The Effects of Augmented Reality: A Comparative Study in an Undergraduate Physics Laboratory Course

Sebastian Kapp[1] [a], Michael Thees[1], Fabian Beil[1], Thomas Weatherby[2], Jan-Philipp Burde[3], Thomas Wilhelm[2] and Jochen Kuhn[1]

[1] *Physics Education Research Group, TU Kaiserslautern, Erwin-Schrödinger-Str. 46, Kaiserslautern, Germany*

[2] *Department for Physics Education Research, Goethe-University Frankfurt, Max-von-Laue-Str. 1, Frankfurt am Main, Germany*

[3] *Physics Education Research Group, Eberhard Karls University Tübingen, Auf der Morgenstelle 14, Tübingen, Germany*

Keywords: Augmented Reality, Physics Laboratory Experiments, Cognitive Load, Split-attention Effect, Inquiry Learning.

Abstract: The use of augmented reality (AR) in inquiry-based learning has become of increasing interest to researchers. Recent studies highlight the benefits of AR in various instructional scenarios concerning knowledge acquisition and cognitive load compared to traditional settings. Particularly in the context of physics laboratory experiments, previous research examined the context of simple electrical circuits. However, results were limited to laboratory studies and showed contrasting impacts on knowledge acquisition. While one study reported a higher knowledge acquisition in a tablet-based AR setting, another study reported a higher knowledge acquisition and a reduction in extraneous cognitive load in a two-dimensional non-AR setting compared to a smartglasses-based AR setting. Consequently, the importance of context specific aspects must be considered more deeply. In this study we present a randomized controlled trial in a graded physics laboratory course evaluating the effects of a smartglasses-based AR environment on cognitive load and conceptual knowledge acquisition compared to a two-dimensional non-AR setting. The sample consists of a total of $N = 56$ students in two groups performing a set of eight traditional inquiry-based experiments exploring the relationships in basic circuit theory. While both groups reported low extraneous cognitive load and achieved a significant knowledge acquisition, no group differences were detected.

1 INTRODUCTION

1.1 Objective and Rationale

Inquiry-based learning can be described as a form of engaged learning in which learners are presented with situations they need to manipulate in order to create the information they are supposed to learn by themselves (de Jong, 2019). Therefore, as most laboratory courses require the learner to find an answer to a given scientific question, they can be taken as an an example of inquiry-based learning (de Jong, 2019). In general, inquiry-based learning does not require the use of a certain technology. However, a specific technology can be an essential part of its implementation (de Jong et al., 2018; de Jong, 2019). Traditional physics laboratory courses, while providing physical interaction and implementing inquiry-based learning, do not automatically guarantee positive learning outcomes (Husnaini and Chen, 2019; Kapici et al., 2019; Wilcox and Lewandowski, 2017). Yet, adding virtual content to such a setup is claimed to be beneficial for learning outcomes in STEM courses (de Jong et al., 2013; Jones and Sharma, 2019). Augmented reality is thereby capable of displaying virtual components together with the real-world physical setup (Azuma, 1997; Billinghurst and Duenser, 2012; Kuhn et al., 2016; Santos et al., 2014; Strzys et al., 2018) and can be used to provide spatial and temporal contiguity between physical and virtual components (Bujak et al., 2013; Altmeyer et al., 2020; Thees et al., 2020a,b). Ibáñez and Delgado-Kloos (2018) additionally emphasized the application of AR in inquiry-based STEM learning. Overall, Garzón and Acevedo (2019) observed a significant learning gain by using AR (medium effect, $d = .68$) in various instruc-

[a] https://orcid.org/0000-0003-1052-8901

Kapp, S., Thees, M., Beil, F., Weatherby, T., Burde, J., Wilhelm, T. and Kuhn, J.
The Effects of Augmented Reality: A Comparative Study in an Undergraduate Physics Laboratory Course.
DOI: 10.5220/0009793001970206
In *Proceedings of the 12th International Conference on Computer Supported Education (CSEDU 2020) - Volume 2*, pages 197-206
ISBN: 978-989-758-417-6

tional scenarios. However, one should keep in mind that learning outcomes are not only governed by spatial and temporal contiguity but also depend on the instructional goal of the lesson as well as the functional relation between virtual and physical components (Ainsworth, 2006; Garzón and Acevedo, 2019; Rau, 2020; Thees et al., 2020a).

Recently, the subject area of electrical circuits in physics laboratory courses was introduced as a new area of research (Altmeyer et al., 2020; Thees et al., 2020a). While Altmeyer et al. reported a benefit of tablet-based AR regarding the learning gain in physics laboratory courses, Thees et al. reported a benefit in learning gain as well as a reduction in extraneous cognitive load for a two-dimensional representation in comparison to a smartglasses-based AR environment. This result could be explained by applying the theory of a spatial contiguity failure (Beege et al., 2019).

With both studies being limited in the form of clinical laboratory studies with voluntary participants, their ecological validity with respect to a real-world scenario is unclear. We present a randomized controlled trial in a graded laboratory course at a German university with $N = 56$ participants in which performance and cognitive load were assessed as dependent variables.

1.2 Multiple Representations, Cognitive Load and Multimedia Learning

As learning environments using technology can contain multiple sources of information like visual elements (e.g. real laboratory setup and visualized data) and verbal information (e.g. numerical data, worksheets etc.; Ainsworth, 2006), one can consider basic principles of multimedia learning to evaluate them (Moreno, 2006).

Combining multiple representations in such a way is particularly prominent in augmented reality setups, as they are able to add multiple visualizations of information and data into the learner's existing physical environment (Azuma, 1997; Billinghurst and Duenser, 2012; Santos et al., 2014). Adding additional information can be described using Cognitive Load Theory (Sweller, 1988; van Merriënboer and Sweller, 2005) and Cognitive Theory of Multimedia Learning (Mayer, 2005, 2009).

Cognitive Load Theory (CLT; Sweller, 1988; van Merriënboer and Sweller, 2005) assumes that the capacity of an individual's working memory is limited. Performing a learning task thereby creates a cognitive load on the working memory. According to Sweller et al. (1998), this load can be described in three categories: intrinsic cognitive load (ICL, caused by the inherent complexity of the task), extraneous cognitive load (ECL; caused by performing task-irrelevant cognitive processes), and germane cognitive load (GCL; caused by constructing conceptual knowledge from information sources). However, in recent years researchers also suggest a realignment towards a two-factor ICL/ECL model in which the effectiveness of the instruction is indicated by the GCL (Sweller et al., 2019).

The Cognitive Theory of Multimedia Learning (CTML; Mayer et al., 1999) states that an individual must actively engage in integrating multimedia information into mental representations of the learning content to achieve meaningful learning. Processing multimedia instructions can be described in terms of three main processes in one's working memory: selection, organization, and integration. In accordance with the cognitive load theory, CTML also assumes that working memory is of limited capacity regarding the amount of information that can be processed simultaneously. To reduce the cognitive load induced by the integration of information and to prevent overload, several design principles that lead to an economic use of a learner's working memory were developed (Mayer, 2009).

Highly important to our work are the principles of temporal and spatial contiguity (Mayer, 2009; Mayer and Moreno, 2003). They suggest that related information should be presented simultaneously as well as spatially close to each other to promote positive learning outcomes and reduce extraneous processing (Ginns, 2006; Schroeder and Cenkci, 2018). The split-attention effect thereby describes the required division in the learner's attention when information is presented spatially and temporally split from each other causing search processes irrelevant to learning. (Mayer and Moreno, 2003; Mayer and Pilegard, 2014; Sweller and Chandler, 1994; Sweller et al., 2019). By presenting information in a spatially and temporally contiguous way this effect can be avoided, and the extraneous load reduced. In their review of different instructional multimedia environments, Schroeder and Cenkci (2018) confirm the advantages of integrated design formats over split-source formats and thus consolidate the split-attention effect as a rather stable finding in multimedia research.

However, under certain circumstances integrated designs fail to lead to higher learning gains and lower cognitive load when compared to split-source formats (Beege et al., 2019; Cammeraat et al., 2020). Such scenarios seem to present a boundary condition for the split-attention effect and Beege et al. (2019) coined the term spatial contiguity failure for these cases. The failure occurs for example when the element interactivity within a text is high before it is integrated into other representations. An element is thereby defined as anything that needs to be learned (Sweller, 2010). In the case of high element interactivity parts of the text can only be understood in combination with each other and learners have to reintegrate the pieces of information that are now split across the instructional material. This reintegration of information can counter the advantages of integrated formats by causing elevated extraneous cognitive load. Therefore, it depends on the instructional material under consideration whether an integrated or a split format leads to better learning outcomes.

1.3 Prior Work

Prior work concerning electric circuits showed that AR-based learning environments can avoid split-attention effects in university STEM laboratory courses by presenting virtual information in close spatial proximity to their physical counterparts:

Altmeyer et al. (2020) used a tablet-based approach to create an AR environment supporting students' knowledge acquisition by visualizing real time measurement data. The visualization was either presented spatially anchored to the real world using tablet-based AR or presented spatially split in the form of a two-dimensional grid on a tablet screen. Their goal was to minimize the learners' extraneous processing while conducting the experiment. In their study, no group-specific reduction of extraneous load was detected and performance scores revealed just minimally higher learning gains in favor of the AR setting.

Based on these results, Thees et al. (2020a) assumed that the setting of Altmeyer et al. (2020) could depict a case of spatial contiguity failure as described by Beege et al. (2019). To support their hypothesis, Thees et al. conducted another study based on the same instructional content as Altmeyer et al.. There, the non-AR group used the same two-dimensional visualization of the measurement data on an tablet as Altmeyer et al.. The AR group, however, used smartglasses to create an integrated, three-dimensional and hands-free visualization. Based on these settings, two types of spatial contiguity were comparable: In the non-AR setting the visualizations were closely related to each other as they were presented in a grid. In the AR setting the visualizations were closely related to the real experimental components by being spatially anchored to them. However, due to the limited field of view of the smartglasses, not all visualizations were visible at all times. Based on these differences in the presentation of the visualizations a difference in the learning outcome of the students was expected. The results showed a lower extraneous load for the non-AR group. Regarding their conceptual knowledge, both groups showed a significant learning gain after the intervention. However, the non-AR group showed a significantly higher gain regarding instruction-related conceptual and transfer tasks compared to the AR group.

1.4 Hypotheses

The purpose of the present study is to validate the hypotheses from Thees et al. (2020a) in practical use during a graded laboratory course. To reach this goal, the same conditions as Thees et al. were used with the instructions and test instruments adapted to the different target audience.

As the instruction primarily requires the learner to identify relations between physical values at different points in the circuit, it is more important for the learner to be able to quickly compare these values between each other instead of precisely locating them in the real setup. With the examined circuits only consisting of a few components, their spatial complexity was low and easy to remember. Therefore a case of spatial contiguity failure regarding the spatial position of the visualizations is assumed to be present (Beege et al., 2019), leading to the first hypotheses:

Hypothesis 1: Working with tablet-based two-dimensional visualizations results in a lower ECL compared to the smartglasses-based AR visualization.

Following the spatial contiguity principle, the close spatial presentation of related information during multimedia learning can support generative processing and therefore increase the efficacy of learning. This efficacy is defined here as the increase in conceptual knowledge (Pundak and Rozner, 2008; Vosniadou, 2007).

Hypothesis 2: Working with tablet-based two-dimensional visualizations results in an increased knowledge gain compared to the smartglasses-based AR visualization.

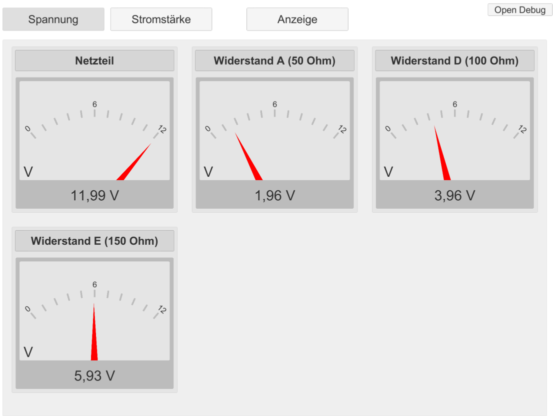


Figure 1: Screenshot of the two-dimensional visualization as presented on the tablet PC.

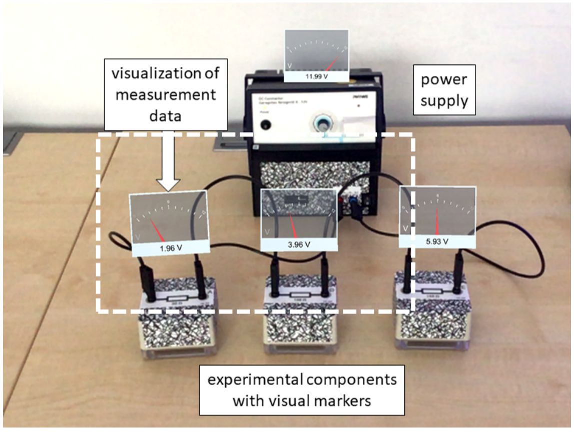


Figure 2: Representation of the AR view as seen through the smartglasses *(limited field of view indicated by white rectangle)*.

2 MATERIALS AND METHODS

2.1 Development of the Technology-enhanced Science Learning Environments

We developed a universal science environment covering electrical circuits to foster the acquisition of conceptual knowledge. The corresponding concept was first presented in Kapp et al. (2019) and also used in Altmeyer et al. (2020) and Thees et al. (2020a). It consists of custom designed measurement nodes integrated into the experimentation components as well as accompanying mobile applications for both a tablet PC (Apple iPad) as well as a pair of smartglasses (Microsoft HoloLens). The measurement nodes communicate their measurements to a tablet in real time, which then either displays them in a two-dimensional representation or sends the data to the smartglasses to be visualized in an AR environment.

Staying as close to traditional physics learning scenarios as possible, the experimentation components offer two sockets to integrate them into the electrical circuit as well as the electronic symbol and details of the integrated component. However, the measurement node added is also kept transparent to the learner using a transparent bottom. This measurement node constantly measures the current and voltage applied to the component and makes the measurements available to the clients using a Bluetooth Low Energy (BLE) service. To measure the voltage and current applied to the circuit by the power supply, a custom 3D printed enclosure also houses an identical measurement node as the components.

In total there were three custom applications used in the present study. All applications were developed using the Unity3D game development engine targeting either an Apple iPad or the Microsoft HoloLens.

The first application targets the two-dimensional visualization using an Apple iPad and is used by the non-AR group. It directly visualizes the measurement data in a grid on the display of the tablet (Figure 1). During the experiment the learner can select between the visualization of voltage (Spannung) or current (Stromstärke) and is able to hide visualizations of unused components.

The second two applications target the use of smartglasses in the experiment and are used by the AR group. They also use an Apple iPad which communicates with the measurement nodes and sends the acquired data to a Microsoft HoloLens for the AR view. The measurement data is first received and evaluated on the Apple iPad to reduce the load on the HoloLens application and to enable multiple smartglasses to take part in the same experiment. This preprocessed data is then sent to the HoloLens over Wi-Fi. The tablet application is only used by the instructor and does not visualize the measurement data to the participant. On the HoloLens the measurement data is spatially anchored to the corresponding experimentation components resulting in an AR view (Fig. 2). The position of the components is identified using visual markers which are recognized by the Vuforia Engine integrated into Unity3D. After an initial localization using the visual markers, the position of the experimentation components is transferred to the tracking capabilities of the Microsoft HoloLens which keeps the visualization stable even when the marker recognition is lost. During the experiment the learner can select between voltage and current visualization using gesture controls provided by the

Microsoft HoloLens. Currently unused components are simply placed outside the experimentation area, which also moves their visualization outside of the field of view of the learner.

To enable the applications to support a multitude of electrical circuits they can be individually configured. In this initial configuration the instructor specifies which experimentation components will be used and initializes the connections to the measurement nodes. After saving the configuration, it is offered to a learner starting the application. Selecting an existing configuration results in the application automatically connecting to all measurement nodes and visualizing the corresponding measurement data. With this setup the participant is immediately able to start the experiment.

2.2 Procedure and Materials

A two-group pretest–posttest design was used in the present study with participants randomly assigned to the AR or non-AR group. The instructional material as well as the visualizations were identical for each group reducing the difference between them to the technology used to present the visualizations and therefore their spatial position.

The tests, design and questionnaires were derived from Altmeyer et al. (2020) and Thees et al. (2020a) while being modified to fit the target group of the physics laboratory course. A conceptual knowledge test was used in both the pre- and posttest to capture the prior knowledge as well as the learning gain. In the posttest we also used a subjective rating scale given to evaluate the cognitive load during the experiment. The usability of the system was also evaluated during the posttest using a subjective rating scale.

2.2.1 Sample

The sample comprises of $N = 56$ biology students taking the second physics laboratory course at the Goethe-University Frankfurt. All students had previously visited the corresponding physics lecture and had taken the first physics laboratory course covering mechanics, optics and thermodynamics.

2.2.2 Procedure

The present study was conducted during a graded laboratory course taking place within the lecture period and included one session per week. In each session, teams of two students conducted one of the experiments in the course. Every week the students were assigned an experiment for the next session and received an instruction manual with which they prepared for it. During the session the total time to work on the tasks from the instruction manual was limited to three hours. After carrying out the experiment, the students were asked to evaluate their results and to prepare an experiment protocol by the next session, one week later. The measures of the study took place exclusively during the execution of the experiment.

At the beginning of the session, the students first answered the pretest consisting of the knowledge test as well as providing demographic data. Afterwards, a short oral examination of the students regarding the content of the experiment as well as a briefing regarding the assistive system was conducted by the supervisor, with both groups having the same supervisor.

Table 1: Demographic data of the sample. *(* voluntary disclosure).*

	non-AR group ($n = 29$)	AR group ($n = 27$)
gender		
female	17	19
male	9	4
no answer*	3	4
age		
M	20.83	20.44
SD	1.42	1.63

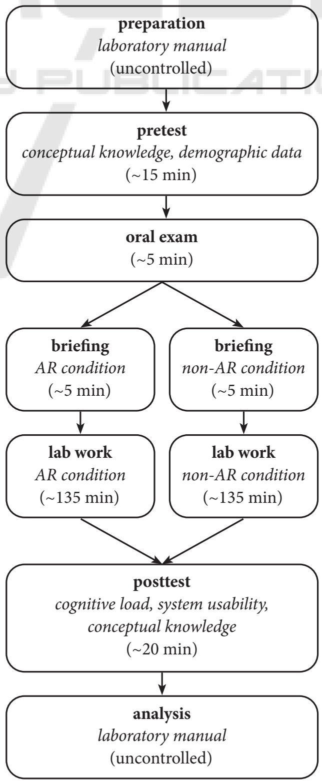


Figure 3: Experimental procedure.

For the group using smartglasses the briefing also included an individual calibration of the device for each student using the default calibration routine. After this, the students independently worked on the experimental tasks as given in the instruction manual. Each task consisted of setting up an electrical circuit, answering qualitative observations and capturing three measurement series regarding the voltage and current at each component. After all tasks had been completed, the supervisor inspected their work ensuring its completeness. Once every task was completed the students answered the posttest consisting of a questionnaire regarding the perceived cognitive load, the usability of the used assistive system, and again the knowledge test.

2.2.3 Instruction Manual and Experimental Tasks

The instruction manual guides the participants through the experiment and the subsequent analysis. It first lists the terminology and theoretical background the student was supposed to prepare before coming to the session. The experiment itself consists of eight electrical circuits in total. The first six experiments cover three serial and three parallel circuits by systematically adding resistors and varying their resistance. For each circuit the students were given a set of qualitative observation tasks and were asked to collect three series of measurements. The last two experiments are combined circuits which include a switch closing either a parallel branch in the first circuit or bridging a resistor in the second. For these circuits only qualitative observation tasks are given, aiming at the understanding of the circuit in both positions of the switch. In the analysis, the students had to substantiate all their qualitative observations using Kirchhoff's circuit laws. They also had to quantitatively evaluate whether their collected measurement series fitted the predictions made using the laws. To ease the quantitative evaluation the tolerances of the measurement nodes used were provided as well as the formula for the propagation of these tolerances.

2.2.4 Conceptual Knowledge Test

To evaluate the learning gain we used a conceptual knowledge test regarding electrical circuits. It consisted of 15 single-choice items and is based on a two-tier test instrument developed by Ivanjek et al. (2020). The focus of the test instrument is to assess students' conceptual understanding of current, voltage and resistance. Each item consists of an item stem describing the circuit as well as a corresponding circuit diagram. The original test instrument uses a two-tier structure in which students not only have to answer a question on the first tier but also provide an explanation for their reasoning on the second tier. However, for the present study we only focused on the first tier of the test instrument. An example item is presented in Figure 4.

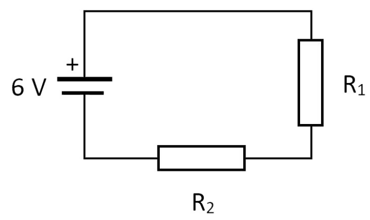
The circuit shown consists of a 6 V battery and two identical resistors.

Compare the current through both resistors.
☐ The current through R_1 is larger than through R_2.
☐ A current flows through R_1 but not through R_2.
☐ The current through R_2 is larger than through R_1.
☐ A current flows through R_2 but not through R_1.
☐ The current through both resistors is the same.

Figure 4: Example item from the conceptual knowledge test used (based on: Ivanjek et al., 2020, translated for this publication).

2.2.5 Cognitive Load

To investigate the intrinsic, extraneous and germane cognitive load of each participant during the experiment we adapted the cognitive load scale (CLS) by Leppink et al. (2013, 2014) to a laboratory setting as presented by Thees et al. (2020b). Thees et al. reported a comparable internal structure as Leppink et al. for the modified six-point Likert scale as well as nearly excellent reliabilities for the three sub scales ($\alpha_{ICL} = .70$, $\alpha_{ECL} = .66$, $\alpha_{GCL} = .86$). The same cognitive load scale was used by Altmeyer et al. (2020).

2.2.6 Usability

We used the System Usability Scale (SUS) by Brooke (1996) to investigate the usability of the assistive systems. There, the participants had to rate their level of agreement with 10 statements concerning the handling and usefulness of the used system on a five-point scale. The SUS was translated into German and the term "system" was specified as "the interaction between the digital assistive system, the corresponding software application, and the experimental equipment". The score is calculated by multiplying the cumulated item scores (value range: 0-4; some items inverted) with the factor 2.5 to achieve a value range between 0 and 100 (best). The same scale was used by Altmeyer et al. (2020) and Thees et al. (2020a).

3 RESULTS

The descriptive results for all dependent variables are shown in Table 2. The assumptions for conducting an ANOVA (independence of samples, normal distribution of residuals, and homogeneity of residuals' variances) and ANCOVA (equal variance across groups, homogeneity of variance) were met for all analyses reported.

Table 2: Standardized Means (M) and Standard Deviations (SD) for Dependent Variables, Separate for the AR and non-AR group.

	non-AR group ($n = 29$)	AR group ($n = 27$)
Cognitive load rating		
ICL	.22 (.15)	.28 (.16)
ECL	.17 (.17)	.17 (.12)
GCL	.67 (.20)	.60 (.19)
Conceptual knowledge		
Pre	.40 (.12)	.39 (.12)
Post	.53 (.21)	.48 (.18)
System usability	87.2 (12.6)	69.9 (17.7)

3.1 Effects on Cognitive Load

The internal consistencies for the cognitive load scale by Leppink et al. (2013) as adapted by Thees et al. (2020b) were acceptable to good ($\alpha_{ICL} = 0.79$, $\alpha_{ECL} = 0.75$, $\alpha_{GCL} = 0.84$).

To evaluate significant group differences for each type of load, an independent-samples t-test was applied to each subscale. No significant differences were detected for ICL ($t(52.6) = 1.38$, $p = .175$), ECL ($t(50.8) = 0.08$, $p = .933$) or GCL ($t(54.0) = -1.39$, $p = .170$).

3.2 Effects on Performance

The students' performance was analyzed with an analysis of variance for repeated measurements (ANOVA-RM) with group as a between-subject factor and time as a within-subject factor. The analyses showed a significant main effect of time ($F(1,54) = 19.64$, $p < .001$) with an effect size of $\eta_p^2 = 0.27$ (large effect). No significance was identified regarding the main effect of group ($F(1,54) = 0.59$, $p = .448$) as well as the interaction between group and time ($F(1,54) = 0.96$, $p = .330$).

Utilizing the performance of the students in the pretest as co-variate shows a significant correlation with their performance in the posttest ($F(1,53) = 9.27$, $p = .004$, $r = .38$). However, controlling for the performance in the pretest using an analysis of covariance (ANCOVA) the group difference in the posttest stays not significant ($F(1,53) = 1.05$, $p = .311$).

3.3 Usability

The internal consistency of the System Usability Scale (Brooke, 1996) was nearly excellent ($\alpha = 0.88$).

Following the scoring method by Brooke the average usability score is shown in Table 2. An independent-samples t-test revealed a significant difference between the two conditions ($t(46.64) = 4.18$, $p < .001$). In compliance with the review by Bangor et al. (2009), the usability of the non-AR condition can be classified as "excellent" and that of the AR condition as "good". These are the second and third best rating-level classifications following "best imaginable".

4 DISCUSSION

The purpose of this study was to validate the results from Altmeyer et al. (2020) and Thees et al. (2020a) in a graded physics laboratory course at a German university. For this, the intervention and control groups from Thees et al. (2020a) were adopted while adapting the instructions and tests to the different target group. Both groups fulfilled the temporal contiguity principle as described by the CTML.

Concerning the performance in the pretest the groups were equally distributed. Both groups showed a low intrinsic (ICL) and extraneous (ECL) cognitive load and a high germane cognitive load (GCL) with no significant differences. Importantly, both reached a significant gain in their conceptual knowledge (large effect) which is not guaranteed per se for inquiry-based learning in a laboratory course (Husnaini and Chen, 2019; Kapici et al., 2019; Wilcox and Lewandowski, 2017). However, we were unable to confirm our hypotheses in form of a reduced extraneous cognitive load and higher group-specific learning gain in favor of the non-AR group (see Table 3).

Due to the limited field of view of the smartglasses, the AR group was unable to register all visualized measurement values at once. The non-AR group in contrast was able to register them simultaneously due to the presentation on a tablet screen. The

Table 3: Summary of main results regarding hypotheses.

Hypothesis (in short)		Dependent Variable	Hypothesis supported
H1)	Less ECL for non-AR condition	Subjective ECL rating score	No
H2)	Higher learning gain for non-AR condition	Conceptual knowledge score	No

predicted split-attention effect found by Thees et al. (2020a) however seems to not show itself in the same aspects in a real laboratory setting.

These results contrast not only the results presented by Thees et al. (2020a) and Altmeyer et al. (2020) but also the reported benefits of the use of AR versus non-AR in other studies (Garzón and Acevedo, 2019; Ibáñez and Delgado-Kloos, 2018; Strzys et al., 2018). This may be due to the interaction between participants in collaboration settings. As Janssen and Kirschner (2020) showed, group interaction could compensate for differences in cognitive load.

The usability score of the assistive systems matched the results found in Thees et al. (2020a) and showed a significantly lower score for the AR group than the non-AR group. Nevertheless, they both can still be categorized as acceptable. However, it is possible that the system usability has a higher influence in a field setting compared to a laboratory setting resulting in the contrasting results.

5 CONCLUSION AND FURTHER RESEARCH

In summary, the present field study did not show differences in extraneous cognitive load or knowledge acquisition between the AR and the non-AR setting. This contrasts previous results for the present subject as well as other findings regarding the superiority of AR in different subjects. However, these results do not exclude a benefit of the use of AR in all cases. Especially referring to the theory of spatial contiguity failure, it is possible that the same digital assistance system results in benefits when addressing different experimental tasks.

Further studies should therefore evaluate the effects of this failure regarding two aspects: The use of the more accessible format of tablet-based AR could reduce the limitation of a small field of view and therefore reduce the split between the visualized measurement values. The same effect might be reached using a future generation of smartglasses with an extended field of view. Additionally, a more complex instruction, e.g. using more components inside the circuits, might require more intrinsic and extraneous processing, which could be addressed by the spatial split-attention effect. A follow up study is intended to evaluate the use of tablet-based AR in the same graded laboratory course. In future studies the aspect of collaboration could also be measured and implemented as further dependent variable.

Although the present study and current research shows the limitations of AR in educational settings, they also show the possibilities of the use of AR. Together with previous research it offers a more detailed look into the requirements for a successful implementation.

ACKNOWLEDGEMENTS

Support from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung; BMBF) via the projects "GeAR" (Grant No. 01JD1811B) and "gLabAssist" (Grant No. 16DHL1022) is gratefully acknowledged.

We thank the Microelectronic System Design Research Group/Technische Universität Kaiserslautern (chaired by Prof. Norbert Wehn), especially Frederik Lauer and Carl Rheinländer, for providing the hardware components for the experiment.

REFERENCES

Ainsworth, S. (2006). Deft: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3):183–198.

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., and Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in stem laboratory courses — theoretical background and empirical results. *British Journal of Educational Technology*. Advance online publication. doi:10.1111/bjet.12900.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385.

Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4(3):114–123.

Beege, M., Wirzberger, M., Nebel, S., Schneider, S., Schmidt, N., and Rey, G. D. (2019). Spatial continuity effect vs. spatial contiguity failure. revising the effects of spatial proximity between related and unrelated representations. *Frontiers in Education*, 4:89.

Billinghurst, M. and Duenser, A. (2012). Augmented reality in the classroom. *Computer*, 45(7):56–63.

Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7.

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., and Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*, 68:536–544.

Cammeraat, S., Rop, G., and de Koning, B. B. (2020). The influence of spatial distance and signaling on the split-attention effect. *Computers in Human Behavior*, 105:106203.

de Jong, T. (2019). Moving towards engaged learning in stem domains; there is no simple answer, but clearly a road ahead. *Journal of Computer Assisted Learning*, 35(2):153–167.

de Jong, T., Ard Lazonder, Margus Pedaste, and Zacharias Zacharia (2018). Simulations, games, and modeling tools for learning. In Fischer, F., Hmelo-Silver, C. E., Goldman, S. R., and Reimann, P., editors, *International Handbook of the Learning Sciences*, pages 256–266. Routledge, Taylor and Francis.

de Jong, T., Linn, M. C., and Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science (New York, N.Y.)*, 340(6130):305–308.

Garzón, J. and Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review*, 27:244–260.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, 16(6):511–525.

Husnaini, S. J. and Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Physical Review Physics Education Research*, 15(1):31.

Ibáñez, M.-B. and Delgado-Kloos, C. (2018). Augmented reality for stem learning: A systematic review. *Computers & Education*, 123:109–123.

Ivanjek, L., Schubatzky, T., Burde, J.-P., Hopf, M., Wilhelm, T., Haagen-Schützenhöfer, C., Dopatka, L., and Spatz, V. (2020). Development of a two-tier instrument on simple electric circuits. Manuscript in preparation.

Janssen, J. and Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: towards a research agenda. *Educational Technology Research and Development*, 36(2):321.

Jones, K. A. and Sharma, R. S. (2019). An experiment in blended learning: higher education without lectures. *International Journal of Digital Enterprise Technology*, 1(3):241.

Kapici, H. O., Akcay, H., and de Jong, T. (2019). Using hands-on and virtual laboratories alone or together—which works better for acquiring knowledge and skills? *Journal of Science Education and Technology*, 28(3):231–250.

Kapp, S., Thees, M., Strzys, M. P., Beil, F., Kuhn, J., Amiraslanov, O., Javaheri, H., Lukowicz, P., Lauer, F., Rheinländer, C., and Wehn, N. (2019). Augmenting kirchhoff's laws: Using augmented reality and smartglasses to enhance conceptual electrical experiments for high school students. *The Physics Teacher*, 57(1):52–53.

Kuhn, J., Lukowicz, P., Hirth, M., Poxrucker, A., Weppner, J., and Younas, J. (2016). gphysics—using smart glasses for head-centered, context-aware learning in physics experiments. *IEEE Transactions on Learning Technologies*, 9(4):304–317.

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

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P., and van Merriënboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30:32–42.

Mayer, R. E., editor (2005). *The Cambridge handbook of multimedia learning*. Cambridge Univ. Press, Cambridge.

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

Mayer, R. E. and Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1):43–52.

Mayer, R. E., Moreno, R., Boire, M., and Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, 91(4):638–643.

Mayer, R. E. and Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In Mayer, R. E., editor, *The Cambridge handbook of multimedia learning*, Cambridge handbooks in psychology, pages 316–344. Cambridge University Press, New York.

Moreno, R. (2006). Learning in high-tech and multimedia environments. *Current Directions in Psychological Science*, 15(2):63–67.

Pundak, D. and Rozner, S. (2008). Empowering engineering college staff to adopt active learning methods. *Journal of Science Education and Technology*, 17(2):152–163.

Rau, M. A. (2020). Comparing multiple theories about learning with physical and virtual representations: Conflicting or complementary effects? *Educational Psychology Review*, 16(3):183.

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., and Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies*, 7(1):38–56.

Schroeder, N. L. and Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: a meta-analysis. *Educational Psychology Review*, 30(3):679–701.

Strzys, M. P., Kapp, S., Thees, M., Klein, P., Lukowicz, P., Knierim, P., Schmidt, A., and Kuhn, J. (2018). Physics holo.lab learning experience: using smartglasses for augmented reality labwork to foster the concepts of heat conduction. *European Journal of Physics*, 39(3):035703.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2):123–138.

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

Sweller, J., van Merriënboer, J. J. G., and Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2):261–292.

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

Thees, M., Altmeyer, K., Kapp, S., Rexigel, E., Beil, F., Klein, P., Malone, S., Brünken, R., and Kuhn, J. (2020a). Augmented reality for inquiry learning in stem laboratory courses: Opportunities and risks, but no simple answers. Manuscript submitted for publication.

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., and Kuhn, J. (2020b). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, 108:106316.

van Merriënboer, J. J. G. and Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2):147–177.

Vosniadou, S. (2007). Conceptual change and education. *Human Development*, 50(1):47–54.

Wilcox, B. R. and Lewandowski, H. J. (2017). Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics. *Physical Review Physics Education Research*, 13(1):65.