

RESEARCH ARTICLE

Open Access



“Connecting concepts helps put main ideas together”: cognitive load and usability in learning biology with an AI-enriched textbook

Marta M. Koć-Januchta^{1*} , Konrad J. Schönborn¹, Casey Roehrig², Vinay K. Chaudhri³, Lena A. E. Tibell¹ and H. Craig Heller⁴

*Correspondence:

marta.koc-januchta@liu.se

¹ Department of Science and Technology (ITN), Linköping University, Campus Norrköping, Norrköping 601 74, Sweden

Full list of author information is available at the end of the article

Abstract

Rapid developments in educational technology in higher education are intended to make learning more engaging and effective. At the same time, cognitive load theory stresses limitations of human cognitive architecture and urges educational developers to design learning tools that optimise learners' mental capacities. In a 2-month study we investigated university students' learning with an AI-enriched digital biology textbook that integrates a 5000-concept knowledge base and algorithms offering the possibility to ask questions and receive answers. The study aimed to shed more light on differences between three sub-types (intrinsic, germane and extraneous) of cognitive load and their relationship with learning gain, self-regulated learning and usability perception while students interacted with the AI-enriched book during an introductory biology course. We found that students displayed a beneficial learning pattern with germane cognitive load significantly higher than both intrinsic and extraneous loads showing that they were engaged in meaningful learning throughout the study. A significant correlation between germane load and accessing linked suggested questions available in the AI-book indicates that the book may support deep learning. Additionally, results showed that perceived non-optimal design, which deflects cognitive resources away from meaningful processing accompanied lower learning gains. Nevertheless, students reported substantially more favourable than unfavourable opinions of the AI-book. The findings provide new approaches for investigating cognitive load types in relation to learning with emerging digital tools in higher education. The findings also highlight the importance of optimally aligning educational technologies and human cognitive architecture.

Keywords: Artificial intelligence, Cognitive load, Usability, Learning, Educational technology, Adaptive textbook, Higher education, Engagement

Introduction

Educational technology aims to make learning more effective, accessible, and engaging for learners. Digital learning environments offer support for learning through glossaries, hyperlinks, multimedia resources and different types of feedback (e.g., Alevén et al., 2003). Ideally, digital learning environments should serve to support learners' cognitive processes by reducing the consumption of cognitive resources while promoting retention and meaningful learning (Bates et al., 2020). However, since learning is impossible without engaging cognitive resources, educational interventions should focus on optimising the overall cognitive effort attributed to making learning more efficient (Eitel et al., 2020; Sweller & Chandler, 1991). In this regard, emerging research is exploring how the integration of artificial intelligence (AI) in educational technology may provide opportunities for improving learning through efficient use of cognitive resources (Zawacki-Richter et al., 2019). Effective learning with digital tools requires optimally channelling cognitive effort (Klepsch et al., 2017) as well as providing students with self-regulated learning skills (Eitel et al., 2020; Ibili & Billinghamurst, 2019). In exploring how to do so, more fine-grained approaches that explore how cognitive load, usability perception and self-regulated learning skills impact learning, are needed.

Our previous work investigated university students' learning and interaction with an AI-enriched digital biology book versus a traditional E-book, during a short educational intervention (cf. Koć-Januchta et al., 2020). One outcome of the study was identifying the need for analysing the specific relationships between cognitive load and usability when engaging with emerging textbook technologies. In response, the current study explores students' differential cognitive effort over a longer period while learning biology with the AI-enriched book. Integration of the AI-book as a learning resource in a real university course setting allowed for an ecologically valid context. Our goal was to explore changes in three sub-types of cognitive load, namely intrinsic, germane, and extraneous load and their relationship with usability perception and self-regulated learning while students learned with the AI-book. In this pursuit, we aim to contribute to knowledge on systematically differentiating between cognitive load types. Specifically, in relation to students' learning with digital environments in higher education.

Theoretical background

Cognitive load theory and learning

Since the 1980's, cognitive load theory (CLT; Sweller & Chandler, 1991) has been established as one of the most applied theories for considering relationships between instructional design and mental problem-solving resources. The theory is concerned with defining the overall mental effort (cognitive load) attributed to working memory resources delegated to accomplishing a task (Kalyuga & Liu, 2015). Cognitive load (CL) comprises three sub-types (van Merriënboer & Sweller, 2005), namely *intrinsic* cognitive load (ICL), *extraneous* cognitive load (ECL), and *germane* cognitive load (GCL).

Intrinsic cognitive load (ICL) is a result of cognitive activities needed for understanding the information inherent in a task and depends largely on the complexity of this information. A high level of ICL is caused by a large number of "task elements" which must be processed simultaneously during learning and also determined by the level of

the learner's knowledge (Sweller, 2010). Extraneous cognitive load (ECL) is experienced by learners when they are forced to invest their cognitive resources in activities that are not immediately relevant to the learning task at hand. The main source of ECL is non-optimal or flawed instructional design, such as an unnecessary complex layout of a digital learning interface (Klepsch et al., 2017). The final sub-type, germane cognitive load (GCL), results from constructing schemas (Sweller et al., 1998) or mental models (Paas et al., 2004) during meaningful learning processes. An example of an activity that can increase GCL is integrating new information with knowledge the learner already has. Therefore, high levels of GCL can be interpreted as a sign of meaningful learning (Klepsch et al., 2017).

From a CLT perspective, learning is defined as constructing and automating schemas in long-term memory (Paas et al., 2004), which involves all three sub-types of cognitive load to some extent or other. To optimise learning, the sum of the load types should not exceed the learner's limited working memory capacity. Hence, assuming that ICL is inherent in the nature of the task, optimising learning should focus on minimising ECL and increasing GCL (see Klepsch et al., 2017). Although the cognitive load construct provides insight into the usefulness and effectiveness of learning with new educational technologies (e.g., Kalyuga & Liu, 2015), it remains challenging to measure (e.g., Klepsch et al., 2017). The literature contains multiple cognitive load measures that have emerged over time, which range from (the most widely-used) self-rating techniques to recent physiological measures (e.g. pupillary responses). From an instructional design standpoint, it is crucial to deduce ways to measure CL differentially by exploring the relative impact of all three load sub-types during learning (Ibili & Billingham, 2019; Klepsch et al., 2017; Mutlu-Bayraktar et al., 2019).

Interactive educational technology and cognitive load

Technology enhanced learning is becoming more and more apparent in higher education bringing with it both hopes and challenges. Amongst the hopes, digital resources that integrate interventions such as AI may help support the learning of complex scientific knowledge such as biology (e.g., Corbett et al., 2010). Interactive technology can also help students to learn more efficiently by offering multimedia resources, interactive glossaries, prompts, answers to questions, help in constructing models and even personalised suggestions for further learning (Alevén et al., 2003; Koć-Januchta et al., 2020; Linn et al., 2014). Nevertheless, new technological opportunities for learning are also associated with multiple challenges. From a motivational perspective, students may experience decreased motivation when learning on their own from a digital learning environment (e.g., DeVore et al., 2017). Consequently, learning with digital tools often requires advanced skills in independent learning, self-regulation and learning strategies (Glover et al., 2016; Means et al., 2009). Additionally, learners may experience cognitive overload when learning with digital technology (Alevén et al., 2003).

As elucidated previously, cognitive overload is often caused by high levels of extraneous cognitive load (Klepsch et al., 2017). While ICL and GCL concern processing of learning elements and promoting meaningful learning, respectively, extraneous load arises mainly from the way information is conveyed. Poorly designed digital learning tools may increase ECL to such an extent that it impairs learning (Moreno & Mayer,

2007). Therefore, where possible, ECL should be reduced by optimising the design of the learning environment. Design elements such as the range and complexity of implemented digital features must be carefully considered, since a complicated interface can render cognitive overload (Scheiter & Gerjets, 2007; Sweller et al., 1998). Lowering extraneous cognitive load frees the availability of cognitive resources that can be directed to germane load and thus stimulate deeper learning (Klepsch et al., 2017). As part of our previous work (Koć-Januchta et al., 2020), we presented students' opinions suggesting that ECL may increase over time when using an AI-enriched book. As part of that study, where we compared an AI-enriched book and a traditional E-book, students from a research university were interviewed about their experiences of using both types of digital book. They pointed out several advantages of the AI-enriched book over the traditional E-book (e.g., obtaining pop-up definitions to terms in real-time) but also reported growing dissatisfaction with the AI-enriched book as usage time progressed. We observed that a longer usage of the book revealed potential design-related disadvantages (e.g., AI-functionalities were sometimes confusing or contained too much information, or made one unsure of their learning). At the same time, the more difficulty students perceived learning with the AI-enriched book, the less positively they assessed its usability (Koć-Januchta et al., 2020).

Usability perception and cognitive load in educational technology

Usability is an essential measure when exploring user experiences of digital educational technologies (Diefenbach et al., 2014). The concept includes subjective and objective components, which consist of perceived usability or satisfaction (how comfortable it is to use a digital tool) and efficiency (the time and effort cost in using the digital tool), respectively (Lewis, 2018). One of the most popular measures of perceived usability is the System Usability Scale (SUS) questionnaire developed by Brooke (1996). The SUS is suitable for measuring satisfaction (e.g., meeting expectations) and ease of using the learning tool.

Many studies show that perceiving a learning system as useful is associated with a reduced cognitive load (e.g., Pantano et al., 2017), whereas feeling confused when using a system leads to increased cognitive load (Kılıç, 2007). Moreover, Costley and Lange (2017) found that an increase in users' intention to use a tool is influenced by effective instructional design. Additionally, optimal instructional design correlates with increased germane load indicating deep learning (Costley & Lange, 2017). Notably, Ibili and Billinghamurst (2019) have stated that perceived usefulness (perceiving a learning tool as improving learning) and perceived ease of use (perceiving a tool as easy to learn with) were strongly correlated with all three types of cognitive load (ICL, GCL and ECL). Specifically, usefulness was negatively correlated with ICL (for females) and with ECL (for males). At the same time, both usefulness and ease of use were strongly positively correlated with germane cognitive load (Ibili & Billinghamurst, 2019). Lastly, cognitive load is strongly connected with self-regulated learning. For example, high cognitive load might originate from students' insufficient self-control skills and low willingness to learn (de Bruin et al., 2020; Eitel et al., 2020).

Self-regulated learning and cognitive strategies

Acquiring self-regulation skills is important for learning and a research topic of high interest when it comes to individual learning with digital tools (Steffens, 2006). Zimmerman (2011) relates self-regulated learning to the degree to which learners participate actively in their own learning at the metacognitive, motivational, and behavioral level. In addition, Paris and Winograd (2003) describe self-regulated learning as a process in which learners approach problems, apply strategies, monitor their performance, and assess the results of their efforts. Self-regulated learners are more likely to improve their academic achievements by selecting and controlling cognitive processes involved in learning (Pintrich & De Groot, 1990). To learn deeply, one should be able to elaborate and organise information and monitor one's learning process (Pintrich & De Groot, 1990; Soenens et al., 2012). In this regard, technology-enhanced learning environments offer an opportunity to support self-regulated learning by helping students to plan, monitor and evaluate the cognitive, motivational, and affective components of their own learning (Steffens, 2006).

Learning biology: conceptual knowledge and cognitive skills

Biology is a natural science concerned with studying structures and processes associated with living organisms (e.g., Sadava et al., 2017). Learning biology involves building a conceptual understanding of the structure of the (bio)molecules of life that include proteins, enzymes, carbohydrates, lipids, and nucleic acids. This learning also includes developing core knowledge about the "unit of life" (the cell) and the plethora of cellular processes such as DNA replication, mitosis, meiosis, and gene expression. In turn, such knowledge must be integrated with understanding physiological functions such as photosynthesis, muscle contraction, neural and endocrine control. Furthermore, all these aspects of biology contribute to understanding populations, ecosystems, and evolution. Moreover, constructing biological knowledge involves making links to other scientific disciplines and reasoning at various levels of spatial and temporal scale.

Cognitive skills associated with successful biology learning include: retaining biological knowledge, integrating knowledge with other concepts (while transitioning different levels of biological organisation), transferring learnt knowledge to novel tasks, as well as reasoning both "locally" and "globally" about a biological concept (Anderson & Schönborn, 2008).

Aims of the study

The objectives of this study are to investigate:

1. Any differences in how students experience the three types of cognitive load (ICL, GCL and ECL) while learning with an AI-enriched biology book at the beginning and the end of the study.
2. Relationships between the three types of cognitive load, usability, self-regulation, cognitive strategy use, and learning gain while interacting with the AI-enriched book.

Methods

Study setting: a reference point and main study

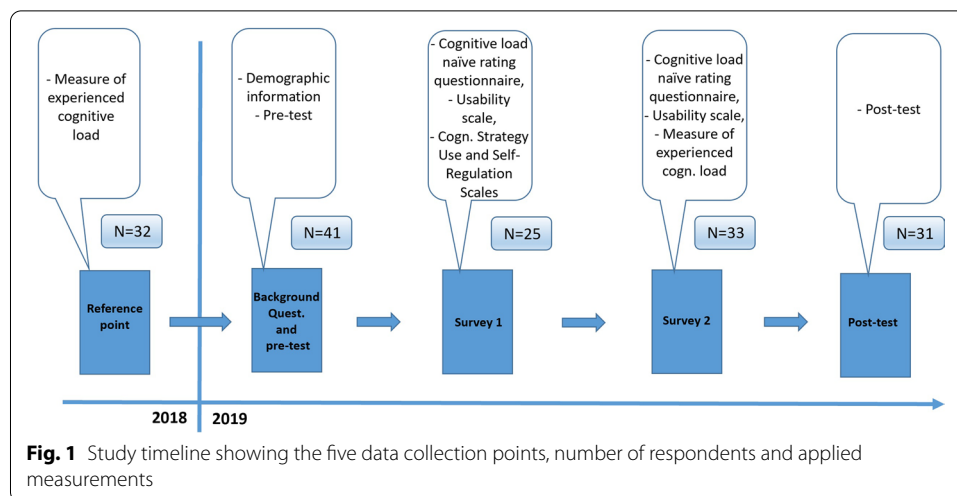
The collaboration described in this paper began in 2018, in planning toward a research intervention involving an AI-enriched digital biology textbook for 2019 at Harvard University. In 2018, we obtained initial information about cognitive load experienced by students learning biology. We asked 32 (53.1% female; age $M = 27.53$; $SD = 6.63$) students attending an introductory biology course, and who used traditional hardcopy and E-books to answer two questions on their experienced cognitive load (Paas, 1992). Doing so generated a reference point for a 2019 main study that investigated cognitive load experienced when students participated in the same course while engaging with an AI-enriched digital textbook.

Data collection and participants: main study

The study was conducted from September to October 2019 while students attended the course “Introduction to Molecular and Cellular Biology” at Harvard University. During this time, students responded to several questionnaires including a written pre-test and demographic questions, followed by two online surveys, and ending with a written post-test.

Demographic questions included gender, age, Grade Point Average (GPA), native language and previously completed science courses. The pre-test (7 multiple choice questions) assessed students’ biological knowledge and was answered during the first lecture. Subsequently, students were provided access to an AI-enriched book integrated in an iPad platform that covered the first ten chapters of the original hardcopy biology textbook (Sadava et al., 2017), and which corresponded to the first seven lectures of the course.

During the study, students had unlimited individual access to the AI-enriched book both on campus and at home. Students could use the book as they wished, while preparing for class, tests, and examinations. During the course, the lecturer provided the students with a pre-class and a post-class study guide for each lecture as supplementary material. The pre-class study guide specified the reading that students were expected to do before class, and the post-class study guide provided the intended learning goals and outcomes of that class. Each study guide also included a set of questions aligned with the learning outcomes for students to assess their knowledge. Students were also provided with explanation materials on how best to use the features of the AI-enriched book in the context of the learning goals (see [Appendix](#)). While using the AI-enriched book students responded to two surveys consisting of cognitive load, cognitive strategy use, self-regulation, and usability scales (the latter included open-ended questions). Altogether, 42 participants participated in the study, of which 69% were female (age range 17 to 44 years, $M = 26.28$ and $SD = 4.87$). Although the study consisted of several measurement points, not all students participated in all data collections. Figure 1 depicts the study timeline, showing the number of participants at each data collection point and associated questionnaires.



Specific instruments and measures

As part of the five measuring points obtained in the study (Fig. 1), Table 1 provides further detailed information about data gathering dates and respective instruments and measurements.

The AI-enriched digital biology textbook

The learning resource in focus is an example of an intelligent textbook that incorporates AI elements. The AI-enriched book is based on a widely used international hardcopy biology textbook (Sadava et al., 2017), and integrated into an iPad. The AI-book has typical electronic book features, such as the possibility to highlight text, enlarge figures, and make notes. However, in applying natural language processing techniques and formal knowledge representation, the AI-enriched book also offers a 5000-concept knowledge base and algorithms that provide the possibility to ask questions and receive answers. Figure 2 demonstrates three different ways of generating/asking questions. Firstly, students can input a question by tapping the “magnifying glass” icon (see 1, Fig. 2). Secondly, students can tap on an underlined term (“dotted word”) to view its short definition and access further information on the topic (e.g., suggested questions referring to the term) by tapping on the button “MORE” placed near the definition (see 2, Fig. 2). Thirdly, highlighting the text produces a note card with suggested, most relevant questions about the highlighted text (see 3, Fig. 2).

The AI-enriched book includes multiple elements of artificial intelligence that incorporate knowledge-acquisition and knowledge presentation processes. Specific AI elements applied in the AI-enriched book comprise a formal knowledge representation of book content, natural language processing to interpret a student’s inputted or selected suggested questions, and natural language mechanisms for generating answers (Chaudhri et al., 2013).

Table 1 Data collection dates and accompanying instruments and measurements

Date	Data collection	Instrument	Specific measurement items
Autumn 2018	Reference point: cognitive load	Cognitive load questions (perceived difficulty and mental effort; Paas, 1992)	Cognitive load questions (2 items), answer scale ranged from 1 point (<i>very, very easy</i> or <i>very, very low mental effort</i>) to 9 points (<i>very, very difficult</i> or <i>very, very high mental effort</i>)
04 Sept. 2019	Background questions and Pre-test	Demographic questions Biology knowledge Pre-test	Demographics (5 items): age, gender, GPA, native language, previous courses Pre-test (7 items): measuring pre-knowledge with multiple choice questions. (Score 1 to 7 points)
04–14 Sept. 2019	Survey 1	Cognitive load <i>naïve rating</i> questionnaire (Klepsch et al., 2017) Usability scale (Brooke, 1996) Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich & De Groot, 1990)	Cognitive load questionnaire (8 items): ICL, GCL and ECL measured on a 7-point Likert scale ranging from 1 (<i>strongly disagree</i>) to 7 (<i>strongly agree</i>) Usability scale (2 Likert scale items and 2 open questions asking for 2 positive and 2 negative aspects of the AI-enriched book. Likert items on a 7-point scale ranging from 1 (<i>strongly disagree</i>) to 7 (<i>strongly agree</i>) Two scales from MSLQ: cognitive strategy use (13 items on organizational, rehearsal and elaboration) and Self-Regulation (9 items on metacognition, comprehension, monitoring and effort management), on a 7-point Likert scale ranging from 1 (<i>not at all true of me</i>) to 7 (<i>very true of me</i>)
09–25 Oct. 2019	Survey 2	Same as Survey 1 but also included Cognitive load questions from Autumn 2018	As described above
Oct.–Nov. 2019	Post-test	Biology knowledge Post-test	Post-test identical to Pre-test

Analytical procedure

Ethical approval for the study was obtained and data analysed only from participants that provided informed consent. Quantitative analyses included descriptive statistics and statistical comparisons. Cronbach's alpha was used to calculate the internal consistency and reliability of the applied measures. We applied t-tests and a General Linear Model (GLM) Repeated Measures procedure to calculate any significant differences between variables. In reducing the risk of a Type I error when using t-tests, we applied a conservative Bonferroni correction. The GLM Repeated Measures procedure provided analysis of variance results when administering the same measurements to the same participants on several occasions (Field, 2013), and to calculate changes between three types of experienced cognitive load (ICL, GCL, ECL) over time (first and second measurements, respectively). We also applied Pillai's trace as a statistic robust to violations of analysis of variance assumptions (Finch, 2005).

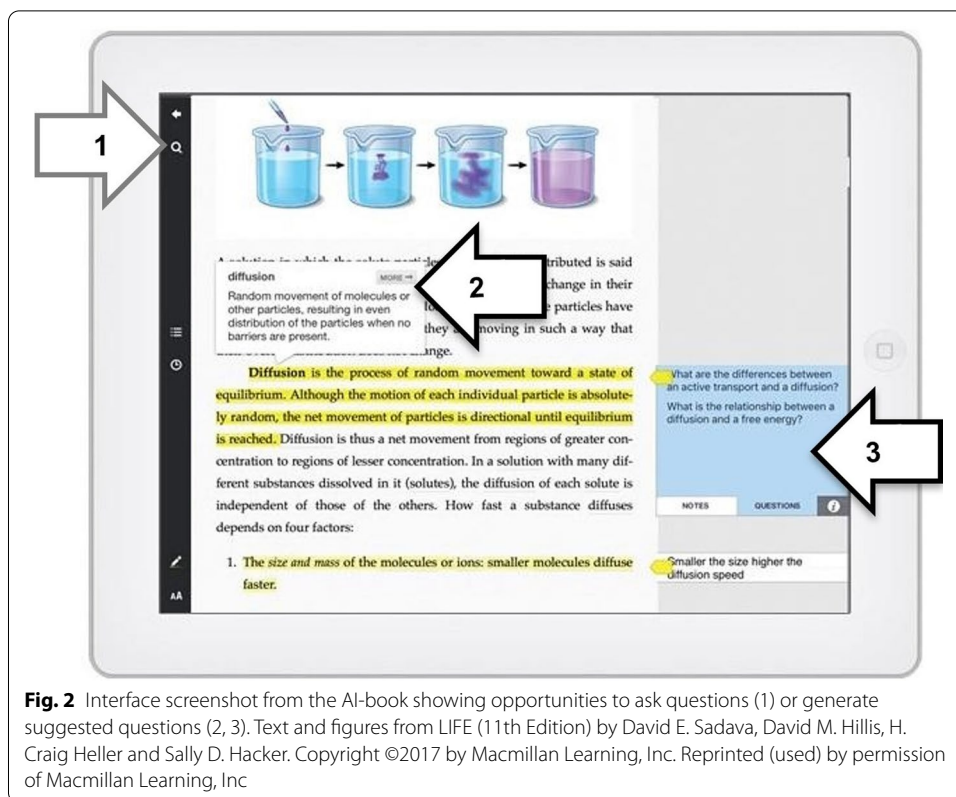


Fig. 2 Interface screenshot from the AI-book showing opportunities to ask questions (1) or generate suggested questions (2, 3). Text and figures from LIFE (11th Edition) by David E. Sadava, David M. Hillis, H. Craig Heller and Sally D. Hacker. Copyright ©2017 by Macmillan Learning, Inc. Reprinted (used) by permission of Macmillan Learning, Inc

Table 2 Reliability of the measurement instruments calculated after first administration

Measurement instrument	Cronbach's alpha	Number of items
Naïve rating questionnaire: intrinsic cognitive load (ICL)	0.90	2
Naïve rating questionnaire: germane cognitive load (GCL)	0.54	3
Naïve rating questionnaire: extraneous cognitive load (ECL)	0.85	3
Usability	0.87	2
MSLQ: cognitive strategy use	0.72	13
MSLQ: self-regulation	0.83	9

Moreover, we calculated several Pearson’s correlations to discern relations between study variables. Finally, we qualitatively analysed students’ open-ended responses on usability in relation to Klepsch et al.’s (2017) definitions for ICL, GCL and ECL.

Results

Reliabilities of applied instruments and measures

The reliability (internal consistency) of the administered scales (Table 1), calculated with Cronbach’s alpha are reported in Table 2.

As shown in Table 2, almost all scales applied in the study had a good or very good reliability, apart from the GCL scale, where reliability was rather low (0.54). However, a similar reliability score was obtained in the original development of the scale (Klepsch et al., 2017).

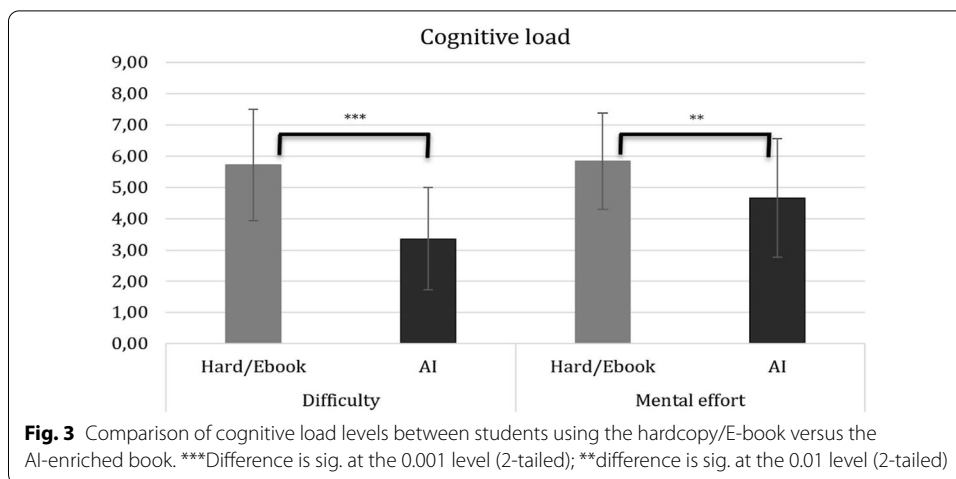


Table 3 Results from a t-test comparing students’ cognitive load when using a hardcopy and/or E-book for learning (2018) with the AI-enriched book (2019)

Cognitive load	Group	N	Mean	Std. deviation	t	p
Difficulty	Hard/E-book	32	5.72	1.78	5.55	p < 0.001
	AI-book	33	3.36	1.64		
Mental effort	Hard/E-book	31	5.84	1.55	2.71	p < 0.01
	AI-book	33	4.67	1.90		

Reference point study: comparison of cognitive load between different book versions

Since the main goal of the study was to investigate cognitive load, we commenced our analyses with establishing a reference point for cognitive load measurement. In this regard, we compared cognitive load experienced by the students during 2018 (when using the hardcopy and/or E-book) with 2019 (using the AI-enriched book) in terms of difficulty and mental effort when learning (Paas, 1992). Figure 3 displays cognitive load levels (in terms of difficulty and mental effort) when students used the hardcopy/ E-book version versus the AI-enriched version of the biology textbook.

Figure 3 shows that both difficulty and mental effort when learning were significantly lower for the AI-enriched book. Table 3 below provides results from a t-test comparison of the two student groups.

An independent samples t-test showed significant higher perceived cognitive load when learning with a hardcopy or/and E-book in comparison to the AI-book. According to Sweller and Chandler (1991) the optimal level of cognitive load depends on the differential type of cognitive load a learner experiences at different points of learning. To specifically investigate such differential cognitive load when learning with the AI-book, the main study focused on exploring the three types of cognitive load (ICL, GCL, ECL) over time.

Main study: identifying and comparing three types of cognitive load

The main study compared three types of cognitive load experienced by students at the beginning and close of the study. To compare levels of three cognitive load types at both measuring points, we analysed the data within a General Linear Model using a repeated measures design framework (Field, 2013). A significant main effect on cognitive load types was revealed, $F(2, 17) = 52.96; p < 0.001; \eta^2 = 0.86$ (Pillai’s Trace) showing that the perceived level of cognitive load differed significantly due to its type (intrinsic, germane or extraneous) (Fig. 4). Detailed pairwise comparisons with Bonferroni correction revealed that:

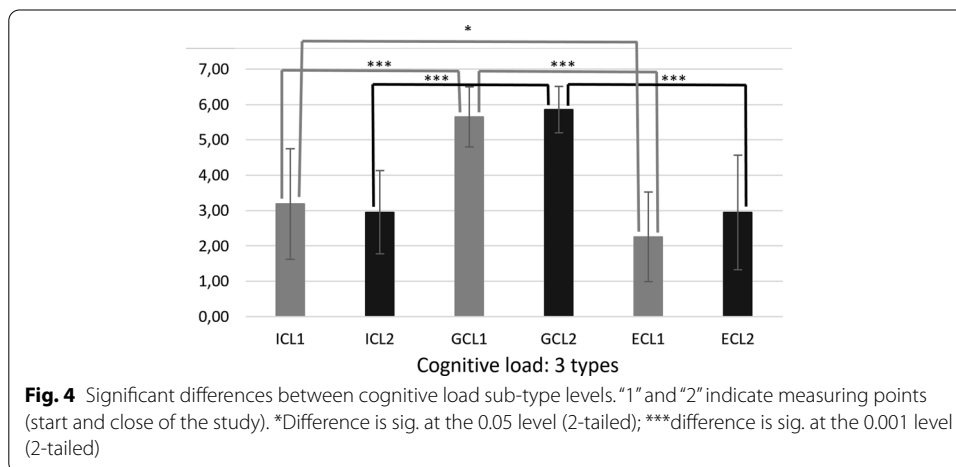
- At the beginning of the study, all three types of cognitive load levels differed significantly. GCL1 was significantly higher than both ICL1 and ECL1. At the same time, ICL1 and ECL1 differed significantly (differences between ICL1, GCL1, ECL1 are depicted in grey in Fig. 4).
- At the end of the study there were significant differences between germane load (GCL2) and the two other cognitive load types (differences between ICL2, ECL2 are depicted in black in Fig. 4). In contrast with the beginning of the study, there was no significant difference between ICL2 and ECL2 ($p > 0.05$) at the end of the study.

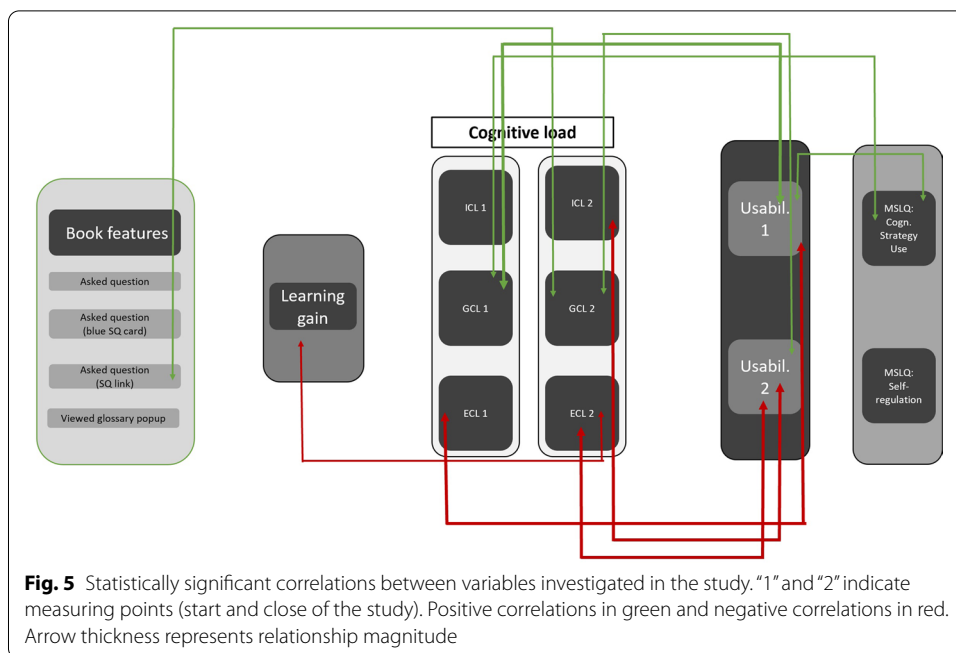
Although non-significant, the t-test comparison between ECL1 and ECL2 at the close of the study suggests a possible increase in ECL, $t(18) = -1.8, p = 0.088; M_{ECL1} = 2.25; SD_{ECL1} = 1.27; M_{ECL2} = 2.95; SD_{ECL2} = 1.62$.

The overall mean level of cognitive load did not differ significantly at the beginning and at the end of the study, and the main effect of the cognitive load level was not statistically significant, $F(1, 18) = 1.53; p = 0.233; \eta^2 = 0.078 (M_1 = 3.68; SD_1 = 0.79; M_2 = 3.82; SD_2 = 0.86)$.

Cognitive load: correlations with other variables

Pearson’s correlation analyses were conducted to discover any relationships between study variables. Figure 5 depicts significant correlations between the three cognitive





load types and usability at the beginning ("1") and close of the study ("2"), with AI-book features, cognitive strategy use, self-regulation and learning gain.

Correlations at the beginning of the study (measurement "1") reveal that germane load (GCL1) correlates positively with both Cognitive Strategy Use ($r=0.50^*$) and Usability1 ($r=0.51^{**}$). High germane load indicates that a learner directs effort to learning deeply, suggesting that at the study commencement, the more advanced students' skills to learn strategically (Cognitive strategy use) the higher the mental effort directed to learning deeply with the AI-book. Moreover, a more positive perception of AI-book usability was linked to higher mental effort. Extraneous load (ECL1) correlated negatively with Usability1 ($r=-0.69^{**}$), which indicates that higher perceived extraneous cognitive load ("undesirable" load arising from "struggling" with a learning environment) accompanied a lower usability perception at the start of the study.

At the close of the study (measurement "2") intrinsic load (ICL2) correlated negatively with Usability2 ($r=-0.50^{**}$), which infers that lower perceived mental effort when learning with the AI-enriched book accompanies a more positive usability perception (Usability2). Furthermore, germane load (GCL2) correlates positively with Usability2 ($r=0.42^*$), which suggests that a more positive usability perception (Usability2) the higher the mental effort directed to learning deeply with the AI-enriched book. Extraneous load (ECL2) correlates negatively with Usability2 ($r=-0.76^{**}$), showing that a higher perceived extraneous load is linked to a lower usability perception. Lastly, results revealed a significant positive correlation between germane load at the end of the study (GCL2) and the number of linked suggested questions accessed for example by tapping the "MORE" button (see Fig. 2 (2)) ($r=0.45^*$). This result indicates that the more often students accessed suggested questions to explore further knowledge, the more mental effort they invested in learning deeply from the AI-enriched book (GCL2).

Learning gain and usability: correlations with other variables

A t-test comparison between post- and pre-test revealed a significant learning gain, $t(30) = 6.45$, $p < 0.001$; $M_{\text{pre-test}} = 3.00$; $SD_{\text{pre-test}} = 1.15$; $M_{\text{post-test}} = 4.74$; $SD_{\text{post-test}} = 1.06$. Learning gain correlated negatively with extraneous load (ECL2; $r = -0.39^*$), which implies that high levels of ECL might have impeded meaningful learning with the AI-book system.

A t-test comparison between Usability1 and Usability2 revealed a near significant decrease in usability perception at the study close, $t(18) = 2.09$, $p = 0.051$; $M_{\text{Usability1}} = 5.82$; $SD_{\text{Usability1}} = 1.15$; $M_{\text{Usability2}} = 5.11$; $SD_{\text{Usability2}} = 1.28$. Additionally, usability at the beginning of the study (Usability1) correlated positively with Cognitive Strategy Use ($r = 0.47^*$), which shows that the more advanced skills in learning strategically (Cognitive strategy use) that students have, the more positive their usability perception of the AI-book.

Students' perceived usability of the AI-book

Thirty-three students rated AI-book usability through Likert scale statements and by being asked to express two positive and two negative aspects of using the book. Generated comments were grouped in relation to the ICL, GCL or ECL definitions in Klepsch et al. (2017) (Table 4). Specifically, we categorized comments regarding mental effort originating from reading, decoding, and memorizing the content as *ICL-related*. Comments mentioning understanding and combining mental information into knowledge were categorized as *GCL-related*. Responses referring to design of the environment were categorized as *ECL-related*. Overall, 12 positive statements referred to ICL, 9 to GCL and 30 to ECL. Four negative comments referred to ICL, 7 to GCL and 33 to ECL. There were also 20 positive and 5 negative unclassified statements that referred to more than one sub-type of cognitive load, which we did not group or analyse in relation to cognitive load.

When perceiving the usability of the AI-book, students revealed 71 positive and 49 negative comments, respectively, with a weighting toward more positive usability perception. Students were very satisfied with the pop-up definitions, explanations, and associated concept maps, which were deemed to provide the meaning of a term without interrupting reading. These positively viewed aspects of the AI-book might help in decoding and memorising information and optimising experienced intrinsic cognitive load. Students also appreciated the search function that allows them to put ideas together, compare terms, and connect concepts to build understanding, and offers answers to complex questions. Such features of the AI-book may promote germane load, since "putting ideas together" underpins the construction of mental models of biological knowledge. Most of the positive comments were classified as design-related, referring to ECL. For instance, students pointed out that the AI-book offered multiple informative and interactive visual features affording possibilities for zooming, rapid navigation and glossary accessibility, as well as ease of use and portability.

When referring to negative aspects of the book in relation to ICL, students mentioned that there were too few pop-up definitions, too much to remember to use the book effectively, and that chapter names did not always match contained content. Among usability

Table 4 Examples of verbatim positive and negative usability comments related to intrinsic, germane and extraneous cognitive load (CL) types

	ICL-related	GCL-related	ECL-related
Positive	<p>"I really liked that whenever I forgot the definition of something, I could just click on the word and be reminded of the definition. That saved me a hell of a lot of time while reading [...]"</p> <p>"I like that the vocabulary is underlined and if you forget a word, you can click on it and be reminded"</p> <p>"Being able to click on various terms while reading to get definitions, explanations, etc. was really helpful"</p> <p>"Not enough inline [dotted] definitions"</p> <p>"Memorizability is still worse with this compared to a paper book due to lack of sense of location when reading"</p> <p>"Too many things to remember while using the AI-book"</p>	<p>"The connection of concepts really helps put main ideas together"</p> <p>"I also really like the feature where you can ask the AI-book to compare topics. This was really helpful [...] when I got confused on the similar terminology"</p> <p>"Connections of concepts to other chapters to build (or review) understanding to maintain learning foundation"</p> <p>"The AI is not good enough at answering basic questions related to the content"</p> <p>"I think the search engine, while amazing, could be further tweaked to really break down the topics, specifically with the comparison feature"</p> <p>"Connections between concepts not always made apparent. More depth would be useful"</p>	<p>"the immense amount of graphs/figures provided"</p> <p>"Media hyperlinks are easier than having to go to a separate device to watch videos"</p> <p>"Easy to use"</p> <p>"The search is [sic] works in a poor and frustrating manner: forcing me to use complete sentences [...] (my habit is for words over sentences as this is the convention of all internet search nowadays [...])"</p> <p>"It seemed to take a long time for the answers to load after i [sic] submitted a query"</p> <p>"Only available on iPads and not online/Android"</p>
Negative			

concerns in relation to GCL, students noted that the AI-based features did not answer basic (or all) posed questions, with some participants expressing a need for more in-depth answers. Usability in connection with ECL mostly included mechanisms in how to ask the book questions. For example, students were unsure why one needed to input complete question phrases rather than single terms alone to generate questions. Students also mentioned delayed loading of answers, problems with the highlighting functionality, lack of page numbers, book availability being constrained to iPads alone, and no possibility to access the book offline.

Discussion

Research suggests that students tend to perceive and learn from print and digital resources in similar ways (Sage et al., 2019; Koć-Januchta et al., 2020). However, this does not mean that the cognitive load associated with these media is similar. For example, studies show that interactively advanced learning tools may cause cognitive overload (Aleven et al., 2003; Scheiter & Gerjets, 2007; van Merriënboer & Sweller, 2005), by imposing design-related, extraneous load (Klepsch et al., 2017). In the reported reference point study, we established that university student participants experienced significantly lower cognitive load, when learning from an AI-enriched book, in comparison to a traditional hardcopy and E-book. Although different, the result did not explain which cognitive load types were differentially lower, when learning from the AI-enriched book. Since cognitive load concerns both the constraints as well as capabilities of human information-processing architecture, cognitive load levels can positively and negatively influence learning. While ECL emerges from design constraints of the learning matter and should be minimised, GCL and ICL are integral parts of the learning process (Moreno & Mayer, 2007; Sweller et al., 1998), where GCL is indicative of meaningful learning and should be promoted (Klepsch et al., 2017).

In the main study, and in contrast with shorter interventions, we measured three types of cognitive load twice, once at the beginning and once at the close of the study after students used the AI-book for 1.5 months. Giving students a longer unrestricted interaction with the AI-book revealed relationships between cognitive load, usability perception, self-regulation, cognitive strategy use, learning gain and book features. Interaction with the book over a longer period during a biology course in an authentic setting also contributed to ecological validity. Overall, students achieved significant learning gains ($M = 1.74$) during the study.

Our previous research suggested that as usage time increased, learners perceived more ECL-related disadvantages of using the AI-book (Koć-Januchta et al., 2020). Indeed, the literature documents a strong relationship between cognitive load and tool usability (e.g., Kılıç, 2007; Pantano et al., 2017), where difficulty in using the AI-enriched book correlated negatively with usability perception (Koć-Januchta et al., 2020). At the same time, a higher ECL level may result from underdeveloped self-regulated learning skills (e.g., de Bruin et al., 2020; Eitel et al., 2020). In reference to these findings, we aimed to investigate differences between the three types of cognitive load and usability, self-regulation, cognitive strategy use, and learning gain when students engaged with the AI-book over

an extended period in an ecologically valid setting. We discuss our findings in relation to the two major aims of the study.

Students' experiences of the three cognitive load types while learning with the AI-book

Results showed that GCL was significantly higher than ICL and ECL throughout the study, indicating that participants tried to learn deeply (Klepsch et al., 2017) for the entire course. At the beginning of the study intrinsic load was higher than extraneous load, while intrinsic load was as high as extraneous load (no significant difference) at the close. While the overall level of the three types of cognitive load did not change significantly, the possible increase between ECL2 and ECL1 could have occurred at the cost of both ICL and GCL. The second measurement revealed the cognitive load pattern after students learned with the AI-book for a longer time. We assume that with more time, students had more possibilities to experience the affordances of the interface design and may have directed their cognitive resources toward such extraneous processing instead of knowledge integration (Mutlu-Bayraktar et al., 2019). In this regard, students identified more ECL-related disadvantages (33) than advantages (30) of the AI-book, and directed most usability comments toward ECL (63) in comparison with ICL (16) or GCL (16).

Since user experience quality impacts engagement with digital educational tools and can induce effective learning (e.g., Bikowski, & Casal, 2018), it is crucial to identify elements that might require modification or improvement. In our study students were unsatisfied with some AI-book features, such as ECL-related usability around the asking question functions, with some experiencing it as inconvenient and frustrating. Moreover, they flagged long processing times for answer generation, problems with highlighting and insufficient visual icon size. While improving "straight-forward" technological and design-related shortcomings can be viewed as a subsequent step for AI-book developers, incorporating changes in response to students' observed critique should also be considered. For example, students were often dissatisfied with the procedure required to ask questions, namely, being required to "use complete sentences" instead of inputting single words. While many conventional search tools do indeed operate with single words, accessing the higher level of questioning offered by the book environment, requires specifying the sought relationship as a more descriptive entry. In this regard, the AI-book question answering interface is designed as a "concept calculator". In the same sense that a numerical calculator performs numerical operations on numerals, the AI-based answering architecture performs describe, compare, and relate "calculations" on concepts. In turn, by posing questions in terms of such "concept calculations", students learn to deploy meaningful questions for accessing, unpacking, and integrating biological knowledge.

Students also expressed many positive aspects of the AI-book such as ease of use, rapid accessibility to features, engaging visuals, and portability. Students often suggested a desire to use the AI-enriched book for a longer period and across different hardware platforms. In line with our previous study (Koć-Januchta et al., 2020), the AI-book seems to afford possibilities for deep, meaningful learning, as evidenced by the significantly higher presence of GCL in comparison with the other load types. At

the same time, the AI-enriched book technology unveils students' expectations of interactive digital tools that appear heavily shaped by experiences with conventional searching tools.

Relationships between the three types of cognitive load, usability, self-regulation, cognitive strategy use, and learning gain while interacting with the AI-book

Significant correlations between cognitive load types and other studied variables were revealed both at the beginning and close of students' interaction with the tool (de Bruin et al., 2020; Eitel et al., 2020; Ibili & Billinghamurst, 2019). In elaborating upon our previous work (Koć-Januchta et al., 2020), this study sheds more detail on relationships between cognitive load and usability when students use the AI-enriched technology to learn.

At the beginning of the study, ICL did not correlate significantly with other variables but correlated negatively with usability at the close. Since ICL originates from the number of interrelated elements for content to be learned (e.g., Klepsch et al., 2017), we assume that demands to learn more complex topics, especially when one has a low level of pre-knowledge, may relate to less favourable perceptions of book usability. The finding that GCL correlated positively with usability throughout the study (e.g., Ibili & Billinghamurst, 2019) and with cognitive strategy use (e.g., de Bruin et al., 2020; Eitel et al., 2020), infers that an increase in germane cognitive load is positively related with favourable perceptions of AI-book usability. Additionally, higher self-regulated learning skills, such as the deployment of organizational, rehearsal and elaboration learning strategies, are positively correlated with deep, meaningful learning. Students appear to appreciate the support offered by the AI-book in constructing mental models needed to acquire knowledge in biology.

Several AI-book features yielded a positive response from students, with the access to pop-up definitions being particularly salient. Students appreciated the possibility to rapidly check the meanings of terms while continuing to read without interruption, which appeared to be a very useful support for understanding biological terminology (Zukswert et al., 2019). In addition, better skills in learning strategically (higher scores on cognitive strategy use) are linked to meaningful, deeper learning (de Bruin et al., 2020; Eitel et al., 2020). ECL correlated only negatively with usability, and most notably, with learning gain. The exhibited link between ECL and usability perception of the tool confirms the nature of ECL (Ibili & Billinghamurst, 2019). Given that ECL is a consequence of expending mental resources on design elements, grappling with the features or technological affordances of the tool may impair learning and decrease satisfaction. Accessing linked suggested questions, by for example, tapping on the "MORE" button (Fig. 2 (2)) correlated significantly with GCL. Herein, "digging deeper" with the AI-book to actively seek and connect biological phenomena is likely linked to meaningful learning. In extension of our previous work (Koć-Januchta et al., 2020), the significant correlation between GCL and the questions feature exposed here begs for further investigation into what it is about the nature of the question feature that promotes meaningful learning.

Limitations of the study

Restriction on the transfer of the findings to other contexts include the following. Firstly, as was also the case in Klepsch et al.'s (2017) original study, the reliability of the GCL scale in the cognitive load questionnaire was unfavourable. We nevertheless adopted the questionnaire since it still manifested good overall item characteristics. Secondly, we acknowledge that the revealed learning gain is unlikely to be attributed to sole use of the AI-enriched book, and we thus only provide evidence of relationships rather than causality. Also, correlation between deep learning with the AI-enriched book and self-regulation might not be attributed to the AI-enriched book exclusively. Students with higher cognitive strategy use would likely tend to learn deeper regardless of textbook resource. Lastly, we did not compare students' use of E-book and AI versions of the resource since the authentic study setting required all students to have equivalent learning conditions.

Conclusions and implications

The initial reference point study showed that the AI-book was associated with lower cognitive load in comparison to hardcopy and E-book counterparts. In the main study that focused on the AI-book, both measurement points (beginning and end), revealed that GCL was always significantly higher than ICL and ECL. This pattern indicates that the AI-enriched book is associated with less mental effort and more meaningful learning. The same result could also be a sign of high learning competences and commitment from the participants, and may further support the positive correlation between GCL and self-regulation skills (Steffens, 2006). Furthermore, the significant correlation between GCL and students' interaction with the linked suggested question feature infers that the book supports deep learning.

While our results show that the AI-book may be less resource-consuming than traditional books in supporting meaningful learning, users' expected inputting mechanisms seem to be shaped by experiences with existing internet search tools. Any conflict with an anticipated way of searching may cause dissatisfaction and decreased engagement with the AI-book. Hence, it is important to reduce ECL by improving the design of the functional features of the book environment. Most importantly, this study showed GCL to be positively connected with cognitive strategy use, usability, and with accessing linked suggested questions. This demonstrates that deep learning is related to seeking additional information and curiosity, and that such AI-enriched books are enthusiastically welcomed (e.g., "I hope this is the future of textbooks").

In moving toward future investigations, the results imply that higher education students need support in acquiring sufficient digital literacy (perhaps from a young age) and self-regulated learning skills, to optimally benefit from emerging AI-based tools. The implication that AI-textbook technology can support deeper learning requires expanded investigation including further learning gain measures, physiological measures of cognitive load as well as controlling for pre-knowledge.

Appendix: Protocol for introducing the AI-enriched book (Intelligent LIFE) in the classroom with accompanying explanation and training materials

1. For the first course lecture, we prepared the following materials to introduce Intelligent LIFE to the students:
 - A video giving an overview of the features of Intelligent LIFE. Available at: <https://www.youtube.com/watch?v=I8swXc3WH1M>.
 - A video illustrating how the students could use Intelligent LIFE in reviewing the learning goals provided in the study guides prepared by the instructor. Available at: <https://www.youtube.com/watch?v=V9j53f4Zfjo>.
 - An in-classroom activity lead by the instructor demonstrated Intelligent LIFE. During the activity, in the process of elucidating the concepts that were part of the lecture, the instructor showed how Intelligent LIFE could help students in their understanding. Activity is available at: <https://docs.google.com/document/d/1bBow2rFUf6hJ5Su8EempsttA-id1p18kv0qH6vBPJw/edit>.
 - We enhanced the post-lecture study guide for Lecture 1 with hints on how students could use Intelligent LIFE. Each hint was presented in the context of a specific learning goal and constituted specific questions that could be posed. Available at: [Lecture 1 post-classroom study guide with question asking hints](#).
2. For the second lecture, we produced an in-classroom activity where students were guided in using Intelligent LIFE. This activity focused on protein enzymes and protein regulation, and it encouraged students to compare different forms of inhibition. For homework, we provided a video that demonstrated how Intelligent LIFE could be used to better understand protein structure. The in-class activity, the protein structure video, and enhanced pre and post classroom study guides available at:
 - [Lecture 2 classroom activity on enzymes and protein regulation](#).
 - [Lecture 2 homework supplementary video on understanding protein structure](#).
 - [Lecture 2 pre-class reading guide with hints to use Intelligent LIFE](#).
 - [Lecture 2 post-classroom study guide with hints to use Intelligent LIFE](#).
3. We enhanced pre and post lecture 3 with hints on how students could use Intelligent LIFE. Each hint was presented in the context of a specific learning goal and constituted specific questions that could be posed. Study guides are available at:
 - [Lecture 3 pre-class study guide with hints to use Intelligent LIFE](#).
 - [Lecture 3 post-class study guide with hints to use Intelligent LIFE](#).
4. No hints were added to the study guides for Lectures 4–6 since the instructor wanted students to use Intelligent LIFE independently. We also envisaged that with the hints provided up to lecture 3, the students would have gained the necessary skills to use the new instructional medium.

5. For Lecture 5, we prepared a classroom activity on feedback mechanisms, and glycolysis regulation through which the use of Intelligent LIFE could be reinforced. Activity available at: [Lecture 5 activity with Intelligent LIFE](#).
6. For Lecture 6 we also provided a classroom supplementary video that illustrated how Intelligent LIFE could be used to prepare for an exam. The video focused on how the question answering capability could be leveraged to reinforce what students had learned in the course. Video available at: <https://youtu.be/10nmrJczDu4>.

Abbreviations

AI: Artificial intelligence; CLT: Cognitive load theory; CL: Cognitive load; ICL: Intrinsic cognitive load; ECL: Extraneous cognitive load; GCL: Germane cognitive load; MSLQ: Motivated Strategies for Learning Questionnaire; SQ: Suggested questions.

Acknowledgements

The authors thank the consenting students for their time and participation in this study.

Authors' contributions

All authors developed the research design and implementation of the study methods. MK-J performed the literature review, data analysis, data interpretation and writing of the manuscript. KS contributed to the literature review, data interpretation and writing of the manuscript. CR performed the on-site data collection and commented on the developing manuscript. VC provided access to the AI-book, provided insights to the data collection and commented on the developing manuscript. LT and CH provided insights and commented on the developing manuscript. All authors read and approved the final manuscript.

Authors' information

MK-J received her Ph.D. in instructional psychology from University of Duisburg-Essen, Germany. She is a researcher at Linköping University in Sweden and investigates learning with multimedia, learning with interactive visualizations, and the influence of cognitive load and cognitive and motivational processes on learning.

KS is professor of visual learning and communication at Linköping University and scientific leader of the Swedish national graduate school in science and technology education (FontD). He explores the role of visualization in STEM education and is an elected Member of the Royal Society of Biology.

CR completed her Ph.D. in the Molecular and Cellular Biology department at Harvard University. Since finishing her degree, she has taught in introductory biology courses at Harvard College and Harvard Extension School and became a Summer Institutes Scientific Teaching Mentor. She is also involved in course development at HarvardX, where she makes massive open online courses.

VC received his Ph.D. in Computer Science from University of Toronto, Canada. He is a recognized expert on artificial intelligence, including knowledge representation and reasoning, question answering, ontologies, and knowledge acquisition. He is currently engaged in promoting logic education for secondary schools, investigating techniques for rapidly acquiring formal knowledge and productizing intelligent textbooks.

LT is professor emeritus of visual learning and communication at the division of Media and Information Technology at Linköping University in Norrköping, Sweden. Her research interest is focused on interactive multimodal digital visualizations of abstract bioscience phenomena that are beyond our direct experiences, in both formal and nonformal educational settings.

CH received his Ph.D. from Yale University in 1970 and has been a professor of Biology at Stanford University since 1972. His research has been in thermoregulation, hibernation, sleep, circadian rhythms, and learning disabilities. He has over 250 publications and is an author of the college biology textbook LIFE: The Science of Biology.

Funding

This research was financially supported by the Marcus and Amalia Wallenberg Foundation (Grant MAW 2014.0107).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Science and Technology (ITN), Linköping University, Campus Norrköping, Norrköping 601 74, Sweden.

²Harvard University, 1 Oxford St., Cambridge, MA 02138, USA. ³Department of Computer Science, Stanford University, Stanford, CA 94305, USA. ⁴Biology Department, Stanford University, Stanford, CA 94305-5020, USA.

Received: 12 July 2021 Accepted: 22 December 2021

Published online: 01 March 2022

References

- Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73(3), 277–320.
- Anderson, T. R., & Schönborn, K. J. (2008). Bridging the educational research-teaching practice gap – conceptual understanding, Part 1: The multifaceted nature of expert knowledge. *Biochemistry and Molecular Biology Education*, 36(4), 309–315.
- Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17, 42.
- Bikowski, D., & Casal, E. (2018). Interactive digital textbooks and engagement: A learning strategies framework. *Language Learning & Technology*, 22(1), 119–136.
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194), 4–7.
- Chaudhri, V. K., Cheng, B., Overholtzer, A., Roschelle, J., Spaulding, A., Clark, P., et al. (2013). Inquire biology: A textbook that answers questions. *AI Magazine*, 34(3), 55–72.
- Corbett, A., Kauffman, L., Maclaren, B., Wagner, A., & Jones, E. (2010). A cognitive tutor for genetics problem solving: Learning gains and student modelling. *Journal of Educational Computing Research*, 42(2), 219–239.
- Costley, J., & Lange, C. (2017). The mediating effects of germane cognitive load on the relationship between instructional design and students' future behavioral intention. *Electronic Journal of e-Learning*, 15(2), 174–187.
- de Bruin, A. B., Roelle, J., Carpenter, S. K., & Baars, M. (2020). Synthesizing cognitive load and self-regulation theory: A theoretical framework and research agenda. *Educational Psychology Review*, 32, 1–13.
- DeVore, S., Marshman, E., & Singh, C. (2017). Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics. *Physical Review Physics Education Research*, 13(1), 010127.
- Diefenbach, S., Kolb, N., & Hassenzahl, M. (2014). The 'Hedonic' in human-computer interaction - History, contributions, and future research directions. *Proceedings of the 2014 conference on international journal of human-computer interaction 1155 designing interactive systems—DIS 14* (pp. 305–314). New York: Association for Computing Machinery.
- Eitel, A., Endres, T., & Renkl, A. (2020). Self-management as a bridge between cognitive load and self-regulated learning: The illustrative case of seductive details. *Educational Psychology Review*, 32, 1–15.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publishing.
- Finch, H. (2005). Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology*, 1(1), 27–38.
- Glover, I., Hepplestone, S., Parkin, H. J., Rodger, H., & Irwin, B. (2016). Pedagogy first: Realising technology enhanced learning by focusing on teaching practice. *British Journal of Educational Technology*, 47(5), 993–1002.
- Ibili, E., & Billingham, M. (2019). Assessing the relationship between cognitive load and the usability of a mobile augmented reality tutorial system: A study of gender effects. *International Journal of Assessment Tools in Education*, 6(3), 378–395.
- Kalyuga, S., & Liu, T. C. (2015). Guest editorial: Managing cognitive load in technology-based learning environments. *Journal of Educational Technology & Society*, 18(4), 1–8.
- Kılıç, E. (2007). The bottle neck in multimedia: Cognitive overload. *Gazi University Journal of Gazi Educational Faculty*, 27(2), 1–24.
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, 8, 1997.
- Koć-Januchta, M. M., Schönborn, K. J., Tibell, L. A., Chaudhri, V. K., & Heller, H. C. (2020). Engaging with biology by asking questions: Investigating students' interaction and learning with an artificial intelligence-enriched textbook. *Journal of Educational Computing Research*, 58(6), 1190–1224.
- Lewis, J. R. (2018). Measuring perceived usability: The CSUQ, SUS, and UMUX. *International Journal of Human-Computer Interaction*, 34(12), 1148–1156.
- Linn, M. C., Gerard, L., Ryoo, K., McElhaney, K., Liu, O. L., & Rafferty, A. N. (2014). Computer-guided inquiry to improve science learning. *Science*, 344(6180), 155–156.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service. Retrieved March 18, 2021 from <http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309–326.
- Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers & Education*, 141, 103618.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics—A cognitive-load approach. *Journal of Educational Psychology*, 84, 429–434.
- Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, 32(1/2), 1–8.
- Pantano, E., Rese, A., & Baier, D. (2017). Enhancing the online decision-making process by using augmented reality: A two country comparison of youth markets. *Journal of Retailing and Consumer Services*, 38, 81–95.
- Paris, S. G., & Winograd, P. (2003). *The role of self-regulated learning in contextual teaching: Principles for teacher preparation: A commissioned paper for the US Department of Education*. Office of Educational Research and Improvement (ED).

- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33.
- Sadava, D. E., Hillis, D. M., Heller, H. C., & Hacker, S. D. (2017). *LIFE* (11th ed.). Macmillan Learning.
- Sage, K., Augustine, H., Shand, H., Bakner, K., & Rayne, S. (2019). Reading from print, computer, and tablet: Equivalent learning in the digital age. *Education and Information Technologies*, 24(4), 2477–2502.
- Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, 19(3), 285–307.
- Soenens, B., Sierens, E., Vansteenkiste, M., Dochy, F., & Goossens, L. (2012). Psychologically controlling teaching: Examining outcomes, antecedents, and mediators. *Journal of Educational Psychology*, 104(1), 108.
- Steffens, K. (2006). Self-regulated learning in technology-enhanced learning environments: Lessons of a European peer review. *European Journal of Education*, 41(3–4), 353–379.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138.
- Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, 8(4), 351–362.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296. <https://doi.org/10.1023/A:1022193728205>
- Van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147–177.
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39.
- Zimmerman, B. (2011). *Handbook of self-regulation of learning and performance*. Routledge.
- Zukswert, J. M., Barker, M. K., & McDonnell, L. (2019). Identifying troublesome jargon in biology: Discrepancies between student performance and perceived understanding. *CBE—Life Sciences Education*, 18(1), ar6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
