


Article

Cognitive Load Implications for Augmented Reality Supported Chemistry Learning

Sebastian Keller ^{1,*}, Stefan Rumann ¹ and Sebastian Habig ² 
¹ Chemistry Education, University Duisburg-Essen, 45127 Essen, Germany; stefan.rumann@uni-due.de

² Chemistry Education, Paderborn University, 33098 Paderborn, Germany; sebastian.habig@upb.de

* Correspondence: sebastian.keller@uni-due.de

Abstract: This paper presents a study about augmented-reality-based chemistry learning in a university lecture. Organic chemistry is often perceived as particularly difficult by students because spatial information must be processed in order to understand subject specific concepts and key ideas. To understand typical chemistry-related representations in books or literature, sophisticated mental rotation- and other spatial abilities are needed. Providing an augmented reality (AR) based learning support in the learning setting together with text and pictures is consistent with the idea of multiple external representations and the cognitive theory of multimedia learning. Using multiple external representations has proven to be beneficial for learning success, because different types of representations are processed separately in working memory. Nevertheless, the integration of a new learning medium involves the risk to hinder learning, in case of being not suitable for the learning topic or learning purpose. Therefore, this study investigates how the AR-use affects students' cognitive load during learning in three different topics of organic chemistry. For this purpose also the usability of AR learning support is considered and the possible reduction of the influence of the mental rotation on learning success will be investigated.

Keywords: augmented reality; cognitive load; chemistry; usability



Citation: Keller, S.; Rumann, S.; Habig, S. Cognitive Load Implications for Augmented Reality Supported Chemistry Learning. *Information* **2021**, *12*, 96. <https://doi.org/10.3390/info12030096>

Academic Editors: Ramon Fabregat, Jorge Bacca-Acosta and N.D. Duque-Mendez

Received: 30 January 2021

Accepted: 17 February 2021

Published: 24 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Augmented Reality as a Tool to Support Learning

Augmented Reality (AR) technology allows to extend real environments with digital elements. This augmentation can be done by the integration of virtual computer-based objects in a real world environment in real time [1]. In the past, typically large head-mounted-displays were needed to use AR, but due to the wide spread of mobile devices, meanwhile it is possible to use AR with common smartphones or tablets [2,3]. Independent of the technical implementation, Azuma [1] defines three characteristics all AR systems have to fulfill. AR systems should present three-dimensional virtual objects on screen. Secondly, it should be possible to interact with these objects in real time on screen, e.g., to manipulate or to rotate them. The last and most important characteristic is that AR combines the real environment with a virtual world. This characteristic allows a differentiation between augmented reality and virtual reality, where the real world environment is completely masked out.

The most popular type of AR is the so-called marker-based AR [4]. In an educational context a marker is typically a two-dimensional image, printed in textbooks or other instructional materials. This image could be an artificial illustration like a QR-code or a symbolic representation like a chemical formula. Moreover, real-world objects can serve as markers [3]. Technically, a mobile device detects a marker with its camera and compares it with a defined database of markers and related virtual 3D objects. In case of a match, the specific virtual object appears on screen above or in any other spatial relation to the marker. Therefore, in marker-based AR approaches the virtual object seems to coexist with the underlying marker [4].

Several studies investigated the modes of AR-use and the benefits and potentials of AR-based learning. The majority of studies report that AR-tools are primarily used on the introductory stage of a new material or knowledge acquisition [3,4]. The aim is to aid students to explore and generalize specific concepts [3,5,6]. One benefit of AR compared to more conventional learning materials is the representation of three-dimensional virtual objects or animations and simulations of processes. This can aid students to recognize phenomena that cannot be seen in reality and to comprehend abstract concepts [2,6]. The possibilities of AR are not only limited to view static objects. The interaction with them is also possible in real time, e.g., to perceive the objects from different sides and angles [7]. Nielsen et al. [2] point out that the combination of multiple representations—namely, the virtual objects with the markers—can facilitate students' ability to experience phenomena that are otherwise impossible or infeasible.

Empirical research projects about different domains found out that students of different ages outperformed reference groups, who did not use AR-support [4,6,8,9]. In addition to learning achievement, positive effects on enhancing students' spatial abilities and learning engagement were also found [8]. What is especially interesting is the finding that AR seems to have a larger impact on low-achieving students than on high-achievers [6]. In addition, higher amounts of learning motivation due to AR-support were measured [2,4]. In interviews students reported to feel supported by the AR-use and the related AR-representations [2].

To obtain such advantages, AR-applications have to fulfil the requirement of usability. Some studies reported usability problems, if an AR-application does not work properly (e.g., the marker recognition) or if it is difficult to use [4,8,10]. Consequently, students may feel frustrated. Both aspects would undermine the principle of learning before technology by Nielsen et al. [2] and the demand of Bacca et al. [4] to take the usability of AR-applications into account.

The question of usability is also related to how the information is presented on screen. It is very important not to distract students by too many virtual objects or animations, which could cause distraction and an informational overflow [4,8]. If too much information appears on screen it can be assumed that the students' cognitive load during learning will increase consequently, leading to split-attention, which is not desirable at all in any learning setting. In addition, besides the usability in a technical sense, a question instructors are faced with during development, is how to design a suitable learning support and simultaneously to avoid cognitive overload [11,12]. For now, most of the studies concerning AR in science education focused on learning outcomes or affective variables like motivation or enjoyment and reported positive results [8,12–14]. On the contrary, only few studies investigated how AR affects the learners' cognitive load during learning [8]. In this case, diverse results are reported, which in some studies AR seems to reduce cognitive load, while other papers measured an increase [8,11,13]. Researchers agree that more research concerning the AR-use in educational settings is necessary [8,12,14].

An interesting subject area for further AR-research is the field of chemistry education, because AR seems to be able to bring out its potential within this field. The combination of different representations like pictures, text and virtual 3D objects can visualize phenomena, which usually cannot be perceived in a more conventional learning setting.

1.2. Representations in Science Education

Representations play an indispensable role in science communication [15–18]. To express chemical phenomena, there is huge variety of representations like wedge-dash-notations, ball-and-stick-models or symbolic representations like chemical equations or chemical structures [16,17,19,20]. Wu and Puntambekar [21] distinguish the following four forms of external representations with the given examples:

- Verbal–textual: This type of representation consists of written text and verbal information.
- Symbolic–mathematical: Examples from the context of chemistry are element-symbols, structural formulas, or ball-and-stick-models.

- Visual-graphical: This type of representations has a reference in their kind of representation to the real-life phenomena on the macroscopic level [20].
- Action-operational: Examples from the context of chemistry are hands-on experiments and inquiry-based actions.

In chemistry, scientists as well as learners deal with multiple external representations—mostly text and pictures—which are typically embedded in instructional materials, or scientific papers. The combination of pictorial and verbal representations has proven to be effective for complex cause-and-effect-systems, which often occur in sciences [22].

To enable learning, the external information of multiple representations have to be transformed into an internal mental model by each learner. A concept, which describes this transfer from the external towards the internal representation, is the Cognitive Theory of Multimedia Learning (CTML) [3]. This theory consists of three main assumptions and different principles to design multimedia learning. First, CTML assumes that visual and verbal/textual representations are perceived and processed by different cognitive channels. The second assumption states that the capacity of the working memory is limited, which is in line with the Cognitive Load Theory by Sweller and colleagues [22]. This means that each of the cognitive channels can only process a limited amount of information simultaneously. The third assumption of the theory foregrounds a constructivist understanding of learning in that the learner must actively process the relevant information for meaningful learning to occur. Therefore, he or she has to select relevant information, organize them into a coherent mental model and integrate this within already existing prior knowledge. After explaining these three main assumptions of the Cognitive Theory of Multimedia Learning, hereafter three selected principles for designing multimedia learning materials that are effective for learning shall be outlined briefly.

The multimedia principle states that presenting two representation types like texts and pictures combined, has proven to be effective for learning. Learners are able to build better mental connections between these two different kinds of representation [11]. This principle is a consequence of the assumption that different kinds of representations are processed in different cognitive channels and each channels' capacity is limited. Close to the multimedia principle there are two other principles that explain how the combination of the different types of media should be orchestrated.

The spatial contiguity principle states that text should be presented near by the corresponding picture or animation that it describes. If done so, learners do not have to invest lots of their limited cognitive capacity for searching information from somewhere else and matching it [11]. The temporal contiguity principle claims that corresponding narrations and animations should be presented at the same time for the same reason, so learners will be better able to make mental connections between words and pictures simultaneously.

In retrospect on Section 1.1, these three principles can be fulfilled by an Augmented Reality tool. The augmentation of the real world itself, provides a second type of representation in the learning setting (multimedia principle). If the AR is designed marker-based, the spatial contiguity—as well as the temporal contiguity principle are already considered by its definition. The marker-based approach specifies that a three-dimensional virtual object or animation is displayed on a mobile device screen, e.g., above a two-dimensional printed image, which serves as a marker. So, the user gets presented both kinds of representation spatially close to each other and at the same time.

The Cognitive Theory of Multimedia Learning shows that in order to integrate the organized mental model from the external representations into existing mental entities, it is necessary to have specific prior knowledge. At this point a problem arises, called the representation dilemma. On the one hand, students have to learn from representations to acquire new knowledge. On the other hand, they also need prior knowledge to learn about these representations themselves [18]. This possible dilemma is very relevant, especially for chemistry learning. Studies found out that the ability of visual model comprehension is a key factor as well as a predictor for learning success in chemistry [16,23].

A vivid example for this representation dilemma is the huge variety of chemical representations, which illustrate chemical phenomena on different levels of abstraction. Obviously to link these different levels of abstraction and to comprehend the chemical contents behind, it is indispensable to have representational competencies to understand each representation of the different abstraction levels and also to see the interconnections in-between [6,17,24]. Otherwise, this would result in insufficient content knowledge.

The dilemma of representational competencies applies also to organic chemical representations, because they convey even three-dimensional spatial information. The reason for this is that the element carbon—which is the key element in organic chemistry—binds itself mostly with four other elements and tries to maximize the distance between these substituents, to reach an energetic minimum. This results in a tetrahedral geometry. Therefore, the use of organic chemical representations requires cognitive processes in three-dimensional spatial domain, like recognizing spatial forms, transformations between different two-dimensional and three-dimensional representations of molecules, and manipulating them mentally in space [16,24]. For students it is crucial to have spatial and mental rotation abilities to be able to comprehend organic chemical representations and to avoid learning errors [16,17]. Research on spatial abilities in the field of organic chemistry showed that students with high spatial abilities outperformed students with low spatial abilities. Students with high spatial abilities were also able to express their mental models by creating and drawing proper representations with spatial information [25–27].

The ability of mental rotation is a variable characteristic that can be developed and advanced over lifetime by appropriate interventions and training [24,28–30]. Therefore, researchers demand chemistry faculties to help their students to become competent in domain specific spatial abilities and in mental rotation [17,31,32]. Harle and Towns [17] especially point out to train the transformation from two-dimensional images towards three-dimensional models. This relates towards the perception and cognitive processing of multiple external representations.

For the training of students in their representational competencies as well as their visuospatial thinking, Wu and Shah [16] suggest five design principles for chemistry-related visualization tools. The first principle demands for providing multiple representations and descriptions of the same information. The verbal explanation of the visual representations enhances students to interpret the visual representations verbally and fosters to use the chemical terminology. As a second design principle, Wu and Shah [16] recommend to make the referential connections between different representation types visible, to avoid students to create incorrect connections. A third principle calls for presenting the dynamic and interactive nature of chemistry. Wu and Shah [16] suggest animations, simulations, or video-clips for this purpose, to support learners in developing dynamic mental models, which have proven to be helpful for imagining dynamic chemical processes. Design principle number four demands to promote the translation between 2D and 3D representations. The problem arises that 2D representations mostly do not provide depth-cues like a structural formula does. To face this problem, features should be included in a visualization tool that facilitate the translation from 2D into 3D representations and vice versa, e.g., the rotation of the 3D model to compare 2D and 3D representations easily. The fifth and last design principle for designing a chemical visualization tool considers the cognitive load while learning. The necessary spatial ability for imagining and rotating three-dimensional chemical models arouses high amounts of cognitive load, so learners with low spatial abilities are in disadvantage. A supportive learning tool should reduce the demands on visuospatial processing resources by presenting visual and verbal information contiguously to enable learners to build systematic connections between both types of representations [16].

The five design principles of Wu and Shah [16] sum up the theoretical circumstances for learning organic chemistry discussed above, especially concerning multiple external representations and visuospatial representational competencies. A very promising approach for a possible learning environment, which could consider all five design principles

of a chemistry-related visualization tool, is the technique Augmented Reality (short AR). As AR is a new medium within the learning process, it has to be researched very properly to ensure that it is beneficial and not hindering for the specific learning purpose of organic chemistry. While designing an AR-application, the design principles deduced from the Cognitive Theory of Multimedia Learning by Mayer [33] as well as the design principles for chemistry-related visualization tools by Wu and Shah [16] should be considered. These principles intend to minimize unnecessary cognitive load in learning and appropriate AR-support can further promote this.

1.3. Aim of This Study and Research Questions

Between the mentioned findings, concerning the educational benefits of Augmented Reality and organic chemistry learning rises a gap in research. For now, it is not yet known how the use of AR affects the cognitive load in cognitive demanding topics of organic chemistry. This paper shall contribute to close the gap and answer the question, whether students perceive lower cognitive load while learning organic chemistry AR-based, and if there is an interrelation towards the perceived AR-usability or the students' mental rotation abilities. Therefore, this paper addresses the following research questions.

RQ1: Do students perceive a lower cognitive load when learning organic chemistry in an AR-based way compared to a control group?

RQ2: Is there an interrelation between the AR-usability and the students' perceived cognitive load?

RQ3: Do students perceive a lower cognitive load when learning organic chemistry in dependence to their mental rotation abilities?

2. Materials and Methods

We try to answer the three research questions as a subsection of a larger pre-/post study in an experimental-/control-group design. For the subject matter of the study, we chose three different topics of organic chemistry. These are stereochemistry, carbonyl-chemistry, and pericyclic reactions. We selected these three topics with care, concerning the research questions, because they demand for sophisticated mental rotation abilities.

2.1. Topics

To elaborate the topic stereochemistry properly, molecules in different ways of chemical representations have to be rotated around themselves at different angles on one or two axes. Depending on the task, they also have to be rotated intramolecularly around bonds between two different atoms. Carbonyl-chemistry is about chemical reactions of carbonyl compounds. For the understanding of these reactions, it is necessary to know at which position and in which angle a reactant can accumulate at another reactant three-dimensionally to initiate a chemical reaction. As the name suggests, also the topic of pericyclic reactions is about chemical reactions. Complex three-dimensional molecules accumulate at other reactants and react towards complex three-dimensional products. These accumulations take place laterally or also one above the other. During these reactions, some parts of the molecules also fold upwards or downwards. So, finally, complex three-dimensional products result from these reactions.

Due to these specific characteristics, it is obvious that the topics of stereochemistry, carbonyl-chemistry and pericyclic reactions demand for sophisticated mental rotation and spatial abilities to understand the contents and concepts behind. Therefore, they offer themselves to investigate the impact of AR while learning these topics.

2.2. Implementation

The sample we used for the study consists of university students at a German university, enrolled in the chemistry Bachelor of Science degree. To minimize the effort for the students to participate in our study, we conducted all parts of the study in the regular lecture session of an organic chemistry lecture. To do so, we collaborated with the lecture's

professor to implement our study in his lecture and coordinated the content. The coordination of content was straightforward because stereochemistry, carbonyl-chemistry, and pericyclic reactions are subject of this lecture regularly.

2.3. Time Table and Sample

In the first week of the semester, we conducted a self-developed pre-test. This test was used to measure the students' prior knowledge in stereochemistry, carbonyl-chemistry, and pericyclic reactions as well as in organic chemistry generally. We also gathered data concerning mental rotation abilities and demographic data.

After the pre-test, we divided the sample in an experimental and a control group. This was done with regard to the research questions to be able to measure possible differences between both groups. We assigned the students to the groups under consideration of their pre-test scores, to ensure to form two comparable groups. A multivariate analysis of variance of the scores of all parameters measured during the pre-test showed no significant differences between both groups.

During the semester, we conducted three interventions—one for each topic (see Figure 1). Some days after the pre-test, which was joined by $N = 61$ students (61.3% male, 37.1% female, 1.6% divers; age: $M = 21.43$; $SD = 2.78$) we intervened 41 students in stereochemistry. After 8 weeks the topic of carbonyl-chemistry followed for 27 students and another 6 weeks later at the end of the semester 21 student elaborated pericyclic reactions. Unfortunately, the number of students dropped sharply during the semester. Each intervention session took place during the lecture's regular schedule slot instead of the regular lecture at each day. During the interventions, group one (nonAR) worked with learning materials for the specific topic and group two worked with the same learning materials and they used Augmented Reality (AR) additionally. There was a working phase of about 60 min for each topic, where the students worked on their own on the given instructional material. Both groups worked at the same time, but in different rooms, to avoid jealousy effects for the AR-use in group one. Subsequent to the 60-min working phase, we measured the students' content knowledge of the specific topic elaborated in the intervention before with a post-test. Therefore, we used the same items from the pre-test to allow a pre-to-post-comparison. Additionally, we asked the students to rate their perceived cognitive load during the working phase. The students from group two, using augmented reality, rated the perceived usability, additionally.

2.4. Instruments

In order to conduct the study as described, we developed three different learning materials and AR-apps, each one for the topic stereochemistry, carbonyl-chemistry, and pericyclic reactions as well as corresponding test items.

All learning materials provide the topics' contents by external representations, namely by a combination of texts and pictures. Concerning the dual coding theory by Paivio [34] and the Cognitive Theory of Multimedia Learning by Mayer [23], we paid much attention on a content wise fitting of texts and pictures in the learning materials, to avoid an increase of extraneous cognitive load. The pictures in the learning materials serve as markers to trigger the AR. The learning material for stereochemistry provides 22 AR-markers, the material for carbonyl-chemistry 24 AR-markers and the material for pericyclic reactions 19 AR-markers.

Additionally to the learning materials, we developed an iOS-based app called Augmented Reality Chemistry (short ARC) with the software Unity-3D and the Vuforia-Framework. In the app, pictures, which are printed two-dimensionally in the learning materials, are linked with 3D models or 3D animations. These 3D models and animations are self-developed with the software Blender. The app was provided in the study on Apple iPads (6th generation). The participants can scan the markers in the learning materials with the iPad's camera, to trigger the AR. All displayed models or animations on screen allow user interaction in real time, e.g., scaling or rotating. All animations are played

in a continuous loop. The AR-models in the app, present the specific three-dimensional information given in the learning materials by text and pictures only two-dimensionally.

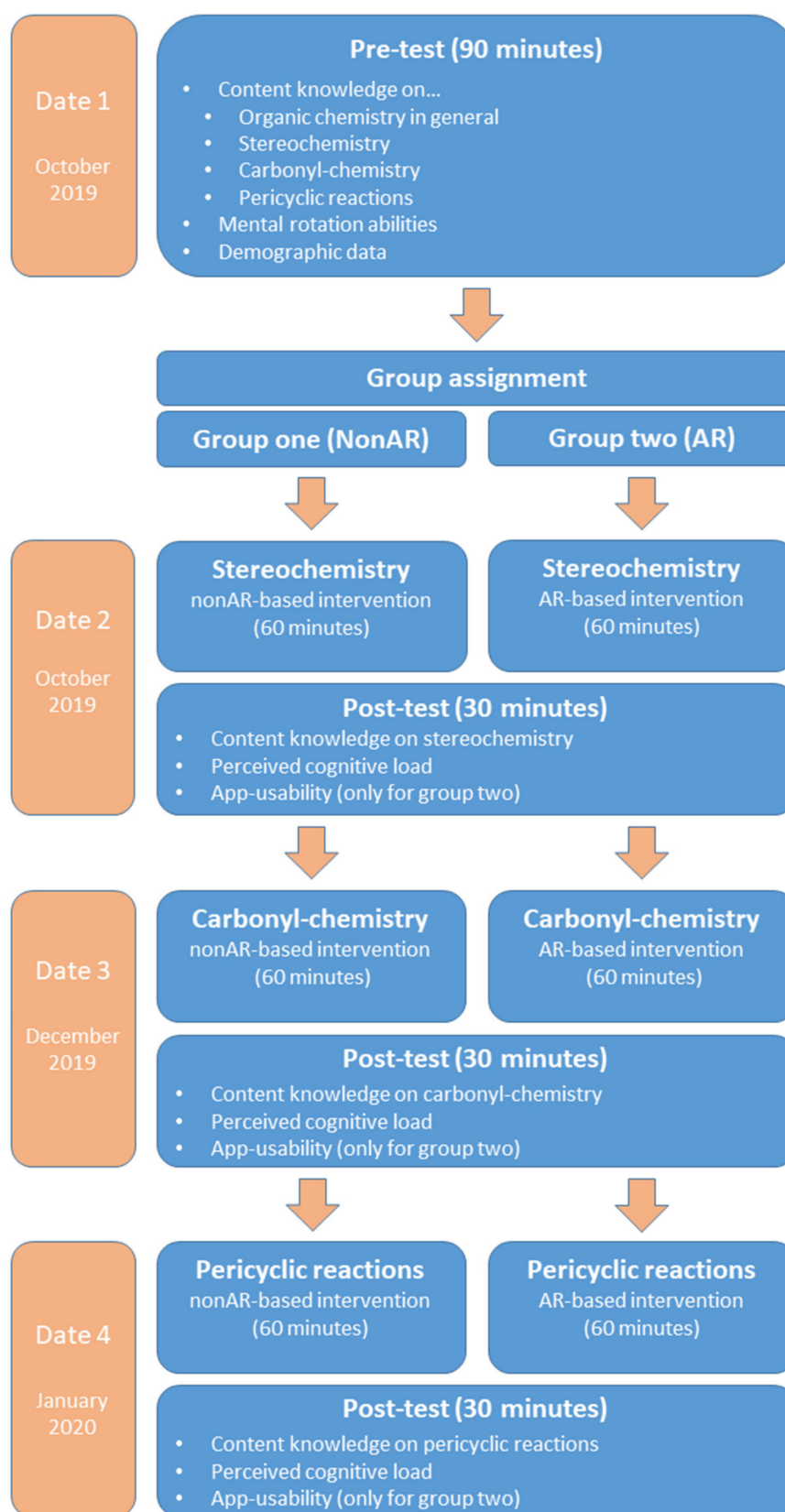


Figure 1. Timetable of the entire study | Used instruments are described in Section 2.4.

Figure 2 shows an example of the self-developed learning material, in this case for stereochemistry. The text describes the differences of organic molecules in eclipsed or staggered formation, which are both illustrated in wedge-dash-notation for the ethane molecule below. The red letters “AR” under the pictures invite the participants to scan this figure with the iPad’s camera to trigger the AR. In this case, a 3D animation of the ethane molecule appears in wedge-dash-notation on screen, which switches in a continuous loop from the eclipsed into the staggered formation and backwards. In this manner the effects of eclipsed and staggered formation can be clarified three-dimensionally.

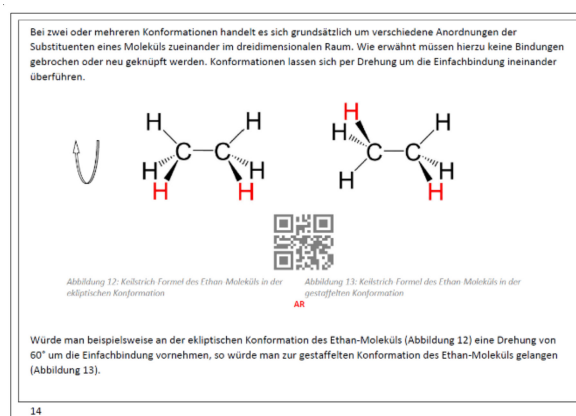


Figure 2. Example of the learning material about stereochemistry.

In Figure 3 the app ARC can be seen in use in combination with the learning material. The ball-and-stick-model of a hypothetical molecule on screen shall be rotated around its axes. This can be done by finger-gestures in real time on screen. The learning goal of this task is to illustrate and self-experience the principle of chirality, where two structural identical molecules behave like image and reflection but do not provide the same chemical characteristics. Figure 3 shows illustrative how the virtual three-dimensional object on screen seems to coexist with the two-dimensional picture on the paper.

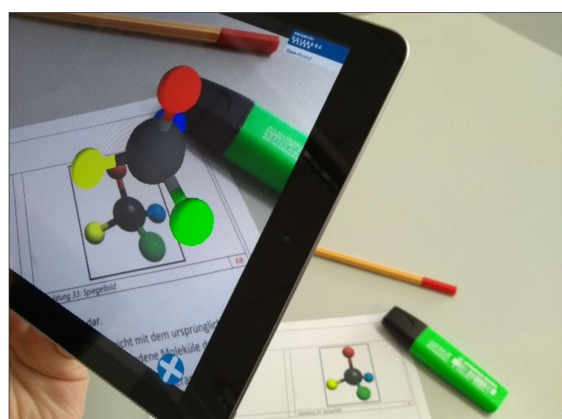


Figure 3. Example of the App Augmented Reality Chemistry (ARC) in use (rotateable 3D model).

As well as the pictures and the text in the learning materials, also the AR-models were developed according to typical textbooks for organic chemistry [35].

To present not only static objects in the app, Figure 4 shows another screenshot of the AR-app in use. This time, the app displays a chemical reaction from the learning material of the topic pericyclic reactions. The three-dimensional virtual objects on the screen play an animation in a continuous loop of the so-called Diels–Alder-reaction. During the 15 s lasting animation, both chemical educts accumulate to the chemical product of the reaction. In order to obtain the reactions’ product, the so-called orbitals of several atoms have to be

considered (red and blue shapes in Figure 4), which is why understanding the reaction is cognitive demanding. The big advantage of the AR-app here is that it does not only present the initial-state as well as the product-state of the reaction, but also the way from the educts to the product, which might be very useful for a reasonable understanding of the reactions concept.

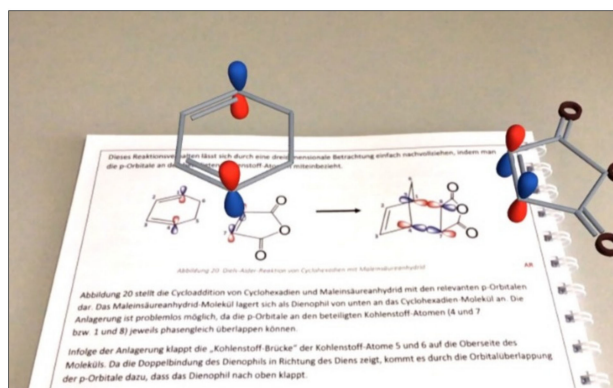


Figure 4. Example of the App ARC in use (3D animation of a chemical reaction).

Additionally to the descriptions of how the app ARC works, Supplemental Materials are offered. Via QR-codes at the end of this paper, two short video clips can be accessed, which illustrate the examples from Figures 3 and 4 in use.

In order to answer research question one, we measured the students' perceived cognitive load. Therefore, we used a 9-item questionnaire on a 6-point Likert-scale by Klepsch and colleagues [36]. This questionnaire is based on a self-assessment and distinguishes the cognitive load in three types. The intrinsic cognitive load describes the amount of load that learners need to invest to actively process the learning subject. In other words, the learning subject brings a specific complexity, which cannot be reduced without removing aspects. The second part of cognitive load is called extraneous cognitive load. It describes the cognitive effort needed to comprehend the learning material by its layout and design. An example for extraneous load is already given by Mayer [23] in his spatial contiguity principle. If a text and a related picture are placed on different pages of a book, a lot more extraneous cognitive load is necessary, to transfer the information from the text towards the picture and vice versa, compared to presenting text and picture spatially close to each other. The third component of cognitive load is the germane load. Under estimation that the cognitive capacity is limited, this is the part of cognitive capacity, a learner can invest in meaningful learning activities individually.

Concerning research question two, the participants rated the app-usability on the System-usability scale by Brooke [37]. This questionnaire was developed to evaluate the usability of several complex technical systems and has proven to be inexpensive, effective and highly reliable [38]. It consists of 10 items, each 5 with positive and negative statements alternating, which ask for agreement on a five-point Likert-scale from strongly agree to strongly disagree. Out of these 10 items a usability score for each participant can be calculated. This score covers a scale from zero, which indicates a very bad usability to 100, which means a perfect usability. Bangor, Kortum, and Miller [38] added an adjective description to the questionnaire, to make the usability score more interpretable. Depending on the peculiarity of the individual usability score, seven adjective ratings range from worst imaginable to best imaginable.

To measure the participants' performance in pre- and post-tests, we developed test items for each of the three topics. All these items are in a multiple choice single-select format. Each item has one correct answer and three distractors for choice.

For the topic of stereochemistry, six items ask to rotate a given molecule around its bonding-axis in wedge-dash-notation or in Newman-projection. In four items the participants have to rotate molecules represented as ball-and-stick-models on one axis and

in another four items on two axes. In another two items each, the same has to be done with molecules presented in wedge-dash-notation. Furthermore, the test includes two items to mirror a molecule and two items to determine the correct so-called absolute configuration of a molecule in the wedge-dash-notation.

We also developed 17 items to test the participants' knowledge in carbonyl-chemistry. The first two items present ball-and-stick-models, where the participants shall name the correct corresponding orbitals and an electrostatic potential map. The next three items are presented as ball-and-stick-models, where possible areas where reactants can accumulate in case of carbonyl chemical reactions have to be selected. Six items ask to identify the correct geometry of typical carbonyl-molecules or intermediates presented in wedge-dash-notations. Another six items show first steps of carbonyl-reactions. The participants shall predict the correct product of the reaction, depending on the three-dimensional geometry, e.g., of the reactants or the intermediates of the reaction.

Six of the 15 items for the topic of pericyclic reactions ask to name the correct orbital matching of two reactants of pericyclic reactions. Three items of them are each in two-dimensional and three-dimensional representation in the wedge-dash-notation. Nine items aim to predict the correct product of pericyclic reactions. Three items of them present two-dimensional structure formulas and another six items show the task in three-dimensional wedge-dash-notations.

For measuring the participants' general prior knowledge in organic chemistry, we used 10 items of the organic chemistry expertise test by Dickmann et al. [24]. The mental rotation abilities were measured with 12 items of the PURDUE Visualization of Rotations Test [39].

3. Results

3.1. Reliability Analysis

Before taking a closer look at the results, we analyzed the internal consistency of all scales used. For the 12 items of the PURDUE Visualization of Rotations Test, a Cronbach's alpha value of $\alpha = 0.719$ ($N = 55$) was measured. Table 1 shows the Cronbach's alpha values for the three subscales of the cognitive load questionnaire by Klepsch et al. [36] for each of the three topics.

Table 1. Reliability analysis of the cognitive load items.

	Intrinsic Load (α)	Extraneous Load (α)	Germane Load (α)
Stereochemistry ($n = 41$)	0.798	0.700	0.708
Carbonyl-chemistry ($n = 27$)	0.888	0.728	0.569
Pericyclic reactions ($n = 21$)	0.771	0.623	0.578

The results of the reliability analysis for the 10 items of the System-usability scale are reported in Table 2.

Table 2. Reliability analysis of the System-usability scale.

	System-Usability Scale (α)
Stereochemistry ($n = 19$)	0.797
Carbonyl-chemistry ($n = 15$)	0.671
Pericyclic reactions ($n = 13$)	0.718

The internal consistency of the cognitive load items seem to be overall acceptable. For the extraneous load of pericyclic reactions as well as for the germane load in carbonyl-chemistry and pericyclic reactions the Cronbach's alpha value is quite poor. The relatively small sample size could be an explanation for this. The reliability of the System-usability scale overall as well as of the PURDUE Visualization of Rotations Test are also acceptable. Therefore, the three scales can be used for further calculations.

3.2. Results for the Topic Stereochemistry

In order to answer the three research questions, we analyzed the perceived cognitive load, the students' rating on the app-usability and the correlations between different parameters. This section focusses on the topic stereochemistry. The means of the cognitive load rating are shown in Figure 5 divided by intrinsic, extraneous and germane load as well as between both groups. Tables 3 and 4 show the correlation matrices, separated for both groups. Included are the students' general abilities in organic chemistry (GenOC), their mental rotation abilities (Mental_Rotation), their scores in the pre-test (..._Pre), their scores in the post-tests (PostScore_...), their perceived intrinsic (InL_...), extraneous (ExL_...) and germane (GeL_...) cognitive load as well as their individual usability rating on the app ARC (SUS_...).

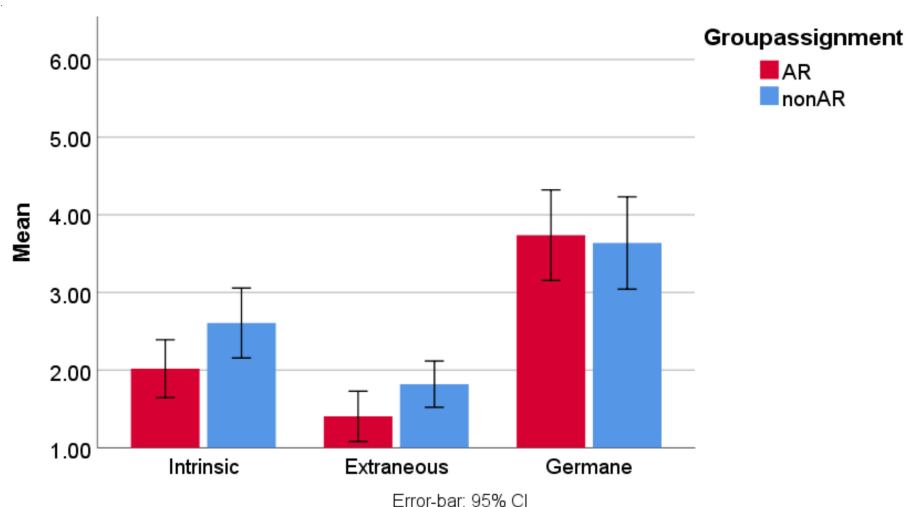


Figure 5. Cognitive load for the topic stereochemistry (n = 41).

Table 3. Correlations for group one (non-Augmented Reality (AR)) for the topic stereochemistry.

	GenOC	Mental_Rotation	S_Pre	PostScore_S	InL_S	ExL_S	GeL_S	SUS_S
GenOC								
Mental_Rotation	0.107							
S_Pre	0.378 *	0.546 **						
PostScore_S	0.478 *	0.683 **	0.845 **					
InL_S	−0.454	−0.035	−0.362	−0.200				
ExL_S	−0.469 *	−0.111	−0.123	−0.163	0.502 *			
GeL_S	−0.449	0.276	−0.203	−0.005	0.301	−0.071		
SUS_S	−	−	−	−	−	−	−	−

* The correlation is significant at the level of $p < 0.05$; ** The correlation is significant at the level of $p < 0.01$.

Table 4. Correlations for group two (AR) for the topic stereochemistry.

	GenOC	Mental_Rotation	S_Pre	PostScore_S	InL_S	ExL_S	GeL_S	SUS_S
GenOC								
Mental_Rotation	0.116							
S_Pre	0.177	0.601 *						
PostScore_S	0.248	0.030	0.818 **					
InL_S	−0.111	0.267	−0.234	−0.597 **				
ExL_S	0.129	−0.565 *	−0.040	−0.657 **	0.757 **			
GeL_S	0.073	0.474	0.153	0.151	0.566 *	0.153		
SUS_S	−0.195	0.306	−0.477	0.112	−0.262	−0.524 *	0.163	

* The correlation is significant at the level of $p < 0.05$; ** The correlation is significant at the level of $p < 0.01$.

The students of group two (AR) rated the usability of our application *ARC* for the topic stereochemistry with the System-usability scale by Brooke [37]. For each student a usability score was calculated. The mean usability score of all students for the topic stereochemistry is $M = 81.32$ ($SD = 14.08$).

Figure 5 shows that students, who worked with AR-support perceived a lower intrinsic as well as extraneous cognitive load, compared to students, who learned without using Augmented Reality on a descriptive level. However, these differences miss statistically significance $F(1, 39) = 1,791$, $p = 0.189$. The values for germane load of both groups are nearly at the same level.

Tables 3 and 4 present the correlations between several parameters. Some results for both groups align with existing findings from the literature. For example, there are strong-sized and significant correlations between the mental rotation abilities and the pre-test scores as well as between the pre-test- and the post-test scores [16,17].

However, there are also some differences between both groups. While in group one the abilities in organic chemistry in general correlate strongly and significantly with the students' results in the post-test, this connection is only small-sized and non-significant for group two. A similar finding appears for the connection between the student's mental rotation abilities and the post-test score. While both parameters correlate strongly and significantly in group one, this connection does not appear for group two. The results show a strong negative and significant correlation between the usability rating of the app *ARC* and the extraneous cognitive load. While in group one there are only small and non-significant negative connections between the post-test score and the intrinsic as well as with the extraneous cognitive load, these negative correlations are strong-sized and significant for group two. Additionally, interesting to point out are group differences for the connections between the organic chemistry abilities in general and the three types of cognitive load. While there are medium-sized negative correlations (for extraneous load even significant) in group one, these cannot be found for group two.

3.3. Results for the Topic Carbonyl-Chemistry

After investigating the cognitive load and several correlations for the topic stereochemistry, this section focusses on the evaluation of the second topic of the study, namely carbonyl-chemistry. The usability rating for the topic carbonyl-chemistry results in a usability score of $M = 89.17$ ($SD = 8.05$).

The three types of cognitive load displayed in Figure 6 differ between group one and two. Students of the experimental group reported a lower intrinsic as well as extraneous cognitive load, compared to the reference group. However, these findings remain non-significant $F(1, 25) = 2,608$, $p = 0.119$. The values for the germane cognitive load are on a comparable level for both groups, in little advantage for the reference group.

Again correlations were analyzed (see Tables 5 and 6). A group difference is visible for the correlation between mental rotation abilities and the post-test score. While in group one this correlation is medium-sized, it does not occur for group two. The correlation between the abilities in organic chemistry in general and the post-test score is a little stronger for group one than for group two. Comparing the correlations between mental rotation abilities and post-test score shows that for group one there is a medium-sized significant connection, which does not come up for group two. Furthermore, by contrasting both groups between mental rotation abilities and the three types of cognitive load, it is visible that the connections for group one are a little stronger, than for group two. As already occurred for the topic of stereochemistry, also for carbonyl-chemistry there is a strong-sized, negative, and significant correlation between the students usability rating on the app *ARC* and their perceived extraneous cognitive load. While for group one the calculation results in a strong-sized and significant correlation between pre-test- and post-test score for the topic of carbonyl-chemistry, this correlation does not appear for group two.

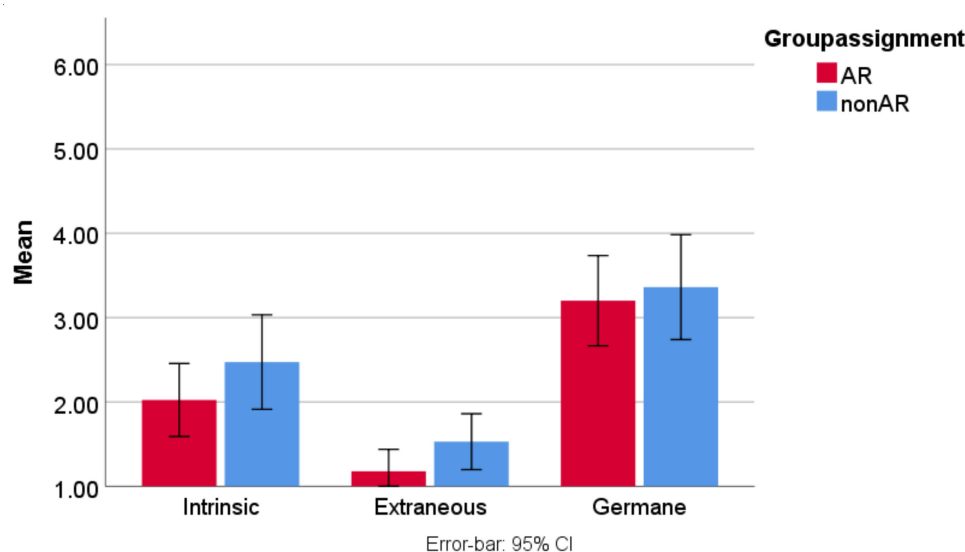


Figure 6. Cognitive load for the topic carbonyl-chemistry (n = 27).

Table 5. Correlations for group one (NonAR) for the topic carbonyl-chemistry.

	GenOC	Mental_Rotation	C_Pre	PostScore_C	InL_C	ExL_C	GeL_C	SUS_C
GenOC								
Mental_Rotation	0.107							
C_Pre	0.590 **	0.406 *						
PostScore_C	0.739 **	0.440	0.644 *					
InL_C	−0.252	0.286	−0.058	−0.312				
ExL_C	−0.227	0.333	0.224	−0.005	0.331			
GeL_C	−0.594 *	0.228	−0.502	−0.535	0.253	−0.110		
SUS_C	-	-	-	-	-	-	-	-

* The correlation is significant at the level of $p < 0.05$; ** The correlation is significant at the level of $p < 0.01$.

Table 6. Correlations for group two (AR) for the topic carbonyl-chemistry.

	GenOC	Mental_Rotation	C_Pre	PostScore_C	InL_C	ExL_C	GeL_C	SUS_C
GenOC								
Mental_Rotation	0.116							
C_Pre	−0.127	−0.053						
PostScore_C	0.600 *	0.014	0.090					
InL_C	0.080	0.124	0.405	−0.330				
ExL_C	0.198	0.252	0.097	0.268	0.335			
GeL_C	0.288	0.118	0.509	0.061	0.183	−0.154		
SUS_C	−0.521 *	0.280	0.095	−0.435	−0.120	−0.588 *	0.099	

* The correlation is significant at the level of $p < 0.05$.

3.4. Results for the Topic Pericyclic Reactions

In this section, the analysis of cognitive load and the correlation calculations focusses on the topic pericyclic reactions (see Tables 7 and 8). The usability score for this topic is $M = 90.63$ ($SD = 8.13$).

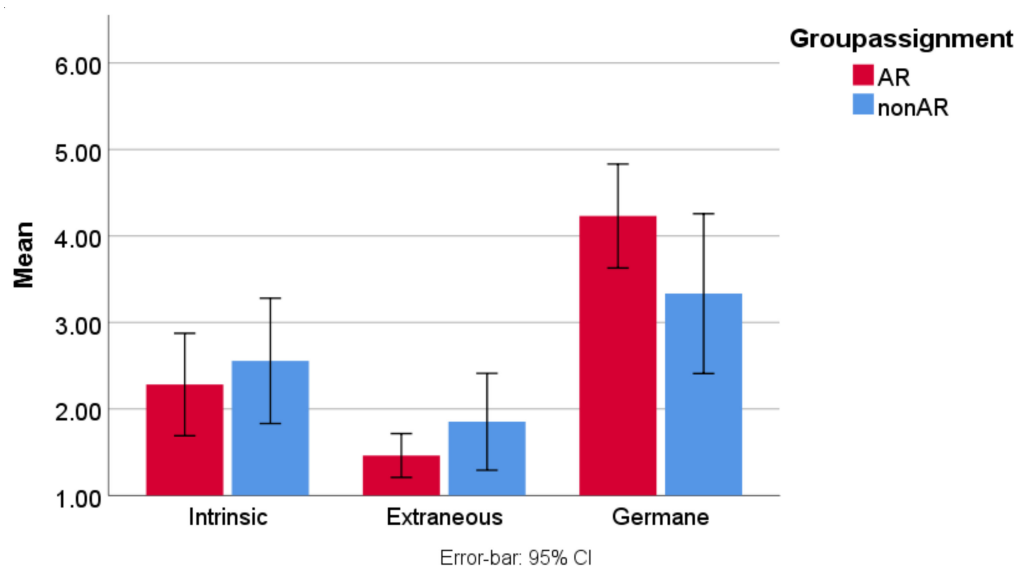
With regard to both other topics, a similar trend of group differences in cognitive load is visible in Figure 7 for the topic of pericyclic reactions. Again the students of group two rated the intrinsic as well as the extraneous cognitive load lower than the students in group one. For the germane load, the experimental group seems to be in small advantage on a descriptive level compared to the reference group $F(1, 20) = 0.083$, $p = 0.776$.

Table 7. Correlations for group one (NonAR) for the topic pericyclic reactions.

	GenOC	Mental_Rotation	P_Pre	PostScore_P	InL_P	ExL_P	GeL_P	SUS_P
GenOC								
Mental_Rotation	0.107							
P_Pre	0.333	0.041						
PostScore_P	0.411	0.870 *	0.452					
InL_P	−0.091	0.482	−0.257	0.571				
ExL_P	−0.089	0.305	0.414	0.147	−0.007			
GeL_P	−0.633	0.333	−0.361	0.346	0.625	0.285		
SUS_P	-	-	-	-	-	-	-	-

* The correlation is significant at the level of $p < 0.05$.**Table 8.** Correlations for group two (AR) for the topic pericyclic reactions.

	GenOC	Mental_Rotation	P_Pre	PostScore_P	InL_P	ExL_P	GeL_P	SUS_P
GenOC								
Mental_Rotation	0.116							
P_Pre	0.014	0.032						
PostScore_P	0.031	0.130	0.078					
InL_P	0.333	0.033	−0.016	0.083				
ExL_P	0.369	0.211	0.181	−0.064	0.310			
GeL_P	0.020	0.473	−0.140	0.248	0.184	−0.121		
SUS_P	−0.505	0.154	−0.393	0.150	0.023	−0.641 *	0.013	

* The correlation is significant at the level of $p < 0.05$.**Figure 7.** Cognitive load for the topic pericyclic reactions (n = 21).

Further, some findings in the correlations shall be mentioned for this third topic. The general abilities in organic chemistry or also the pre-test score correlates positively for group one with the post-test score. Moreover, a medium-sized and significant connection between the mental rotation abilities and the post-test score is visible for group one. In comparison with group two, these correlations do not appear at all. The correlations between the pre-test score or also the post-test scores with the three types of cognitive load turn out to be larger for group one, than for group two. Furthermore, by considering the correlations between the mental rotation abilities and the intrinsic as well as extraneous cognitive load, group two shows smaller correlations, compared to group one. Again, the app-usability correlates negatively and strong-sized with the extraneous cognitive load of the students.

4. Discussion

4.1. Research Question One

In this paper we raised three different research questions. The first question is: Do students perceive a lower cognitive load when learning organic chemistry AR-based, compared to a control group? The presented results show that students, who used AR-support perceived lower intrinsic as well as extraneous load in all of the three tested topics of organic chemistry, than the students in the control group did on a descriptive level. This means that they needed to invest a lower cognitive effort for understanding the topics itself (intrinsic load) and also to comprehend the instructional designs, including the representations of organic chemistry (extraneous load). This finding is expectable, with regard to the results of the meta-study by Ibáñez and Delgado-Kloos [8]. Properly designed learning materials including additional tools like AR, which, e.g., avoid split-attention effects could cause a difference in extraneous load, because learners only have to effort a lower amount of extraneous cognitive load in comprehending the materials, comparing to a reference group, like already reported by Altmeyer and colleagues [11]. The reduction of extraneous cognitive load is also a proof that it was successful, to add an additional representation type in the investigated learning setting, without causing a cognitive overload. However, what is pretty worth to point out, are the group differences also in intrinsic cognitive load. Usually it would be expectable that the intrinsic cognitive load, which describes the complexity of a learning topic itself, can hardly be influenced by any kind of learning support like the AR-use. Here in all three topics, learners of the experimental group perceived a lower intrinsic cognitive load, than the students of the reference group did. It has to be kept in mind that the only difference between both groups during elaborating the topics was using AR or not.

The amounts of intrinsic cognitive load of both groups do not change from the first to the second topic. The intrinsic cognitive load of group one is also for the third topic comparable with the first and the second one. However, group two reports a noticeable increase of intrinsic cognitive load for the topic pericyclic reactions. This increase shall be interpreted with regard to the germane cognitive load. For stereochemistry and carbonyl-chemistry, where the amounts of intrinsic cognitive load do not change relevantly, there are no group differences in germane cognitive load. However, for the topic of pericyclic reactions, where the AR group perceived an increase in intrinsic cognitive load, they outperformed group one clearly in germane cognitive load.

Keeping in mind that the topic of pericyclic reactions contains very abstracts concepts and is the most complex one of the three investigated topics in this study, this finding leads to the assumption that the impact of AR-based learning increases with the complexity of the learning topic. This assumption corresponds to the statement of Bacca and colleagues, “that AR is effective for teaching abstract or complex concepts” [4]. Nevertheless, the results of lower intrinsic and extraneous cognitive loads of group two, which are in line with existing literature, certify that the AR-use for stereochemistry and carbonyl-chemistry is not useless at all. However, significant group differences in affective and cognitive parameters, also with regard to the learning gains seem to be expectable for advanced topics only, even if Nechypurenko and colleagues found out in their review study that the most chemistry-related AR-tools are used on the introductory stage [3]. This open contradiction, whether AR-use is more effective on an introductory or on an advanced stage, demands for further research.

4.2. Research Question Two

As Bacca and colleagues [4] recommended to investigate the usability in AR-based learning settings, research question two asked for interrelations between the AR-usability and the students’ perceived cognitive load in our study.

The measured usability scores for all three topics are satisfying. On the scale from zero (bad) to 100 (best imaginable) the usability scores for the three topics fit into the best twenty percent. Applying the adjective rating scale by Bangor, Kortum, and Miller [38] the

usability score of stereochemistry can be interpreted as “good” and the usability scores for carbonyl-chemistry and pericyclic reactions can be interpreted as “excellent”. Although limited to the specific AR-application, these usability ratings are comparable to the results of a usability investigation with an AR-based learning setting in physics by Altmeyer and colleagues [11].

The further calculations showed significant and negative bivariate correlations for all dates of measurement between the usability scores, which represent the app-usability, and the perceived extraneous cognitive loads. These findings can be interpreted as the better the student rate the app-usability, the lower they perceive extraneous cognitive load. These connections underline the concerns of Bacca and colleagues [4] as well as Nielsen and colleagues [2] that inverted would mean that a severe designed AR-setting will cause a cognitive overload while learning. Fortunately, due to the good or even excellent usability scores as well as the low extraneous cognitive loads, this is not the case in this study. In addition, the results suggest that a separate measurement of extraneous, intrinsic, and germane cognitive load is useful to evaluate the effects of comparable learning supports. The fact that the usability scores consistently and exclusively correlate with the perceived extraneous load can be seen as a validity criterion for the separate measurement.

4.3. Research Question Three

Research question three asked for the influence of mental rotation abilities on cognitive load, when learning AR-based or not, respectively. The results of the correlations differ between the three topics. For stereochemistry, the correlations between mental rotation abilities and the intrinsic as well as extraneous cognitive load appeared to be contradictory with the existing literature. Ibáñez and Delgado-Kloos reported in their systematic review that most of the science-related AR-tools foster the learners in elaborating spatial abilities [8]. According to this it would be expectable that the correlations between mental rotation abilities and intrinsic or extraneous cognitive load would be larger for group one than for group two. The relatively small sample size (group one: $N = 22$; group two: $N = 19$) could be an explanation, why the mental rotation abilities have a greater influence on the cognitive load when elaborating stereochemistry AR-based, than in a more traditional way. For carbonyl-chemistry as well as for pericyclic reactions the correlations between the parameters were smaller for students, who worked AR-based, compared to group one. Here the supporting character of AR already found in other publications is apparent [6,8,10], because the influence of mental rotation abilities on the cognitive load decreased due to AR-use.

However, the question remains still open to why the results for carbonyl-chemistry and pericyclic reactions align with the insights of earlier research, but for the topic stereochemistry, they do not. The differences between the three topics could be explained by considering the sample of the study. The students in the lecture “Organische Chemie II” (organic chemistry II) typically succeeded already in the basic-lecture of organic chemistry. Typically, stereochemistry is a topic on the introductory level of organic chemistry and is sometimes even part of general chemistry lectures. With regard to Tuckey and colleagues [25], who stated that mental rotation abilities could be advanced by training, it can be assumed that some kind of training effect has already taken place due to participating in the basic-lecture of organic chemistry. In contrast, carbonyl-chemistry and pericyclic reactions typically are advanced topics in organic chemistry, including complex and quite abstract concepts. It can be assumed that the students of the investigated sample did not meet those concepts before. In light of the discussion of research question one, there was the assumption that AR-based learning becomes more effective with increasing complexity, a similar trend becomes apparent here. While the correlation between mental rotation abilities and intrinsic cognitive load increases for group one for carbonyl-chemistry as well as for pericyclic reactions, this correlation does not increase for group two. In conclusion on research question three, the results indicate that using AR can reduce the influence of

mental rotation abilities on the learners' cognitive load, especially if topics are discussed, the students are faced with for the very first time.

4.4. Discussion on Further Findings

In addition to the three research questions, the correlation analyses also showed some other interesting findings. Next to the investigation of mental rotation abilities and their connection to cognitive load, it seems also worth to examine their connection to the post-test scores. The calculations showed that there are medium to strong-sized correlations between the mental rotation abilities and the post-test scores for group one. For stereochemistry and pericyclic reactions, these are even significant. In contrast, these correlations do not appear for the students of group two. So, besides the influence on cognitive load, augmented reality seems also be able to reduce the influence of mental rotation abilities on the learning gains [16,17]. This finding is very promising to overcome difficulties for low-achievers in mental rotation when learning complex chemical concepts and also, to overcome possible sex-differences in mental rotation abilities, as pointed out by Halpern and Collaer [20,33]. For further research it would be very interesting to conduct a follow-up test on mental rotation abilities after several AR-uses, to investigate, whether the AR-learning improved the learners mental rotation abilities sustainably [16,17].

Another aspect, which is also interesting to consider, are the organic chemistry abilities in general and their effect on the post-test scores. As already found by Dickmann [24], also in our study for the group one, the prior knowledge was a relevant predictor for post-test success, while it was not for group two. So, it could be possible that AR-based learning also seems to be able to reduce the influence of prior knowledge in organic chemistry on the learning success.

4.5. Limitations and Outlook

Although the findings of this study are promising, some limitations have to be mentioned. The most important limitation of our study is the relatively small sample size we based our analyses on. This small sample size is also a result of the huge dropout, which unfortunately took place from the pre-test (62 students) over the semester towards the third intervention (27 students). Larger sample sizes always enable to make more meaningful statements. Therefore, we aim to re-run the study gain a much larger sample size. This will allow to replicate and validate the obtained results. Once the data collection will be completed, we plan to measure also learning gains from pre- to post-test and also possible group differences in post-test scores to investigate, whether the AR-support will affect also the students learning gains. For a replication study, it is also conceivable to extend the AR-based learning to other topics of organic chemistry.

4.6. Conclusions

Our self-developed learning settings fulfil the five design principles for chemistry-related visualization tools by Wu and Shah [16]. Therefore, our AR-app offers itself as supplemental medium to convey the three investigated topics and also to train representational competencies as well as visuospatial thinking. The obtained results indicate that learning the topics stereochemistry, carbonyl-chemistry, and pericyclic reactions AR-based, reduces the intrinsic as well as extraneous cognitive load compared to a reference group. The usability of the AR has proven to be a key-factor on extraneous cognitive load and is closely linked with it. Therefore, a good usability should be ensured, when developing AR-based learning settings. By using AR while learning chemistry topics, also the influence of prior knowledge in organic chemistry as well as mental rotation abilities on the cognitive load as well as on the post-test results could be reduced. Our results suggest that especially for complex and abstract concepts, the AR-use becomes more and more effective.

Supplementary Materials: The following are available online. The scaling and rotating of virtual three-dimensional models on screen explained in Figure 3, can be accessed as a short video clip via <https://youtu.be/wx4vqcsIaE4>, accessed on 23 February 2021. To get an impression of how anima-

tions of complex three-dimensional chemical reactions, like the Diels–Alder-Reaction explained in Figure 4, are included in the AR-app please click on the following link: <https://youtu.be/-EkdaVhvCQs>, accessed on 23 February 2021.

Author Contributions: Conceptualization, S.K. and S.H.; methodology, S.K. and S.H.; software, S.K. and S.H.; validation, S.K. and S.H.; formal analysis, S.K.; investigation, S.K.; data curation, S.K. and S.H.; writing—original draft preparation, S. K. and S.H.; writing—review and editing, S.H. and S.R.; visualization, S.K.; supervision, S.H. and S.R.; project administration, S.H.; funding acquisition, S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (DFG), grant number 415026237.

Institutional Review Board Statement: Ethical review and approval were waived for this study, because all subjects participated in the study voluntarily and without further constraints. To avoid disadvantages, all participants were offered to review the learning material after the end of the study.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is contained within the article. For further questions, please contact the authors.

Acknowledgments: We thank Alper Öztürk and Christoph Pelka (both University of Duisburg-Essen) for assistance with the app-development and technical support. We would also like to thank the reviewers whose comments have greatly improved the manuscript.

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

References

1. Azuma, R.T. A survey of augmented reality. *Presence: Teleoperators Virtual Environ.* **1997**, *6*, 355–385. [\[CrossRef\]](#)
2. Nielsen, B.L.; Brandt, H.; Swensen, H. Augmented reality in science education—affordances for student learning. *Nord. Stud. Sci. Educ.* **2016**, *12*, 157–174. [\[CrossRef\]](#)
3. Nechypurenko, P.P.; Starova, T.V.; Selivanova, T.V.; Tomilina, A.O.; Uchitel, A.D. Use of augmented reality in chemistry education. In Proceedings of the 1st International Workshop on Augmented Reality in Education, Kryvyi Rih, Ukraine, 2 October 2018; Kiv, A.E., Soloviev, V.N., Eds.; CEUR: Kryvyi Rih, Ukraine, 2018.
4. Bacca, J.; Baldiris, S.; Fabregat, R.; Graf, S. Augmented Reality trends in education: A systematic review of research and applications. *J. Educ. Technol. Soc.* **2014**, *17*, 133–149.
5. Bujak, K.R.; Radu, I.; Catrambone, R.; Macintyre, B.; Zheng, R.; Golubski, G. A psychological perspective on augmented reality in the mathematics classroom. *Comput. Educ.* **2013**, *68*, 536–544. [\[CrossRef\]](#)
6. Cai, S.; Wang, X.; Chiang, F.K. A case study of augmented reality simulation system application in a chemistry course. *Comput. Hum. Behav.* **2014**, *37*, 31–40. [\[CrossRef\]](#)
7. Núñez, M.; Quirós, R.; Núñez, I.; Carda, J.B.; Camahort, E. Collaborative Augmented Reality for Inorganic Chemistry Education. In Proceedings of the EE'08, 5th WSEAS/IASME International Conference on Engineering Education, Heraklion, Greece, 22–24 July 2008; Point, S., Ed.; World Scientific and Engineering Academy and Society: Heraklion, Greece, 2008; pp. 271–277.
8. Ibáñez, M.; Delgado-Kloos, C. Augmented reality for STEM learning: A systematic review. *Comput. Educ.* **2018**, *123*, 109–123. [\[CrossRef\]](#)
9. Cai, S.; Liu, E.; Yang, Y.; Liang, J.C. Tablet-based AR technology: Impacts on student's conceptions and approaches to learning mathematics according to their self-efficacy. *Br. J. Educ. Technol.* **2019**, *50*, 248–263. [\[CrossRef\]](#)
10. Radu, I. Augmented reality in education: A meta-review and cross-media analysis. *Pers. Ubiquitous Comput.* **2014**, *18*, 1533–1543. [\[CrossRef\]](#)
11. Altmeyer, K.; Kapp, S.; Thees, M.; Malone, S.; Kuhn, J.; Brünken, R. The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *Br. J. Educ. Technol.* **2020**, *51*, 611–628. [\[CrossRef\]](#)
12. Tzima, S.; Styliaras, G.; Bassounas, A. Augmented Reality Applications in Education: Teachers Point of View. *Educ. Sci.* **2019**, *9*, 99. [\[CrossRef\]](#)
13. Lai, A.F.; Chen, C.H.; Lee, G.Y. An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *Br. J. Educ. Technol.* **2018**, *50*, 232–247. [\[CrossRef\]](#)
14. Lasica, I.E.; Meletiου-Mavrotheris, M.; Katzis, K. Augmented Reality in Lower Secondary Education: A Teacher Professional Development Program in Cyprus and Greece. *Educ. Sci.* **2020**, *10*, 121. [\[CrossRef\]](#)
15. Lemke, J. Multiplying meaning: Visual and verbal semiotics in scientific text. In *Reading Science: Critical and Functional Perspectives on Discourses of Science*; Martin, J.R., Vell, R., Eds.; Routledge: New York, NY, USA, 1998; pp. 87–113.
16. Wu, H.K.; Shah, P. Exploring visuospatial thinking in chemistry learning. *Sci. Educ.* **2004**, *88*, 465–492. [\[CrossRef\]](#)

17. Harle, M.; Towns, M.A. Review of spatial ability literature, its connection to chemistry, and implications for instructions. *J. Chem. Educ.* **2011**, *88*, 351–360. [[CrossRef](#)]
18. Rau, M.A. Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educ. Psychol. Rev.* **2017**, *29*, 717–761. [[CrossRef](#)]
19. Hoffmann, R.; Laszlo, R. Representation in chemistry. *Angew. Chem.* **1991**, *30*, 1–16. [[CrossRef](#)]
20. Habig, S. Who can benefit from augmented reality in chemistry? Sex differences in solving stereochemistry problems using augmented reality. *Br. J. Educ. Technol.* **2020**, *51*, 629–644. [[CrossRef](#)]
21. Wu, H.K.; Puntambekar, S. Pedagogical affordances of multiple external representations in scientific processes. *J. Sci. Educ. Technol.* **2012**, *21*, 754–767. [[CrossRef](#)]
22. Castro-Alonso, J.C.; Ayres, P.; Sweller, J. Instructional visualizations, cognitive load theory, and visuospatial processing. In *Visuospatial Processing for Education in Health and natural Sciences*; Castro-Alonso, J.C., Ed.; Springer: Cham, Switzerland, 2019; pp. 111–143.
23. Dickmann, T.; Opfermann, M.; Dammann, E.; Lang, M.; Rumann, S. What you see is what you learn? The role of visual model comprehension for academic success in chemistry. *Chem. Educ. Res. Pract.* **2019**, *20*, 804–820. [[CrossRef](#)]
24. Tuckey, H.; Selvaratnam, J.; Bradley, J. Identification and rectification of student difficulties concerning three-dimensional structures, rotation, and reflection. *J. Chem. Educ.* **1991**, *68*, 460–464. [[CrossRef](#)]
25. Carter, C.S.; LaRussa, M.A.; Bodner, G.M. A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *J. Res. Sci. Teach.* **1987**, *24*, 645–657. [[CrossRef](#)]
26. Pribyl, J.R.; Bodner, G.M. Spatial ability and its role in organic chemistry; A study of four organic courses. *J. Res. Sci. Teach.* **1987**, *24*, 229–240. [[CrossRef](#)]
27. Hegarty, M.; Waller, D.A. Dissociation between Mental Rotation and Perspective-Taking Spatial Abilities. *Intelligence* **2004**, *32*, 175–191. [[CrossRef](#)]
28. Mohler, J.L. Computer Graphics Education: Where and how do we develop spatial Ability? In Proceedings of the Eurographics, Vienna, Austria, 4–8 September 2006; Education Papers; Eurographics: Lausanne, Switzerland, 2006; pp. 79–86.
29. Stieff, M. Mental rotation and diagrammatic reasoning in science. *Learn. Instr.* **2007**, *17*, 219–324. [[CrossRef](#)]
30. Terlecki, M.; Newcombe, N.; Little, M. Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Appl. Cogn. Psychol.* **2008**, *22*, 996–1013. [[CrossRef](#)]
31. Small, M.Y.; Morton, M.E. Research in college science teaching: Spatial visualization training improves performance in organic chemistry. *J. Coll. Sci. Teach.* **1983**, *13*, 41–43.
32. Halpern, D.F.; Collaer, M.L. Sex Differences in Visuospatial Abilities: More Than Meets the Eye. In *The Cambridge Handbook of Visuospatial Thinking*; Shah, P., Miyake, A., Eds.; Cambridge University Press: New York, NY, USA, 2005; pp. 170–212.
33. Mayer, R.E. *The Cambridge Handbook of Multimedia Learning*; Cambridge University Press: Cambridge, UK, 2014.
34. Paivio, A. *Mental Representations: A Dual Coding Approach*; Oxford University Press: Oxford, UK, 1990.
35. Schmuck, C. *Organische Chemie*, 2nd ed.; Pearson-Studium: Munich, Germany, 2018.
36. Klepsch, M.; Schmitz, F.; Seufert, T. Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Front. Psychol.* **2017**, *8*, 1–18. [[CrossRef](#)]
37. Brooke, J. SUS-A quick and dirty usability scale. In *Usability Evaluation in Industry*; Jordan, P.W., Thomas, B., Weerdmesster, B.A., McClelland, A.L., Eds.; CRC-Press: New York, NY, USA, 1996; pp. 18–194.
38. Bangor, A.; Kortum, P.; Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.* **2007**, *4*, 114–123.
39. Bodner, G.M.; Guay, R.B. The purdue visualization of rotations test. *Chem. Educ.* **1997**, *2*, 1–17. [[CrossRef](#)]