# Learning Anatomy with Augmented Reality

## – learning design and app design for optimal learning

**Cita Nørgård**
*Specialkonsulent*
SDU Universitetspædagogik.

**Lotte Dyhrberg O'Neill**
*Lektor*
SDU universitetspædagogik.

**John Chemnitz**
*Lektor*
IMM – neurobiologisk forskning, SDU

**Gunver Majgaard**
*Lektor*
MMMI – Embodied systems for Robotics and learning, SDU

# Abstract

An Augmented Reality (AR) application was developed to help students at SDU to learn the anatomy of the human body (mediastinum). This research project intended to evaluate whether AR strengthened the students' self-efficacy and motivation, improved learning, and provided a good learning experience.

This study focusses on how AR can help students to translate two-dimensional into three-dimensional understanding and evaluates formats of the app (the use of quizzes) together with the didactic design of the teaching sessions with AR. The objectives were to examine A. the effectiveness of using AR on student's short, long term, and transfer learning outcome compared with traditional teaching, B. the effect of quizzes, and C. app design in relation to didactic design.

In the AR world students saw a standardized hologram body combined with selected images from a computed tomography (CT) scan. The CT images were presented at the proper spatial positions in a hologram body. The user could select which structures to be shown at the body (e.g. skin, the vascular system, etc.). They could approach the body, circumvent it, study structures in details and compare the concurrent presentation with the different CT images.

During class one group received traditional teaching and two groups studied CT scans in Augmented Reality. One AR-group complemented the AR app with quiz questions and received corrective feedback. The other AR group did not use quiz questions but studied content on their own hand. A teacher was present and ready to help students in all groups.

The aim of this paper is to share lessons learned from this intervention and suggest solutions for app design and learning design to facilitate deeper learning processes and scaffold learners' needs with AR as a learning resource.


Keywords: Augmented Reality, app design, Quizzes, spatial conception, anatomy, presence, motivation, self-efficacy, learning design, role of the teacher

# Introduction

## Background

At the University of Southern Denmark (SDU), we decided to explore what additional value Augmented Reality (AR) could offer for medical students' understanding of anatomy. AR is here understood as the physical environment with at digital layer on top. HoloLens provides a 3-dimentional digital layer so-called holograms (Majgaard, 2018).

In a recent review by Akçayır and Akçayır (2017) it is discussed that academic achievement has been examined in studies with AR as a tool for learning. Also, other perspectives like the novelty of technology have been a subject for research. Yet exemplars of teaching apps and didactic design is scarce because AR technology is covering the range from hand held devices to head mounted devices (HMD), and because technology evolves rapidly. When we examine the effects of AR on learning we are interested in students' development of self-efficacy amongst other parameters. Self-efficacy is the belief that one can execute needed steps to achieve a goal (Bandura, 1977). Other factors such as the student's motivation and self-efficacy have over the last years been investigated in relation to virtual lab simulations; for instance, by Makransky, Thisgaard, and Gadegaard (2016) and in Bacca, Baldiris, Fabregat, Graf, and Kinshuk (2014). Despite the obvious differences there are many similarities between virtual labs, VR and HMD-AR learning situations and publications for these areas are of great interest across these technologies. Makransky et al. (2016) showed promising results for preparing in virtual labs.

Developing an app is expensive and time consuming and a suitable application for this HoloLens experiment was developed by a group of engineering students. The special interest in this study was the stimulation of two-dimensional into three-dimensional understanding and comparing teaching setups with and without quizzes between the intervention groups with reference to a control group. The anatomical content was found to be relevant to the HoloLens spatial presentation and correlated the two-dimensional CT images to a 3D body hologram. A small part of a third semester course ”Circulation and Respiration” focuses on the anatomy of the mediastinum, a midline partition of the thoracic cavity from sternum anteriorly to the thoracic vertebrae posteriorly, and from the superior thoracic aperture to the inferior thoracic aperture. The students should learn to use the correct adverbs of anatomical orientation and descriptive terminology and anatomical nomenclature according to Federative international program on anatomical terminologies ”Terminologia Anatomica”.

The limitation of plain radiographic techniques is the 2-dimensional representation of 3-dimensional structures. Generally, CT-studies are performed with the patient/person supine and images are obtained in the transverse or axial plane. 2-dimensional axial CT-images are conventionally rendered so that the view is as though looking up at it from the persons feet. The accepted conventional viewing of axial CT-images helps the students' spatial orientation concerning the orientation of left-right and anterior-posterior.

CT images are very thin and appear as very "flat" axial images which does not give the student the impression of spatial structures above and below the actual transection. Interpretation of CT images include viewing the consecutive axial sections up and down (superior-inferior). This gives the students the last dimension for understanding the 3-D appearance, by superimposition of images of structures outside the area. In our study we have used axial CT images of the thoracic cavity representing consecutive sections and the idea was to supplement student's scrolling up and down the images with the Hologram body view to facilitate spatial learning. Thus, content was added to HoloLens to examine if the virtual world added the desired spatial understanding and improved learning, motivation and self-efficacy for students. These were factors indicated suitable for measuring AR-learning in Bacca et al. (2014) and Makransky et al. (2016) and we consider these as in good correspondence with the underlying principle for active teaching and learning at SDU (Underlying Principles).

In this study we intended to allow students to "see the unseen" by looking into the body at a 3D hologram (Dunleavy, 2014), explore feedback functionalities by quizzes build in the app (Bacca et al., 2014), and search for applicable learning designs. We identified students' motivation, self-efficacy, and learning outcome as well as parameters on app quality (by measures of presence and evaluations of students' user experience).

To agree on the rationale programmers, course teachers, learning researcher, and project-leaders were collaborating on the project from the start to align APP production with learning goals and project research. The way the app should be used in the teaching scenario was chosen from a small catalogue of possible teaching setups created by the Centre for Teaching and Learning at SDU though partly limited by the already scheduled course curriculum. Other limitations were the limited number of HoloLens available, teacher resources, and students' unfamiliarity with HoloLens. Regarding the learning situation we aimed at testing the effect of using quizzes in the app due to research on effectiveness of feedback and the types of feedback (Moreno, 2004). Explanatory feedback would have been preferable, but because of the contextual limitations mentioned above, we agreed on corrective feedback. During the development of the app, we discussed highlighting the correct organ system of the 3D body for clueing the right answer, but we argued that such clueing would circumvent important reflections for students.

### Research objectives

The objectives were to examine

- the effectiveness of using AR on student's short term, long term, and transfer learning outcome compared with traditional teaching.
- the effect of quizzes
- app design in relation to didactic design

## Methods

### Teaching setup

When students arrived at the exercise classes they divided themselves into 4 groups. Each group consisted of approximately 15 students. The 15 students circulated between 4 stations with a different exercise at each station. Each station lasted for 35 minutes.  One of the stations was the mediastinum station where the intervention took place. Thus, content at the station in question was only a smaller part of the course content.
The purpose of the mediastinum station was to train students to recognize and identify anatomical structures on CT images of the mediastinum. In each rotation 7 students arriving at the station were randomly selected to join the station version with HoloLens and the rest (8 students) received traditional teacher led introduction to the CT images with a PowerPoint show. CT images in the HoloLens app and the PowerPoint was the same. The CT images was a selection of 10 pictures from a healthy male individual who had given his permission to use the scan for educational purposes. The exercises were run over two days. The first day the students at the HoloLens station used the quizzes in the HoloLens app. The second day students were not allowed to use the quizzes. This resulted in the following three comparable groups (fig. 1). Students who didn't attend the teaching session also received the questionnaire thus forming a fourth group.

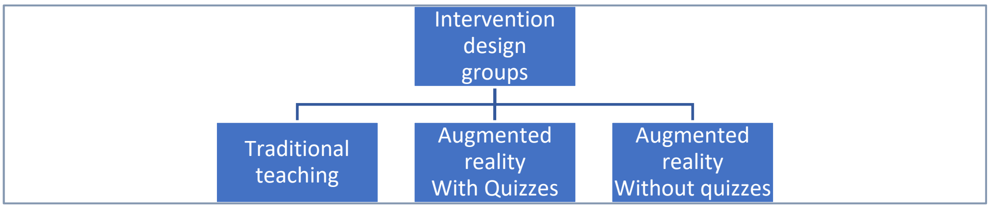


*Figure 1. Intervention design groups*

Seven pairs of HoloLens were available and because this was a new experiment with only 35 minutes at the stand each HoloLens was supported by a student technician to avoid waste of time due to HoloLens crashes. Two crashes appeared day one and none day two. Importance if technicians preventing valuable time consumption during teaching as described in literature (Munoz-Cristobal et al., 2015).
The teacher (a student instructor) was made familiar with the HoloLens app before the exercise. Every set of HoloLens had a wireless connection to a computer showing the students view to enable

the instructor to follow the students' progress and to enable her and the specific student to communicate with direct reference to the student's view inside the virtual room.
Upon arrival at the mediastinum HoloLens stand students were made familiar with the expected learning outcome. Goal was to build up an awareness of named structures and orientation shown at the CT images. They received instruction to the HoloLens app facilities, the clicker control, and were encouraged to study the 3D body in combination with the CT images.

## App design for the test

In the HoloLens app the selected CT images was shown together with a 3D hologram of a whole male body. The 3D body showed all organ systems relevant to the CT image including those relevant for the quiz question. A picture with quiz questions asking to identify structures shown by arrow A and B was shown left in the virtual world. Next to the 3D body the "quiz CT image" is repeated (red) and a CT image of the students' own choice (green) is shown right to the body to make comparison easy. Students could use the green picture to scan up and down the body to view continua of different structures. The 3D body is visualizing the "unseen" and function as reference for studying structures and recognizing spatial orientation of the CT images (fig. 2).

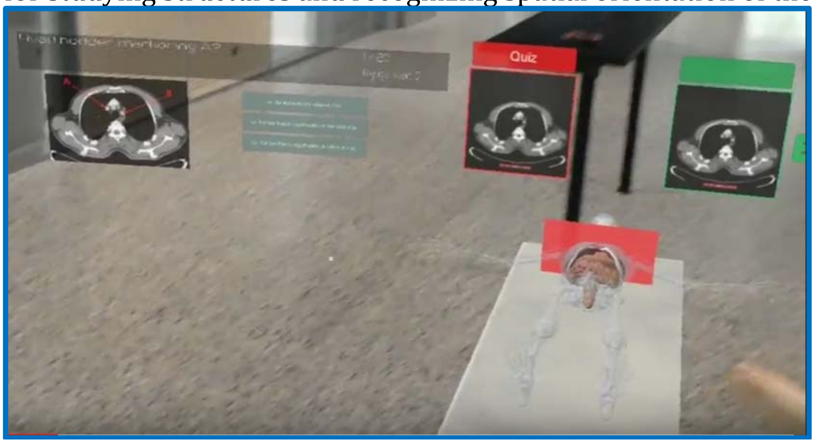


*Figure 2. The quiz is shown to the left. The transection regarding the quiz is shown in red color and the movable transection for comparison in green to the right of 3D male body*

The Multiple-Choice Question (MCQ) quizzes contain three possible answers to each question and students receive corrective feedback. Left picture turns red if answer is wrong and green if answer is correct. A list at the top sums up the correct and wrong answers. When student answer correctly system turns on to the next question.
On the body the red and green positions are repeated for sensation of corresponding CT image positioning (fig. 3). When students approach the body, they see the structures in more details.

 ISSN: 1903-248X

*Figure 3. a closer look at the 3D male body. Quiz transection (red) and movable transection (green)*

The tools at the right hand in the augmented room (fig. 4 and 5) represented opportunities to either turn off, highlight, shadow, or reset different organ systems in the 3D body to get a clearer view of different systems.

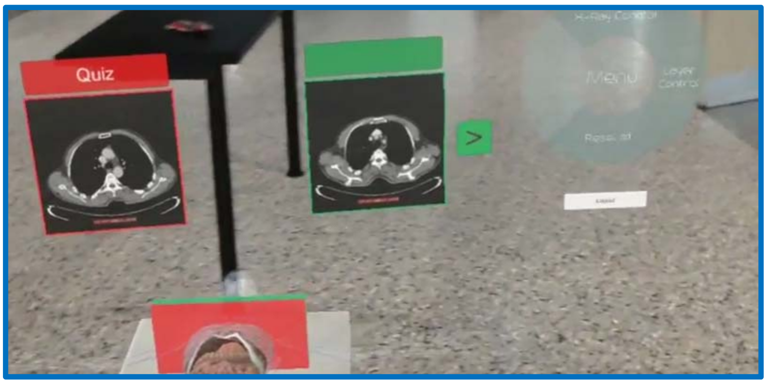


*Figure 4. At the right in the virtual room students find the menu for selecting anatomical structures to be shown in the 3D body hologram*

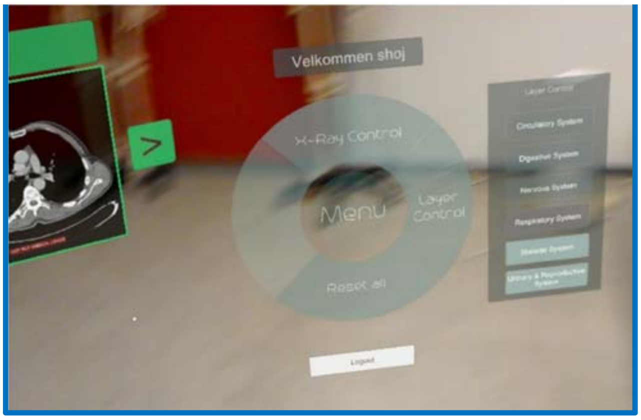


*Figure 5. The menu for selection of organ systems to be shown at the 3D body*

Intervention group without quizzes (ref fig 1) were told to switch of the quizzes by a voice command. This resulted in the same augmented room as the group with quizzes, except from the left part of the screen with the quizzes which were removed. Thus, students could shift between the CT-scans, the 3D body hologram, and the organ structures of their own choice in the 3D body.

## Data collection

Five days after the teaching session students received a questionnaire regarding their self-efficacy, motivation, presence in the augmented room, Self-efficacy and motivation were measured quantitatively with a selection of 7 and 5 questions respectively as a derived version of questions from Printritch et al (1991). Transformation was made to fit course subject. Presence vas measured with an extract and adjustment of 5 relevant questions from (Makransky, Lilleholt, & Aaby, 2017). Students' experiences with HoloLens learning situation, and how they used the quizzes was open answer questions. In addition, all students answered a test with the same 20 questions used in the teaching session at this exact station. They also got three additional transfer questions on the content presented in a format which were new to students. And finally, students' scores in the mediastinum questions in the exam 2 month later were collected to examine the long-term memorization of content. Table 1 shows an overview of collected data.

| Ordinary /traditional teaching | AR without quizzes | AR with quizzes | Absentees Students who didn't attend the teaching |
|---|---|---|---|
| motivation | motivation | Motivation | motivation |
| Self-efficacy | Self-efficacy | Self-efficacy | Self-efficacy |
| test | test | Test | test |
| transfer | transfer | Transfer | transfer |
| | presence | Presence | |
| | open answer, HoloLens use | open answer, HoloLens use | |
| | | open answer, quizzes | |
| Exam score | Exam score | Exam score | Exam score |

*Table 1. Data collected for each intervention group*

## Quantitative methods

Internal consistency was calculated for all measures using Cohen's alpha (α). Correlations between measures were examined with a correlation matrix using Pearson's r, and Cohen's convention was used to determine effect size categories (small, moderate or large) (Cohen, 1988). Group differences were examined with one-way analysis of variance (ANOVA). Assumptions of equal variance in ANOVA analyses were checked by calculation and inspection of Levene's robust test statistic and pairwise group differences were examined with Tukey's post hoc test. All statistical analyses were performed with STATA IC 15.

### Qualitative methods

Qualitative data sources were questionnaire open answers, and observations during the exercises by project leader and teachers. Two authors looked at the open answer qualitative data separately and no remarkable inter-rater differences were found. Open answer survey responses were analysed for frequent themes, patterns, contradictions, and unexpected information (Uddannelses- og forskningsministeriet 2017). Furthermore, a determined search was conducted searching for typical answers regarding a. students' overall opinions, b. students mentioning spatial learning stimulated by the app in general and in the quiz related comments.
Teachers experiences was registered as reflections on action after the teaching sessions.

### Limitations

Students were not tested for predisposition for spatial understanding and possible learning setup was limited by the preplanned course curriculum.

## Results

### Learning, motivation and self-efficacy

The number of students in each group were: Ordinary students n=33, AR with quizzes n=35, AR without quiz n=29, Absentees n=13. The total number of respondents was 110, and the number of non-respondents was 84 corresponding to a response rate of 0.56, which is arguably somewhat on the low side. This could have affected the degree to which responses collected represented the intended student cohort. The descriptive statistics of the variables examined is presented in table 2.

| | $n_{obs}$ | Min-max | Mean | SD | α |
|---|---|---|---|---|---|
| Motivation | 109 | 14-25 | 20.89 | 2.75 | 0.82 |
| Self-efficacy | 108 | 8-35 | 23.06 | 4.68 | 0.87 |
| Presence | 64 | 6-23 | 16.47 | 3.74 | 0.80 |
| Test score | 84 | 0-19 | 8.87 | 5.18 | 0.88 |
| Transfer score | 84 | 0-3 | 1.13 | 2.11 | 0.79 |
| Exam score | 156 | 0-15 | 10.56 | 3.58 | 0.66 |

*Table 2. Variables*

There were no significant group differences in motivation, test scores, transfer or exam scores, but the mean self-efficacy was significantly higher for the AR without quizzes-group compared to the ordinary teaching group (table 3).

| | n | Mean | SD |
|---|---|---|---|
| Ordinary teaching | 32 | 21.38[a] | 4.62 |
| Absentees | 13 | 22.46 | 4.79 |
| AR with quiz | 34 | 23.53 | 5.19 |
| AR without quiz | 29 | 24.62[a] | 3.53 |

*Table 3. Self-efficacy scores by student groups. [a] p=0.033 for the pairwise difference in mean self-efficacy between these means. All other pairwise differences in mean self-efficacy scores were not significant.*

| | Motivation | Self-efficacy | Test score | Transfer score | Exam score |
|---|---|---|---|---|---|
| Motivation | | | | | |
| Self-efficacy | **0.32***** | | | | |
| Test score | 0.11 | 0.14 | | | |
| Transfer score | 0.06 | -0.02 | **0.51***** | | |
| Exam score | 0.16 | 0.06 | 0.15 | -0.03 | |

*Table 4. Correlation matrix of outcomes. *** p<0,001*

We found significant correlations between self-efficacy and motivation and between transfer scores and test scores respectively (table 4). Since none of the variables suffered from decidedly low α-values (table 2), it seems less likely that unreliability is the main cause of the insignificant correlations presented in table 4.

The app does not give significantly different learning outcome regarding the subjects we measured. No significant differences were shown between groups in learning outcome. This means that AR in this teaching setup was neither worse nor better than traditional teaching for these students' learning outcome. However, based on the qualitative results many students seemed to appreciate AR as a modern resource for learning. Ability to transfer content is equally good between groups and related to learning outcome rather than teaching method. Transfer competence correlates with test scores for all students regardless of teaching method. Long term learning measured through the exam scores is not dependent on the teaching method and no significant difference between groups were measured.

## AR as a learning resource for motivation

Figure 6 shows the students' estimated appreciation of the AR learning app based on interpretation of the free text answers. Many students evaluate AR as a positive learning experience in the open answer question. Students with quizzes estimated the AR resource higher than students without quizzes. But generally, students were glad about the AR app facility and welcome the initiative (fig 6). Also, many students were impressed by the virtual presentation of body and CT images. Positive critique and recommendations for improvements are frequently included in the answers. Students welcome this new method and the possibility to study at their own pace.

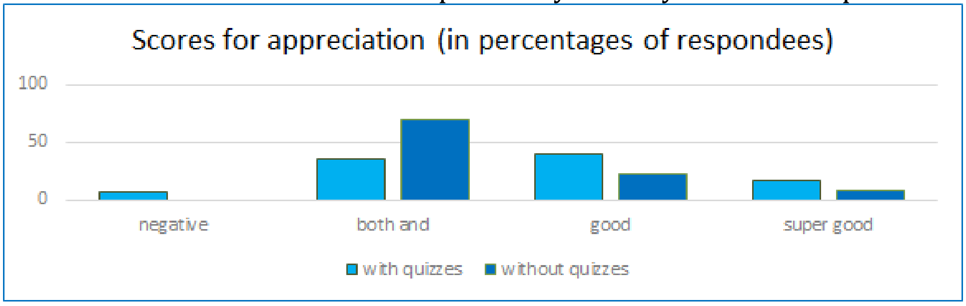


*Figure 6. Student's appreciation of using AR for learning. Grading of students' qualitative answers. Negative (students who only express negative evaluation); Both and (students with both positive and negative comments); Good (students who have only positive comments); Super good (students who have overwhelming positive comments and no negatives)*

## The AR learning experience

Based on the qualitative data AR was an appreciated variation in learning resources by more than 50 percent of students with quizzes and by approx. 35 percent students without quizzes (fig 6). However, there are challenges. In the open answer question "How did you experience using HoloLens as part of the teaching?" many students explained that they guessed a lot trying to solve the quizzes. They explained that they didn't reach professional depth in subject knowledge because they lacked the prerequisites from an introductory lesson or other preparations on the mediastinum structures. Students without access to quizzes described guessing to an even higher degree and appeared to be more critical to the app (fig. 6) and the teaching design. Furthermore, some students without quizzes felt "left alone" even though we explained the learning goals. They felt frustrated until they started to collaborate on learning the app content together by quizzing each other. This underpins the need for an outlined learning design with study-instructions corresponding with the app design. Based on the qualitative answers we conclude that students without quizzes felt to a higher degree alone during the exercise without being able to gain profound depth in knowledge and had problems with receiving feedback on their interpretations of app content.
There was however no significant difference in the monitored learning outcome. This led us to check for differences in sensation of "presence" in the virtual room to examine the recorded parameter related to the app quality (table 5). There were no significant differences between the two groups (AR with quizzes compared to AR without quizzes p=0.5235). We conclude that results show that the app felt equally real to students despite the differences in learning experiences.

| Intervention group | Mean | Std. dev. | Frequency |
|---|---|---|---|

| | | | |
|---|---|---|---|
| AR with quizzes | 16.74 | 4.10 | 35 |
| AR without quizzes | 16.13 | 3.29 | 29 |
| Total | 16.47 | 3.74 | 64 |

*Table 5. Summary of sum scores for the five presence questions*

## Quizzes, feedback and learning design

The second open answer question "How did you use the quizzes?" was only addressed to the students who had access to quizzes and these students were generally more positive than students without quizzes. Their critique mostly addresses recommendations for improvements of the app. They recommend a continue button to enable them to dwell on a question for deep learning. Students generally want to find out more about the right answer before they continue. This indicates a sound constructivist approach to learning content among students.

Corrective feedback in the quizzes is not evaluated to be good enough. Students recommend explanatory feedback and suggests visual highlights of the correct structures on the 3Dmodel as part of the feedback.

## Reaching spatial understanding

When we ask students how they use the quizzes we expected students to refer to their strategy such as frequency for answering the same question, use of the virtual body for answering etc. Our assumption was that that using the app would help students gain a spatial perception, but the frequency of students discussing the spatial learning is higher when we ask them about their use of quizzes. So unexpectedly this question gave rise to students explaining how quizzes stimulated the spatial 3D understanding of structures because they start scrolling up and down the hologram seeking for the right answers and transformation of 2D into 3D using the hologram. Based on this part of the results we conclude that spatial understanding might be stimulated by the quizzes.

## Learning of terminology

If the anatomical terminology is not in the app (students without quizzes) students want to use the book besides the app as a feedback for correct understanding which is not easy with the glasses on. Students accentuate good potential if using the app was coupled with traditional lectures. Many students, both students with quizzes and those without quizzes, explain that they would prefer a subject review first.

For the learning of terminology quizzes are essential in the actual app design. Students without quizzes explain the lack of anatomical terminology which the other group met through the quizzes. The correlations of learning terminology and the access to quizzes of course corresponds with the larger need for teacher assistance when students do not have the access to quizzes. They need the teacher's explanations and teacher's feedback.

## The role of the teacher

Observations whilst teaching showed that the role of the teacher suffered from many challenges in this setting. The teacher could follow each students' HoloLens view in the app on the computer screen, but the instructor found it hard to help students and especially students without quizzes felt that they lacked the teaching situation. These students reported that teacher was busy helping other students, waiting time was unacceptable, and students felt they lacked feedback.
Generally, it seems that students with quizzes had a better and more self- driven learning experience. We consider this as a further indication that the connection between the learning design and the app design is outmost essential. App design must supplement the learning design.

# Discussion

## Learning, motivation and self-efficacy

Self-efficacy is a central concept in social cognitive theories of learning (Bandura, 1978, 1997). It can be defined as a person's belief in his/her ability to carry out a specific task or activity. This *task specificity* distinguishes self-efficacy from more global psychological concepts of self and motivation. Positive self-efficacy beliefs have been found to be associated with: increased motivation, choice of activity, participation, and persistence (Zimmerman, 2000). In addition, students' levels of self-efficacy have also been shown to predict students' learning and academic performances in concordance with Self-efficacy Theory (Zimmerman, 2000). In this study, we did find measures of self-efficacy and motivation to be significantly correlated (as hypothesized in Self-efficacy Theory), but we were unable to find significant associations between self-efficacy and motivation for anatomy on the one hand and subsequent performances as measured by either test scores or exam scores (table 4). Had we found significant differences in performances amongst the four groups, it would of course have been important to be able to control for the mediating effects of general anatomy motivation and self-efficacy. A clear-cut indication that using AR was associated with increased levels of self-efficacy was not present either, since only one of the AR groups (AR without quizzes) displayed significantly higher levels of self-efficacy than the ordinary teaching group. It is entirely possible that 'the dose' of AR administered to the AR groups in this study was simply too small to elicit sufficiently large effects on students' subsequent performances. After all, we investigated the effect of only 35 minutes of AR training on a single occasion and in one course sub-topic only. However, it is well-known that students' performances are generally very context specific (i.e. quite depended on the specific anatomy topic in the session too) (Eva, 2003; van der Vleuten, 2014). We would therefore strongly suggest to other developers and researchers interested in examining the effects of AR on learning to consider the possibility of increasing the 'dose' of AR, to try to counteract such context specificity bias. This can be done by exposing students to a sufficiently high number of different AR learning situations (perhaps 8-12). This would most likely help the generalizability of measures of AR learning by simultaneously increasing the 'signal' and decreasing the 'noise'. Increased generalizability of measures of AR learning would in turn increase the chances of being able to find significant differences between AR learning and other types learning. On the plus side, our results indicate that the AR learning on this occasion did not lead to poorer learning than the conventional learning. This may be quite important for teachers' perceived barriers to experimenting with AR in classes.
Besides that, we know from other studies that students' personal preferences for learning resources differs (Midtiby et al 2017). Indicating that some students can be motivated and engaged

by studying with AR. As shown in figure 6, a major proportion of students are quite positive about AR as a learning resource and liked discovering anatomy in the AR room. Though of course students have supplementary comments they are generally positive if certain adjustments were made in the app. The engagement for using AR in students' responses are very promising for future learning resources in AR and VR. Meeting the AR app may have been a source for learning in a liminal blend as discussed in Enyedy et al. 2015, but we cannot conclude on that since we haven't questioned students on added value of the different sources that Enyedy describes (interaction analysis, materially anchored blends, conceptual blends, across time) nor was the research designed to clarify those factors. We recognize a novelty effect in some students' answers. Akçayır and Akçayır (2017) describes that students endorse the modern technology but loose the engagement with the technology as it becomes more common. We wish to circumvent that by continuing our work building relevant learning elements in AR and continue research on additional learning, self-efficacy, and motivation among students.

## AR as a learning resource for connecting 2D and 3D - spatial learning

Introducing new resources for learning, we are always looking for the add on's that is offered by the new resources. Especially introducing AR and VR facilities because it is expensive to develop the specific apps.

The add on for learning in this situation is facilitating the spatial conception of the "flat" two dimensional CT-images and the transformation into understanding the 3D, spatial appearance as explained in the introduction. The intension to help students connect traditional cross-sectional imaging and identification of anatomical structures with a spatial relation in the thoracic cavity seemed successful. Viewing the consecutive CT-images, up and down, the students seem to notice the interpretation of the anatomical structures in a 3-dimensional perspective scanning up and down the hologram body. Examples of students' reports are *"At first, I looked at the CT image that I should quiz. Then I scrolled up and down to create a sense of space"* and *"The first quizzes I used to find out if I had the right orientation and the right spatial conception of the CT-scans. Then I used the quizzes to check my knowledge of positions of the different anatomical structures in the chest in relation to each other."* Such answers lead us to the interpretation that quizzes facilitate spatial understanding.

We therefore have no doubt that students actively use the possibility of seeing the "unseen" (Dunleavy, 2014; Majgaard et al, 2017) and that the hologram function as reference for studying structures and recognizing spatial orientation of the CT images (fig. 2)

## Learning design

### *The quiz worked as a guide in AR*

Building the learning design comprises educational goals, meaningful activities, meaningful experiences, and assessment (Him & Hippe, 1997). For activities to be meaningful, they should guide de students to achieve the educational goals. Even though learning goals were explained to students at the beginning of the lesson they were in doubt about objectives. In the quiz scenario, the quizzes guided the students to explore the essential layers of the Hologram. In the scenario without the quiz, the students were more disoriented towards how to link the education goals to exploring the content of the hologram. One student explained it like this: *"…I needed some professional guidance and information about what I was looking at".* Some of the students without the quizzes

clearly needed some directions. *"I think it was a good and fun experience which improved as we joined small groups and quizzed each other in the different structures that we saw..."*
Besides quizzing, directions could be given orally, as an assignment or instructions given in AR. Mayer (2001) characterize successful multimedia learning scenarios as visual and verbal/text material presented simultaneously. The text can give away information about the visual content and additionally guide the students in the desired direction. Mayers (2001) approach on instructional design support our idea of developing the learning design further in future AR scenarios.

### Feedback and the Role of the Teacher

Quizzes help students to self-evaluate and another important factor of self-evaluation of own learning is discussions with teacher and peers. The teacher, of course, is an important source for feedback when AR is used in face2face teaching. The students need feedback either from the teacher (who in this version can't go directly into the virtual world together with the students), through instant feedback incorporated in the app (which can be solved in many ways), through collaborative setup during exercises, or otherwise.
We recommend explanatory feedback rather than corrective feedback as also concluded in Moreno (2004). A student describes: *"I used them [the quizzes] to get an overview of the CT images. However, it was a pity that you didn't get the right answer when your answer was false to learn from your mistakes."* In this study we tried a teaching scenario where AR was included in face2face teaching. Explanatory feedback was not included in the app which only offered corrective feedback. The teacher who could follow each students' challenges on the adjacent computer screen was intended to circulate and discuss results with students and challenge students' reflections. This unfamiliar teacher-role offers special challenges interrogating students whilst working in the app. Especially students without quizzes needed feedback. Individual help is time consuming, and students experienced long waiting time to get feedback.
If the app is intended for face2face teaching in the future setup's, we would recommend a collaborative AR solution where teacher can flex directly and virtually between students' AR rooms and discuss online in the AR room. Solutions for giving students feedback inside the augmented room should include anatomical terminology together with using the advantages of the technology like highlighting structures on the 3D body or other intelligent solutions thus supporting the learning goals for the teaching session (terminology and spatial understanding). A more profound explanation of structures for intended learning with human speaks could be worth developing maybe even using artificial intelligent solutions because students explain the desire for reflective study processes in the app.

### Peer feedback

Students without quizzes used each other in a process of scaffolding knowledge, and the level of communication raised in the room for students without quizzes contrasted by very silent, intense, concentrated, individually oriented atmosphere when students was working alone in the quiz-version. Such collaborative element is also mentioned in other studies ex. Dhu and Klopher, 2013. The recent years a lot of research have been published on learning from collaboration in VR (ex. Shao-Chen & Hwang, 2017 and Cho, 2017). We expect that collaborative learning in the app would have positive impact on the learning outcome. Giving and receiving peer feedback would have been easier if more students were in the same AR room together. Students without quizzes experienced this during the exercises." *I think it was a good and fun experience which improved as we joined small groups and quizzed each other in the different structures that we saw..."* (citation from student). We estimate that bringing two or more students into the same AR room and make them collaborate directly on quizzes would furthermore stimulate discussions between students and offer an even better learning situation and furthermore a teacher might enter the AR-room simultaneously with students.

As soon as technically possible we recommend development of a multiple user version for collaborative learning since collaborative learning seems to be a favorable learning situation also in the virtual world (Wu, Lee, Chang, & Liang, 2013 and Cho 2017). This would also solve the problem with the absence of a teacher in the virtual room during face2face teaching.
Thus, collaborative AR learning environments would promote teaching setups with peer feedback and even open the door for online AR team based learning (TBL) for the future.

### *Deep and self-directed learning*

Quizzes seems important for both terminology training (knowledge dimension) and for spatial understanding or 3D awareness therefore we recommend that quizzes are made permanent in the standard app-setup. "*The first quizzes I used to find out if I had the right orientation and the right spatial conception of the CT-images. Then I used the quizzes to check my knowledge of positions of the different anatomical structures in the chest in relation to each other*."
Quizzes needs to challenge students at the right level. Quiz questions therefore need to be aligned with the learning objectives which can limit the use of the app to specific courses or instructional settings.
A study process where students try, reflect and self-evaluate is required. Therefore, two extra buttons (a "feedback" and a "continue" button) would apply the app with proper flexibility for individual learning, timely feedback and self-regulated progression. For stimulating deeper learning processes explaining the expected learning outcome is important in an AR app for higher education to help students into self-directed learning processes.
Examples for upcoming app versions could be:
Level 1: After using the app make sure that you know the mediastinum terminology and understand the spatial interpretation of the pictures. In this app the explanatory feedback can be given in formats like highlighting the right answer on the hologram or showing and explaining the right answer on the screen in written or spoken language.

Level 2: Reflect and compare how different diseases displays on CT images. Quiz questions would then change and invite reflections: explain how … (a given diseases) … would display on a CT scan (choose between three relevant possibilities for answers). Or the app could include alternative sets of CT scans displaying specific diseases with questions like: Determine whether this patient suffers from … (choose between a, b or c).

### *Implementing e-learning tools in the teaching scenario*

Three main problems including the app in the teaching session was detected in the qualitative answers.

First, students were not familiar with using HoloLens and only had access to the app for 35 minutes (so far, this was the only HoloLens app developed for their study in medicine). A guide with a clear description of the intended way to use and navigate in the app is important. Secondly, students must be aware of the learning outcome and the learning design including preparation for this as for any other learning challenge along their study. Therefore, we recommend implementing the 5-stage model of Salmon (2011) in the learning design.

Thirdly, students show up unprepared. They report guessing and correlates it with their lack of preparation because they tend to read after the exercises, not before. This study habit doesn't apply to the actual app design and the learning design in any way. Students who first attend the classroom and then study the material themselves will become completely lost and lack both anatomical orientation and terminology.
If we could reach the point where students are sufficiently familiar with the technology to grab a HoloLens and prepare for lessons with a diversity of apps, we would be close to the optimal learning scenarios with HoloLens for learning anatomy. Therefore, development of a gateway system for teachers to easily upload relevant content to the app (pictures, quizzes and feedback) would provide students with a relevant learning resource which they would meet regularly and get familiar with. This need for plasticity in app design was also noted by Wu et al. 2013.

### Technical challenges

Interaction in HoloLens' are in our case done by gestures or clickers. The gestures must be very precise for them to work. Some of the students found it difficult to interact with the system especially in the beginning. The mouse click function was replaced with pressing together thumb and pointer finger in a specific angle in front of the glasses. A guide for interaction could have eased learning this new way of interaction.
When students came too close to the hologram body it started to flicker due to a technological limitation – in future version the body will be larger to make it possible to observe details without moving too close.
On the top of students' wish list for improvements were a bigger hologram, ability to rotate the hologram, a zoom function to the CT-images, explanatory feedback, and more informative text in AR.
HoloLens is a new platform and not yet entirely mature. During the development of the software, the programming platform changed several times and the development group had to reprogram aspects of the features more than once. This was very challenging for the engineering students never to know what would change when.

## Conclusion

Based on this project AR seems to have potential as a learning resource for teaching anatomy in Higher Education. It is our suggestion that studying axial CT-images combined with their positions on a 3-dimensional anatomical model in the AR room with the time necessary for training, will stimulate the spatial understanding of anatomical structures on axial CT-images and thereby improve the transformation into understanding the 3-D appearance of the anatomical structures.

The present study answers to the research questions in the following way

a. The effectiveness of using AR on student's short term, long term, and transfer learning outcome compared with traditional teaching: the study shows that this is an equally good learning resource for learning mediastinum content, transferring content and long-term memory.
b. The effect of quizzes seems to be that quizzes stimulate students' spatial conception (the transformation of content from the 2D scan pictures into the 3D corpus structures) because quizzes stimulates students to go back and forth through scans and compare with the hologram. Quizzes also seems to be important for both feedback and terminology training (knowledge dimension)
c. The present research gives detailed of information on the need for correlation between app design and learning design: in general there is a tendency to find a choose an app and then create a learning design that fits the app. Here we tried to design an app that fitted the intension of the learning design. We met certain challenges but the learning outcome was evenly good as with the traditional learning and students found the AR program usefull.

The learning design include specific formulations of the learning objectives for studying in the AR room, precise instructions for students' preparations before attending AR teaching sessions, and clear instructions to students for the use of AR HMD as a resource for learning.
This demands a prominent level of self driven use of the app and customization for students to use HoloLens in general.
App design can be sharpened to facilitate deeper learning processes and further scaffolding of learners' needs through specific app development. Future goals are to include a short guide for HoloLens use, explanatory feedback rather than corrective feedback – not only by text but also by technological features, self-pace progression, and collaborative facility for peer and teacher feedback. As discussed through this article it is our opinion that the app could be revised with some minor corrections with the purpose of extending the use for students' preparations and repetition in their process of learning.

# References

Akçayır, G., & Akçayır, M. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1-11.

Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk. (2014). Augmented Reality Trends in Education: A Systematic Review of Research and Applications. *Journal of Educational Technology & Society, 17*(4), 133-149.

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy, 1*(4), 139-161.

Bandura, A. (1997). *Self-efficacy: The exercise of control*: Macmillan.

Cho, Young Hoan; Lim, K. Y. T (2017) Effectiveness of collaborative learning with 3D virtual worlds *British Journal of Educational Technology*, 01/2017, Volume 48, Nummer 1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale N.J: Lawrence Erlbaum Associates.

Dunleavy, M. (2014). Design Principles for Augmented Reality Learning. In (Vol. 58, pp. 28-34). Washington: Springer.

Eva, K. W. (2003). On the generality of specificity. *Medical Education, 37*(7), 587-588.

Enyedy, N; Danish, J. A; Deliema, D (2015). Constructing liminal blends in a collaborative augmented-reality learning environment, *International Journal of Computer-Supported Collaborative Learning*; Vol. 10, (1), 2015, 7-34.

Federative international programme on anatomical terminologies ”Terminologia Anatomica” (2nd edition, Thieme, Stuttgart/New York, 2011)

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior, 72*, 276-285.

Makransky, G., Thisgaard, M. W., & Gadegaard, H. (2016). Virtual Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills. *PloS One,* 11(6).

Majgaard, G., Larsen, L. J., Lyk, P., & Lyk, M. (2017). Seeing the unseen: Spatial visualization of the Solar System with physical prototypes and Augmented Reality. *International Journal of Designs for Learning*, 8(2), 95-109.

Midtiby H.S., Noergaard C., Kjaer C. (2017) Students' Benefit from Video With Interactive Quizzes In A First-Year Calculus Course, *INTED2017 Proceedings*, pp. 2290-2299.

Moreno, R. (2004). Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science,* 32(1/2), 99-113.

Munoz-Cristobal, J. A., Jorrin-Abellan, I. M., Asensio-Perez, J. I., Martinez-Mones, A., Prieto, L. P., & Dimitriadis, Y. (2015). Supporting Teacher Orchestration in Ubiquitous Learning Environments: A Study in Primary Education. *IEEE Transactions on Learning Technologies,* 8(1), 83-97.

Nørgård C., O'Neill L., Nielsen K.G., Juul S.H., Chemnitz J. (2018) Learning Anatomy with Augmented Reality *EDU Learn 2018 Conference article - in press*

PRR Pintrich, D Smith, T Garcia, W McKeachie, ”A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*". Ann Arbor Michigan.* 48109: 1259 1991

Salmon, G. (2011). *E-Moderating: the key to teaching and learning online* (3. ed.). New York: Routledge.

Shao-Chen Chang & Gwo-Jen Hwang (2017) Development of an effective educational computer game based on a mission synchronization-based peer-assistance approach, Interactive Learning Environments, 25:5, 667-681.

Uddannelses- og forskningsministeriet, 2017, *Kvalitative studier, en håndbog for kvalitative studier I Uddannelses- og forskningsministeriet*, december 2017

Underlying principles at SDU; https://www.sdu.dk/en/om_sdu/institutter_centre/c_unipaedagogik/baerende_principper

Van der Vleuten, C. P. M. (2014). When I say … context specificity. *Medical Education,* 48(3), 234-235.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education,* 62, 41-49.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82-91.