THE EUROPEAN EDUCATIONAL RESEARCHER
DOI: 10.31757/euer.632

http://www.eu-er.com/

Paper-Based vs. Digitalized Glossaries in Laboratory Scripts

Jenna Koenen
Technical University of Munich, Germany

Lars Mariot
Primo-Levi-Gymnasium, Germany

Rüdiger Tiemann
Humboldt-Universität zu Berlin, Germany

Abstract: In the future, learning will be essentially characterized by the ability to regulate the learning process and monitor success independently from a teacher. The technical possibilities offer better access to learning contents, precise and more individualized feedback, and learning phases adapted precisely to the needs of the learner in terms of scope and pace. In this study, we investigate an important aspect of the digitization of teaching/learning processes using the example of laboratory scripts for chemistry students at university. The focus is on looking up terms and concepts in preparation for the lab internships, firstly in a paper-based glossary and secondly in a digital glossary. During a two-day study, a total of 16 students prepared for experiments on two topics with completely identical materials. We then studied the influence of content knowledge, motivation, and cognitive load. While all students show significant learning achievements, there are no significant differences between the groups. Furthermore, results show that pure digitization of information has no effect, despite the theoretically assumed advantages.

Keywords: *Digitalization; Glossary; Laboratory scripts; Self-regulated learning*

Introduction

Laboratory internships are the central components of chemistry courses. The goals of internships usually lie in cognitive, psychomotor, and affective areas (e.g. Hucke, 2000; Niedderer et al., 2003; Gokhale, 1995). This makes them complex learning environments, characterized e.g. by the integration of knowledge, skills and attitudes, or the transfer of what is already learned to new settings (Seery, Agustian, & Zhang, 2019). In general, learning in chemistry laboratories is based on laboratory scripts describing specific experiments. These primarily address the cognitive area, while at the same time representing instructions for psychomotor implementation. With the transition into digitization, more and more of these scripts are being converted into e-learning formats. This refers to forms of learning that a) use digital media for the representation and distribution of learning materials and/or b) support interpersonal communication (Kerres, 2001). Although there are many theories about the design of learning materials such as the theory of cognitive load (Sweller, Ayres, & Kalyuga, 2011; Sweller, 1988) or the theory of multimedia learning (Mayer, 2009), these are often not considered in the transformation.

Theoretical Background

Cognitive load theory

The cognitive load theory (CLT, Sweller, 1988; Sweller et al., 2011) has preoccupied many scientists for several decades. This theory has yielded important insights into the human learning process. CLT assumes that human memories and experiences, as well as new information, are processed in the working memory and stored in long-term memory. Long-term memory can theoretically store information indefinitely in the form of schemes that serve to interconnect knowledge (Paas & Sweller, 2014; Sweller & Sweller, 2006). However, new information is first processed and systematized in the working memory, which has a limited capacity (Cowan, 2001). Miller (1956) described that only seven plus-minus two units of information can be held simultaneously present in the working memory. As a result of this assumption, a kind of bottleneck effect occurs in the construction of new schemas (Ayres & van Gog, 2009). Pre-existing schemas in long-term memory, however, reduce the load on working memory and can be retrieved with little effort (van Gog, Ericsson, Rikers, & Paas, 2005).

Based on the aforementioned limitations, learning opportunities should be adequately designed. Processing new information causes a cognitive load in the working memory. This load can be divided into three categories. The intrinsic cognitive load: the content of what is to be learned and the pre-existing prior knowledge condition (Kalyuga, 2007; Paas, Renkl, & Sweller, 2003; Paas & Sweller, 2014). Additionally, the learning environment itself causes extrinsic cognitive load by the way it is designed. If the content to be learned is clearly presented and a restriction is only made to the information that is truly necessary, then the cognitive system hardly has to expend cognitive resources to distinguish the truly relevant information from the irrelevant information (Paas et al., 2003; Paas & Sweller, 2014). Since the number, complexity, and interconnectedness of the information of the learning content cause the intrinsic cognitive load, it is usually beyond the teacher's control. The teacher, however, has an opportunity to positively influence the learning process through the presentation of this very information, i.e., the way it is conveyed. This is because the processing of the information also requires resources which are referred to as germane cognitive load. Thus, if the intrinsic cognitive load cannot be changed by the learning content yet the extrinsic cognitive load can be kept small through clever mediation by the teacher, then with an overall limited capacity of working memory, more resources are left for the actual processing of the information - the learning process. Conversely, this means that when the capacity of working memory is exceeded by too much new information, cognitive overload occurs, resulting in a diminished learning effect (Paas et al., 2003). Cognitive load theory thus provides clues and principles for how instruction can be designed and structured to be as effective as possible. Many of these principles have already been studied in more detail, such as the worked example effect, goal-free effect, and modality effect (for an overview Gretsch & Holzäpfel, 2016; Sweller, 2004). Those that are particularly relevant to the present work are briefly outlined below.

- The expertise reversal effect (Sweller, Ayres, Kalyuga, & Chandler, 2003) describes that instructional material designed for novices produces no or even negative learning success when presented to experts for processing.

- The split-attention effect (Kalyuga, 2007; Sweller, 2004; Sweller et al., 1998) describes that tasks with multiple sources of information that are incomprehensible on their own require mental integration before this information can be understood.
- Equally important is the redundancy effect (Chandler & Sweller, 1991; Renkl & Atkinson, 2003), whereby redundant information is counterproductive to learning success.

The cognitive aspects of learning described above strongly influence learner performance. Additionally, motivational aspects contribute to learning success, too. For example, Paas and van Merriënboer (1994) state that influencing cognitive load shows only a small effect when learners are not motivated to process the instructions. A link between an individual's motivation and the cognitive load has not yet been extensively researched. Preliminary results by Paas, Tuovinen, van Merriënboer and Aubteen Darabi (2005) show that experienced learners put less effort into solving tasks designed for novices because of their motivation, and vice versa. Also, it is not the amount of experience on a particular topic that is necessary for the acquisition of expert knowledge, but rather the voluntary willingness to acquire it (van Gog et al., 2005). In conclusion, learning environments that have a motivating effect on the learner which also reduce extrinsic cognitive load, such that there is sufficient capacity in working memory for intrinsic and learning-related cognitive load, are effective (in terms of learning success) and efficient (in terms of learning time) (Schüßler, 2016; van Gog et al., 2005).

Cognitive theory of multimedia learning

For the design of learning materials, Mayer (2009) provides guidance with his cognitive theory of multimedia learning. It combines the cognitive load theory (Sweller, 1988) and the dual coding theory (Paivio, 1986). In general, the cognitive theory of multimedia learning describes the processing of multimedia learning materials, from verbal and visual presentations to the storage of acquired knowledge in long-term memory. New information is processed through a visual and an auditive/verbal channel (Mayer, 2009; Paivio, 1986). A number of principles for multimedia learning can be derived from this approach. For example, the presentation of information in the form of a text and a picture is more conducive to learning than if only one form of presentation is chosen (principle of dual coding). Also, the joint presentation of information as text and picture is more effective if both forms of presentation take place in close spatial proximity (principle of spatial proximity). Likewise, this approach can be used to describe how information that may be interesting and exciting yet unimportant and irrelevant to the learning success - whether presented visually or auditorily - can hinder the learning success (principle of coherence). Thus, a well-intentioned picture or further information can quickly become rather counterproductive when designing instructional materials.

The role of motivation

So far cognitive load theory and the theory of multimedia learning mainly focus on the cognitive aspects of learning. In general, digital learning environments often aim at offering motivating materials in order to foster a deeper learning engagement. Therefore, it can be assumed that motivation plays an important role in the efficacy of digital learning environments as well. The role of motivation in the present study is best described by Heckhausen's "person-situation model" (Heckhausen & Heckhausen, 2006). According to this model, a person has certain needs, motives and goals that lead to an action and a result in a certain situation that provides opportunities and incentives to pursue these needs, motives and goals. The "situation" factor in particular can be influenced significantly by the technical possibilities, since, for example, access to information can be made very low-threshold, or possible incentives can be offered in a very wide variety of ways. On the one hand, emotions accompany motivated, i.e. goal-oriented, behavior. On the other hand, (positive) emotions are themselves the target of motivated behavior. This highlights the close, reciprocal relationship between these two constructs. Although they can be theoretically and empirically separated, they are strongly related to each other. Although the importance of emotions and motivation for the learning process has been known for a long time, considerations of this in learning processes in the laboratory have so far received little attention. Galloway, Malakpa and Bretz (2016) are able to show that here, too, affective experiences influence cognitive and psychomotor performance in the laboratory. In this context, Weiner's (1986) attribution model shows the influence of attributing an action to oneself or to another person. Due to the digital possibilities, technology-supported learning environments offer opportunities to strengthen self-efficacy (Deci & Ryan, 1993) and to positively influence emotions and motivation.

Glossaries as aids to self-regulated learning

In classrooms, the teacher often still intervenes in the learning process in a regulating way, for example by monitoring the learning progress and selecting appropriate materials or applying strategies and giving feedback. But situations increasingly arise in which a learner has to regulate his learning process by himself. Planning, monitoring, and evaluating the learning process are key skills to master this regulation (Flavell, 1979). If, for example, a learner notices in their own learning process that their prior knowledge is insufficient to access the content to be learned, they must be able to look for help in order to bridge this difference between prior knowledge and the subject matter (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003). Aleven et al. (2003) describe that such help can take the form of cues, feedback, opportunities for reflection, or through further information. Wallen, Plass, & Brünken (2005) use the example of text comprehension to show the effectiveness of appropriate, additional notes and explanations. Specialized books often have such aids in the form of glossaries which are intended to enable the reader to better understand a text, if necessary. Such glossaries contain explanations, definitions, or formulas of terms from a text. Learners identify technical terms (Göpferich, 1998) in a text and refer to the glossary. With the help of such glossaries, readers can be supported in their learning process. Gaps in concepts and principles can be filled and misunderstandings can be corrected if necessary.

However, the use of a glossary can also lead to divided attention, since the technical term being searched for must first be found in the glossary. During the search process, other technical terms are briefly considered and other information irrelevant to the actual comprehension process come into focus, at least briefly. Attention shifts away from the content of the script to searching for the technical term in the glossary. A split-attention effect may occur (Kalyuga, 2007; Sweller et al., 1998; Sweller, 2004). This effect increases cognitive load and should be avoided in terms of good instructional design. Therefore, to reduce this divided attention, an e-glossary (electronic glossary - digital version of a glossary for a text) was created in the present work. Technical terms in the laboratory script are directly linked to the respective entry in the e-glossary. This reduces the search time needed to find the technical terms and again the connecting point in the script. In addition, the principle of proximity can be applied here, since the information in the glossary appears on the same screen.

Furthermore, while looking up entries, technical terms in the glossary may hold information already known to participants, resulting in increased extrinsic load and a redundancy effect (Chandler & Sweller, 1991; Renkl & Atkinson, 2003). With the e-glossary, this is avoided by direct linking. Within the e-glossary, it is not necessary to scroll through all the technical terms, which may increase intrinsic load.

Summary

The fundamental description of learning processes in laboratories as processes in complex learning environments leads on the one hand to the question of how the cognitive load can be addressed in such a situation. On the other hand, affective components also play an essential role, since they also influence the students' performance. Both strands can be combined in the approach of multimedia learning and their respective characteristics can be well taken into account, especially by adding components of self-regulated learning to increase self-efficacy and thus motivation.

We have implemented these considerations in the digital glossaries as learning aids. Glossaries for technical terms and thus tools for content knowledge because we have placed the focus in this project on content knowledge as a key indicator for successful learning in almost all learning environments - including those of the laboratory.

Research Questions

During the digitization process, many learning materials that were previously paper-based are transferred into a digitized format. However, the advantages and disadvantages in terms of learning achievements, motivation, and cognitive load are seldom discussed within authentic laboratory settings in higher education. In addition, the differences that result from the transformation into the digitized format have not yet been analyzed in such a specific setting, neither from a theoretical nor an empirical point of view. Therefore, this study aims at discussing these differences between laboratory scripts that are paper-based with the paper-based glossary on one side and digitized laboratory scripts including the e-glossary on the other. Content knowledge, acquired through instruction strategies has been identified in many studies as an essential prerequisite for learning how science "works" and thus the declared main goal of laboratory work (Seery, Agustian & Zhang, 2016; Zhan, Kirschner, Cobern & Sweller, 2021). The strategies for acquiring content knowledge are closely linked to the Cognitive Load Theory, since the complexity of the learning situation can be reduced via the intrinsic load. Thus, content knowledge and the Cognitive Load Theory are key factors and in the focus of our research questions.

Hence, the following questions guided the research for this study:

a) To what extent does learning with two different formats of learning materials - paper-based laboratory scripts and digitalized scripts - affect the students' content knowledge?

b) To what extent do motivation, cognitive load and learning time differ from each other when students are learning with paper-based or digital laboratory scripts?

c) To what extent does the use of glossaries by the students vary between the different versions of the script?

Methodology

Study Design

To answer the questions raised, an intervention study in a pre-post design within a lab setting was carried out. The study included four intervention groups that were based on two different formats of the laboratory scripts which were a) paper-based and b) digital. Since the study was based on a physical chemistry lab course, two different topics from the course were chosen. The two different topics were: 1) Determination of the temperature independence, the heat capacity, as well as the entropy and enthalpy of a solid and 2) Spectroscopic determinations of dipole moments in the electronic state using the example of Coumarin 7. Each laboratory script consisted of a detailed description of the experiment including a task, a description of the relevant prior knowledge, a description of the learning goals, content specific knowledge descriptions, the detailed descriptions of the execution of the experiments, as well as a glossary. The laboratory scripts' digitalized version differs from the paper-based version only in the way that the e-glossary could directly be assessed via hyperlinks within the descriptions. That means that the students only needed to click on the unknown expression, for example Coumarin, and were directly transferred to the glossary entry describing the molecular structure. In contrast, the paper-based formats contained a glossary at the very end of the script, so they needed to thumb through the paper to find the relevant expression. However, in terms of content, both formats are completely identical; only the implementation in the respective medium varied.

Both laboratory scripts were implemented in the two different formats in a rotated design, which resulted in four different intervention groups (see Table 1). The study took place over two consecutive days, ensuring that every student could invest the same mental effort to study each laboratory script. Every student worked individually on the two different topics and received one script in the paper-based format and the other one in the digitalized format. The order of the two different formats as well as the topics were counterbalanced between the groups to reduce sequence effects (Dunn, 2013). Learning took place individually and independed from the lab course itself. The study was designed as a laboratory study. Meaning, the number of limiting variables was reduced as much as possible. For example there was no interaction between the students or feedback from the instructor. Therefore, it can be assumed that possible effects can primarily be attributed to the different modes of instruction, meaning paper-based or digital.

Table 1

Study Design

	Group **1**	Group **2**	Group **3**	Group **4**
Day 1	Script 1 **paper-based**	Script 2 **paper-based**	Script 1 **digital**	Script 2 **digital**
Day 2	Script 2 **digital**	Script 1 **digital**	Script 2 **paper-based**	Script 1 **paper-based**

Participants

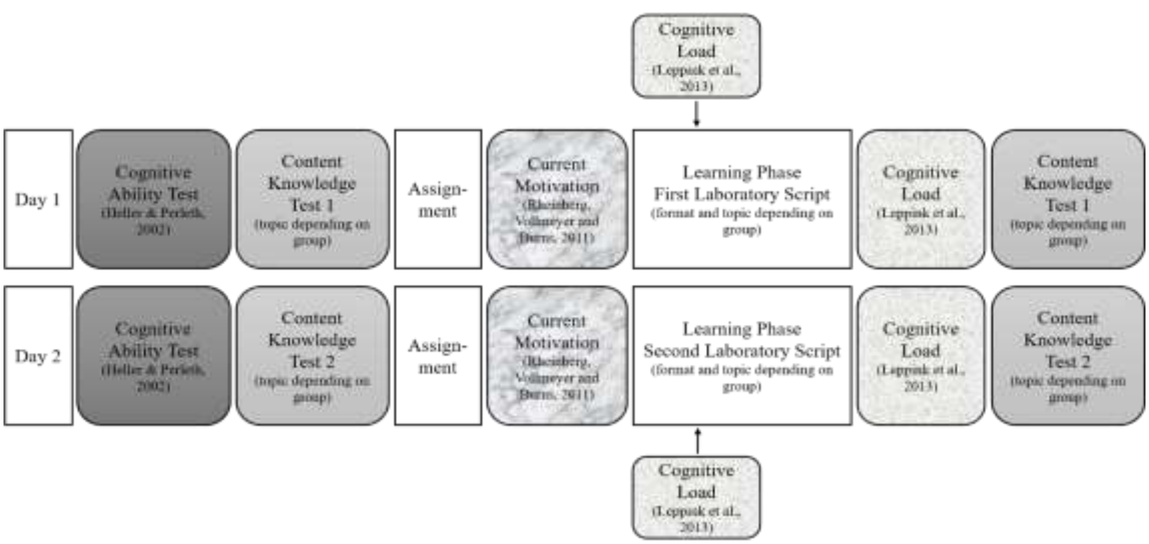
The study took place in a German University in Berlin. A total of 16 students participated in this study. All of them were preservice teachers studying chemistry as their first subject. Their age average was 26 years (SD = 3.48) and 62.2% were female. They were randomly assigned to their intervention group. Every student worked on both laboratory scripts which were implemented in the two different formats. This procedure ensured that work with the laboratory scripts was independent of the students' specific characteristics like cognitive ability.

Data Collection and Instruments

To investigate the effects of the two different formats of the implemented laboratory scripts on students' learning achievements - in terms of content knowledge as well as cognitive load and motivation - different measurement tools were implemented at different points in time (Figure 1). Within this study the students only studied with the laboratory scripts. They did not conduct the experiments themselves within the laboratory. Prior to learning with the first laboratory script, the cognitive ability test was administered.

Figure 1

Administered Test Instruments and Test Schedule



The cognitive abilities were measured with the KFT test's N2 scale used by Heller and Perleth (2002). Within this scale, the students had to complete 25 figure analogies within eight minutes. Only one scale was administered, because if the person shows high cognitive abilities in one domain it can be assumed that the person is also able to use these cognitive abilities within another domain (Cattel, 1987). In addition, domain-independent cognitive abilities are measured with these kinds of figure analogies (Baumert, Lüdtke, Trautwein & Brunner, 2009).

To assess the students' content knowledge, two specific content knowledge tests based on the content of the laboratory scripts were developed. Each content knowledge test referred to one script and consisted of 10 items. Each test was directly implemented before and after the students studied the respective laboratory script. Six of them were multiple-choice single select items and four items were open-ended questions. The multiple-choice items referred to basic content knowledge for the experiments, whereas the open-ended questions referred to laboratory relevant methods for conducting the experiments. The answers were evaluated and rated based on a sample solution. This ensured the objectivity of the questionnaire. Every item was worth one point, meaning 10 points in total could be achieved per content knowledge test. The reliability of the developed content knowledge questionnaires was checked for the pre- as well as the post-measurement point of time. Cronbach's Alpha values for Topic 1 are $\alpha = .536$ for the pre-test and $\alpha = .298$ for the post-test. For Topic 2, α-values of $\alpha = .533$ for the pre-test and $\alpha = .632$ for the post-test were obtained. Except for the value for the post measurement point of time for the test of Topic 1 the reliability is acceptable (Nunnally, 1967). In general, one reason for these values could be the small number of items as well as participants. However, one needs to hold in mind that due to this outlier the expressiveness of the results is limited.

Before learning with every script, but after being given the assignment, the current motivation of the participants was ascertained. The instrument used was adopted from that of Rheinberg, Vollmeyer, and Burns (2001). It is based on the four subscales: interest (2 items), probability of success (4 items), fear of failure (5 items), and challenge (4 items), which were also considered separately in the evaluation. Every item needed to be answered on a 7-point Likert scale ranging from 1 meaning "statement does not apply" to 7 "statement does apply". Before learning with the two laboratory scripts the students' motivation was measured. The instrument shows acceptable reliability for those topics (Topic 1: $\alpha = .465$; Topic 2: $\alpha = .468$).

After the assessment of the current motivation, the students started to learn with laboratory scripts. Seven minutes after the start of the learning phase and directly after the end of the learning phase, the cognitive load was measured. To do so, the survey of Leppink, Paas, van der Vleuten, van Gog and van Merriënboer (2013) was used. It consists of three different scales: intrinsic cognitive load (3 items), extrinsic cognitive load (3 items) and germane cognitive load (4 items). All items need to be answered on an 11-point Likert scale ranging from 0 meaning "statement does not apply" to 11 meaning "statement does apply". Each scale was complemented with a 9-point Likert item, where the answer format is based on phrases. These items are from instruments developed by: Ayres (2006) for the intrinsic load, Cierniek, Scheiter and Gerjets (2009) for the extraneous load and Salomon (1984) for the germane load. This procedure was necessary to ensure that sufficient information about the cognitive load of the students could be generated. The reliability for the three scales as well as the two points of measurement and the two topics was checked. The Cronbach's α values all lie in an acceptable to good range (see Table 2).

Table 2

Reliability Measures of the Cognitive Load Scales

	Topic 1		Topic 2	
	Measurement 1	Measurement 2	Measurement 1	Measurement 2
Intrinsic load	.696	.737	.748	.661
Extraneous load	.859	.855	.435	.725
Germane load	.826	.698	.520	.681

These two measurement points allow one to observe changes within the cognitive load during learning and the retrospective summative assessment after the end of the learning phase. This procedure allows to observe variations within the cognitive load (van Gog, Kirschner, Kester & Paas, 2012). In addition, the procedure seems to be necessary as the summative assessments at the end of the learning phase cannot be considered as a spontaneous assessment, because it is based on information already stored in the long-term memory (Schmeck, Opfermann, van Gog, Paas & Leutner, 2015, van Gog et al., 2012).

In addition, the students were filmed while studying. The focus was either on the computer screen or the paper-based materials. The films allow for counting the number of looked up technical terms in the glossary. Furthermore, the learning time was monitored by letting the students write down the start and the end time of their learning phase.

Data Analysis

This study looks at the impact of the two different formats of the implemented laboratory scripts on the assessment scores of the students' content knowledge and their motivation and cognitive load. Therefore, the students' scores were evaluated between the different intervention groups and the implemented formats. To do so, mean values for every scale were calculated. This was true for the content knowledge tests as well as the Likert scale based questionnaires. The calculation of the mean and the standard deviation for the Likert scales assumes that the distances are equal between the points of the scale and that the students perceive this in the described way (Bühner, 2011). For the cognitive load measures, each scale contained several 11-point Likert items as well as one 9-point Likert item, because they were retrieved from different literature. To combine those different items to one scale every subscale is normalized to one. Then the scales are combined. This procedure resulted in a standardized and comparable value for each cognitive load scale.

To compare the different groups and different formats with each other comparative data analysis procedures were implemented. Due to the small sample size, normal distribution was not given for all implemented measures. Therefore, non-parametric analysis procedures, like the Mann-Whitney U test and Kruskal-Wallis tests, were implemented to compare different groups with each other. The Wilcoxon test for depended samples was conducted to calculate significant differences from pre to post. This test is based on the group rank differences and is significantly more robust compared to small samples (Field, 2013). The whole analysis was done using IBM SPSS Statistics Version 28 (IBM, 2021).

Results

Effects of Cognitive Abilities

Before having a specific look at the learning achievement and the effects on motivation and cognitive load, it is necessary to have a specific look at the assets control variable. In this case, it is the cognitive abilities of the students. Overall, the test reveals no significant differences between the students' cognitive abilities ($H(3) = 3.624$, $p = .305$) and, therefore, they don't have any impact on the results regarding content knowledge, cognitive load, and motivation.

Effects on Content Knowledge

In this study learning success is defined as the difference for one topic between the post-intervention measurement and the pre-intervention measurement regarding the test score in the content knowledge test. Therefore, for every student the learning gain is calculated first. The results indicate that there is a significant learning gain from the pre- to the post-measurement time in both topics (Topic 1: $T = 120.00$, $p < .001$, $r = .215$; Topic 2: $T = 91.00$, $p = .001$, $r = .201$) (see Figure 2 and 3).

Figure 2

Learning Gains in the Content Knowledge Test for Topic 1

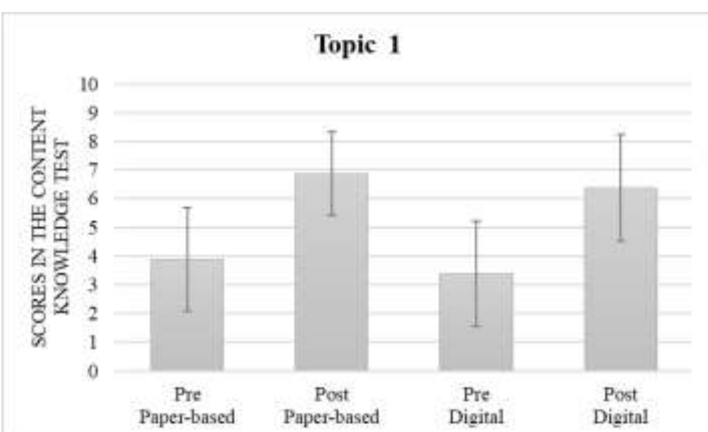
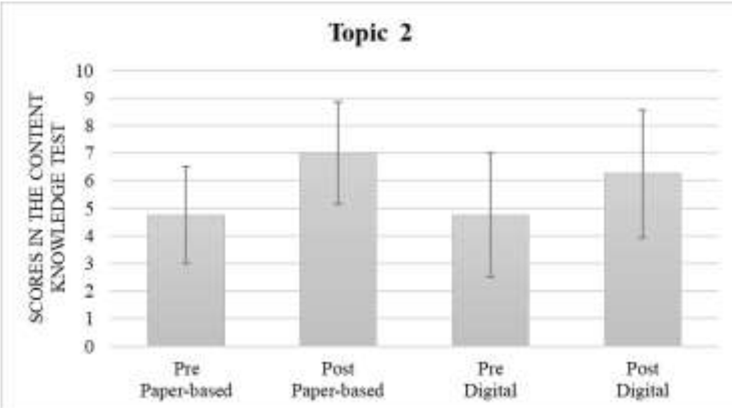


Figure 3

Learning Gains in the Content Knowledge Test for Topic 2



In addition, we found no differences between the four intervention groups (Topic 1: $H(3) = 1.394$, $p = .707$; Topic 2: $H(3) = 2.320$, $p = .509$) and the two formats of presentation. Likewise, we didn't find any sequencing effects. The Kruskal-Wallis test with the position of the format (paper-based – digital (Group 1, 2) vs. digital – paper-based (Group 3, 4)) as the independent variable did not reveal any significant differences regarding the increase in content knowledge for both topics (Topic 1: $H(1) = .192$, $p = .661$; Topic 2: $H(1) = .840$, $p = .359$). The Kruskal-Wallis test with the topic (Topic 1 – Topic 2 (Group 1, 3) vs. Topic 2 – Topic 1 (Group 2, 4)) as the independent variable did likewise not reveal any significant differences regarding the development of the content knowledge (Topic 1: $H(1) = .769$, $p = .380$; Topic 2: $H(1) = .073$, $p = .787$). The learning achievement seems therefore to be independent of the order of the topic as well as of the format. Hence, in the following analyses the two formats - paper-based and digital laboratory scripts - were directly compared to each other. In terms of design, this means that Groups 1 and 4 and Groups 2 and 3 were combined.

To see whether the learning achievement depends on the format in which the laboratory script is presented, a Mann-Whitney-U test was conducted for each topic. Results indicate that in the sample no significant differences regarding the formative presentation occurred (Topic 1: $U = 38.00$, $z = .658$, $p = .511$; Topic 2: $U = 21.00$, $z = -1.186$, $p = .236$).

Effects of Motivation

In the following, we only differentiate between the two formats, paper-based or digital, and show the results for all four subscales. To check for differences between the two formats for each subscale a Mann-Whitney-U test was conducted. For the subscales interests (Topic 1: $U = 24.50$, $z = -.803$, $p = .422$; Topic 2: $U = 21.50$, $z = -1.118$, $p = .264$), probability of success (Topic 1: $U = 30.00$, $z = -.212$, $p = .832$; Topic 2: $U = 35.00$, $z = .316$, $p = .752$) and challenge (Topic 1: $U = 29.00$, $z = -.317$, $p = .751$; Topic 2: $U = 30.50$, $z = -.159$, $p = .874$) no significant differences between the two formats were found.

However, significant differences regarding the laboratory script for Topic 2 were found in the fear of failure subscale. The Mann-Whitney-U test revealed that the fear of failure is significantly higher in the group working with the digital laboratory script (Mdn = 4.00) compared to the group working with the paper-based script (Mdn = 1.70) version (Topic 2: $U = 57.00$, $z = 2.632$, $p = .009$, $r = .658$). Unfortunately, these results could not be confirmed for Topic 1 (Topic 1: $U = 22.00$, $z = -1.059$, $p = .290$).

Cognitive Load while Learning

Cognitive load is considered to be very central during learning processes. Therefore, it is interesting to see whether the two different formats of implementation, paper-based and digital, have an impact on the students' perceived cognitive load.

Yet, similar to the results regarding the students' motivation the cognitive load during the learning process (measurement point 1) does not seem to vary much between the two implemented formats. No significant differences were found for the intrinsic (Topic 1: $U = 38.00$, $z = .633$, $p = .527$; Topic 2: $U = 43.00$, $z = 1.155$, $p = .248$) and the germane load (Topic 1: $U = 30.50$, $z = -.158$, $p = .874$; Topic 2: $U = 22.50$, $z = -1.002$, $p = .316$). However, for the extraneous load the data indicates that there is a significant difference between the two formats. Students studying with the laboratory script of Topic 1 in the paper-based format seem to have a significantly higher extraneous cognitive load while learning than the students studying with the digital laboratory script of Topic 1 (Topic 1: $U = 11.00$, $z = -2.207$, $p = .027$, $r = -.553$). However, this result was not confirmed by the data of the laboratory scripts for Topic 2 ($U = 43.50$, $z = 1.210$, $p = .226$).

In addition, the cognitive load was also measured right after the end of the intervention phase. Now the students rated the cognitive load retrospectively to their learning process. For this point in time for none of the three forms of cognitive load significant differences were found (intrinsic load: Topic 1: $U = 26.50$, $z = -.578$, $p = .563$; Topic 2: $U = 38.00$, $z = .630$, $p = .529$; extraneous load: Topic 1: $U = 16.00$, $z = -1.684$, $p = .092$; Topic 2: $U = 26.50$, $z = -.578$, $p = .563$; germane load: Topic 1: $U = 42.00$, $z = 1.052$, $p = .293$; Topic 2: $U = 36.00$, $z = .421$, $p = .674$).

Effects of Learning Time

Learning time was assessed while students were studying with the laboratory scripts. For Topic 1, students studied on average 16.4 ($SD = 3.46$) minutes with the paper-based scripts and 15.9 ($SD = 5.46$) minutes with the digitized version. For Topic 2 the study time was longer, but did not vary much between the two formats (paper-based: 22.4 ($SD = 3.85$) minutes; digital: 20.9 ($SD = 5.06$) minutes). In general, these differences in learning time are not significant as Mann-Whitney U tests show (Topic 1: $U = 24.00$, $z = -.850$, $p = .396$; Topic 2: $U = 25.50$, $z = -.687$, $p = .492$).

Use of the glossary

In both formats, the laboratory scripts were supplemented by a glossary including the central technical and content-relevant terms. Due to the digital script's easier accessibility, it was assumed that more technical terms would be looked up in the digitized version. However, this assumption cannot be confirmed. In the paper-based format an average of 8.62 terms (*SD* = 6.28) was looked up and in the digital form an average of 8.75 terms (*SD* = 4.45). Which terms were looked up and how many of them seems to be highly individual, as indicated by the high standard deviations.

In addition, the number of terms that were looked up does not vary between the two topics (Topic 1: 8.38 terms, Topic 2: 9.00 terms). On a qualitative level it can be noted that the central technical terms of each topic are looked up more often than less relevant terms. This indicates that the students are able to identify more relevant terms while learning.

Summary and Discussion

Summary of findings

The study investigated the differences in learning achievement, motivation, and cognitive load for university students studying with two different versions of laboratory scripts. A paper-based version of the script including a glossary was compared to the digitized format of the script including an e-glossary. In general, the content was comparable, only the accessibility of the glossary varied between the two formats. Results indicate no differences with regard to learning achievement. Learning took place independently of the implemented formats. In addition, no differences in the use of the glossary could be observed. With regard to motivation and cognitive load only small differences were found. Students rated the fear of failure significantly higher in the digitized format than in the paper-based format for Topic 2. In addition, they described the extrinsic cognitive load during learning in the paper-based format significantly higher than in the digitized format for Topic 1. Besides that, no other significant differences were identified.

Discussion of findings

Overall, the results show that there seem to be no systematic differences between the digitized and the paper-based format for both topics. The reported higher extrinsic cognitive load is in line with the underlying cognitive load theory. This is because looking up technical terms in the paper-based glossary might actually cause a higher cognitive load, since students need to scroll through the glossary in an alphabetical order to find the relevant technical term. By doing this, they need to at least partly process information that is actually not relevant for the learning process itself. This can lead to a split-attention effect, resulting in a higher extrinsic load which might hinder learning. However, even though a higher extrinsic cognitive load is reported the learning achievement is not minimized, which indicates that even though this effect might occur it is not big enough to actually hinder learning. In addition, one needs to recognize that even though the result seems to be in line with the cognitive load theory, the theory itself can also be discussed critically. De Jong (2010) identified a number of problematic issues within the framework of cognitive load theory, which are from a conceptual, methodological and application-related nature. This is for example on a conceptual level the question of the distinction between intrinsic, extraneous and germane cognitive load or the question, if these types of formative note can simply be added (for details see de Jong, 2010). In addition, it needs to be recognized that the fundamentals of the theory are based on instructional design settings (e.g. Mayer & Moreno, 2003; Sweller et al., 1998). In literature, it is often presupposed that these considerations also hold true for more realistic teaching and learning situations, as we did in this article as well. Some of these issues have been addressed within the literature of the last 10 years (e.g. Choi, van Merriënboer, & Paas, 2014, Skulmowski & Xu, 2022). Also considering these boundaries and the development of the cognitive load theory in the last years, it is still one of the major theories used in educational research literature.

In addition, the finding only occurs for one laboratory script and not for the other as well. This might also be an indication that, at least for university students in this setting, the difference between the paper-based glossary and an e-glossary in terms of different cognitive conditions might not be big enough. So, they actually can cope with the little extra extrinsic load within the paper-based version that one would assume, based on the theory.

For one topic, students report a higher fear of failure, i. e. when studying with the digitized laboratory script including the e-glossary. This might indicate that unfamiliar learning formats may cause stress and, therefore, may hinder learning. One could argue that the digitized laboratory script with the e-glossary is an unfamiliar learning format for the students because, so far, they had only studied with the paper-based versions. However, this finding could not be replicated for the other topic. Therefore, the study cannot indicate whether this is a random finding or not. In addition, this does not seem to have an effect on the actual learning achievement, which somehow seems to indicate that students are able to deal with this kind of fear of failure as well, especially in settings where no negative effects occur when failing, like in this study.

So far, the results were interpreted from the cognitive load perspective. However, they should also be noted against the background of the current debate about digitalization, which is increasing due to the pandemic situation. In the last two years, more and more paper-based materials, especially in schools and universities, are simply transferred to a digital format in a one-to-one mode. Most of the time this does not go together with a change in the instructional design. Following the SAMR-Modell (Substitution, Augmentation, Modification, Redefinition; Puentedura, 2006), activities at schools and universities seems to stay mainly on the level of substitution or augmentation. Considering the results from the study, they seem to indicate that this one-to-one transformation, that we also performed here with only minor changes within the glossary based on the possibility of hyperlinks, neither hinders nor improves learning. For university students, no differences in terms of learning achievement were observed. Of course, this could be different for other kinds of students, other kinds of learning materials, or other settings. However, it does not seem to be the case that just changing the mode of presentation already results in a better learning opportunity for students. Probably one should think more about considering the instructional design in general and not just the mode of presentation (digital vs. paper-based). Therefore, these results indirectly confirm the statement made by Reeves, Herrington and Oliver (2004), according to which different media also require different instructional designs. Ultimately, the digitalization allows for many other types of implementation that might increase learning achievement.

Therefore, Bowen, Lack, Chingos and Nygren (2012) state that online learning may result in equivalent, if not better, learning outcomes. However, the online context does present unique obstacles for learning (Dunn, 2014). The technical possibilities offer better access to learning contents, precise and more individualized feedback, and learning phases adapted precisely to the needs of the learner in terms of scope and pace. If this is consistently implemented, then in the future learning will also be essentially characterized by the ability of students to regulate and monitor their learning processes successfully and independently from the teacher. This seems to be especially important, because of the growing diversity of learning opportunities and environments, also due to the increased use of digital tools, in the future. In addition, learning in the future will be more and more individualized due to the available adaptive learning settings. To be able to deal with all the different possibilities for learning it is even more important than today to be able to be self-responsible for the own learning process, to monitor it and to be able to adapt to the individual needs to achieve a good learning outcome. Therefore, it can be expected that self-regulation will also be important in the future. This is especially true for complex learning environments like chemistry laboratories as well as the underlying materials that support those. In this study, the ability to self-regulate the learning process was not monitored based on the assumption that students at the end of their studies at the university level have such high abilities in self-regulated learning that these abilities don’t really matter anymore for learning success, especially in such a short intervention phase. Therefore, in this case it is not possible to comment on the influence of self-regulated learning. However, in the future it will be valuable to investigate this as well, especially for younger students that do not have such self-regulated learning abilities as it is the case here.

Limitations

One of the major limitations of this study is the sample size. With 16 students in total divided into four groups, statistical analysis can of course be considered critical. However, due to the possibility to merge the four groups based on the different intervention methods, the group size increases, but still remains small. Nevertheless, the non-parametric analysis methods performed yield validity (Field, 2013). In addition, a descriptive comparison leads to similar impressions on the basis of a case study (Robson, 2002). Although the small sample size is problematic, it also results from regulations regarding the inclusion of participants that help strengthen the research design. In addition, the highly controlled setting ensures that the results can be mainly attributed to the different mode of instruction.

Besides the small sample size, the intervention itself was very short. Participants only studied two times for approximately 20 minutes with the laboratory scripts. It is also possible that due to the short intervention period, effects regarding learning achievement, motivation, and differences in cognitive load as they can be assumed from theory are just not visible, because differences might be so small that they do not play an important role within the short learning periods. In addition, the reliability of the newly developed content knowledge tests is relatively low. Therefore, for future research the tests will need revision to ensure their reliability.

Although results indicate that the different implementation formats - paper-based versus digitized laboratory scripts - complemented by a glossary or an e-glossary, in general do not have a significant effect on learning achievement, motivation and cognitive load. The limitations put into question the extent to which the results are applicable to another population that is larger or characterized by other students' specific measures, like content knowledge or proficiency with self-regulated learning. All these kinds of characteristics can have an impact on the learning behavior and therefore the results of the study. However, due to limitations in the study time these characteristics were not controlled for.

Therefore, further studies should definitely focus on clarifying the effects of paper-based laboratory scripts with a glossary in comparison to the digitized format with the e-glossary on learning achievement, motivation and cognitive load. This is also important, because preparing students for learning in the laboratory is beneficial for the learning success in the complex environment (Seery, Agustian, & Zhang, 2019) and laboratory scripts are the typical instrument to do so. Further research needs to focus on longer intervention periods, including other important student characteristics and possibly also other kinds of learning achievement measures, for example, the performances in the lab course itself, which was not part of the study.

Acknowledgements

In particular, we thank Wolfgang Christen from Humboldt-Universität zu Berlin for allowing us to use two of his laboratory scripts for the study. We also thank the participating students, because without them the study would not have been possible.

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. https://doi.org/10.3102/00346543073003277

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *16*(5), 389–400.

Ayres, P., & van Gog, T. (2009). State of the art research into Cognitive Load Theory. *Computers in Human Behavior*, *25*(2), 253–257.

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*(3), 165–176. https://doi.org/10.1016/j.edurev.2009.04.002

Bowen, W., Lack, K., Chingos, M,. & Nygren, T. (2015). Interactive Learning Online at Public Universities. New York. https://doi.org/10.18665/sr.22464

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion [Introduction to test and questionnaire construction]*. PS Psychologie. München: Pearson Studium. Retrieved from http://ebooks.ciando.com/book/index.cfm/bok_id/203986;B:CIANDO

Cattell, R. B. (1987). *Intelligence: Its Structure, Growth and Action*. Advances in psychology: Vol. 35. s.l.: Elsevier textbooks. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10259213

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

Choi, H.-H., van Merriënboer, J. J. G. & Paas, F. (2014). Effects of the Physical Environment on Cognitive Load and Learning: Towards a New Model of Cognitive Load. *Educational Psychology Review, 26*(2), 225-244. https://doi.org/10.1007/s10648-014-9262-6

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?. *Computers in Human Behavior*, *25*(2), 315–324.

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

Deci, E.& Ryan, R. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik [The self-determination theory of motivation and ist importance for pedagogy]. Zeitschrift für Pädagogik, *39* (2), 223-238.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38*(2), 105-134. https://doi.org/10.1007/s11251-009-9110-0

Dunn, D. S. (2013). *Research Methods for Social Psychology*. Hoboken: Wiley.

Dunn, K. (2014). Why Wait? The Influence of Academic Self-Regulation, Intrinsic Motivation, and Statistics Anxiety on Procrastination in Online Statistics. *Innovative Higher Education, 39*(1), 33-44. https://doi.org/10.1007/s10755-013-9256-1

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. and sex and drugs and rock 'n' roll. Los Angeles, London, New Dehli, Singapore, Washington DC: SAGE Publications.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-9011. https://doi.org/10.1037/0003-066X.34.10.906

Galloway, K. R., Malakpa, Z., & Bretz, S. L. (2016). Investigating affective experiences in the undergraduate chemistry laboratory: students` perception of control and resposibility. *Journal of chemical education, 39*, 227-238.

Gokhale, A. A. (1995). Collaborative Learning Enhances Critical Thinking. *Journal of Technology Education, 7*(1), 22–30.

Göpferich, S. (1998). *Interkulturelles Technical Writing: Fachliches adressatengerecht vermitteln; ein Lehr- und Arbeitsbuch [Intercultural technical writing: conveying technical information in a way that is appropriate to the target audience; a textbook and workbook]. Forum für Fachsprachen-Forschung: Vol. 40*. Tübingen: Narr.

Gretsch, P., & Holzäpfel, L. (Eds.). (2016). *Lernen mit Visualisierungen: Erkenntnisse aus der Forschung und deren Implikationen für die Fachdidaktik [Learning with visualizations: Findings from research and their implications for subject didactics]*. Münster, New York: Waxmann Verlag. Retrieved from http://ebooks.ciando.com/book/index.cfm/bok_id/2180772

Heckhausen, J., & Heckhausen, H. (Eds.). (2006). *Motivation und handeln [Motivation and action]*. Springer-Verlag

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeits-Test (Rev.) für 5.-12. Klasse (KFT 5-12+R) [Cognitive Ability Test (Rev.) for grades 5-12 Class (KFT 5-12+R)]*. Göttingen: Beltz-Testgesellschaft.

Hucke, L. (2000). *Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums [Action regulation and knowledge acquisition in traditional and computer-supported experiments in physics internships]*. Berlin: Logos.

IBM Corp. Released 2021. *IBM SPSS Statistics for Windows, Version 28.0*. Armonk, NY: IBM Corp

Kalyuga, S. (2007). Enhancing Instructional Efficiency of Interactive E-learning Environments: A Cognitive Load Perspective. *Educational Psychology Review*, *19*(3), 387–399. https://doi.org/10.1007/s10648-007-9051-6

Kerres, M. (2001). *Multimediale und telemediale Lernumgebungen: Konzeption und Entwicklung [Multimedia and telemedia learning environments: conception and development]*. München, Wien: Oldenbourg.

Leppink, J., Paas, F., van der Vleuten, C. P. M., van Gog, T., & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072.

Mayer, R. E. (2009). *Multimedia learning*. New York NY u.a.: Cambridge Univ. Press.

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.

Niedderer, H., Tiberghien, A., Haller, K., Hucke, L., Sander, F. & Fischer, H. (2003). Talking Physics in Labwork Contexts. A Category Based Analysis of Videotapes. In D. Psillos & H. Niedderer (Ed.), *Teaching and Learning in the Science Laboratory* (Vol. 16, pp. 31–40). Dordrecht: Kluwer Academic Publishers.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Paas, F., & Sweller, J. (2014). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Ed.), *Cambridge handbooks in psychology. The Cambridge handbook of multimedia learning* (pp. 27–42). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.004

Paas, F., & van Merriënboer, J. (1994). 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.

Paas, F., Tuovinen, J. E., van Merriënboer, J., & 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

Paivio, A. (1986). *Mental representations: A dual coding approach*. *Oxford psychology series: Vol. 9*. New York NY u.a.: Oxford Univ. Press.

Puentedura, R. (2006). Transformation, technology, and education [Blog post]. Retrieved from http://hippasus.com/resources/tte/.

Reeves, T. C., Herrington, J., & Oliver, R. (2004). A development research agenda for online collaborative learning. *Educational Technology Research and Development, 52*(4), 53-65.

Renkl, A., & Atkinson, R. K. (2003). Structuring the Transition From Example Study to Problem Solving in Cognitive Skill Acquisition: A Cognitive Load Perspective. *Educational Psychologist*, *38*(1), 15–22.

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen [FAM: A questionnaire to record current motivation in learning and performance situations]. *Diagnostica, 48,* 57–66.

Robson, C. (2002). *Real world research*. Oxford: Blackwell Publishing.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*(4), 647–658.

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: Differences between immediate and delayed ratings. *Instructional Science*, *43*(1), 93–114. https://doi.org/10.1007/s11251-014-9328-3

Schüßler, K. (2016). *Lernen mit Lösungsbeispielen im Chemieunterricht. Einflüsse auf Lernerfolg, kognitive Belastung und Motivation [Learning with solution examples in chemistry lessons. Influences on learning success, cognitive load and motivation]*. Universität Duisburg-Essen, Duisburg.

Seery, M. K., Agustian, H. Y. & Zhang, X. (2019). A framework for learning in the chemistry laboratory. *Israel Journal of Chemistry*, *59,* 546-553. https://doi.org/10.1002/ijch.201800093

Skulmowski, A. & Xu, K. M. (2022). Understanding Cognitive Load in Digital and Online Learning: a New Perspective on Extraneous Cognitive Load. *Educational Psychology Review, 34*(1), 171-196. https://doi.org/10.1007/s10648-021-09624-7

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. (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., Ayres, P., Kalyuga, S., & Chandler, P. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Sweller, J., & Sweller, S. (2006). Natural Information Processing Systems. *Evolutionary Psychology*, *4*(1), 147470490600400. https://doi.org/10.1177/147470490600400135

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. New York. Springer.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

van Gog, T., Ericsson, K. A., Rikers, R. M. J. P., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development*, *53*(3), 73–81. https://doi.org/10.1007/BF02504799

van Gog, T., Kirschner, F., Kester, L., & Paas, F. (2012). Timing and Frequency of Mental Effort Measurement: Evidence in Favour of Repeated Measures. *Applied Cognitive Psychology*, *26*(6), 833–839. https://doi.org/10.1002/acp.2883

Wallen, E., Plass, J. L., & Brünken, R. (2005). The function of annotations in the comprehension of scientific texts: Cognitive load effects and the impact of verbal ability. *Educational Technology Research and Development*, *53*(3), 59–71. https://doi.org/10.1007/BF02504798

Weiner, B. (1986). *An attributional theory of motivation and emotion*. Springer Series in Social Psychology

Zhang, L., Kirschner, P. A., Cobern, W. W., & Sweller, R. (2022). There is an evidence crisis in science education policy. *Educational Psychology Review, 34,* 1157-1176. https://doi.org/10.1007/s10648-021-09646-1

Corresponding Author Contact Information:

Author name: Prof. Dr. Jenna Koenen

Department: TUM School of Social Sciences and Technology, Department Educational Studies, Chemistry Education

University, Country: Technical University of Munich, Germany

Email: jenna.koenen@tum.de

Please Cite: Koenen, J., Mariot, L. Tiemann, R. (2023). Paper-Based vs. Digitalized Glossaries in Laboratory Scripts. *The European Educational Researcher*, *6*(3), 79-99. DOI: https://doi.org/10.31757/euer.632

Conflict of Interest: The authors declare no conflict of interest.

Received: January 19, 2023 ▪ Accepted: September 25, 2023