Smartglasses in STEM laboratory courses—the augmented thermal flux experiment

Martin P. Strzys,¹ Michael Thees,¹ Sebastian Kapp,¹ and Jochen Kuhn¹

¹*Department of Physics, Physics Education Research Group,
Technische Universität Kaiserslautern, D-67653 Kaiserslautern, Germany*

Augmented reality (AR) learning scenarios with see-through smartglasses create a wearable education technology providing active access to various additional information without distracting from the physical interaction with reality. We already have introduced such an AR version of a standard physics experiment of introductory lab courses on heat conduction in metals, using real physical data from external sensors for analyzing and displaying thermal phenomena in real-time. Besides a direct feedback, ensuring that students get an immediate impression of the effects of the experimental parameters, this scenario is also able to visualize invisible physical processes, using false-color representations to show the temperature of the apparatus. In a previous study conducted in an introductory STEM laboratory, we were able to show that such an AR learning environment indeed is suitable to foster learners' conceptual understanding of thermal phenomena. In the current paper we focus on the question how learners can benefit from such a scenario by influencing cognitive load (CL). In a second study we use the cognitive load scale (CLS) to discriminate the different types of learners' CL. We confirmed the structure of the scale by a factor analysis, finding three factors corresponding to the three types of CL, each with high reliability. Moreover, we were able to show that with our AR scenario extraneous load could significantly be reduced, compared to a non augmented traditional setup.

I. INTRODUCTION

Learning, knowledge acquisition and schemata construction is an extremely individual process, strongly depending on learners' level of expertise, self-directed action and motivation. Especially during laboratory courses in STEM university education all channels of knowledge construction are important as students have to synchronize their prior theoretical understanding with new experimental hands-on experience, causing an intense interaction between theory and experiment. Realizing AR scenarios with see-through smartglasses like HoloLens, in contrast to using other modern digital media like smartphones or tablet PCs, creates a wearable education technology providing learners with active access to various kinds of additional information while simultaneously ensuring a clear hands-free scenario [1, 2]. This finally yields the possibility of implementing a real-time supporting system without distracting students from their physical interaction with the traditional experimental hardware setup.

Support may range from detailed instruction or interactive tutorials, over safety guidelines, to various real-time representations of measurement data. Diagrams with fitting curves, important numerical values or animations visualize the consequences of students' haptic action in the real world, e.g., varying basics parameters of the setup, just in time. This enables them to evaluate their own measurements in more detail, to draw conclusions for further investigations or functional correlations between physical quantities; all of this is possible while conducting the experiment.

By displaying information like real-time representations of measurement data, diagrams with fitting curves or important numerical values, directly in the user's field of view, the combination of the real world and virtual objects creates an AR learning environment that obeys basic psychological design principles of cognitive theory of multimedia learning (CTML) for visual content and learning with multiple repre-

sentations [3, 4]. In particular, the smart technology allows to prepare data in real-time and to connect its representation to corresponding components of the experimental setup within 3D space. Therefore, especially the spatial and temporal contiguity principles of CTML [4] are automatically obeyed by such a setup. In terms of cognitive load theory (CLT) [5] this corresponds to avoiding the so called split-attention effect [6] by ensuring a strong spatial and temporal connection between different sources of information to reduce extraneous load, i.e., load caused by inappropriate design.

In a previous quasi-experimental 2×2 -study [2] we were able to show that, compared to a traditional setting, such an AR learning environment indeed improves students' gain in conceptual understanding in the context of thermal conduction (Cohen's $d = 0.43$). This, however, does not answer the question about the origin of the improvement of the learning performance. Thus, in this paper we present a follow-up study focusing on the comparison of students' CL between the traditional non-AR-setup for the control group (CG) and the AR-scenario using HoloLens, which we call a *holo.lab* setup, for the treatment group (TG). Using the CLS [7] we were able to validate the scale and to show a significant reduction of the extraneous load for the students of the TG.

II. THEORETICAL BACKGROUND

Today modern digital media, such as smartphones and tablet PCs, with their various precise sensors for different physical quantities may be efficiently be used as portable mini-labs to perform experiments in almost all branches of physics [9–14]. However, in an experimental setup using such smart devices important design principles of CTML, like spatial and temporal contiguity [4], can only be partially obeyed, since students have to actively integrate the information from the real-world experiment and the digitally processed repre-

sentations of the experimental data. Therefore, they are urged to split their attention temporally or spatially between these different sources of information: the physical experimental setup and the smart measurement device. This so-called split-attention effect of CLT [5, 6] increases the level of extraneous load, i.e., CL caused by inappropriate design of the learning environment.

Since according to CLT total working memory capacity is limited, instructional multimedia design principles thus aim at reducing extraneous load to free as much working memory as possible in favor of intrinsic CL, caused by the natural complexity of the task, and germane CL, stemming from active information processing, schema construction and automation [6]. This finally can be achieved using AR technologies which have become more and more important in the educational context [1, 15, 16]. Today, smartglasses allow virtual 3D objects to be embedded into the real-world environment, leading to a true immersive AR experience. In such a digitally enhanced surrounding, virtual and real objects do not only co-exist, but are also able to interact with each other in real-time. Therefore, the benefit of such a holo.lab setup using smartglasses, as introduced in [1, 2], is the possibility of simultaneously keeping track of the real experimental setup and digital representations of the data.

According to CTML it can be assumed that temporal and spatial contiguity of the holographic projection directly onto the real object effectively supports the learning process by avoiding the split-attention effect that would occur according to CLT, if other display types, like smartphones, tablet PCs or classical computer screens, would be used for the virtual augmentation. This expectation is based on the assumption that basic principles of multimedia learning theory can straightforwardly be extended to the case of AR scenarios if real objects take over the role of pictures, while virtual augmentations substitute written text [15]. Therefore, real-time real-world annotations ensure both, temporal and spatial contiguity, thereby reducing CL of the students [15]. Thus, a larger fraction of the short term memory may be used during the cognitive processes of learning with the experiment.

In a holo.lab scenario the free capacities may especially be used for processing the immediate feedback-loop, enabling a more effective interrelation of theory and experiment, as experimental processes can be followed in real-time through different representations of the data and be directly related to theoretical predictions, which—in classical laboratory class setups—is only possible in hindsight while analyzing experimental data outside the lab.

III. LABORATORY CLASS STUDY

A. Hypothesis

We hypothesize that AR learning environments using smartglasses, ensuring a hands-free experimental setup and a real-time real-world annotation using different representa-

tions of the actual experimental data, like our holo.lab approach, obey the temporal and spatial contiguity principles, avoid the split-attention effect and therefore reduce the extraneous load (EL) of the students.

B. Sample

Our sample was made up of students attending the first year's physics laboratory class (mechanics and thermodynamics) at TU Kaiserslautern in spring 2018. It consists of 35 participants (30 male) aged 18–33 (mean 19.9 years); all of them are Physics majors, 34 were in their first semester (one in the second).

C. Study Design

In a quasi-experimental pre-post-test study with TG and CG the students completed a pre-pre-test collecting personal data, experience with AR/VR and other digital technologies, as well as stable affective variables approximately one month before the starting of the laboratory classes. As during the laboratory classes two students always conduct each experiment together. Groups of two were randomly distributed among the two groups of the study design (CG with traditional setup, $N_C = 16$ and TG with holo.lab setup, $N_T = 19$). Our evaluation in a 2×2 -design was integrated into the standard processes of the laboratory course: After arrival at the lab, students first have to pass a short introductory test, consisting of few oral questions by their advisor, determining that they have the background knowledge required to use the experimental apparatus. After that, both groups were introduced to the details of the experimental setup; for the TG this also included a short instruction to HoloLens and a calibration of the device. Immediately before they started to experiment, all students completed a pre-test including items about affective variables and a concept test consisting of HTCE-items [8] as in [2]. Having finished their experimental work they completed the post-test, which additionally contained the CLS items as well as the System Usability Scale (SUS). In this paper we exclusively report the results related to CL, further results of this study will be part of an upcoming publication.

The CG performed the experiment in the framework of the laboratory course working with the traditional experimental setup equipped with a hand-held IR camera and a PC, while the participants of the TG used the holo.lab setup with HoloLens support, as described in [1, 2] in more detail (a short recap of the setup will be given in the following section). The laboratory course is three hours long, which meant participants had only 2 1/2 hours to complete the experiment.

The 10 items of the CLS from [7] have been translated into German and slightly reformulated and sharpened with a strong focus on the very process of experimenting, as the original items only address the general “activity”. To achieve this, two items had to be doubled to reflect both,

FIG. 1. a) holo.lab setup with AR experience; augmented representations: false-color representation of temperature along the rod, numerical values at three points above the rod, temperature graph; b) Experimental setup and user wearing a HoloLens.

the (hands-on) process of data acquisition and the use of different representations of the data during the experiment. Therefore, Item 2 and Item 7 of [7] have been replaced by

Item 2a: The process of collecting data in the experiment I perceived as very complex.

Item 2b: The representations used during the experiment I perceived as very complex.

Item 7a: The process of experimenting really enhanced my understanding of the topics covered.

Item 7b: The process of experimenting really enhanced my understanding of collecting data.

Thus, the factors corresponding to intrinsic load (IL) and germane load (GL) each were expected to contain one more item, compared to [7], namely four and five items, respectively. To smoothly fit the rest of items in our survey we moreover switched to a six-point Likert scale. All surveys were presented on tablet PCs using an online assessment tool.

D. Experimental Setup

The heat conduction experiment used in our laboratory courses consists of cylindrical metal sample rods (cf. Fig. 1); the sample is heated at one end with a cartridge heater and cooled at the other with a standard CPU fan. The temperature data is extracted from thermal images taken either with a hand-held IR camera (CG) or with a stationary IR camera (TG). Students have to determine the stationary state of the temperature distribution of this setup and use it to calculate physical quantities like the thermal conductivity of the metals. The students of the CG use the IR camera to manually take thermal images of the sample. These images first have to be transferred to a provided PC, where the temperature data is extracted using commercial software of the camera manufacturer. After that the data can be exported to a csv-file for further processing and plotting. In the HoloLens setup used by the TG the temperature values are passed directly to the HoloLens via WiFi in real-time, where the visualization is done. The corresponding HoloLens-App provides three different representations of the data: false-color image of the

TABLE I. Factor loadings of the confirmatory factor analysis of the CLS items, reliability (Cronbach's α or α_i if Item i is dropped), mean $\mu_{C/T}$ and SD of the (normalized) score per group (CG/TG), t -test statistics (between groups).

Factor/Item	Loading	α/α_i	$\mu_C(SD)$	$\mu_T(SD)$	t	p
Factor GL		0.83	0.55(0.21)	0.61(0.13)	-1.07	0.30
Item 9	0.90	0.76	0.52(0.29)	0.65(0.17)		
Item 7b	0.80	0.78	0.56(0.30)	0.65(0.21)		
Item 8	0.54	0.80	0.49(0.26)	0.56(0.23)		
Item 10	0.53	0.79	0.59(0.22)	0.63(0.20)		
Item 7a	0.48	0.82	0.58(0.19)	0.55(0.17)		
Factor IL		0.83	0.37(0.21)	0.41(0.18)	-0.69	0.49
Item 3	0.93	0.74	0.42(0.24)	0.48(0.25)		
Item 1	0.78	0.83	0.49(0.26)	0.55(0.22)		
Item 2b	0.56	0.78	0.29(0.25)	0.33(0.20)		
Item 2a	0.53	0.80	0.26(0.27)	0.29(0.22)		
Factor EL		0.66	0.41(0.20)	0.28(0.17)	2.12	0.04
Item 6	0.67	0.51	0.42(0.27)	0.26(0.24)		
Item 4	0.55	0.50	0.35(0.31)	0.26(0.23)		
Item 5	0.38	0.66	0.45(0.20)	0.32(0.23)		

temperature values, projected directly onto the sample cylinder, numerical values at three pre-defined points and a temperature graph as a function of the position along the rod, hovering above the setup, cf. Fig. 1. Additionally the data can directly be exported to a csv-file.

Therefore, the main difference between the two setups is the absence of spatial and temporal contiguity of processed data and experimental setup in case of the CG, i.e. this group does not get any kind of real-time feedback in form of different representations; for further details see [1, 2].

IV. RESULTS

We performed a confirmatory factor analysis with three factors on the CLS items, using an oblique rotation (i.e. Oblimin) to also take correlations among the factors into account; in general correlations between the three factors are expected, since correspondence to a certain type of CL also depends on the pre-experiences of the participants and therefore has some intrinsic variance. The resulting factor loadings can be found in Table I.

The general factorial structure with three factors could be confirmed (RMSR = 0.08); the correlation between the factors is relatively low and given by $r(GL, EL) = 0.22$, $r(GL, IL) = -0.20$ and $r(EL, IL) = 0.16$, respectively, which also is in agreement with the values reported in [7].

The results of a reliability analysis are also presented in Table I. It turns out that the items forming one factor indeed form a homogeneous scale and that dropping items would not result in a higher value of Cronbach's α . Moreover, the Cron-

FIG. 2. Mean and SE of the CL score (normalized) per factor (GL, IL, EL) for each group (CG/TG).

bach’s α -values of the three factors are also in agreement with the values reported in [7].

Finally we compared the three different types of CL between the two groups. Thus, we calculated mean $\mu_{C/T}$ and SD of the normalized CL scores per factor and group, which may also be found in Table I. Additionally, Figure 2 yields a graphical representation of these quantities. For our setup, it turns out that perceived GL and IL of the TG is slightly higher (statistically non significant), while the EL is significantly smaller for the TG (cf. Table I for t -test statistics). This is also true if usability scores from SUS, which have a medium correlation with the EL score, are taken into account as covariates in an analysis of covariance (ANCOVA), yielding a significant ($F = 6.02, p = 0.02$) corresponding medium effect size ($d = 0.56$). Over all CL including all three factors did not differ significantly between the two groups.

V. CONCLUSION

In our experimental study accompanying the development of holo.lab experimentation scenarios, we were able to validate the CLS proposed in [7] and to confirm the reliability of the scale. With small modifications mentioned in Section III C to putting a strong focus on the process of experiment-

ing, the CLS can be used to clearly discriminate the three types of CL in laboratory scenarios.

The results clearly show a significant reduction of EL for the TG using the AR setup with HoloLens assistance. This supports our hypothesis assuming that EL could be decreased by avoiding the split-attention effect via applying instant real-world annotation. This is remarkable, as introducing multiple real-time representations to the learning environment also has a distractive potential. The use of smartglasses as technical solution to create AR environments may be assumed to introduce additional EL [17]. This, however, might not be captured by the CLS, as all items related to the factor EL refer to “instructions and/or explanations”. This is of special interest, as one might argue that AR support in this case achieves a clear improvement of the instructions during the experiment, e.g., by providing the correct representations, which might help to trigger correct observations and to eliminate a possible lack of clarity in the written instructions. To clarify this further, finer investigations will be needed, e.g. capturing the effects of real-time representations.

Moreover, although we are not able to provide statistically significant results here, the tendency that GL and IL score lower in case of the CG is still clearly visible. It will be important to verify this observation in upcoming studies and to achieve a deeper understanding of its primary causes.

Despite the unanswered questions, the results presented here encourage us to pursue the development of holo.lab learning scenarios and to accept the challenges that go hand in hand with it. The positive effects on the distribution of CL together with the hints to improved conceptual understanding from our prior study [2] indicate that efficient learning scenarios using smartglasses, of which STEM laboratory learning settings can clearly benefit from, are already within reach.

ACKNOWLEDGMENTS

Support from the German Federal Ministry of Education and Research (BMBF) via the projects *Be-greifen* [18] and *gLabAssist* is gratefully acknowledged.

-
- [1] Strzys, M. P. *et al.*, *Phys. Teach.* **55**(6), 376 (2017).
 - [2] Strzys, M. P. *et al.*, *Eur. J. Phys.* **39**, 035703 (2018).
 - [3] Mayer, R. E. and Fiorella, L. In *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press (2014).
 - [4] Moreno, R. *Curr. Dir. Psychol. Sci.* **15**(2), 63–67 (2006).
 - [5] Sweller, J. *Instructional design in technical areas*. Camberwell, ACER Press, (1999).
 - [6] Mayer, R. E. *Multimedia Learning*. Cambridge (2010).
 - [7] Leppink, J. *et al.*, *Behav Res* **45**(4), 1058–1072 (2013).
 - [8] Thronton, R. K. and Sokoloff, D. R. www.physport.org/assessments/assessment.cfm?I=16&A=HTCE, (2001).
 - [9] Klein, P. *et al.*, In *Multidisciplinary Research on Teaching and Learning*, 270–288. Palgrave Macmillan (2015).
 - [10] Kuhn, J. *et al.*, *Phys. Teach.* **52**, 351–356 (2014).
 - [11] Schwarz, O. *et al.*, *Phys. Teach.* **51**, 312–313 (2013).
 - [12] Vogt, P. *et al.*, *Phys. Teach.* **49**, 383–384 (2011).
 - [13] Hochberg, K. *et al.*, *J. Sci. Educ. Technol.* **27**, 385–403 (2018).
 - [14] Kuhn, J. *Am. J. Phys.* **82**, 94 (2014).
 - [15] Santos, M. E. C. *et al.*, *IEEE Trans. Learn. Technol.* **7**(1), 38–56 (2014).
 - [16] Kuhn, J. *et al.*, *IEEE Trans. Learn. Technol.* **9**(4), 304–317 (2016).
 - [17] Baumeister, J. *et al.*, *IEEE Trans. Vis. Comput. Graph.* **23**(11), 2378–2388 (2017).
 - [18] <https://www.physik.uni-kl.de/en/kuhn/forschungsprojekte/aktuelle-projekte/be-greifen/>.