology (Reed, 2012). Processing information in a one-on-one time manner in a working-memory limited capacity (Baddeley, 1992; Moreno, 2006) may result in memory overload (Keysar et al., 2012) due to increased element interactivity (Plass et al., 2010; Sweller, 2023), which hampers performing cognitive activities (Mayer & Moreno, 2003; Paas, Renkl et al., 2003); therefore; the core principle of CLT (Sweller, 2010b; Van Merriënboer & Sweller, 2005) is that instructional design should reduce the WM overload to free up capacity for learning-related processing (Anmarkrud et al., 2019). Our WM undergoes certain types of loads, one of which overloading WM impedes learning (Clark et al., 2006; Paas, Renkl et al., 2003); as a result, various techniques such as worked examples (Bannert, 2002) have been devised to manipulate CL to test their effects whether they help or hurt learning (Ginns, 2005), and facilitate learning, it is necessary to keep CL at a manageable level and prevent WM from overloading (Ayres, 2006).

Sweller (2003) states that if learners have no access to previously acquired knowledge structures helping them to organize incoming new information, they have to randomly integrate elements to determine their usefulness in solving the problem. As a result, it requires instructors' attention in designing effective learning procedures, without which learners must select alternatives that may be slow, effortful, and time-consuming. Learners' prior knowledge in long-term memory, gradually acquired randomly, acts as a central executive; instruction inspired by CLT, however,

can take over its role in organizing information. Properly designed instruction decreases the random generation of knowledge and facilitates its accommodation in LTM. As an external executive, instructional guidance is an alternative to prior knowledge that allows learners to develop schemata and directs further cognitive processing (Sweller, 2005). Increased prior knowledge lowers ICL (intrinsic cognitive load) due to reduced elements in the learning materials (Shin, 2020).

To reduce WM load, we need to explicitly teach the domain-specific skills that are the subject of CLT (Sweller & Paas, 2017). We have not evolved to acquire biologically secondary skills, yet we have evolved the cognitive system capable of acquiring a virtually unlimited range of secondary skills. A significant proportion of general cognitive skills, which are biologically primary, are fundamental to human existence and can be effectively addressed through instructional procedures, as humans have evolved to acquire these skills automatically (Sweller et al., 2011). They are learnable but not teachable. Unlike primary knowledge, domain-specific skills are taught in educational institutions and are usually consciously learned (Sweller, 2008). Biologically primary skills, acquired automatically and unconsciously, are generic-cognitive and are associated with learning, thinking, and solving problems. Biologically secondary knowledge is culturally determined and taught in education and training contexts. It is considered a single, unified system segment and requires learners' conscious effort and explicit instruction. Learning to listen and read may differentiate the two processes (Sweller et al., 2019); however, to reach their cognitive capacity limit, learners need to assign their available resources to the primary and the secondary tasks (Van Gog et al., 2011), which may hinder effectual monitoring of the secondary task and task performance of the primary task or both. Conversely, task demands can be efficiently accommodated in low complexity due to sufficient resources.

CLT assumes that individuals dealing with complex cognitive tasks cannot manage unlimited elements in their WM-limited processing capacity (Kirschner et al., 2009b; Sweller, 2011, 2012). A CLT refined view (Kalyuga, 2011) differentiates ICL and ECL, in which element interactivity (Sweller & Chandler, 1994) together with the expertise of a learner determines the amount of perceived ICL (Larmuseau et al., 2020; Sweller et al., 2019). Germane cognitive load (GCL) is part of ICL and is considered an allocated resource to ICL that contributes to learning; therefore, the total CL is lower than the one in the triple-factor model unless GCL is zero (Leppink, 2017; Van Merriënboer et al., 2006). WM's allocated resources reflect only ECL in its former view (Kalyuga, 2011). It proposes element interactivity as a task complexity measure (Leahy & Sweller, 2008) and aims to devise instructional designs to reduce WM load through techniques for determining informational complexity. Element interactivity estimates the number of elements a learner processes simultaneously in WM (Sweller, 2010b). ICL, a task's inherent complexity, is defined as element interactivity. Element interactivity itself, in turn, is affected by the task and learner's expertise (Kirschner et al., 2018). Authentic learning tasks are based on real-life tasks (Van Merriënboer & Kirschner, 2001) to integrate knowledge, coordinate constituent skills, and transfer learned elements to real-life settings (Van Merriënboer et al., 2003).

ECL, imposed unrelated information elements to the learning task, is also determined by element interactivity (Ayres, 2006; Chen et al., 2023; Paas, Renkl et al., 2003; Sweller & Chandler, 1994) produced and varied by instructional procedures (Sweller, 2011). Multiple sources of information may overload the visual and auditory processors of WM. Shifting some of the CL to either part may support learning by relieving the load (Van Merriënboer & Sweller, 2005). Decreased ECL is considered insignificant if element interactivity is low while performing a task as sufficient cognitive resources are available. It even fosters learning complex tasks with a high level of element interactivity (Carlson et al., 2003; Sweller & Chandler, 1994) by freeing up processing resources devoted subsequently to learning. More straightforward tasks impose lower ICL than more complex tasks (Bannert, 2002), and fewer elements and interactions are processed at once in WM (Pollock et al., 2002).

Low element interactivity of material, hence low in ICL, employs fewer WM resources because elements do not interact and are learned as single elements. High element interactivity requires elements to be processed simultaneously by engaging WM resources (Hochberg et al., 2020; Van Merriënboer & Sweller, 2005) that may exceed its capacity, leading to learning problems; as a result, it is a significant challenge for instructional designers (Sweller et al., 2011). Element interactivity cannot be manipulated by instructional design (Paas, Renkl et al., 2003; Pollock et al., 2002); however, developing sufficient command over appropriate schemata decreases ICL as the interactions are learned and incorporated into domain-specific schemas, and treated as single elements (Ayres, 2006). Recent studies indicate that intrinsic cognitive load can be reduced either by scaffolding information or adjusting the guidance with the level of learners' expertise (Van Merriënboer & Sweller, 2005; Vogel-Walcutt et al., 2011).

As WM includes semi-independent visual and auditory components (Penney, 1989), the overloading sub-processor of either one may impose higher ECL. GCL is closely related to ICL (Kirschner et al., 2018; Sweller, 2010b) and is associated with relevant processes in better schema construction, transfer of learning, and automation intrigued

by the variability of problem situations, which requires the conscious involvement of the learners in the learning process. Optimizing GCL can be obtained through either low ICL or low ECL, keeping the total amount of CL within limits, and it is realized by elaborating more deeply on learning material (Bannert, 2002; Paas & Van Merriënboer, 1994b). Instructional manipulations to control ECL release cognitive resources and allow learners to invest mental effort in learning processes. Thus, the CLT additivity hypothesis claims that learning is impaired as the total amount of ICL, ECL, and GCL outreaches the WM limits (Plass et al., 2010), and the prime focus of CLT is to curtail ECL and enhance GCL within cognitive processing limits (Clark et al., 2006; Sweller, 2003, 2004, 2010b; Sweller et al., 2019) by managing ICL determined by learners' level of expertise (Van Merriënboer & Ayres, 2005).

In the classical CLT, learners' CL was not determined by the instructional designer but rather by the complexity of the material and learners' prior knowledge levels (Kalyuga & Renkl, 2010). Task-relevant prior knowledge in LTM fosters effective learning (Kalyuga et al., 1998) by lowering the learner's CL unless it interferes with the information already available, increasing the experienced CL and, thus, reducing their performance (Janssen & Kirschner, 2020); therefore, learners should take active roles by funneling cognitive resources towards relevant aspects to accomplish tasks (Moos & Pitton, 2014). CL, overrun the cognitive capacity limit counteracts learning (Mayer & Moreno, 2003), restricts our cognitive capacity to accommodate incoming task demands (Brunken et al., 2003). The amount of three additive loads (Kirschner et al., 2009b; Paas et al., 2005; Paas, Tuovinen et al., 2003) should stay within the WM processing capacity in an effective instructional design. Worked examples may reduce ECL (Atkinson et al., 2000); however, a balance is required between ICL and GCL in such a way that processing of all interactive elements should leave a sufficient processing capacity available for GCL to engage in schema acquisition and automation (Leppink, 2017; Paas et al., 2004; Paas, Renkl et al., 2003). Segmenting and sequencing information elements and interactions in the part-whole approach can be used to manage ICL (Van Meriënboer et al., 2006) and increase the variability of learning tasks to enhance GCL (Paas & Van Merriënboer, 1994a). CL is influenced by various elements. Firstly, the nature of the task—whether it's complex or simple—plays a significant role. Secondly, learner characteristics, such as prior knowledge and cognitive abilities, impact how much cognitive load they experience. Lastly, effective instructional designs, including scaffolding and appropriate feedback, can mitigate cognitive load and enhance learning outcomes. Learners are gradually introduced to small increments of tasks by adjusting the size and grouping to help learners control the rate at which they access instruction; it helps learners manage and reduce ICL and accommodate it within WM limits.

A low ICL and a high ECL are considered less harmful as the total CL remains within WM limits. Learners' performance may be further improved by encouraging them to engage in controlled and conscious processing relevant to the construction of schemas. Redirecting learners' attention from irrelevant processes to relevant ones toward the conscious construction of schemas can be attained by decreasing ECL and increasing GCL (Sweller et al., 2019). Understanding takes place once the interacting elements are assimilated into a higher-order schema.

The study holds potential to contribute to the existing literature on cognitive load, particularly in the context of learners' descriptive and expository essay writings. By manipulating task conditions, the research aims to uncover how specific variables impact learners' cognitive processes during writing. Understanding how different conditions (such as topic complexity, time constraints, or scaffolding) affect cognitive load can provide valuable insights for educators and instructional designers. Additionally, the focus on variables reflected in learners' essays—such as coherence, organization, and overall quality—adds depth to the examination of cognitive load. Investigating how learners manage these variables while composing essays sheds light on their cognitive architecture and resource allocation. The inclusion of semi-structured interviews allows us to explore learners' subjective experiences and perceptions, complementing quantitative measures. Aligning findings with the writing tasks ensures practical relevance and bridges theory with real-world application.

The Research Questions

1. *To what extent does the manipulation of task conditions modulate cognitive load during the essay writing process among learners?*
2. *How do learners' subjective experiences and perceptions, as captured through semi-structured interviews, elucidate the cognitive load experienced during essay composition?*

METHODS
PARTICIPANTS

Eighty-three university students at IAU, Babol Branch, Iran, volunteered to participate in the study using convenience sampling. Before their admission, they were taught how to write descriptive and expository essays. Their mean age was 20.9 years. Their age varied between 18 and 22 years. Fifty-eight participants, 17 males, and 41 females, were selected with intermediate to high-intermediate commands in English based on their writing expertise levels in pre-writing tests and their scores in the Cambridge English Language Proficiency Test Online. The selection is a relatively homogeneous group of EFL learners in their instructional background to use English inside and outside their language class. To address ethical considerations, explicit informed consent was obtained from all participants. Additionally, they were assured of strict confidentiality and complete anonymity.

INSTRUMENTS

Cambridge English Language Assessment online (https://www.cambridgeenglish.org/test-your-english/) was administered to assess the participants' proficiency level. The test consists of 25 multiple-choice items measuring the test-takers' overall English language proficiency. Prior to conducting the experiment, participants were administered two preliminary writing tasks, expository and descriptive writing modes, in order to assess their writing proficiency. The selected topics were based on the IELTS General Writing module and taken from the essay prompts of the IELTS writing tasks. Besides, the participants were given a popular social topic of the General Writing Module of IELTS to work on for the experimental and control conditions. Twelve participants attended semi-structured interviews to help us delve into participants' impressions of the task.

PRE-WRITING TESTS

To obtain a valid and reliable writing proficiency test score of the participants, two pre-writing tests, comprising a descriptive pre-test and an expository pre-test, were administered. The LIBRO software and CO-METRIX web tool (http://tool.cohmetrix.com/) analyzed the two tests based on the assigned variables. The scores of these two tests and the learners' scores on the Cambridge English Language Assessment were applied to obtain an average writing proficiency test score for each participant. Before the commencement of the experimental writing task, participants were randomly allocated to two distinct experimental conditions: the Topic-plus Argument/Counter-argument group, the Topic-plus Mechanics group, and a control group (Topic-Only). This allocation was based on their attained L2 writing proficiency scores, ensuring the formation of homogeneous groups for the study.

PROCEDURES

In this study, six variables pertaining to cognitive and psycholinguistic processes involved in reading were analyzed. These variables are as follows:

1. Coh-Metrix L2 Readability: Reflects readability in a second language, considering linguistic features such as sentence length, word frequency, and syntactic structures.
2. Flesch-Kincaid Grade Level: Estimates the approximate reading grade level required to understand a text based on factors like sentence length and word complexity.
3. Gunning-Fog Index: Evaluates text difficulty by assigning it a specific grade level, considering sentence length and the percentage of complex words.
4. Shannon-Wiener H' Index: Assesses lexical diversity by analyzing the richness of unique words and their distribution.
5. Lexical Density-Average Word Length: Reflects the proportion of content words (nouns, verbs, adjectives) and average word length, contributing to overall text complexity.
6. SMOG Index: Measures the understandability of writing, considering sentence length and polysyllabic words.

Descriptive and expository writing tasks were incorporated into the experimental phase, and participants were required to complete them within 60 min. They were recommended to complete the descriptive task in 20 minutes and the expository one in 40 minutes. Three task groups, namely, the Topic-Only, Topic-plus Argument/Counter-argument, and Topic-plus Mechanics, represented the corresponding conditions in the study.

In this study, a sequential explanatory design, a mixed method type of research, was used. The design is straightforward, with the primary objective being to employ quantitative methods initially, followed by qualitative methods. This sequential approach aims to provide deeper insights into the quantitative results. After the research interventions, a cohort of 12 participants (comprising five boys and seven girls) was randomly selected. These participants engaged in retrospective semi-structured interviews conducted in their native language, Persian. The purpose of these interviews was to delve deeper into their feelings and attitudes regarding the writing experience, while also prompting them to assess the conditions they had encountered. Following the research interventions, 12 (five boys and seven girls) participants, randomly selected, attended retrospective semi-structured interviews to explore further their feelings and attitudes on the writing experience, and they were prompted to assess the conditions they had experienced. In the qualitative dimension of the study, to capture cognitive alterations associated with the methods applied in the test, face-to-face interview sessions were scheduled after the test, and multi-fact data were collected on the manipulated conditions of descriptive and expository writings, reflecting imposed cognitive load. Participants shared their thoughts and feelings per the directions in the sense that the set of questions guided the interview. At the same time, the interviewer was allowed autonomy to clarify the participants' comments. The following questions, revised by the experts' feedback on the applicability of content and structure, focused on interviewees' mental effort perceptions enforced by the methods applied: (a) How do you feel about the writing conditions (the pre-writing task and the writing task)? (b) Was the intervention of any help? How?, and (c) Which writing task were you involved in more? And why? Interviews lasted for six-ten minutes for each participant. They were individually conducted in a quiet office between the interviewees and the interviewer, acting as a facilitator, encouraging them to voice their opinions and helping them to stay on track. The participants' outputs were audio-recorded with their permission, transcribed verbatim, and coded. To conduct in-depth data analysis, MAXQDA 2022 was used to handle the data regarding interviewees' intrinsic cognitive loads. The qualitative analysis was drawn from the responses to the interviews, in which the respondents were asked to provide feedback on the pre-writing and writing tasks based on their experiences, whence well-refined themes were collaboratively developed (Corbin & Strauss, 2008).

RESULTS
DESCRIPTIVE WRITING

These variables were analyzed in learners' descriptive and expository writings across three task conditions: Topic-Only, Topic-plus Argument/Counter-argument, and Topic-plus Mechanics. Data from these conditions were subjected to a single-factor between-subjects analysis of covariance (ANCOVA) using JASP 0.18.3.0, an open-source statistics program, with the variables serving as covariates.

There were no statistically significant differences among group means on Coh-Metrix L2 Readability [$F(2, 54) = 2.555$, $p = .087$, partial $\eta^2 = .086$], Flesch-Kincaid Grade Level [$F(2, 54) = .632$, $p = .535$, partial $\eta^2 = .023$], the Gunning-Fog Index [$F(2, 54) = 2.555$, $p = .087$, partial $\eta^2 = .038$], Lexical Density–Average Word Length [$F(2, 54) = 0.789$, $p = 0.460$, partial $\eta^2 = .028$], and the SMOG Index [$F(2, 54) = 0.161$, $p = 0.852$, partial $\eta^2 = .006$].

However, there was a significant difference in Shannon-Wiener H' at the $p<.001$ level for the three conditions [$F(2, 54) = 4.179$, $p = .021$, partial $\eta^2 = 0.134$]. Post hoc comparisons, conducted using the Tukey Honestly Significant Difference (HSD) test (Table 1), revealed that the mean score for the Topic-plus Mechanics condition was significantly different from the Topic-Only condition (Mean Difference = 0.8, $SE = 0.2$, Cohen's $d = 0.92$, indicating a very large effect size, and lesser diversity); nonetheless, the Topic-plus Argument/Counter-argument condition did not significantly differ from the Topic-plus Mechanics condition (Mean Difference = -0.4, $SE = 0.2$) and the Topic-Only condition (Mean Difference = 0.3, $SE = 0.2$). Comparing the estimated marginal means, it was observed that the highest Shannon-Wiener H' (the higher index, the lesser diversity) goes to Topic-plus Mechanics (mean = 9.5) compared to Topic-plus Argument/Counter-argument, and Topic-plus Mechanics (mean = 9.1, 8.7), respectively.

Table 1.

Shannon-Wiener H' Variable in Descriptive Writing - Post Hoc Comparisons

		Mean Difference	95% CI for Mean Difference		SE	t	Cohen's d	95% CI for Cohen's d		ptukey	pbonf
			Lower	Upper				Lower	Upper		
(Topic-Counter-/Argument)	(Topic-Mechanic)	-0.488	-1.195	0.219	0.293	-1.662	-0.548	-1.372	0.277	0.229	0.307
	(Topic-Only)	0.335	-0.361	1.030	0.289	1.159	0.376	-0.430	1.181	0.482	0.754
(Topic-Mechanic)	(Topic-Only)	0.822	0.134	1.510	0.285	2.880	0.923	0.101	1.745	0.015 *	0.017 *

* p < .05

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

EXPOSITORY WRITING

The three conditions significantly affected the Coh-Metrix L2 Readability [$F(2, 54) = 6.903$, $p = .002$, partial $\eta^2 = 0.204$]. Subsequent to the Tukey (HSD) test (Table 2), it was evident that the mean score for the Topic-plus Mechanics condition significantly differed from the Topic-Only condition at the $p < .01$ level. (Mean Difference = -6.9, *SE* = 1.8, Cohen's *d* = -3.71, indicating a huge effect size); however, no difference was observed between the Topic-plus Argument/Counter-argument and Topic-plus Mechanics conditions (Mean Difference = 3.5, *SE* = 1.8), nor between the Topic-plus Argument/Counter-argument condition and the Topic-Only condition (Mean Difference = -3.3, *SE* = 1.9). Comparing the estimated marginal means adjusted by Bonferroni CI showed that the highest Coh-Metrix L2 Readability goes to Topic-Only (mean = 21.5) compared to Topic-plus Argument/Counter-argument and Topic-plus Mechanics (mean = 18.1, 14.6), respectively.

Table 2.

Coh-Metrix L2 Readability in Expository Writing - Post Hoc Comparisons

		Mean Difference	95% CI for Mean Difference		SE	t	Cohen's d	95% CI for Cohen's d		ptukey	pbonf
			Lower	Upper				Lower	Upper		
(Topic-Counter-/Argument)	(Topic-Mechanic)	3.570	-0.962	8.101	1.880	1.898	0.618	-0.200	1.436	0.149	0.189
	(Topic-Only)	-3.390	-7.984	1.203	1.906	-1.779	-0.587	-1.415	0.240	0.186	0.243
(Topic-Mechanic)	(Topic-Only)	-6.960	-11.477	-2.443	1.874	-3.714	-1.205	-2.057	-0.354	0.001 **	0.001 **

** p < .01

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

A significant difference was observed in the Flesch-Kincaid Grade Level for the three conditions [$F(2, 54) = 68.929$, $p < .001$, partial $\eta^2 = 0.719$]. Following the Tukey (HSD) test (Table 3), it became evident that the mean score for the Topic-plus Argument/Counter-argument condition was significantly different from the Topic-plus Mechanics condition (Mean Difference = 8.4, *SE* = 0.7, Cohen's *d* = 3.83, indicating a huge effect size), and Topic-Only condition (Mean Difference = 2.7, *SE* = 0.7, Cohen's *d* = 1.26, indicating a very large effect size); furthermore, Topic-plus Mechanics condition significantly differed from the Topic-Only condition (Mean Difference = -5.6, *SE* = 0.7, Cohen's *d* = -2.56, indicating a huge effect size). Comparing the estimated marginal means adjusted by Bonferroni CI showed that the highest Flesch-Kincaid Grade Level goes to Topic-plus Argument/Counter-argument (mean = 13.4) compared to Topic-plus Mechanics and Topic-Only (mean = 5, 10.6), respectively.

Table 3.

The Flesch-Kincaid Grade Level in Expository Writing - Post Hoc Comparisons

			95% CI for Mean Difference				Cohen's d	95% CI for Cohen's d			
		Mean Difference	Lower	Upper	SE	t		Lower	Upper	ptukey	pbonf
(Topic-Counter-/Argument)	(Topic-Mechanic)	8.466	6.709	10.223	0.729	11.612	3.830	2.608	5.053	< .001 ***	< .001 ***
	(Topic-Only)	2.796	1.074	4.518	0.715	3.913	1.265	0.412	2.119	< .001 ***	< .001 ***
(Topic-Mechanic)	(Topic-Only)	-5.670	-7.463	-3.878	0.744	-7.623	-2.565	-3.597	-1.534	< .001 ***	< .001 ***

** p < .01, *** p < .001

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

The Gunning-Fog Index significantly differed in the three conditions [$F(2, 54) = 4.591$, $p = .014$, partial $\eta^2 = 0.145$]. Subsequent to the Tukey (HSD) test (Table 4), it became evident that the mean score for the Topic-plus Mechanics condition was significantly different from the Topic-Only condition (Mean Difference = 2.4, $SE = 0.8$. Cohen's $d = 0.98$, indicating a very large effect size); however, it did not significantly differ from the Topic-plus Argument/Counter-argument condition (Mean Difference = -0.9, $SE = 0.8$). There was no significant difference between the Topic-plus Argument/Counter-argument and the Topic-Only conditions (Mean Difference = 1.5, $SE = 0.8$). Comparing the estimated marginal means, it was observed that the highest Gunning-Fog Index goes to Topic-plus Mechanics (mean = 11.2) compared to Topic-plus Argument/Counter-argument and Topic-Only (mean = 10.3, 8.8), respectively.

Table 4.

The Gunning-Fog Index in Expository Writing - Post Hoc Comparisons

			95% CI for Mean Difference					95% CI for Cohen's d			
		Mean Difference	Lower	Upper	SE	t	Cohen's d	Lower	Upper	ptukey	pbonf
(Topic-Counter-/Argument)	(Topic-Mechanic)	-0.901	-2.898	1.096	0.829	-1.087	-0.358	-1.176	0.460	0.526	0.845
	(Topic-Only)	1.571	-0.377	3.519	0.808	1.943	0.624	-0.183	1.431	0.137	0.172
(Topic-Mechanic)	(Topic-Only)	2.472	0.472	4.472	0.830	2.978	0.982	0.134	1.829	0.012 *	0.013 *

* p < .05

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

There was likewise a significant difference in Lexical Density–Average Word Length for the three conditions [$F(2, 54) = 4.426$, $p = .017$, partial $\eta^2 = 0.141$]. Post hoc comparisons, conducted using the Tukey (HSD) test (Table 5), revealed that the mean score for the Topic-plus Argument/Counter-argument condition was significantly different from the Topic-Only condition (Mean Difference = 0.3, $SE = 0.6$, Cohen's $d = 0.97$, indicating a very large effect size); nonetheless, it did not significantly differ from the Topic-plus Mechanics condition (Mean Difference = 0.1, $SE = 0.1$); furthermore, no significant difference was observed between Topic-plus Mechanics condition and the Topic-Only condition (Mean Difference = 0.1, $SE = 0.1$). Comparing the estimated marginal means adjusted by Bonferroni CI showed that the highest Lexical Density–Average Word Length goes to Topic-plus Argument/Counter-argument (mean = 4.9) compared to Topic-plus Mechanics and Topic-Only (mean = 4.7, 4.5), respectively.

Table 5.

Lexical Density–Average Word Length in Expository Writing - Post Hoc Comparisons

		Mean Difference	95% CI for Mean Difference		SE	t	Cohen's d	95% CI for Cohen's d		ptukey	pbonf
			Lower	Upper				Lower	Upper		
(Topic-Counter-/Argument)	(Topic-Mechanic)	0.185	-0.117	0.487	0.125	1.478	0.481	-0.331	1.294	0.309	0.436
	(Topic-Only)	0.375	0.071	0.678	0.126	2.975	0.975	0.133	1.817	0.012 *	0.013 *
(Topic-Mechanic)	(Topic-Only)	0.190	-0.120	0.500	0.129	1.475	0.493	-0.341	1.328	0.311	0.438

* p < .05

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

There was a significant difference in Shannon-Wiener H' for the three conditions [$F(2, 54) = 89.934$, p <.001, partial $\eta^2 = 0.769$]. Subsequent to the Tukey (HSD) test (Table 6), it was evident that the mean score for the Topic-plus Argument/Counter-argument condition was significantly different from the Topic-plus Mechanics condition, indicating its higher diversity than the Topic-plus Mechanics (Mean Difference = -7.9, *SE* = 0.7, Cohen's *d* = -3.64, indicating a huge effect size); however, it did not significantly differ from the Topic-Only condition (Mean Difference = 0.4, *SE* = 0.7); furthermore, Topic-plus Mechanics condition significantly differed from the Topic-Only condition (Mean Difference = 8.4, *SE* = 0.7, Cohen's *d* = 3.85, indicating a huge effect size and its lesser diversity than the Topic-Only condition). Comparing the estimated marginal means, it was observed that the highest Shannon-Wiener H' (the higher index, the lesser diversity) goes to Topic-plus Mechanics (mean = 17.8) compared to Topic-plus Argument/Counter-argument and Topic-Only (mean = 9.8, 9.3), respectively.

Table 6.

Shannon-Wiener H' in Expository Writing - Post Hoc Comparisons

		Mean Difference	95% CI for Mean Difference		SE	t	Cohen's d	95% CI for Cohen's d		ptukey	pbonf
			Lower	Upper				Lower	Upper		
(Topic-Counter-/Argument)	(Topic-Mechanic)	-7.995	-9.712	-6.278	0.712	-11.224	-3.642	-4.822	-2.462	< .001 ***	< .001 ***
	(Topic-Only)	0.477	-1.222	2.176	0.705	0.677	0.217	-0.578	1.012	0.778	1.000
(Topic-Mechanic)	(Topic-Only)	8.472	6.772	10.172	0.705	12.009	3.859	2.646	5.072	< .001 ***	< .001 ***

*** p < .001

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the tukey method).

There was similarly a significant difference in the SMOG Index for the three conditions [$F(2, 54) = 25.362$, p <.001, partial $\eta^2 = 0.484$]. Following the Tukey (HSD) test, it became evident that the mean score for the Topic-plus Argument/Counter-argument condition was significantly different from the Topic-plus Mechanics condition (Mean Difference = -7.7, *SE* = 1.2, Cohen's *d* = -1.96, indicating a very large effect size); however, it did not significantly differ from the Topic-Only condition (Mean Difference = 0.2, *SE* = 1.2); additionally, Topic-plus Mechanics condition significantly differed from the Topic-Only condition (Mean Difference = 8, *SE* = 1.2, Cohen's *d* = 2.03, indicating a huge effect size). Comparing the estimated marginal means revealed that the highest SMOG Index goes to Topic-plus Mechanics (mean = 12.8) compared to Topic-plus Argument/Counter-argument and Topic-Only (mean = 5.1, 4.8), respectively.

INTERVIEW FINDINGS

Interviews help generate hypotheses or refine research questions. Researchers can explore unexpected themes or patterns that emerge during interviews (Corbin & Strauss, 2008). MAXQDA 2022 was used to handle the data and conduct in-depth data analysis. Utilizing MAXQDA as interview transcription and analysis software allows researchers to employ a wide range of procedures for quantitative material evaluation. Researchers can sort sources

based on document variables, compare quantities using frequency tables and charts, and work with diverse data sets for advanced coding, retrieval, transcription, and visualization.

Interviewees expressed that they experienced lesser mental load in the descriptive writing task than the expository one and even more when undergoing the experimental conditions. As seen in Fig. 1 (generated by MAXQDA 2022), the cited mental effort impacts among G-E1, G-C, and G-E2 in descriptive writing were 1, 0, 1, and 5, 11, 1 in expository writing in the pre-test stage, respectively; they, nevertheless, significantly changed to 9, 3, 9 in descriptive, and 14, 15, 11 in expository writing, respectively in the post-test stage. Participants felt more cognitive load in expository writing than in descriptive writing in both the pre-test and post-test stages. Fig. 2 illustrates the comparison of groups using the two-case models generated by MAXQDA 2022.

Fig. 1.
The Interviews Code Matrix Browser

Note: An infographic chart—Coded items (on the right) and their corresponding frequencies underneath the three groups, namely Topic-plus Argument/Counter-argument, Topic-Only, and Topic-plus Mechanics conditions, respectively.
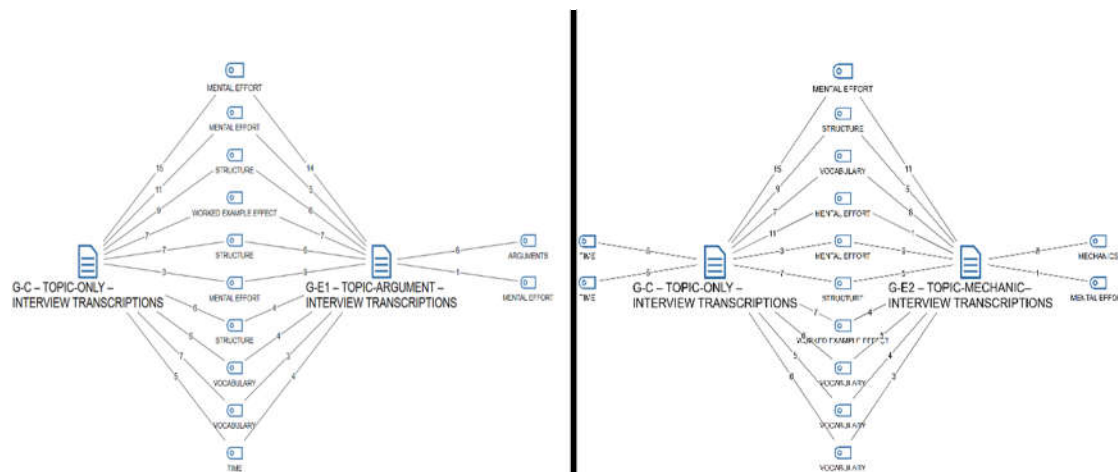
Fig. 2.
Two-Cases Model

Fig. 2. *Two-Cases Model*

Note: An infographic of coded items—on the left: the comparison of the Topic-Only (G-C) and Topic-plus Argument/Counter-argument conditions (G-E1) and their coded items with corresponding frequencies, and on the right: Topic-Only and Topic Mechanics conditions (G-E2)

DISCUSSION
EXPOSITORY AND DESCRIPTIVE WRITING TESTS

This study aimed to address two primary research questions:
1. To what extent does the manipulation of task conditions modulate cognitive load during the essay writing process among learners?
2. How do learners' subjective experiences and perceptions, as captured through semi-structured interviews, elucidate the cognitive load experienced during essay composition?

In the descriptive writing test, a less demanding task, participants' scores were not significantly different, except for the Shannon-Wiener H' variable between the Topic-Only and Topic-plus Mechanics conditions. The Topic-Only condition had the lowest score, indicating the highest diversity among the other two conditions. This suggests that low intrinsic cognitive load (ICL), due to low element interactivity and an imposed extraneous cognitive load (ECL), predominantly results in total cognitive load (CL) within working memory (WM) limits. However, the imposed ECL inversely affected the Topic-plus Argument/Counter-argument and Topic-plus Mechanics conditions.

In contrast, expository writing, a more demanding task, showed significant changes in all variables under the Topic-Only, Topic-plus Argument/Counter-argument, and Topic-plus Mechanics conditions. It appears that when ICL is high, imposing ECL may result in memory overload due to increased element interactivity and the unavailability of attentional resources (Plass et al., 2010), which impedes the performance of cognitive activities (Paas, Renkl et al., 2003). The methods applied in the test increased WM loads, incurring significant effects on learners' task performance. To address the first research question, the results demonstrate that task condition manipulation significantly modulates cognitive load during the essay writing process.

The results obtained are consistent with research conducted by Sweller (2019) and Sweller, Van Merriënboer, and Paas (2019), who emphasized the role of Cognitive Load Theory (CLT) in informing instructional design and enhancing teaching practices. According to CLT, learning is most effective when it aligns with the human cognitive architecture. Our cognitive structures enable us to perceive, think, and solve problems by integrating the elements into meaningful ones. Learning involves modifying schematic structures in long-term memory, progressing from initial error-prone performances to a state of smooth and effortless mastery.

During this developmental process, effective instructional materials can reduce the load on working memory, facilitating schema acquisition through practical applications such as worked examples and goal-free problems. Although CLT has primarily been applied to technical subjects, its principles are also relevant to language-based discursive areas. Therefore, by addressing cognitive load in instructional design, optimizing instructional materials, and considering human cognitive architecture, educators can create efficient learning environments that foster knowledge acquisition and skill development.

The interviews coded in MAXQDA 2022 confirmed that increased cognitive task demands imposed by the task conditions significantly affected learners' attentional pools, as they predominantly recorded higher mental effort in expository writing than in descriptive writing. Regarding the Topic-plus Argument/Counter-argument or Topic-plus Mechanics conditions, the applied methods acted as further loads in the post-test stage. The cited argument cognitive loads were 6, 0, and 0 in G-E1, G-C, and G-E2, respectively; for Mechanics, they were 0, 0, and 8, respectively. In expository writing, directing attentional resources toward specific characteristics incurred a cognitive overload when test-takers felt an insufficient attentional pool. These findings address the second research question, highlighting how learners' subjective experiences and perceptions elucidate the cognitive load experienced during essay compositions.

The results obtained from the quantitative writing analysis were consistent with the coded interview records. This study highlights the impact of cognitive load on learners' writing. Participants expressed their challenges when confronted with increased constraints, while dealing with more demanding writing. They identified the vital points that require attention. While working on specific areas of writing strategies and other facets of effective pedagogy, adaptive instruction requires understanding learners' cognitive resources to appropriately direct their attention.

This study aligns with the findings reported by Liu et al. (2021), Moos and Pitton (2014), Szarkowska et al. (2016), Turan, Meral, and Sahin (2018), and Yu (2021) expressed that manipulating task demands significantly influenced subjects' cognitive load and performance. Additionally, Kastaun et al. (2021) and Klepsch & Seufert (2021) also supported the importance of managing cognitive load in writing tasks. However, it is essential to acknowledge potential sources of bias in the study, such as students' self-censorship and their tendency to tailor their responses to the interview context.

No sources were found to directly contradict the findings of this study; however, the possibility of varied results under varying conditions cannot be excluded. Some studies, such as Schnotz and Kürschner (2007), indicate potential conceptual flaws and suggest that a reduction in CL can impair learning. Akin and MurrellJones (2018) examined existing practices and proposed alternative instructional methods. In addition, Nawal (2018) suggests that thinking in L2 directly, rather than translating from L1, can reduce CL and improve writing performance. The results of this research align with CLT, yet the above-mentioned complexities suggest that the CL-writing performance relationship can be more intricate than initially assumed.

CONCLUSION AND IMPILICATIONS

Although the study's empirical findings suggest potential avenues for future research, it was limited to some degree by the small number of participants; consequently, the study may have required more power. The approach was limited to examining two types of writing modes (descriptive and expository) under the two manipulated conditions. The limited sample consisted of 58 first-year university students at proficiency levels ranging from intermediate to upper-intermediate, who attended a course related to English writing. There were also only a certain number of variables and conditions in the study; therefore, further studies can explore the impact of other conditions on cognitive load using varying scales and methods. Moreover, the structure of the interviews limited the data collection, involving only 12 participants at a certain level of English proficiency. Therefore, further studies using more comprehensive and varied interview methods are required. Implementing additional reliable measurements of cognitive constructs will contribute to the development of effective instructional designs and help students improve their learning performance (Korbach et al., 2018).

Future research dealing with the exploration of cognitive load (CL) should encompass diverse demographic factors, including age groups, proficiency levels, and task performer variables. To enhance the generalizability of the findings, investigations should incorporate tasks spanning various levels of complexity. Moreover, the integration of subjective self-report measures, such as perceived workload and mental effort (Kastaun et al., 2021; Klepsch & Seufert, 2021; Krell et al., 2022) and objective physiological indicators, such as heart rate variability and pupillometry, as explored by Ayres et al. (2021), would enrich our understanding of CL dynamics. Addressing the limitations

identified in this study and expanding the scope of future research will contribute to a deeper understanding of CLT in the educational context.

The implications of this study are multi-faceted. By considering CL in instructional design, educators can develop more effective instructions to minimize ECL and optimize ICL. This helps in setting tasks at appropriate CL levels, which can enhance learning and writing performance. This includes the development of clear and concise materials to mitigate unnecessary complexities and foster schema acquisition. These findings can also inform assessment practices. The imbalance between learners' cognitive capacity and task complexity makes the test too challenging, which can overload their WM and reduce their performance. Therefore, this study encourages instructors to create more effective learning environments to support learners' cognitive development. The integration of technological equipment, such as MAXQDA, used in the study, and physiological measures for CL measurement can help researchers develop innovative approaches to studying and managing CL.

REFERENCES

Akin, I., & MurrellJones, M. (2018). Closing the gap in academic writing using the cognitive load theory. *Literacy Information and Computer Education Journal, 9(*1), 2833–2841. https://doi.org/10.20533/licej.2040.2589.2018.0373

Anmarkrud, Ø., Andresen, A., & Bråten, I. (2019). Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. *Educational Psychologist*, *54*(2), 61–83. https://doi.org/10.1080/00461520.2018.1554484

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Psychology of Learning and Motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181–214. https://doi.org/10.3102/00346543070002181

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology*, *20*(3), 287–298. https://doi.org/10.1002/acp.1245

Ayres, P., Lee, J. Y., Paas, F., & van Merriënboer, J. J. G. (2021). The validity of physiological measures to identify differences in intrinsic cognitive load. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.702538

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bannert, M. (2002). Managing cognitive load—Recent trends in cognitive load theory. *Learning and Instruction*, *12*(1), 139–146. https://doi.org/10.1016/S0959-4752(01)00021-4

Berndt, M., Strijbos, J.-W., & Fischer, F. (2022). Impact of sender and peer-feedback characteristics on performance, cognitive load, and mindful cognitive processing. *Studies in Educational Evaluation, 75*, 101197. https://doi.org/10.1016/j.stueduc.2022.101197

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 53–61. https://doi.org/10.1207/S15326985EP3801_7

Carlson, R., Chandler, P., & Sweller, J. (2003). Learning and understanding science instructional material. *Journal of Educational Psychology*, *95*(3), 629–640. https://doi.org/10.1037/0022-0663.95.3.629

Chalmers, K. A., & Freeman, E. E. (2019). Working memory power test for children. *Journal of Psychoeducational Assessment*, *37*(1), 105–111. https://doi.org/10.1177/0734282917731458

Chen, O., Paas, F., & Sweller, J. (2023). A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review, 35*(2). https://doi.org/10.1007/s10648-023-09782-w

Clark, R. C., Nguyen, F., Sweller, J., & Baddeley, M. (2006). Efficiency in learning: Evidence-based guidelines to manage cognitive load. *Performance Improvement*, *45*(9), 46–47. https://doi.org/10.1002/pfi.4930450920

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research (3rd ed.): Techniques and procedures for developing grounded theory*. SAGE Publications, Inc. https://doi.org/10.4135/9781452230153

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (1st ed., pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*(4), 313–331. https://doi.org/10.1016/j.learninstruc.2005.07.001

Hochberg, K., Becker, S., Louis, M., Klein, P., & Kuhn, J. (2020). Using smartphones as experimental tools—a follow-up: Cognitive effects by video analysis and reduction of cognitive load by multiple representations. *Journal of Science Education and Technology*, *29*(2), 303–317. https://doi.org/10.1007/s10956-020-09816-w

Janssen, J., & Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: Towards a research agenda. *Educational Technology Research and Development*, *68*(2), 783–805. https://doi.org/10.1007/s11423-019-09729-5

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, *23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7

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. https://doi.org/10.1518/001872098779480587

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science*, *38*(3), 209–215. https://doi.org/10.1007/s11251-009-9102-0

Kastaun, M., Meier, M., Küchemann, S., & Kuhn, J. (2021). Validation of cognitive load during inquiry-based learning with multimedia scaffolds using subjective measurement and eye movements. *Frontiers in Psychology, 12.* https://doi.org/10.3389/fpsyg.2021.703857

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*(6), 661–668. https://doi.org/10.1177/0956797611432178

Kirschner, F., Paas, F., & Kirschner, P. A. (2009a). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, *21*(1), 31–42. https://doi.org/10.1007/s10648-008-9095-2

Kirschner, F., Paas, F., & Kirschner, P. A. (2009b). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, *25*(2), 306–314. https://doi.org/10.1016/j.chb.2008.12.008

Kirschner, P. A., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior*, *27*(1), 99–105. https://doi.org/10.1016/j.chb.2010.06.025

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano R., J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, *13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, *48*(1), 45–77. https://doi.org/10.1007/s11251-020-09502-9

Klepsch, M., & Seufert, T. (2021). Making an effort versus experiencing load. *Front. Educ., 6*(645284), 1-14. https://doi.org/10.3389/feduc.2021.645284

Korbach, A., Brünken, R., & Park, B. (2018). Differentiating different types of cognitive load: A comparison of different measures. *Educational Psychology Review*, *30*(2), 503–529. https://doi.org/10.1007/s10648-017-9404-8

Krell, M., Xu, K. M., Rey, G. D., & Paas, F. (2022). Editorial: Recent approaches for assessing cognitive load from a validity perspective. *Front. Educ., 6.* https://doi.org/10.3389/feduc.2021.838422

Kuldas, S., Ismail, H. N., Hashim, S., & Bakar, Z. A. (2013). Unconscious learning processes: Mental integration of verbal and pictorial instructional materials. *SpringerPlus*, *2*(1), 105. https://doi.org/10.1186/2193-1801-2-105

Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., & Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. *British Journal of Educational Technology*, *51*(5), 1548–1562. https://doi.org/10.1111/bjet.12958

Leahy, W., & Sweller, J. (2008). The imagination effect increases with an increased intrinsic cognitive load. *Applied Cognitive Psychology*, *22*(2), 273–283. https://doi.org/10.1002/acp.1373

Leppink, J. (2017). Cognitive load theory: Practical implications and an important challenge. *Journal of Taibah University Medical Sciences*, *12*(5), 385–391. https://doi.org/10.1016/j.jtumed.2017.05.003

Leppink, J., & Van Den Heuvel, A. (2015). The evolution of cognitive load theory and its application to medical education. *Perspectives on Medical Education*, *4*(3), 119–127. https://doi.org/10.1007/S40037-015-0192-X

Liu, Q., Yu, S., Chen, W., Wang, Q., & Xu, S. (2021). The effects of an augmented reality based magnetic experimental tool on students' knowledge improvement and cognitive load. *Journal of Computer Assisted Learning, 37*(3), 645–656. Portico. https://doi.org/10.1111/jcal.12513

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Moos, D. C., & Pitton, D. (2014). Student teacher challenges: Using the cognitive load theory as an explanatory lens. *Teaching Education*, *25*(2), 127–141. https://doi.org/10.1080/10476210.2012.754869

Moreno, R. (2006). When worked examples don't work: Is cognitive load theory at an Impasse? *Learning and Instruction*, *16*(2), 170–181. https://doi.org/10.1016/j.learninstruc.2006.02.006

Nawal, A. F. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies, 3*(1), 385–402. https://doi.org/10.1080/23752696.2018.1513812

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994a). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*(4), 351–371. https://doi.org/10.1007/BF02213420

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994b). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*(1), 122–133. https://doi.org/10.1037/0022-0663.86.1.122

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, *32*(1/2), 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 27–42). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.004

Paas, F., & Sweller, J. (2021). Implications of cognitive load theory for multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 73–81). Cambridge University Press. https://doi.org/10.1017/9781108894333.009

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F., Tuovinen, J. E., Van Merriënboer, J. J. G., & Aubteen Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, *53*(3), 25–34. https://doi.org/10.1007/BF02504795

Paas, F., Van Gog, T., & Sweller, J. (2010). Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, *22*(2), 115–121. https://doi.org/10.1007/s10648-010-9133-8

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, *17*(4), 398–422. https://doi.org/10.3758/BF03202613

Plass, J. L., Moreno, R., & Brünken, R. (Eds.). (2010). *Cognitive load theory* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511844744

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, *12*(1), 61–86. https://doi.org/10.1016/S0959-4752(01)00016-0

Reed, S. K. (2012). *Cognition: Theories and applications*. CENGAGE learning.

Rikers, R. M. J. P., Van Gerven, P. W. M., & Schmidt, H. G. (2004). Cognitive load theory as a tool for expertise development. *Instructional Science*, *32*(1/2), 173–182. https://doi.org/10.1023/B:TRUC.0000021807.49315.31

Shin, S.-S. (2020). Structured query language learning: Concept map-based instruction based on cognitive load theory. *IEEE Access*, *8*, 100095–100110. https://doi.org/10.1109/ACCESS.2020.2997934

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J. (1999). *Instructional design in technical areas*. ACER Press.

Sweller, J. (2003). Evolution of human cognitive architecture. In *Psychology of Learning and Motivation* (Vol. 43, pp. 215–266). Academic Press. https://doi.org/10.1016/S0079-7421(03)01015-6

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, *32*(1/2), 9–31. https://doi.org/10.1023/B:TRUC.0000021808.72598.4d

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (1st ed., pp. 19–30). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.003

Sweller, J. (2008). Evolutionary bases of human cognitive architecture: Implications for computing education. *Proceedings of the Fourth International Workshop on Computing Education Research*, 1–2. https://doi.org/10.1145/1404520.1404521

Sweller, J. (2010a). Cognitive load theory: Recent theoretical advances. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory* (1st ed., pp. 29–47). Cambridge University Press. https://doi.org/10.1017/CBO9780511844744.004

Sweller, J. (2010b). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2011). Cognitive load theory. In J. Mestre & B. Ross (Eds.), *Psychology of Learning and Motivation* (Vol. 55, pp. 37–76). Elsevier. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues.* (pp. 295–325). American Psychological Association. https://doi.org/10.1037/13273-011

Sweller, J. (2019). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68(1*), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review, 35*(4), 1-20. https://doi.org/10.1007/s10648-023-09817-2

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*(3), 185–233. https://doi.org/10.1207/s1532690xci1203_1

Sweller, J., & Paas, F. (2017). Should self-regulated learning be integrated with cognitive load theory? A commentary. *Learning and Instruction*, *51*, 85–89. https://doi.org/10.1016/j.learninstruc.2017.05.005

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Szarkowska, A., Krejtz, K., Dutka, Ł., & Pilipczuk, O. (2016). Cognitive load in intralingual and interlingual respeaking – a preliminary study. *Poznan Studies in Contemporary Linguistics, 52*(2). https://doi.org/10.1515/psicl-2016-0008

Turan, Z., Meral, E., & Sahin, I. F. (2018). The impact of mobile augmented reality in geography education: achievements, cognitive loads and views of university students. *Journal of Geography in Higher Education, 42*(3), 427–441. https://doi.org/10.1080/03098265.2018.1455174

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology*, *25*(4), 584–587. https://doi.org/10.1002/acp.1726

Van Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development*, *53*(3), 5–13. https://doi.org/10.1007/BF02504793

Van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*, *20*(3), 343–352. https://doi.org/10.1002/acp.1250

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13. https://doi.org/10.1207/S15326985EP3801_2

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Van Merriënboer, J. J., & Kirschner, P. A. (2001). Three worlds of instructional design: State of the art and future directions. *Instructional Science*, *29*, 429–441. https://doi.org/10.1023/A:1011904127543

Verhoeven, L., Schnotz, W., & Paas, F. (2009). Cognitive load in interactive knowledge construction. *Learning and Instruction*, *19*(5), 369–375. https://doi.org/10.1016/j.learninstruc.2009.02.002

Vogel-Walcutt, J. J., Gebrim, J. B., Bowers, C., Carper, T. M., & Nicholson, D. (2011). Cognitive load theory vs. constructivist approaches: Which best leads to efficient, deep learning?: CLT vs. constructivism for learning. *Journal of Computer Assisted Learning*, *27*(2), 133–145. https://doi.org/10.1111/j.1365-2729.2010.00381.x

Young, J. Q., & Sewell, J. L. (2015). Applying cognitive load theory to medical education: Construct and measurement challenges. *Perspectives on Medical Education*, *4*(3), 107–109. https://doi.org/10.1007/S40037-015-0193-9

Yu, Z. (2021). The effect of teacher presence in videos on intrinsic cognitive loads and academic achievements. *Innovations in Education and Teaching International*, 59(5), 574–585. https://doi.org/10.1080/14703297.2021.1889394

Zavgorodniaia, A., Duran, R., Hellas, A., Seppala, O., & Sorva, J. (2020). Measuring the cognitive load of learning to program: A replication study. *United Kingdom & Ireland Computing Education Research Conference.*, 3–9. https://doi.org/10.1145/3416465.3416468