Karlsruhe Institute of Technology

INSTITUTE OF INFORMATION SYSTEMS AND MARKETING

Educating Data Science with Mixed Reality: Developing an Exploratory Learning System

Teamproject Software Development - Report

Remy Carosi, Maily Doan, Isabella Mebus Kishi de Oliveira, Julian Strietzel, Luca Wolfinger

2287850, 2284748, 2291878, 2285252, 2304032

At the Department of Economics and Management
at the Institute of Information Systems and Marketing (IISM)
Information Systems I

Reviewer:	Prof. Dr. Alexander Mädche
Supervisor:	Dr. Peyman Toreini
Advisor:	M.Sc. Shi Liu

Duration: 12th of May 2022 – 30th of September 2022

Contents

Abstract

From text, over printing to the internet, progress in technology has changed the way we access information, teach and learn. The use of technology has proven to improve access, motivation, and impact. As computers, videos and smartphones have changed the way we learn already, Mixed reality is promising to be the next step in digital teaching. In this team project we explore how the Microsoft HoloLens 2, as a first-generation mixed reality device, can be used to enhance the learning experience. In a first approach we tried to teach basics of Data Science on the example of decision tree theory to university students. Doing so we explored how we can prototype for Mixed Reality and used Unity with Microsoft MRTK for MR-Development.

List of Abbreviations

MRTK .... Mixed Reality Toolkit
DT .... Decision Tree
IISM .... Institute of Information Systems and Marketing
MR .... Mixed Reality
UCD .... User Centered Design

1. Introduction

The need for data scientists and data science concepts is increasing, due to the emerging amount of data and new data sources. That's why Data Science is becoming an increasingly important part of education. Data Science is also quite an interdisciplinary subject. The work with data is carried out in almost all industries today. For students, however, the concepts can sometimes be quite unfamiliar and the learning process therefore quite demotivating. Especially for students who do not come from the STEM field. So there needs to be a way to make the learning process more motivating and engaging.

Mixed reality has become an increasingly relevant topic in recent years. In that time, scientists have been researching some application areas in which mixed reality can offer an increase in value. Gasques et al. (2021), for example, has developed a mixed reality system to assist inexperienced surgeons during operations. But also in the field of education, research was conducted for possible applications, reviewed by Maas and Hughes (2020) for example. Lafargue (2018) found that mixed reality has a positive impact on student engagement and student motivation. These applications of Mixed Reality showed great promise to aid in our initial Problem.

The goal of this project is therefore to develop a learning experience, design and implement a software artefact that contains this learning experience, and evaluate with students whether the desired motivation-enhancing effects occur. This way we wanted to answer the question how mixed reality can enhance current data science education.

The mixed reality device we chose was the Microsoft HoloLens 2. The HoloLens is a head-mounted mixed reality headset and is capable of projecting digital holograms over a display in front of the eye as if they were physically there. Furthermore, it allows interaction with these holograms via various gesture controls. In our case, mainly the hand tracking feature was used. We decided to work with the HoloLens mainly because of its availability to us.

2. Methods

In the following, we will go into more detail about the methods used during the project, as well as their benefits for its successful implementation.

2.1. User Centered Design (UCD)

The learning success of e-learning systems depends, among other things, on how easy the system is to use (Selim, 2005). The user should not learn how to use the system, but the content that is to be taught to the user. We therefore decided to focus on the user and his interaction with the system from the very beginning and followed the user centered design (UCD) process (Norman & Draper, 1986). In total, we had two UCD iterations, each of which concluded with a user evaluation. The first iteration primarily involved the design of the learning experience itself and how to implement it in mixed reality. The result was a paper-based prototype that we were able to perform tests on. The second iteration was concerned with the implementation of our protoype and how to provide a good user experience when using the HoloLens and its features. In this iteration we were able to complete a first software artifact.

We used a variety of tools and methods to perform each of the UCD process steps, which are introduced in the following sections.

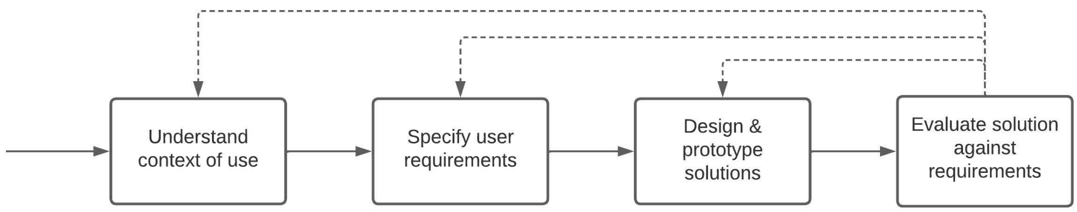


Figure 2.1.: User Centered Design Process

2.2. Brainwriting

After we had dealt with the two main topics of our project ”Mixed Reality” and ”Decision Trees” and were able to develop a basic understanding of them, we started to collect our ideas in a structured way and to develop them further in the group. For this purpose, we primarily used the framework ”Brainwriting” (VanGundy, 1984).

Figure 2.2.: Brainwriting

In this framework, each participant starts by writing down three basic ideas in a separate column. These can be basic program structures (e.g. module structure) or implementation proposals (e.g. use of gesture control for a special use case). These should be described and explained in as much detail as possible so that another team member can then take up the suggestions and ideas and develop them further. Each of these idea collection rounds lasted five minutes. Thus, with a group size of four participants (one team member absent due to illness), the overall process took 20 minutes.
We took a lot of time for the subsequent discussion of all the ideas collected, as well as for ranking these ideas, etc., because we were aware that this was the cornerstone of the entire project.

2.3. Prototyping & Bodystorming

After we had created and revised our system design, it was now time to have the developed ideas tested in practice by real users. However, already at this point our first problem arose. We would need a suitable prototype to adequately test the real and virtual objects and their interaction.
But what should such a prototype design look like? How could we represent virtual components in a pure real prototype and consequently test their functionality? In consultation with our supervisors, we agreed to create a paper prototype that we would then present to an IISM employee for a first test run.
In order to create a unified prototype, we decided to move forward with the entire design process together (rather than individually) in one day. One part of the team dealt with the planning of the test procedure, for which the necessary system components had to be planned on paper. It was clear to us that we wanted to prototype the entrance to the learning experience as well as one of the first learning sections first. This was not only for causal reasons, but also for practical ones. It was necessary that the user would first get a feel for working with MR and receive a short concise briefing on how DT works and then dive into its further study.
In order to implement this in our paper prototype, we decided to model "Module 1" as well as "Module 2" in our prototype.
Once all the design decisions were made, we started with the implementation and creation of the prototype. We quickly realized that the dynamics (programmed) in the final artifact would be extremely difficult to implement in an analog way. Therefore, we decided for the prototype exclusively to model the DT given to us and, if necessary, to force the decision on this during the implementation with the participants. This way, we could ensure the correct functioning of the prototype on the one hand, which comes closest to the final user

experience, and on the other hand, limit the complexity of the prototype and consequently minimize any problems that might arise.

The creation of the prototype took two working days including all the steps (writing the script, designing and drawing paper objects, creating/printing icons and applying them on the balls). Now our prototype was ready to be tested and evaluated on different user groups.

Even before conducting all the user tests, we had to deal with the question of how we were going to ask users about their opinions and experiences with our artifact after the tests. We decided to create a detailed questionnaire that not only asked about the experiences of the individual participants but also about their general level of knowledge in order to relate them to each other and to analyze correlations.

We decided to use a Likert scale. In addition, we added several open-ended questions at the end to give the user the opportunity to describe the user experience in his own words via general further comments and thus not to restrict the feedback in advance.

In order to be as user-centered as possible during the prototype tests, we decided to follow a Bodystorming approach *Bodystorming for movement-based interaction design - ProQuest* (n.d.) in order to be able to analyze intuitive gestures of the users already in this phase and to be able to adapt the subsequently developed first artefact to them.

2.4. Scrum

In addition to UCD, we used a SCRUM-like process because we were convinced that an agile team-based organization would be best with a team of our size. To do this, we initially set up all the necessary software e.g. Teams, Trello, GitLab etc. We then used the GitLab Issues feature as a product backlog defining missing features, bugs, etc. as an Issue and interconnecting related Issues. We then met with our advisor every week and discussed the results, as well as distributed the new tasks for the week. During this week we had two to three digital meetings in which we reported our progress and helped each other with any issues that occurred. In this way, we usually had a product increment every two to three weeks, which we could present and compare with the initial user requirements. Based on this, we were able to include further improvement tasks in the product backlog.

3. Design

For the rough design, we decided to structure the learning experience into modules that sequentially describe and explain different features of a DT and let the user interactively apply what they have learned. In this way, we could ensure that users with existing knowledge would not lose motivation and could start with a later module and allow learning breaks, which could enhance the user's concentration. All modules can be reached via the main menu, which in turn can be called up at any time within the modules. The user can also seamlessly switch to the next module at the end of a module without entering the main menu.

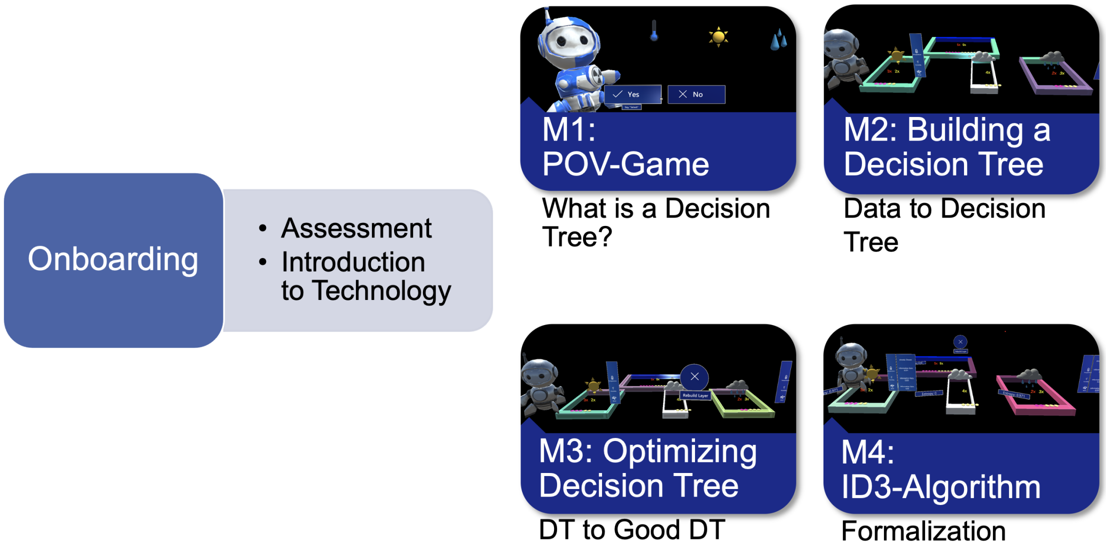


Figure 3.1.: Structure of the Modules

3.1. Onboarding

When the user starts the application, he first enters a small onboarding, structured into two components. The first component deals with an evaluation of the user's previous knowledge. After answering a few questions, the main menu will highlight the module recommended for the user's level. The second component goes over some basic mechanics of the HoloLens and explains them. These mechanics are intended to minimize confusion for the user as the app progresses. The Onboarding also introduces KAI, a robot figure which can also be seen in Figure 3.1, which guides the user through the first module and is then replaced by Alice. The idea behind KAI and Alice was to give the user a guide that can lead him through the modules and teach certain theoretical aspects. In this way, the user is provided with a familiar and traditional learning method, namely that of a teacher, and is thus not completely confused by the new technology.

3.2. Module 1: The Basics

The first module is designed to explain the fundamentals of DTs. It starts with the user being asked to decide whether or not to play tennis based on weather data. The goal here is to intuitively show the user that a decision has a certain structure and depends on certain parameters that influence the result. The decisions the user makes are then presented in the form of a tree. Based on this tree, the terminology and usage of a DT is explained. Thus, based on his answers, the user is taught how to visualize the structure of his decisions and gets an overview of the most important terminology. We have decided not to include a lot of interaction in the first module to give room for communicating the basic idea of Data Science and Decision Trees.

3.3. Module 2 & 3: Building a Tree

In module two and three, the user should build his first own tree. For this purpose, he receives a data set in the form of tennis balls and can now decide in each layer of the tree after which category he would like to split a node. Module two primarily explains the mechanics of how a tree is built within the app. The user can then build a first tree himself. In module three he gets the possibility to reset different levels of the tree. He is then motivated to build the best possible tree. The idea is that the user will think algorithmically during the process about how he can do this best and how he can possibly improve. Ideally, after a few brute force attempts, the user thinks about what makes a good split and how to achieve one from the information in the data. This then guides the user to module four.

3.4. Module 4: Formalization

Module four is designed to introduce the user to the concepts of entropy and information gain. The user builds a DT, just as in modules two and three. This time, however, he receives additional information about entropy and information gain. With this, he can now build an "optimal" DT. The goal was that the user goes through the steps of the ID3 algorithm once and thus understands the functionality behind it and learns that splitting by category, given entropy or information gain, is relatively straightforward. However, an existing problem of this module is that the user never has to calculate the entropy himself and it is also not questioned whether the user has understood the concept at any point. In an extension of the program such an evaluation should still be built in.

4. Implementation

The second iteration of our UCD Process was dominated by the implementation of the prototype and design ideas we had developed. With our module based design approach to the learning experience, we could easily split up the work as each team member could concentrate on specific modules and functionalities. We focused our efforts on building an easily accessible introduction, suitable also for the learner who is unfamiliar with the device, a module to introduce the basic decision tree idea and the process of learning a decision tree from given data. The final step was distributed over the last 3 modules as it is the key learning goal. It contained a dynamic implementation of a tree-like structure, which adopted the structure specified by the learner.

4.1. Technologies

Implementing for mixed reality on HoloLens 2 is dominated by the Unity Game Engine qianw211 (n.d.) using C# in combination with the Microsoft MRTK.polar kev (n.d.) As this is the most spread, and best documented solution we decided to use this stack, too.

4.1.1. Engine: Unity & C#

Unity is a development environment specifically for 3d games. It can also be used for app development on several platforms including the Universal Windows Platform. It has some useful features like Holographic Remoting and Microsoft Visual Studio support which enhance the experience building and testing directly from Unity to HoloLens 2. This is very useful in the development process, as some aspects of the application are not exactly represented in Unity game preview as on HoloLens mixed reality.
Unity is structuring an application with game objects. Those represent each text, block or player object on the screen. They are organized in Scenes which represent one specific screen in the application. Scripts, imports, objects and more are categorized as assets and can be attached to game objects. Scripts are coded in C#, which is an object-oriented programming language familiar to C, C++ and Java.(BillWagner, n.d.)

Advantages

Coming from Game Development Unity gives a lot of build-in functionality to build 3d environments and structure the game experience. Unity has native support for Mixed Reality development and most features still work as expected. Using a specific Scene for each of our modules helped us not only to structure our environment but also helped a lot while collaborating via GitLab. On top of that, there is an active community and tons of assets available to use (like our beautifully animated Retrobot. (*Retrobot, Game and App Character, Robot Mascot | 3D Characters | Unity Asset Store*, n.d.))

4.1.2. HoloLens 2

HoloLens is a head-worn holographic computer, that is able to place 3d content in the environment and build a mixed-reality environment around you.(scooley, n.d.) As it is one of the first consumer ready devices in this field, it still has it flaws, but the technology is promising and our project was all about testing its application in the university's learning environment. HoloLens 2 has a lot of exiting features, we briefly want to introduce the key features we used in our application:(polar kev, n.d.)

1. World-anchoring: HoloLens is able to place Holograms persistently at a specific place in space. We used this to place our Decision Tree permanently on the same spot.
2. Hand tracking: The user can interact with holographic objects intuitively by grasping and touching. It also detects the rotation of the hand, which we used to show a so called hand menu for easy navigation and help.
3. Spatial awareness: HoloLens recognizes surfaces and objects in the environment. We used this to attach virtual objects (our self-built decision tree) to the table.

4.1.3. MRTK

MRTK is a project providing integration of HoloLens Features for Unity.(polar kev, n.d.) It is therefore necessary to include it in every Unity Scene using any of the specific MR features. MRTK includes so called solvers, which have to be added to every game object using specific HoloLens functionality like surface magnetism, gaze following and object manipulation via hand gestures.

4.2. Features

Designing and implementing our learning experience we took advantage of the usage of HoloLens 2 and a digital teaching approach in general. This enabled us to create an application that will not only motivate the learner more than conventional methods, but also improves immersion and hopefully learning success (Lafargue, 2018).

Dynamic decision tree

A main part of our learning experience is building and optimizing a decision tree structure from given data. We explain this concept by representing data points with physical tennis balls. Those are supposed to be placed in frames representing nodes of a decision tree, and are moved by the learner to understand the way a decision tree works. The combination of physical data points and virtual boxes is a key advantage of our mixed reality experience opposed to conventional learning applications.
A first approach was the other way around: Having physical objects to build the decision tree structure from and virtual data points that flow through it. During prototyping and reading the documentation we decided to use the advantage of a virtual tree, as it is hard to track smaller sized objects with the HoloLens and we would not be able to scan the tree structure. Our implementation gives us the possibility to dynamically change the tree, adding icons and colors, and in general improve immersion.
This also created challenges, though. As general tree visualizing algorithms mostly rely on the input of a predefined structure and have nodes with a fixed size, they were not applicable for our dynamically building tree. In each layer we had to rethink about spacing and scaling the frames. The final implementation uses a tree data structure internally to get information about how many data points are ”more left” in this layer, from the previous one. It than estimates the width of the not yet built tree parts, by the number of data points in it, and scales itself so that all the tennis balls will fit in the frame.

Surface magnetism

It is important for the decision tree to anchor on physical surfaces, as we are using a physical representation of data points. The surface magnetism solver of MRTK (Microsoft, n.d.) using Scene Understanding is able to place any Objects relative and fixed to tables, floors or walls. We still need the user to choose a place for the decision tree in the possibly changing environment, though, which is done in the beginning of a module and can be redone on request. For example if the user notices a wrong placement during the module.

Generalization

As we are using a digital approach, we are able to generalize our learning experience to different scenarios and data sets. Therefore, we are using a json data structure to import the data and general information about the scenario, as categories, name and number of data points. It can easily be switched or modified and the application adopts to it, with some limits. In future work this should be improved further, so that all buttons, their labels, the number of categories and the story in general can be adopted to a specific scenario, for example in an employee training program.

Hand Menu

We implemented a so called hand menu, which appears when the user looks at the palm of his hand. This is used for easy access to the main menu, a scene dependent hint ,and the repositioning of the decision tree.

4.3. Challenges

Unity is described as easy to learn and use. Nevertheless, starting a project in a new development environment is always challenging and our team had to take its time to learn all the unity specific and not always intuitive functionality, as well as the specific packages used in C# for Unity. Especially that Unity contains thousands of specific features and settings, and they somehow all interconnect, was a challenge for us. Moreover, the power of Unity also kind of overwhelmed the power of our not for 3d simulations built laptops.
Placing and scaling objects in virtual reality using Unity is a bit of a challenge, as the field-of-view on HoloLens is not well represented and distances are hard to estimate. Therefore, it takes some iterations till it all works as expected. We solved this by fast iterations and frequent testings on Holographic Remoting with HoloLens.
Even though Unity is very well documented in general, in combination with MRTK there often is a lack of available information on the relevant sites. Even the official MRTK-Documentation does not provide all the necessary information. Amongst other topics, when it comes to how exactly all the solvers (Microsoft, n.d.) using HoloLens Sensors work, the documentation falls short and we sometimes had to trust our intuition and testing.
HoloLens 2 is able to recognize surfaces and larger objects. It has no functionality to track smaller objects, like our tennis balls, though. At first, we wanted to track each on of them for giving specific feedback on the movements made by the user and augmenting it with virtual icons. Limited by technology we decided to print the icons on our tennis balls. We also use small indicators on our frames to give feedback on how many objects have to be placed in a frame. The learner uses them to self-check on his movements.
The HoloLens it self made some problems, too. Using Holographic Remoting and/or live streaming from it, we faced overheating and connection errors on a regular basis. The problem worsened in warm conditions and with long sessions, having to recharge, too. Particularly during evaluation, where we invited several testers to the Lernfabrik, this was a mayor issue. Wearing a 2.878 kg(scooley, n.d.) overheating computer on your head drastically reduced the level of comfort and concentration while using our application. Optimization in code, reducing animations and deactivating unused features of HoloLens will be tested in future work to improve this issue.
Further challenges of us can be categorized by the 8 key barriers in Authoring AR/VR Applications by Ashtari, Bunt, McGrenere, Nebeling, and Chilana (2020) and are explained there in more detail.

5. User Evaluation

Two user tests were conducted on this project. The first test was for the basic design decisions (evaluation of the paper prototype), the second test (user study), for the general learning experience, interactions with the constructed artefact, and the overall user experience. These were timed into the project according to the UCD framework (one test per iteration).
Thus, the evaluation consists of two questionnaires that users are asked to fill out after completing the tests. Both contain categorized questions to be answered on a Likert scale, as well as open-ended questions to give the user the opportunity to describe the user experience in his own words via general further comments and thus not to restrict the feedback in advance.

5.1. Experimental Design

In the following, the experimental design of the user test conducted in each of the UCD iterations is described and explained. All participants gave their verbal consent to processing the tests, analyzing the data obtained and possible image and video recordings, and using them in project-related publications before the tests were started.

5.1.1. Paper Prototype

In the first UCD iteration, we conducted a test of the built paper prototype as described in the previous text. The purpose of this was to analyze the design decisions we made for suitability, usability, and simplicity.
The feedback received is essential for us as well as for the further course of the project, since we wanted to develop as user-centered as possible and implement any form of suggestions for improvement in order to achieve the best possible end result.
Six participants from different disciplines took part in the user test of the first UCD iteration (see Figure 5.1). Accordingly, their previous experience in the MR environment as well as with the subject of DTs varied.

Figure 5.1.: Participants of the user test in the first UCD iteration

The test was divided into two sections (see Figure 5.2). In the first section, the practical ”system application”, which lasted about 20 minutes, users were asked to work through predefined modules (”Module 1” and ”Module 2”) step by step. Since the built paper prototype did not have the possibility to include explanatory texts to the necessary extent, one team member guided the participants through the prototype while another team member recorded notes on usage behavior and possible problems as well as their elimination.

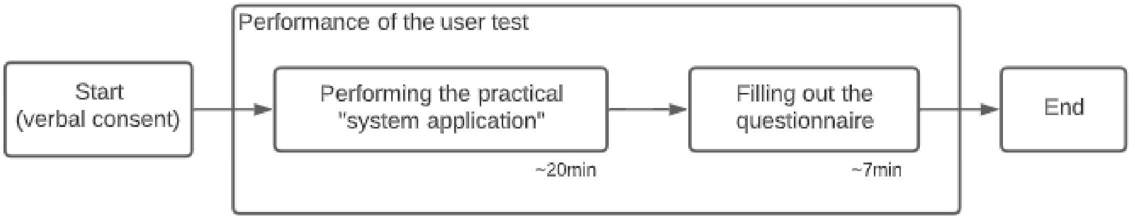
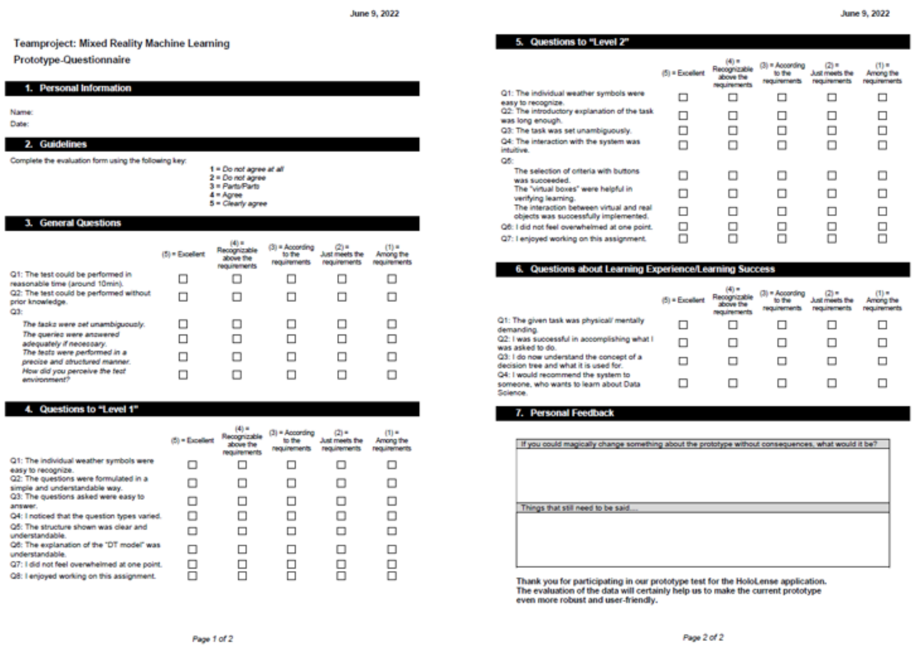


Figure 5.2.: First User Test: Prototyping - Sections

Participants were then asked to complete a 2-page questionnaire (see Figure 5.3) based on their impressions of the user experience. The questionnaire had seven sections that built on each other with a total of 29 questions to be answered. Most of the questions were alternately positively and negatively worded statements that had to be answered on a 5-point Likert scale. The final two questions were open-ended, so as not to restrict the participants in their feedback and to explore possible areas for improvement. The duration of the participants was not limited in time, but took an average of seven minutes. Thus, this user test could be completed in its entirety in 30 minutes.



June 9, 2022

Teamproject: Mixed Reality Machine Learning
Prototype-Questionnaire

1. Personal Information

Name:
Date:

2. Guidelines

Complete the evaluation form using the following key:

1 = *Do not agree at all*
2 = *Do not agree*
3 = *Parts/Parts*
4 = *Agree*
5 = *Clearly agree*

3. General Questions

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The test could be performed in reasonable time (around 10min).	☐	☐	☐	☐	☐
Q2: The test could be performed without prior knowledge.	☐	☐	☐	☐	☐
Q3:					
The tasks were set unambiguously.	☐	☐	☐	☐	☐
The queries were answered adequately if necessary.	☐	☐	☐	☐	☐
The tests were performed in a precise and structured manner.	☐	☐	☐	☐	☐
How did you perceive the test environment?	☐	☐	☐	☐	☐

4. Questions to “Level 1”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q3: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q4: I noticed that the question types varied.	☐	☐	☐	☐	☐
Q5: The structure shown was clear and understandable.	☐	☐	☐	☐	☐
Q6: The explanation of the "DT model" was understandable.	☐	☐	☐	☐	☐
Q7: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q8: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 1 of 2

June 9, 2022

5. Questions to “Level 2”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The interaction with the system was intuitive.	☐	☐	☐	☐	☐
Q5:					
The selection of criteria with buttons was succeeded.	☐	☐	☐	☐	☐
The "virtual boxes" were helpful in verifying learning.	☐	☐	☐	☐	☐
The interaction between virtual and real objects was successfully implemented.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

6. Questions about Learning Experience/Learning Success

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The given task was physical/ mentally demanding.	☐	☐	☐	☐	☐
Q2: I was successful in accomplishing what I was asked to do.	☐	☐	☐	☐	☐
Q3: I do now understand the concept of a decision tree and what it is used for.	☐	☐	☐	☐	☐
Q4: I would recommend the system to someone, who wants to learn about Data Science.	☐	☐	☐	☐	☐

7. Personal Feedback

If you could magically change something about the prototype without consequences, what would it be?

Things that still need to be said...

Thank you for participating in our prototype test for the HoloLense application. The evaluation of the data will certainly help us to make the current prototype even more robust and user-friendly.

Page 2 of 2

Figure 5.3.: Questionnaire - Prototyping

5.1.2. User Study

The user test conducted in the second UCD iteration, in the form of a user study, was also structured in two sections (see Figure 5.4), which were intended to analyze the general learning experience, the interactions with the built artefact, and the overarching user experience, or to identify potential for improvement. Five participants, from a wide variety of disciplines, again took part. Four of these participants already took part in the user test in the first iteration.

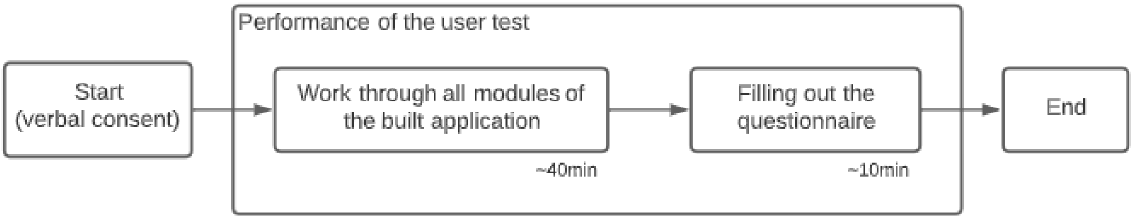


Figure 5.4.: Second User Test: User Study - Sections

In the first section, the interaction with the built artefact, which took about 40 minutes (net time without considering possible battery charging times, cooling down times, as well as fixing connection problems with the screen to observe the interaction), the participants were asked to master all modules on their own, without prior instruction and/or assistance, based on the created interaction possibilities (see Figure 5.5). The two team members who were present during the user tests took care of possibly occurring problems with the HoloLens 2 and observed the user interaction of the participants on the connected screen in order to obtain additional information about the usage behavior and to identify potentially occurring problems as well as potentials for improvement (see Figure 5.6). Then

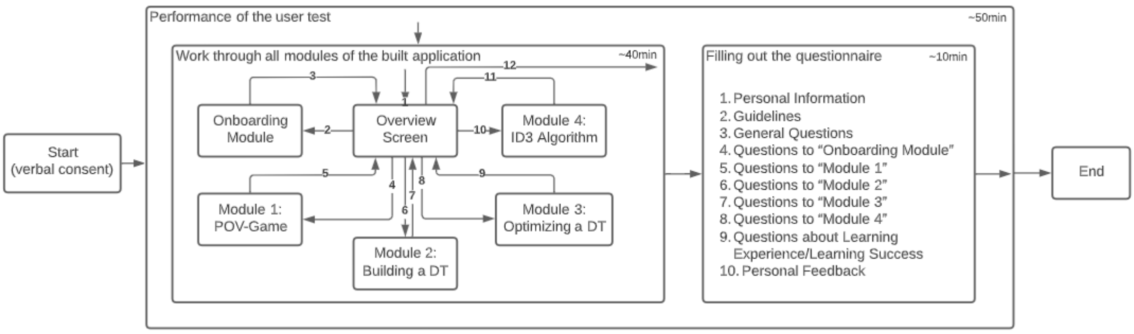
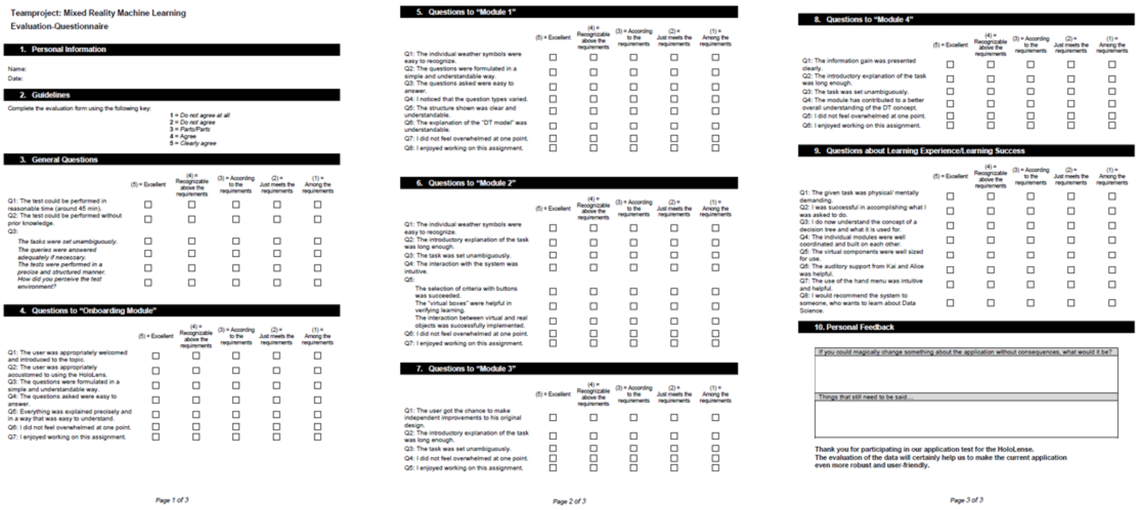


Figure 5.5.: User Study - Detailed Flow Chart

participants were asked to complete a questionnaire (see Figure 5.7), this time three pages long, with 51 questions from ten categories (see Figure 5.8). Most of the questions were again phrased alternately positively and negatively and were to be answered on a 5-point Likert scale. This questionnaire also provided the opportunity to add personal feedback and suggestions for improvement via open-ended questions at the end. There was no time limit for the participants to complete the questionnaire, but it took an average of ten minutes. This user test could therefore be completed in its entirety in 50 minutes.

Figure 5.6.: User Study - Observation



Teamproject: Mixed Reality Machine Learning
Evaluation-Questionnaire

1. Personal Information

Name:
Date:

2. Guidelines

Complete the evaluation form using the following key:

1 = *Do not agree at all*
2 = *Do not agree*
3 = Parts/Parts
4 = Agree
5 = *Clearly agree*

3. General Questions

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The test could be performed in reasonable time (around 45 min).	☐	☐	☐	☐	☐
Q2: The test could be performed without prior knowledge.	☐	☐	☐	☐	☐
Q3:					
The tasks were set unambiguously.	☐	☐	☐	☐	☐
The queries were answered adequately if necessary.	☐	☐	☐	☐	☐
The tests were performed in a precise and structured manner.	☐	☐	☐	☐	☐
How did you perceive the test environment?	☐	☐	☐	☐	☐

4. Questions to "Onboarding Module"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The user was appropriately welcomed and introduced to the topic.	☐	☐	☐	☐	☐
Q2: The user was appropriately accustomed to using the HoloLens.	☐	☐	☐	☐	☐
Q3: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q4: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q5: Everything was explained precisely and in a way that was easy to understand.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 1 of 3

5. Questions to "Module 1"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q3: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q4: I noticed that the question types varied.	☐	☐	☐	☐	☐
Q5: The structure shown was clear and understandable.	☐	☐	☐	☐	☐
Q6: The explanation of the "DT model" was understandable.	☐	☐	☐	☐	☐
Q7: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q8: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

6. Questions to "Module 2"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The interaction with the system was intuitive.	☐	☐	☐	☐	☐
Q5:					
The selection of criteria with buttons was succeeded.	☐	☐	☐	☐	☐
The "virtual boxes" were helpful in verifying learning.	☐	☐	☐	☐	☐
The interaction between virtual and real objects was successfully implemented.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

7. Questions to "Module 3"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The user got the chance to make independent improvements to his original design.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q5: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 2 of 3

8. Questions to "Module 4"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The information gain was presented clearly.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The module has contributed to a better overall understanding of the DT concept.	☐	☐	☐	☐	☐
Q5: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q6: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

9. Questions about Learning Experience/Learning Success

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The given task was physical/ mentally demanding.	☐	☐	☐	☐	☐
Q2: I was successful in accomplishing what I was asked to do.	☐	☐	☐	☐	☐
Q3: I do now understand the concept of a decision tree and what it is used for.	☐	☐	☐	☐	☐
Q4: The individual modules were well coordinated and built on each other.	☐	☐	☐	☐	☐
Q5: The virtual components were well sized for use.	☐	☐	☐	☐	☐
Q6: The auditory support from Kai and Alice was helpful.	☐	☐	☐	☐	☐
Q7: The use of the hand menu was intuitive and helpful.	☐	☐	☐	☐	☐
Q8: I would recommend the system to someone, who wants to learn about Data Science.	☐	☐	☐	☐	☐

10. Personal Feedback

If you could magically change something about the application without consequences, what would it be?
Things that still need to be said…

Thank you for participating in our application test for the HoloLense.
The evaluation of the data will certainly help us to make the current application even more robust and user-friendly.

Page 3 of 3

Figure 5.7.: Questionnaire - User Study

categories	number of questions	covered topics
General Questions	6	duration, prior knowledge, unambiguousness, helpfulness of the team, structuredness, test environment
Questions to “Onboarding Module”	7	introduction to MR and HoloLens, use of the HoloLens
Questions to “Module 1”	8	icon size, simplicity of questions, structuring, comprehensibility
Questions to “Module 2”	9	icon size, explanation style, handling, interaction with buttons, interaction of physical and virtual components, frames for self-monitoring
Questions to “Module 3”	5	functionality, simplicity of use, handling
Questions to “Module 4”	6	representation of the information gain, extended DT understanding
Questions about Learning Experience/Learning Success	8 (cleared 7)	DT concept understanding, coordination of modules, auditive support by interlocutor, hand menu, recommendation of the artefact
Personal Feedback	2 (open questions)	additional feedback and suggestions for improvement

Figure 5.8.: Questionnaire Overview

5.2. Results

All results mentioned below are based on the data and information collected during the user tests, as can be seen in Figure 5.9. Observations were made and recorded using MS Word and MS OneNote. In addition, the answers to the questionnaires were analyzed by collecting their results in an Excel file and then statistically analyzing them.

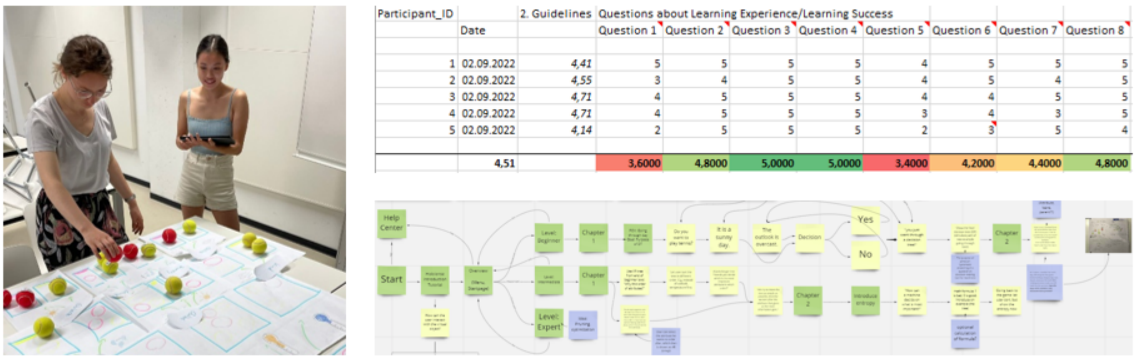


Figure 5.9.: Logging and Analysis

All used services can be found in figure 5.10.

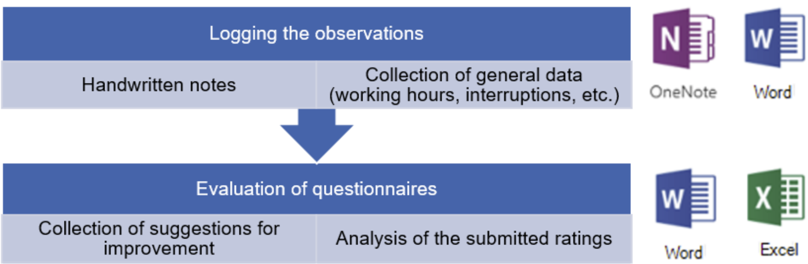


Figure 5.10.: Services for logging

5.2.1. Paper Prototype Results

Observations

In the course of the first user test, or rather during the team-internal preparations for it, we noticed that a compact explanation of what was happening, the intended goals, the functions available for use, as well as assistance (visual or auditive) would be invaluable for the interaction with the prototype and later also for the interaction with the artefact. Thus, not only did the idea of including "hints" (cf. hand menu) in each module emerge, but our design decision to include a virtual interlocutor (cf. Kai and Alice) to guide the user through the modules was also reinforced.
Furthermore, we realized that the interaction of the participants with the provided "paper buttons" for selecting options (e.g. criteria selection) proved to be very intuitive.
The frames provided to check if the positioning of the tennis balls was correct was also useful, as each participant used this permanently for self-monitoring. This ensured a better learning flow, as users did not permanently approach our team members with queries, which will be crucial for the final artefact (self-study artefact).
It was clear that the appropriate sizing of the individual elements (virtual buttons, virtual explanatory texts, weather icons on the physical tennis balls) as well as their coordination with each other (positioning, spacing, etc.) would be crucial for a complication-free learning experience using the built artefact.

Questionnaire Feedback

The evaluation of the questionnaire confirmed and extended our observations.
In figure 5.11 below, you can see the participants' ratings broken down by module, as well as the overall rating of the paper prototype. Here it is noticeable that with an overall rating of 4.6728 out of 5, as well as seven 5.0 ratings (out of all questions, which represents 25.9% of the questions), the prototype was very well received. The individual modules are rated higher overall than the general learning experience, which can be attributed to the fact that a cohesive learning experience cannot yet be provided using a paper prototype and all modules are viewed more as individual pieces.
The open-ended questions asked did not produce any implementable suggestions for improvement. The only request was for a more digital prototype, but the team decided against this from the outset in consultation with the supervisors due to the simplicity and rapid implementation of a paper prototype.

In summary, then, all of the design decisions made for the final artefact, such as interaction via virtual buttons, the inclusion of visual and auditive explanations and assistance, and the inclusion of options for self-monitoring and undoing actions, are justified and desired by the user.

5.2.2. User Study Results

Observations

Some problem points were also quickly apparent in the user study. For example, the limitations of the HoloLens 2 itself. The device tended to overheat very quickly, making it difficult for the participants and us to have a consistent user experience. In addition, the battery life was quite low which led to the same problem.
In general, it was also observed that most participants were overwhelmed by the technology itself, so some time of familiarization and trial and error had to be granted at the beginning. Increased effort or overwhelm due to too small/too many virtual components could also be observed among some participants.
The implemented reposition buttons were often used (mostly several times per module) due to the unintentional far interaction movements of the frames and proved to be extremely useful.
In addition, we observed that participants who already participated in the first user test positively perceived the progress as well as the new features of the application (cf. hand menu, hints, overview screen) compared to the paper prototype, but did not have a significant "usage advantage" compared to participants who came into contact with the project for the first time. For us, this represented a clear indication of ease of use. It shows that no prior knowledge (technical or thematic) is necessary to participate in the learning experience.

Questionnaire Feedback

The evaluation of the questionnaire confirmed and extended our observations.
The questionnaire ratings of the user study were also presented in the mapping scheme used previously (see figure 5.11).
It can be seen that with an overall rating of 4.51 out of 5, as well as eight 5.0 ratings (out of all questions, which represents 16.3% of the questions), the prototype was also very well received. Looking only at responses that averaged more than a 4.75 rating across all participants (top 5% responses), the percentage is as high as 34.7% (17 out of 49 questions).
The average rating of the modules (Onboarding, M1, M2, M3, M4) of 4.5314 is now also not seriously different from the adjusted rating of the coherent learning experience of 4.5142.
This speaks for a consistent and coherent user experience across the entire application.
As already noted in the observations on the user study, the participants also reported back via the open questions that in some cases the size ratio of some virtual components to one another and the coordination of virtual and physical components to one another did not fit. This deficiency has already been remedied. In addition the problem of the unintended far interaction movements of the frames was noted, which could be likewise solved.
The lines (clarification of the tree structure) between the frames in the modules M2-M4, desired by the participants, were consciously not implemented from the beginning, since the advantages in things clarification of the hierarchy would have been inferior to the disadvantages in things unclearness (by too many virtual components).

The desired German-language version of the artefact was also not implemented, but could be quickly implemented by making minor adjustments in JSON files provided for this purpose.

In summary, despite a few technical difficulties (appropriate sizing of physical and virtual components, random movement of frames due to unintended far interaction movements), which we quickly got under control, the artefact was very well received by the participants, who found the interaction intuitive and easy, and was sufficiently supported visually and auditorily (cf. observations in the first user test). All in all, no serious changes had to be made and the participants enjoyed using the artefact, which furthermore fulfilled its task of teaching DT concepts.

	Prototyping		User Study	
	Rating (out of 5)	*Duration of the modules (in min)*	*Rating (out of 5)*	*Duration of the modules (in min)*
Questionnaire Sections				
General Questions	4,6389	-	4,5000	-
Questions to "Onboarding Module"	-	-	4,4571	~ 5
Questions to "Module 1"	4,7917	~ 2	4,5250	~ 7
Questions to "Module 2"	4,7778	~ 18	4,4000	~ 9
Questions to "Module 3"	-	-	4,6000	~ 9
Questions to "Module 4"	-	-	4,7667	~ 10
Questions about Learning Experience/ Learning Success	4,2500	-	4,4000	-
	4,8333*	-	4,5142*	-
Avg. Rating of modules/ Avg. Duration of modules	*4,6728*	*(~ 10)*	*4,5314*	*~ 8*
People				
Best Rating	4,8519	-	4,7143	-
Worst Rating	4,4444	-	4,1429	-
Questions				
Best Rating (5,0)	7 out of 27	25,9%	8 out of 49	16,3%
Top 5% Rating (> 4,7500)	15 out of 27	55,6%	17 out of 49	34,7%
Worst Rating	2,5000 (1x)	-	3,4000 (1x)	-
	4,3333* (2x)	-	3,4000* (1x)	-

Non-informative durations are bracketed.
"*" - Adjusted value, because a ambiguous question would have led to a bias in the results.

Figure 5.11.: Analysis of the Questionnaires

6. Discussion

In this section, we briefly recap the results of the evaluation and address the limitations of this project.

6.1. Results

The results of this work were very promising, especially in the area of user experience. The users enjoyed working on the tasks, which led to increased motivation throughout the entire learning experience. The users also had the feeling of having understood the concepts and being able to apply them. Due to the high focus on generalizability, the data set can be exchanged at any time and can be made as simple or complex as desired.

6.2. Limitations & Future Work

This work comes with some limitations which are outlined in the following.

First, the evaluation groups consisted of only 5 participants. This number was intentionally chosen to be small because the focus of the project was the development of the software artifact and there was insufficient time to evaluate larger test groups. A group of this size is hardly sufficient to make meaningful statements about general acceptance, but it is enough to estimate a general picture and to support the design decisions made. In addition, all of these participants had an academic background. Within this project, we have specialized in students as a target group, which is why the purely academic background is not a problem, but no statements can be made about users with other backgrounds. An interesting topic for later would be to conduct evaluations with bigger groups from different backgrounds.

Furthermore, there are some limitations related to the software artifact itself. The biggest problem is the overheating of the HoloLens. As already explained, the constant overheating leads to a lower comfort level and interruptions that interfere the learning experience. The problem is probably that the app is not optimized enough. A future work could address these optimization issues and find out if the issue could be the app or maybe even the HoloLens itself. Another limitation of our app is that there are no verifications within the modules whether the user has properly sorted the tennis balls. It is assumed that the user himself compares his results with the given solutions. An automatic check would improve both the error-proneness of the process and the user experience and could therefore also be part of a later work. There is also a lack of evaluation of the overall learning progress within the app. Some questions or exercises at the end of modules could give the user feedback on his learning process and thus positively influence the learning process. This could also be addressed in a future work.

7. Conclusion & Outlook

In this project, we explored how mixed reality can be used to enhance learning experiences. We designed and implemented a learning application for basic concepts of data science on HoloLens 2 using the Microsoft MRTK for Unity.
During the project we learned that developing for brand new device in a new field is tricky. Issues with the documentation, as well as the technology, and our own unfamiliarity with it have caused problems and delays. On top of that, tools for prototyping do not exist like they do for web- or app-development, so we had to switch to paper prototyping.
In the end, the result of "Learning Data Science in MR" is a working MVP of a learning environment in our topic. It explains the idea and use of decision trees and how a computer can learn them from given data. Even though a lot of work still has to be done, it shows that learning apps can be transferred successfully to mixed reality.
In our evaluation, we learned, that most of our students were more motivated to use our app and learn something about the topic of data science, than with regular methods. The experience stayed in the mind of most of our testers and they were eager to test it. However, it is questionable to what extent this plus of motivation is only due to the use of a new technology, and whether this effect is sustainable. It still has to be evaluated, whether the learning result was adequate, too.
So far, it seems like for mixed reality learning applications to succeed, they really need to use the chances and advantages of the technology. Just showing holographic videos will not enhance the experience. But, for example in the fields of biology and astronomy, there are promising example projects. Otherwise, the overload of using a new and not yet finally developed technology may, so far, even out weight the advantages even for experienced user.

8. Declaration

Ich versichere hiermit wahrheitsgemäß, die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, den 15. Oktober 2022 ____________________

Remy Carosi, Maily Doan, Isabella Mebus Kishi de Oliveira, Julian Strietzel,
Luca Wolfinger

Appendix

More information, files and images about our project are provided below.

A. Questionnaires

A.1. Questionnaire - Prototyping

A.2. Questionnaire - User Study

B. Images

June 9, 2022

Teamproject: Mixed Reality Machine Learning
Prototype-Questionnaire

1. Personal Information

Name:
Date:

2. Guidelines

Complete the evaluation form using the following key:

1 = *Do not agree at all*
2 = *Do not agree*
3 = *Parts/Parts*
4 = Agree
5 = *Clearly agree*

3. General Questions

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The test could be performed in reasonable time (around 10min).	☐	☐	☐	☐	☐
Q2: The test could be performed without prior knowledge.	☐	☐	☐	☐	☐
Q3:					
The tasks were set unambiguously.	☐	☐	☐	☐	☐
The queries were answered adequately if necessary.	☐	☐	☐	☐	☐
The tests were performed in a precise and structured manner.	☐	☐	☐	☐	☐
How did you perceive the test environment?	☐	☐	☐	☐	☐

4. Questions to "Level 1"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q3: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q4: I noticed that the question types varied.	☐	☐	☐	☐	☐
Q5: The structure shown was clear and understandable.	☐	☐	☐	☐	☐
Q6: The explanation of the "DT model" was understandable.	☐	☐	☐	☐	☐
Q7: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q8: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 1 of 2

Figure A.1.: Questionnaire - Prototyping - Page 1

June 9, 2022

5. Questions to "Level 2"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The interaction with the system was intuitive.	☐	☐	☐	☐	☐
Q5:					
The selection of criteria with buttons was succeeded.	☐	☐	☐	☐	☐
The "virtual boxes" were helpful in verifying learning.	☐	☐	☐	☐	☐
The interaction between virtual and real objects was successfully implemented.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

6. Questions about Learning Experience/Learning Success

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The given task was physical/ mentally demanding.	☐	☐	☐	☐	☐
Q2: I was successful in accomplishing what I was asked to do.	☐	☐	☐	☐	☐
Q3: I do now understand the concept of a decision tree and what it is used for.	☐	☐	☐	☐	☐
Q4: I would recommend the system to someone, who wants to learn about Data Science.	☐	☐	☐	☐	☐

7. Personal Feedback

If you could magically change something about the prototype without consequences, what would it be?
Things that still need to be said....

Thank you for participating in our prototype test for the HoloLense application. The evaluation of the data will certainly help us to make the current prototype even more robust and user-friendly.

Page 2 of 2

Figure A.2.: Questionnaire - Prototyping - Page 2

Teamproject: Mixed Reality Machine Learning
Evaluation-Questionnaire

1. Personal Information

Name:
Date:

2. Guidelines

Complete the evaluation form using the following key:

1 = *Do not agree at all*
2 = *Do not agree*
3 = *Parts/Parts*
4 = Agree
5 = *Clearly agree*

3. General Questions

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The test could be performed in reasonable time (around 45 min).	☐	☐	☐	☐	☐
Q2: The test could be performed without prior knowledge.	☐	☐	☐	☐	☐
Q3:					
The tasks were set unambiguously.	☐	☐	☐	☐	☐
The queries were answered adequately if necessary.	☐	☐	☐	☐	☐
The tests were performed in a precise and structured manner.	☐	☐	☐	☐	☐
How did you perceive the test environment?	☐	☐	☐	☐	☐

4. Questions to "Onboarding Module"

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The user was appropriately welcomed and introduced to the topic.	☐	☐	☐	☐	☐
Q2: The user was appropriately accustomed to using the HoloLens.	☐	☐	☐	☐	☐
Q3: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q4: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q5: Everything was explained precisely and in a way that was easy to understand.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 1 of 3

Figure A.3.: Questionnaire - User Study - Page 1

5. Questions to “Module 1”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The questions were formulated in a simple and understandable way.	☐	☐	☐	☐	☐
Q3: The questions asked were easy to answer.	☐	☐	☐	☐	☐
Q4: I noticed that the question types varied.	☐	☐	☐	☐	☐
Q5: The structure shown was clear and understandable.	☐	☐	☐	☐	☐
Q6: The explanation of the "DT model" was understandable.	☐	☐	☐	☐	☐
Q7: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q8: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

6. Questions to “Module 2”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The individual weather symbols were easy to recognize.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The interaction with the system was intuitive.	☐	☐	☐	☐	☐
Q5:					
The selection of criteria with buttons was succeeded.	☐	☐	☐	☐	☐
The "virtual boxes" were helpful in verifying learning.	☐	☐	☐	☐	☐
The interaction between virtual and real objects was successfully implemented.	☐	☐	☐	☐	☐
Q6: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q7: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

7. Questions to “Module 3”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The user got the chance to make independent improvements to his original design.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q5: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

Page 2 of 3

Figure A.4.: Questionnaire - User Study - Page 2

8. Questions to “Module 4”

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The information gain was presented clearly.	☐	☐	☐	☐	☐
Q2: The introductory explanation of the task was long enough.	☐	☐	☐	☐	☐
Q3: The task was set unambiguously.	☐	☐	☐	☐	☐
Q4: The module has contributed to a better overall understanding of the DT concept.	☐	☐	☐	☐	☐
Q5: I did not feel overwhelmed at one point.	☐	☐	☐	☐	☐
Q6: I enjoyed working on this assignment.	☐	☐	☐	☐	☐

9. Questions about Learning Experience/Learning Success

	(5) = Excellent	(4) = Recognizable above the requirements	(3) = According to the requirements	(2) = Just meets the requirements	(1) = Among the requirements
Q1: The given task was physical/ mentally demanding.	☐	☐	☐	☐	☐
Q2: I was successful in accomplishing what I was asked to do.	☐	☐	☐	☐	☐
Q3: I do now understand the concept of a decision tree and what it is used for.	☐	☐	☐	☐	☐
Q4: The individual modules were well coordinated and built on each other.	☐	☐	☐	☐	☐
Q5: The virtual components were well sized for use.	☐	☐	☐	☐	☐
Q6: The auditory support from Kai and Alice was helpful.	☐	☐	☐	☐	☐
Q7: The use of the hand menu was intuitive and helpful.	☐	☐	☐	☐	☐
Q8: I would recommend the system to someone, who wants to learn about Data Science.	☐	☐	☐	☐	☐

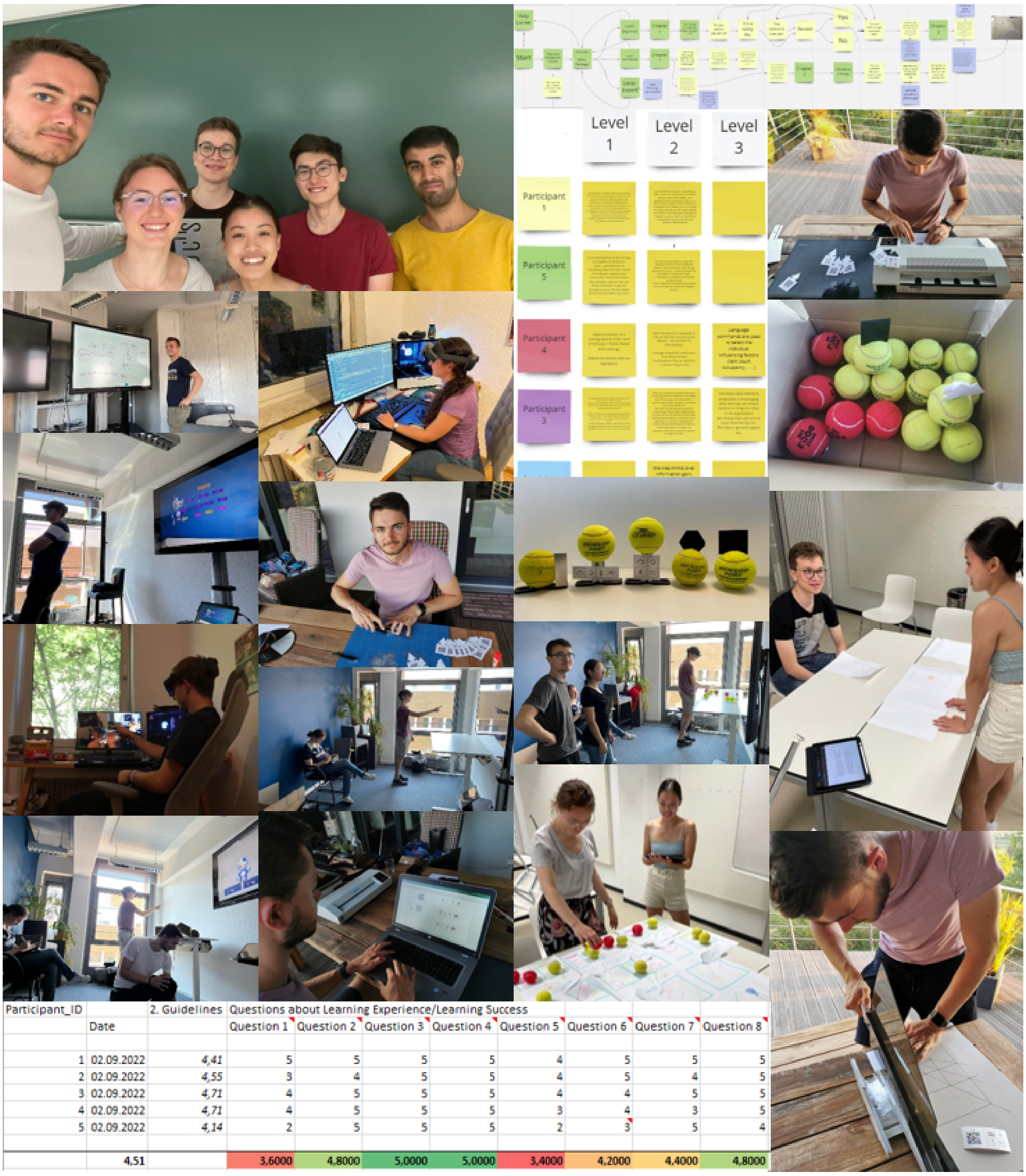
10. Personal Feedback

If you could magically change something about the application without consequences, what would it be?
Things that still need to be said....

Thank you for participating in our application test for the HoloLense.
The evaluation of the data will certainly help us to make the current application even more robust and user-friendly.

Page 3 of 3

Figure A.5.: Questionnaire - User Study - Page 3



Participant_ID		2. Guidelines	Questions about Learning Experience/Learning Success							
	Date		Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8
1	02.09.2022	4,41	5	5	5	5	4	5	5	5
2	02.09.2022	4,55	3	4	5	5	4	5	4	5
3	02.09.2022	4,71	4	5	5	5	4	4	5	5
4	02.09.2022	4,71	4	5	5	5	3	4	3	5
5	02.09.2022	4,14	2	5	5	5	2	3	5	4
	4,51		3,6000	4,8000	5,0000	5,0000	3,4000	4,2000	4,4000	4,8000

Figure B.1.: Image collection of the project

References

Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., & Chilana, P. K. (2020). Creating augmented and virtual reality applications: Current practices, challenges, and opportunities. , 1–13. Retrieved from https://doi.org/10.1145/3313831.3376722 doi: 10.1145/3313831.3376722

BillWagner. (n.d.). *Überblick über c# – Übersicht.* Retrieved 2022-09-22, from https://learn.microsoft.com/de-de/dotnet/csharp/tour-of-csharp/

Bodystorming for movement-based interaction design - ProQuest. (n.d.). Retrieved 2022-10-13, from https://www.proquest.com/openview/6cc8d3eb7f6fbe7535ae46ae18727da5/1?pq-origsite=gscholar&cbl=5514779

Gasques, D., Johnson, J. G., Sharkey, T., Feng, Y., Wang, R., Xu, Z. R., ... Weibel, N. (2021). ARTEMIS: A collaborative mixed-reality system for immersive surgical telementoring. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–14). Association for Computing Machinery. Retrieved 2022-09-19, from https://doi.org/10.1145/3411764.3445576 doi: 10.1145/3411764.3445576

Lafargue, D. (2018). *The influence of mixed reality learning environments in higher education STEM programs: A study of student perceptions of mixed reality self-efficacy, engagement, and motivation using augmented and virtual reality* (D.E.). Retrieved 2022-09-22, from https://www.proquest.com/docview/2207526133/abstract/960AC407B744121PQ/1 (ISBN: 9781392041963 Publication Title: ProQuest Dissertations and Theses)

Maas, M. J., & Hughes, J. M. (2020). Virtual, augmented and mixed reality in k–12 education: a review of the literature. *Technology, Pedagogy and Education*, *29*(2), 231-249. Retrieved from https://doi.org/10.1080/1475939X.2020.1737210 doi: 10.1080/1475939X.2020.1737210

Microsoft. (n.d.). *Solver overview - MRTK 2 | microsoft learn.* Retrieved 2022-09-27, from https://learn.microsoft.com/de-de/windows/mixed-reality/mrtk-unity/mrtk2/features/ux-building-blocks/solvers/solver?view=mrtkunity-2022-05

Norman, D. A., & Draper, S. W. (1986). *User centered system design; new perspectives on human-computer interaction.* USA: L. Erlbaum Associates Inc.

polar kev. (n.d.). *MRTK2-unity developer documentation - MRTK 2.* Retrieved 2022-09-22, from https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/

qianw211. (n.d.). *Choosing your engine - mixed reality.* Retrieved 2022-09-27, from `https://learn.microsoft.com/en-us/windows/mixed-reality/develop/choosing-an-engine`

Retrobot, game and app character, robot mascot | 3d characters | unity asset store. (n.d.). Retrieved 2022-09-22, from `https://assetstore.unity.com/packages/3d/characters/retrobot-game-and-app-character-robot-mascot-214185`

scooley. (n.d.). *HoloLens 2 hardware.* Retrieved 2022-09-22, from `https://learn.microsoft.com/en-us/hololens/hololens2-hardware`

Selim, H. M. (2005, Nov). *Critical success factors for e-learning acceptance: Confirmatory factor models.* Pergamon. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0360131505001338`

VanGundy, A. B. (1984). BRAIN WRITING FOR NEW PRODUCT IDEAS: AN ALTERNATIVE TO BRAINSTORMING. , *1*(2), 67–74. Retrieved 2022-10-13, from `https://doi.org/10.1108/eb008097` doi: 10.1108/eb008097