

Intelligent, socially oriented technology VI:

Projects by teams of master level students in cognitive science and engineering

Editors

Birger Johansson

Agneta Gulz

Magnus Haake

Mattias Wallergård

Jens Nirme

Eva-Maria Ternblad

Betty Tärning

Johansson, B., Gulz, A., Haake, M., Wallergård, M., Nirme, J., Ternblad, E. M., & Tärning, B. (Eds.) (2023). Intelligent, socially oriented technology VI: Projects by teams of master level students in cognitive science and engineering. *Lund University Cognitive Studies*, 185.

ISSN 1101-8453
ISRN LUHFDA/HFKO—5070--SE

Copyright © 2023 The Authors

Table of Contents

The Mystery of Ecoville	5
<i>Ebba Axelsson Nord, Emma Nicklasson, Arvid Bengtsson, Gustaf Eklund</i>	
Släpp en robot fri	31
<i>Daniel Carlström Schad</i>	
Ansiktsigenkänning och åldersestimering på en Epi-robot	45
<i>William Do, Stella Johannsen, Johannes Rosenfrost</i>	
Social Distraction in Virtual Reality	61
<i>Victoria García, Daniel Harrysson, Emil Johansson, Louise Pham</i>	
To Be a Robot - Simulating Robot Perception in Virtual Reality	83
<i>Mia Huovilainen, Vanessa Nguyen, Alexander Wåhlander</i>	
Under dina fötter	107
<i>Madeleine Loostrand, Charlotta Ekelöw Nilsson, Anna Dahlström</i>	
Arkimedes övertygelse: Inlärning och övertygelse av verkliga fenomen i virtuella lärmiljöer	115
<i>Daniel Zander, Sonja Holmer, Jacob Lundell, Joel Waye, Kaspian Jakobsson</i>	

The Mystery of Ecoville

Developing a digital learning resource for targeting data literacy in
12-year-olds

Ebba Axelsson Nord
ebbaxelsson@hotmail.com

Emma Nicklasson
em1656ni-s@student.lu.se

Arvid Bengtsson
ar5780be-s@student.lu.se

Gustaf Eklund
gu1444ek-s@student.lu.se

2023-01-17

This report details a project developing a digital learning resource prototype with the aim to strengthen 12-year-olds data literacy skills. Data literacy refers to the ability to critically collect, read, analyse and apply data. The learning resource is conceptualised in the form of a mystery game with embedded learning tasks providing clues needed to solve the mysterious events that occur in the town of Ecoville. A theoretical background is given in regards to the concepts of data literacy, feedback, game development, and design principles. Through an interactive work process, a high-fidelity prototype was created. The prototype was evaluated with participants in the targeted age group in a classroom setting. A central aspect of the educational game is a modified use of a teachable agent (TA) in the form of a socio-motivational peer-agent.

1 Introduction

In the digital age, data is everywhere. From an educational aspect, it therefore seems necessary that students are given the support needed to strengthen their understanding of data, and the ability to navigate it, i.e., students need to develop data literacy. In addition, being data literate also entails recognising that conclusions based on data can be effective and useful, while also sometimes deceptive and misleading.

This report describes how a digital learning resource called *The Mystery of Ecoville* was developed with the aim to increase the understanding of data literacy for students in the sixth grade (12-year-olds). The vision of the digital learning resource project was to complement the Swedish national curriculum relying on strategies for learning, such as the influence of teachable agents, the role of feedback and the use of gamification. The novelty of this project resides in exploration of data literacy as a relatively new concept (in educational settings) and more specifically in combination with the targeted age group that appear to have been understudied in relation to this topic. To develop the actual prototype, GDevelop was used as an IDE and an evaluation was done parallel to the implementation. A possible future

application for the learning resource is for it to be used at the Science Centre in Malmö that has expressed interest in the project.

The structure is divided into six sections. First comes this introductory section. In section two, we review literature relevant to this project. Next, in section three, we provide a detailed description of the methodology and development process for the prototype. A user evaluation is described in section four, followed by discussion in section five. Finally, future directions and ideas for the project can be found in section six. Throughout this report, we use the terms ‘digital learning resource’ and ‘(educational) game’ interchangeably.

2 Background

2.1 Data literacy

Our contemporary way of having a technology-based lifestyle comes with the implication of increasing data dependence. Through expanding amounts of screen time, children are exposed to data in different forms from a younger age, and the amount of information they are faced with is greater than ever before (Rideout, 2016). From this stems the need for bringing attention and action towards strengthening their ability to understand and interact with data. Data literacy is a relatively new concept discussed in corporate, educational and research settings. Our reliance on data is visible in many contexts, the Oceans of Data Institute (ODI) suggest that there are three main reasons for the importance of data literacy (2015, p. 2):

1. Our world economy and our jobs are increasingly defined by data and by the knowledge and skills required to use them effectively.
2. We are all perpetually producing streams of data, which we need to shape and manage to ensure our privacy and personal security.
3. Effective use of data empowers us to make objective, evidence-based inferences and fundamental decisions af-

fecting our lives, both as individuals and among societies.

It is evident that many aspects of everyday-life are centred around data. Thus, it is important that citizens have a good understanding of how to manage this data. This implies that data literacy is imperative for the society as a whole — not only for the rising generation. A knowledge synthesis report published by Ridsdale et al. (2015, p. 11) generated the concise definition of data literacy as the “the ability to collect, manage, evaluate and apply data, in a critical manner.” With the collective understanding that data is essentially everywhere and is continuing to increase in quantity, it also needs to be acknowledged that the same thing cannot be said about the quality of the produced data. Arguing with data is part of the definition of data literacy (Ridsdale et al., 2015) and involves using data to support a larger narrative intended to communicate a message to a particular audience (D'Ignazio & Bhargava, 2016). The ability to critically think about data is central to all aspects of data literacy. This entails not only to be critical towards the source or production of it, but also in regards to how the data is presented and visualised. Hence, being data literate means to have a good understanding of several features, including the collection, analysis, visualisation and implications of data.

The growing need for data understanding suggests that it should be reflected in educational systems. Being data literate is, indeed, a civic skill that is needed for all current and future citizens of the digital world (Inverarity, Tarrant, Forrest, & Greenwood, 2022; Oceans of Data Institute, 2015; OECD, 2019; Pangrazio & Sefton-Green, 2020; Wolff, Gooch, Cavero Montaner, Rashid, & Kortuem, 2016). Furthermore, data literacy is even increasingly considered to be a life skill (Wolff et al., 2016) and central for being well-informed members acting in democratic societies.

Data literacy in education

In 2019, the organisation for economic co-operation and development (OECD) published a framework with the aim to set out an aspirational vision for the future of education, called the Learning Compass 2030. In this framework, data literacy is included as being part of the cognitive foundations which is one of three identified core foundations for 2030 together with health foundations, and social and emotional foundations (OECD, 2019). The framework further states that the skill is needed for using mathematical tools and reasoning in everyday life, including in digital environments (OECD, 2019). Thus, it appears to be a key skill and central to the future of education.

Together with the curriculum, the Swedish elementary school (grundskolan) publishes a mission statement (Lgr22, 2022, p. 7, translation by the authors) in which it is expressed that:

“Students should be able to navigate and act in a complex reality with a vast flow of information and increased digitalisation at a fast rate of change.

Skills and methods to interact with- and use new knowledge are therefore important. It is also necessary for students to develop their ability to critically review information, facts, and relationships in order to understand the consequences of different outcomes.”

Related to this, the goal with data literacy in the educational systems should be to strive towards making sure that students feel comfortable interacting with data, trust in their ability to evaluate data, and distinguish “good” and useful data use from “bad” that could be misinformative. The reason for this is because not being data literate comes with the risk of taking biased interpretations of data as facts, which could lead to students being misinformed (Wolff et al., 2016).

However, in the Swedish national curriculum there are no explicitly expressed requirements for data literacy. Still, criteria can be found embedded in various subjects including mathematics as well as the social- and natural sciences. In mathematics, under the topic of probability and statistics, students in grade four through six are expected to “use tables and graphs with and without digital tools and have the ability for interpretation of them” (Lgr22, 2022, p. 57, translation by the authors). Students are also expected to know about the three measures of central tendency in statistics: mean, mode and median, as well as how these measures are used in statistical inquiries. That includes having an understanding of the implications that can be derived from these different central tendency measures, in a statistical sense. In the social sciences, students are expected to “know the media’s role in regards to information and opinion as well as be able to critically analyse how digital (and non-digital) media can be used responsibly from social, ethical and legal aspects” (Lgr22, 2022, p. 198, translation by the authors). Here, media refers to newspapers, radio and TV, but also social media and other online sources. In addition, students are expected to have knowledge about different perspectives of current social issues and “determine how message, sender and purpose can be reviewed critically, and in relation to its source” (Lgr22, 2022, p. 198, translation by the authors). The curriculum for social science repeatedly emphasises the foundations of democracy and a democratic society through all grades. Similarly, the curriculum for natural sciences also includes the ability to critically analyse information. In addition, students are expected to “know about planning, executing and evaluating results and how to present these using words, images and graphs” in relation to all three natural sciences (Lgr22, 2022, p. 156, 163, 169, translation by the authors).

Lgr22 (2022), the current version of the Swedish national curriculum, is newly updated and brought into action in the beginning of the academic year 2022-2023. This current version (Lgr22, 2022) emphasises the importance of understanding what characterises a fact and how all subject knowledge has an inherent value. This being explicitly stated is a change from the previous version

from 2011 (Lgr11, 2011, revised 2019). Another change can be found in the social sciences, where the updated version of the curriculum has added a new learning objective that focuses on reviewing current civic matters. In addition, understanding of the media's role is further strengthened in Lgr22 (2022). The updates reflect a call for students to have more advanced civic skills in regards to physical and non-physical (online) aspects of the community.

How should data literacy be taught?

There are many different pedagogical approaches and methods of teaching data literacy. Some studies emphasise the importance of a creative process (e.g., Bhargava, Kadouaki, Bhargava, Castro, & D'Ignazio, 2016) or having students collect their own data for further inquiries (e.g., Williams, Deahl, Rubel, & Lim, 2014; Wolff, Wermelinger, & Petre, 2019) to facilitate better understanding. The different approaches to teaching data literacy vary due to it being a multifaceted skill and therefore, more difficult to target. Consequently, data literacy, as any literacy, is probably best understood as a process of learning rather than simply being learned (Pangrazio & Sefton-Green, 2020). Another challenge when dealing with data is its partly non-transparent nature. Once students are presented with processed data, they are often without insight to how it was collected or without opportunity to see the actual data points. This is a challenge for educational contexts and a reason for the importance of an action-oriented approach where the focus is both on using data and understanding data (Van Audenhove, Van den Broeck, & Mariën, 2020). It is also worth noting that there is much work to be done and that cohesive research targeting data literacy pedagogy is currently lacking (Wolff et al., 2019).

2.2 Feedback

Feedback is a critical component in nearly any educational setting and is used to improve learning. In line with Shute (2008), feedback can be defined as any information that an agent (e.g., a teacher, peer, or oneself) provides regarding how well (or poorly) one is performing on a task. Shute (2008) and Hattie and Timperley (2007) both reviewed studies that reported feedback to be more effective when it includes a corrective (or critical) part as well as a constructive part. This kind of feedback is encapsulated by the term critical constructive feedback (CCF), a term which is used, for instance, by Tärning et al. (2020). The critical element signals discrepancies between one's current performance and the goal of the task i.e., the desired performance. It addresses the accuracy of the selected response and identifies what mistakes or other shortcomings were made while the constructive element identifies areas of improvement by providing guidance that can help the student attain the learning objectives. This guidance can, for example, consist of relevant details, such as giving cues about the task so the student can revise their answer (Shute, 2008).

The challenge related to feedback is to get students to not only read the feedback, but also to act upon it. In a study by Silvervarg, Wolf, Blair, Haake, and Gulz (2021) where CCF is investigated it is suggested that the influence of a teachable agent (see section *teachable agents*) might have a positive effect on the students' use of feedback. One suggestion for this is that feedback often is the result of failure of some sort and having something or someone (e.g., a teachable agent) as a buffer between the learner and the task at hand might take some possible insecurity away as the responsibility can be shared (Silvervarg et al., 2021). The issue with students not using the feedback provided was further investigated by Tärning et al. (2020). The study used a new approach where feedback was divided into five stages: noticing, decoding, making-sense, acting upon and using feedback. The findings from the study suggest that the feedback that is "surviving" each stage is steadily decreasing. In addition, this means that just because the feedback text is read does not equal it being acted upon or used. Hence, providing feedback that is used by students through all five stages is difficult and something that needs to be given much thought (Tärning et al., 2020).

2.3 Game Development

The following subsections describe the underlying theory used in the project in regards to game development.

Gamification

Imagine you need to learn 10 new verbs in French this week. How would you go about it? You could do it through a more traditional approach, such as re-reading the words from a paper. Or you could perhaps ask your mother, partner, or dog to test you over and over again. Or you could download a gaming app designed to make learning French verbs exciting, fruitful and sexy (at least according to the advert). Although not a new phenomenon, taking inspiration from video games and applying it to new environments has grown extensively over the last decade (Nacke & Deterding, 2017). The term gamification encapsulates this idea. Deterding, Dixon, Khaled, and Nacke (2011, p. 10) define gamification as "[...] the use of game design elements in non-game contexts."¹

To get further insight into gamification the definition needs to be unpacked. Deterding et al. (2011, p. 12) restrict gamification to "[...] the description of elements that are characteristic to games – elements that are found in most (but not necessarily all) games, readily associated with games, and found to play a significant role in gameplay." In this view, game elements are identified as the building blocks of games. The focal point of gamification is the *design*

¹Not all researchers would, however, agree with this definition of gamification (e.g., Huotari & Hamari, 2012). This can be attributed to the fact that the term has been used inconsistently within the fields of Human-Computer Interaction (HCI) and game studies (Deterding et al., 2011).

of a game (Deterding et al., 2011). That is, gamification describes the design process. This separates it from closely related terms, such as “serious game”, “educational game”, “applied game”, and “funware”. The fundamental difference is that these parallel terms describe an end product (Laine & Lindberg, 2020). It is worth noting that gamification does not have to be bound to any digital medium. Although not as common, gamification design can be applied to practices that occur in the offline world, for example, in a classroom (Deterding et al., 2011). Another aspect the definition highlights relates to the application area, namely a *non-game context*. Unlike entertainment games, the objective does not lie in the entertainment value (Deterding et al., 2011). One non-game context where research has seen the potential of adopting gamification techniques is in education (Kapp, 2012; Schaaf & Mohan, 2014).

Gamification in education

The goal of using gamification strategies in an educational context is to encourage learning. This is achieved by seeking to integrate game design elements that motivate learners through engaging learning experiences (Kapp, 2012). Explaining the effect of gamification requires a theoretical understanding. The question concerns how we can understand the motivational mechanisms that give rise to the desired behavioural outcomes, e.g., engagement and learning. Studies within the field of gamification often fail, however, to provide theoretical grounding (Nacke & Deterding, 2017; Zainuddin, Chu, Shujahat, & Perera, 2020). For those studies that do offer a theoretical model the most popular is Self-Determination Theory (SDT) (Nacke & Deterding, 2017).²

According to SDT, motivation can either be *intrinsic* or *extrinsic*. Intrinsically motivated behaviours are constituted by the inherent value that the individual either finds from within or from the relation to the activity, e.g., finding the activity meaningful and getting enjoyment from it. In contrast, the driving force of extrinsically motivated behaviours has to do with its instrumental value. In this sense, one’s motivation for a behaviour or engaging in activities depends on the external rewards that can be gained in doing so (Ryan & Deci, 2000). Let’s look at an example to illustrate the difference. On one hand, a college student may be intrinsically motivated to read the course literature because of a genuine interest in the subject. On the other hand, a college student might do it for the sake of a passing grade and to avoid negative consequences (for example, flunking out of college or disappointing the parents). In the latter case, the college student is extrinsically motivated.

Systematic reviews found both positive and negative effects when applying gamification to education. General findings from Hamari, Koivisto, and Sarsa (2014) and Zainuddin et al. (2020) suggest that gamification has a positive influ-

ence on learning outcomes. Zainuddin et al. (2020) found several studies that show how game design elements such as points, badges, levels, leaderboards, and avatars can be used as extrinsic motivators to increase engagement among learners. There were, however, articles that reported contradictory findings. While some of the reviewed papers concluded that both extrinsic and intrinsic motivation could be influenced by gamification strategies, others questioned the effectiveness of using external rewards to create intrinsic motivation (Zainuddin et al., 2020). Laschke and Hassenzahl (2011) stress the importance of including a narrative in a gamified system. They criticise gamification on the account that it heavily relies on external reward mechanisms, such as badges and points, rather than tapping into intrinsic motivation by creating a meaningful experience. Almeida, Kalinowski, and Feijó (2021) conducted a systematic mapping of game design elements’ potential negative effects. They found badges, competitions, leaderboards, and points to have either a lack of or a negative impact on learning. This was attributed to the competitive learning environment these game design elements may cause. According to Almeida et al. (2021), gamification designers should be aware of the potential negative effects certain game design elements may have on users. If not, these elements run the risk of undermining their intended purpose (resulting in, for example, decreased motivation or performance loss).

Furthermore, Hamari et al. (2014) found studies that ascribed the results of gamification to a novelty effect, i.e., the effect decreased over time. This highlights the need for future work to study the long-time effects of gamification. So far, only one study (Barata, Gama, Jorge, & Gonçalves, 2017) was found to have collected data over an extended period of time (Zainuddin et al., 2020).

New developments in gamified learning

Recent research has shifted the focus from points, badges, and leaderboards to other game design features (Nacke & Deterding, 2017). Laine and Lindberg (2020) conducted a systematic review of previous literature to look at how gamification strategies could help design educational games. The study identified game motivators and game design principles meant to increase an educational game’s motivational pull and offer support to the learning process. Laine and Lindberg (2020) targeted game motivators meant to, at first hand, address intrinsic motivation. The authors, however, emphasised that these motivators could influence both intrinsic and extrinsic motivation. This depends on which game design elements are used and in what context they operate. It is noteworthy that Laine and Lindberg (2020) do not propose using all design principles in the same educational system. Rather, this taxonomy should be seen as a guideline for educational game designers.

Among others, Laine and Lindberg (2020) list the following game motivators: *challenge*, *control*, *feedback*, *goals*, *storytelling* and *fantasy*, and *social interaction* (or *social*

²Other established theories in gamification research are flow theory and the goal-setting theory. See Zainuddin et al. (2020) for an overview of learning theories applied in gamification research.

play). In short, *challenges* can be a motivational factor when they are congruent with the player's skill level. *Control* can be a motivator if the player experiences that he or she can manipulate the game world. Furthermore, *control, feedback, goals*, and *challenge* are closely intertwined with each other. The challenges in an educational game often consist of different tasks designed to teach the player a particular subject (e.g., data literacy). For this to be successfully achieved, the player must understand what to do in the learning task and how to do it, i.e., what the goal is. The tasks must also include feedback, so the player becomes aware of what he or she is doing right and what can be done differently. In other words, the educational game must deliver both the goals and feedback in a clear and timely fashion. If this is done correctly, a sense of control emerges. Another way for games to be engaging is by creating a meaningful experience through storytelling. This is what Laine and Lindberg (2020) pinpoint with the game motivator *storytelling and fantasy*. When storytelling is used as an instruction method, a narrative-driven game can make the learning content more memorable in addition to immersing the player in the game world. Moreover, immersion and the sense of control can be strengthened if the player experience that s/he can influence the narration of the story. Laine and Lindberg (2020) emphasise the importance of having the game's story be relatable and appropriate for the targeted users (a storyline suitable for a younger audience may, for instance, not appeal to older players). Lastly, *social interaction* has been found to play a key motivational role in learning. One suggestion from Laine and Lindberg (2020) is for educational games to offer the player the possibility to interact socially. One way this can be implemented is to have the player interact with *non-player characters* situated in the game.

Another avenue that is beginning to be explored within the educational domain is the use of tailored gamification. The basic idea of tailored education is the use of an adaptive gamification approach, where the gamification experience is adapted based on, for instance, the students' character profile, e.g., personality traits. There is a growing interest within the research community to better understand when and how gamification can positively influence learning (Oliveira et al., 2022). For example, Barata et al. (2017) studied how students' individual differences can affect how they experience and are engaged by gamified components. The authors draw the conclusion that it may be beneficial if the gamification experience can be adapted depending on the learner's needs and characteristics. Oliveira et al. (2022) did a literature review on tailored gamified education. Being such a recent topic, they conclude that it is, at this time, not obvious if tailored gamified educational environments will benefit learning outcomes. On one hand, preliminary findings from Oliveira and colleagues (2022) indicate that a tailored educational system can benefit learning. Students appear to spend significantly more time in this sort of system compared to what they did in non-tailored systems. Non-tailored systems, on the other hand, appeared to be better at motivating students. In the end, even if research

has to some extent shown promising results, there is still a lack of empirical and longitudinal studies on the subject.

Teachable agent

A teachable agent (TA) is a character (computer agent) whose purpose in a digital learning resource is to be taught something by the student playing, and promote learning through the *learning-by-teaching principle* (Sjödén, Tärning, Pareto, & Gulz, 2011). With the use of artificial intelligence, a TA can seemingly learn and reason about the things that it has been taught. An important feature of a TA is that it is both a reflection of what it has been learned from the player but also that it provides the illusion of having a mind of its own (Chase, Chin, Oppezzo, & Schwartz, 2009). The benefits of having a TA in the context of educational technology include support for increased engagement and motivation from students since the task is to learn for the sake of someone else, rather than just for themselves. This is called *the protégé effect* (Chase et al., 2009) which, in addition to engagement and motivation, also is beneficial for learning. Thus, many positive effects of TAs are in regards to the social aspects of learning. For instance they can act as an *ego-protective buffer* between the student playing and the knowledge it targets. If, for example, the TA is failing a test, it can take some of the blame, or the failure can be attributed to poor teaching rather than some "internal" factor within the student. The ego-protective buffer seems to be extra beneficial for low-achieving students with low self-efficacy (Chase et al., 2009).

Kirkegaard (2016) and Tärning, Silvervarg, and Gulz (2018) have investigated how a TA's personality traits can influence how the students experience a TA. Although Kirkegaard (2016) did not establish substantial evidence of an effect, the results indicated that the TA's personality could benefit students differently depending on the student's level of self-efficacy. A student with high self-efficacy benefited more from a TA that expressed a more questioning nature, while students with low self-efficacy preferred a TA that was more compliant. Tärning et al. (2018) reported that students with low self-efficacy performed better with a teachable agent that exhibited the same level of self-efficacy, i.e., a lower level of self-efficacy.

Studies investigating TAs also suggest that a strong social bond between the student and agent is to be desired and can be strengthened through spoken or written conversation (Gulz, Haake, & Silvervarg, 2011). The conversations could be related to the task but it is also important with off-task topics. Findings presented by Gulz and colleagues (2011) suggest that socially oriented off-task conversation can contribute in a positive sense to the player's experience and improve learning. In addition, off-task conversation strengthens the social relationship between player and agent which seem to have positive effects on reaching learning goals in the longer term, such as learning resources with an embedded TA (Gulz et al., 2011). In this context, off-

task conversations could be small talk, reassuring, or mutual self-disclosure. All of which are aspects of off-task sociability. There appears to be many benefits to having a TA in a digital learning resource, however, findings also suggest that it is important that the agent is introduced in a way that makes it clear for the student what is to be expected from the TA (Silvervarg & Måansson, 2018).

Teachable agents can also influence how students respond to feedback. Critical constructive feedback (CCF) refers to feedback that both make students aware of mistakes (*critical*) and help them improve and make progress (*constructive*) (Silvervarg et al., 2021). Being faced with critical feedback is never pleasant, and the inclusion of a teachable agent can improve students' experience by being able to share the failure with the TA (the *ego-protective buffer* effect). It has been observed that lower-achieving students are less likely to register and interact with feedback, possibly as a result of trying to protect oneself from critique. However, findings demonstrated by Silvervarg and colleagues (2021) suggest that having a TA present can decrease the feedback neglect.

2.4 Design guidelines

Fitts' law

Fitts' law determines the time needed for a user to interact with their target, by using the relationship between the size and the distance of the target (Budiu, 2022). If a set of controls are frequently used in sequence it is recommended to put these close to each other in order to optimise task time by minimising the distance (Budiu, 2022).

Norman's design principles

Norman (2013) defines seven principles of design that can be used in all types of design to improve usability and make the user experience more intuitive. The seven principles are:

Feedback in regard to the interaction design is an important tool when making a system or product. With feedback a system can tell the user that it is reacting to and is processing an action. Without any response it can be hard to tell if what the user does has any impact and can leave them wondering what to do. A simple button in a GUI can light up in colour or play a sound when clicked to indicate that the input has been registered.

Mapping is the relationship between controls and their corresponding effect. For example the arrow keys on a keyboard can have the corresponding effect of moving something on screen in the direction of the arrow shown on the key.

Discoverability is a concept used to weigh how easy it is to determine what actions are possible and what the state of the device is. The easier this is to determine the better the discoverability is.

Conceptual models are abstract and psychological representations of how tasks are to be performed. They are frequently used subconsciously and intuitively to systemise processes.

Affordances exist all around us and are essentially possible actions that objects offer or afford us, what the user is capable of doing with the object.

Signifiers are a way to optimise affordances, through optical cues the designer can indicate to the user that an action is possible and how to take that action.

Constraints are a way to lead the user into taking a certain course of action by removing other choices. For example in a GUI some buttons may be greyed out and unusable, leaving the user with only the chosen affordances. As the Hick-Hyman law tells us, the more options a user is offered, the more time is spent making a decision (Soegaard, 2020).

Design patterns

Design patterns are used as a way to make design choices that can easily be recognised by users. Preece, Rogers, and Sharp (2015) define them as a solution to a problem in a certain context that has been found to work effectively. It is also described as an extension of common practice. A problem with using common practice is, however, that they are not necessarily good practices. Bad practice of design patterns are known as *anti-patterns* and it is critical when implementing a design to make sure to avoid them. The reason that anti-patterns exist and are used despite this can be attributed to technological changes. An example of anti-patterns is using a pattern that works well on a desktop computer, but translates poorly to a mobile device.

3 The Development Process

The aim for this project was to create an educational (mystery) game with the possibility to be implemented and used by the Science Centre in Malmö. In addition, the game of Ecoville could be used as a resource in research for mapping and exploring how students (in grades 5-7) can be supported in their progress of data literacy. Prior to the start of this project, a technical and conceptual outline was provided that included a possible storyline and ideas to some of the learning tasks in the game. This basis was available for the group to use and interpret as wished. The following methodology section describes the process from idea to the final version of the prototype.

3.1 The iterative process

During the course of the project the work has followed a clear workflow with three main iterative cycles. The iterat-

ive cycles has been carried out in four steps. The first step is *ideation*, during which the ideas and underlying concepts are defined. This is followed by the step of *creating alternatives*, during which the ideas and concepts are used to specify more tangible alternatives for what direction the project is going to take. After the alternatives are created, they are summarised and become the basis for the third step, *creating prototypes*. During this step, a prototype is created according to the alternatives defined in the previous stage. The last step in the iterative process is the *evaluation* of the finished prototype. The iterative cycle is then repeated until the initial design goal is reached. The resulting feedback and thoughts that come from the last *evaluation* step are discussed in the first stage to generate new ideas which, hopefully, improve the product.

3.2 The Storyboard

The storyboard that existed for this project outlined the general story of the digital learning resource (targeting data literacy) that was setup around learning tasks. The outline involved the player solving a mystery as to why a small town (called Ecoville) was having issues with the environment or, more precisely, a groundwater shortage, and what was causing it. The storyboard also included the idea of tossing a coin in the town's famous wishing well as being part of the introduction to the game, while also being a motivation to find out what happened to the groundwater. A decision was made to use this general idea from the storyboard.

3.3 Initial decisions and Prototype 1

Having read the background information and the pre-existing storyline suggestion, the first decision that was made was to focus on the initial scenes, including the first learning task. Based on discussions and brainstorming sessions, the basic idea was to make a more detailed version of the game's beginning and expand on this as the project went along. This meant that focus was not put on the larger mystery as a whole, but rather the introduction of the game. Next followed a decision to make our own version of a storyboard in the format of a low-fidelity prototype.

The characters

During the first iterative cycle, primary choices were made regarding the characters. This initial idea was that the player would have an avatar as its *player character* and would act as a side-kick to a *non-player character* in the game. Another decision (made rather quickly) was the non-player character to be both a teachable agent and an in-game guide. The rationale behind this double role was the delicate balancing act that needs to be considered in game design, namely how to guide the player through the game world without breaking immersion, that is, the feeling of being deeply involved in the game world (Wirtz, 2022). Furthermore, it is equally important that the player

feels some sense of control over how the game progresses, something that can be especially tricky when the game has a linear narrative, where the player follows an already set story progression that ultimately ends in the unfolding of the mystery. In other words, the game does not allow the player to influence the storyline by, for example, skipping ahead. The reason for having the non-player character take on the role of guiding the player was, thus, primarily decided as a way to structure how the game's story elements unfold and to pace the game well. This is to help the player stay immersed and in a flow state. An improved example on how to tackle the problem of having the player experience control while having a linear narrative is presented in subsection *Prototype 2*. Why we wanted the non-player character to *also* take on the role of a teachable agent was because of the many benefits of teachable agents found in research, particularly in terms of their positive effects on student's learning and motivation (see subsection *Teachable agent*). Since the overall objective with the game is for players to learn data literacy, we wanted to explore every possible avenue that could potentially support the learning process. Teachable agent was one such avenue.

The pre-defined storyboard included a biologist as the character the player would first encounter in the story, and this idea carried over into the project's first prototype. The main reason was that this character (which we named *the Biologist*) worked in the overarching story. In addition, the authors felt that a biologist as a character would fit the idea of a teachable agent as well as being on theme for environmental issues and sustainability.

There was at first a large discussion over what the player's avatar would be which was then settled in a few brainstorming sessions, with a set of different options being defined. These options included the player simply getting to choose between two avatars before the game begins, the avatar being a tourist, a detective, or an animal. In the end, the chosen alternative was for the player's avatar to be in the form of a cat (most probably based on the project group's preference).

To summarise, these initial decisions led to the first prototype having two distinct characters, the player, and the teachable agent.

Other features

One feature that came up in an early brainstorming session was the idea of providing a game map. The map can be used as a tool to help the player navigate and keep track of the progression (the player's status) in the game. The map is meant to depict the town of Ecoville and its surroundings. In the beginning of the game, the idea is to have the game map showing clear signs of a drought as a result of the water shortage in Ecoville. As the player gathers more clues about the mystery and completes the learning tasks, different sections of the game map will be unlocked and turned from a yellow to a green landscape. When the player has completed the game by solving the mystery the

whole game map depicting Ecoville will, ultimately, have turned to a flourishing, green landscape. This use of a game map is supported by Laine and Lindberg (2020) as a gamification strategy that can be used to raise curiosity, and by extension keep the player engaged and motivated.

The learning tasks

As already discussed, the idea is that the player follow a clearly marked path and go through the learning tasks in a pre-specified order. The reason for this is partly due to having a linear narrative where the mystery unfolds over time, and partly as a way to be certain that the player develops the right competence, i.e., develops their skill in *data literacy*. The learning tasks will become more challenging as the player progress in the game, but in a way that is relative to the player's skill level. For this to be achieved, we need to be certain that the players have obtained the knowledge that is required to be able to move on and face more difficult tasks. As previous literature has suggested (e.g., Laine & Lindberg, 2020; Norman, 2013), challenge can be a powerful motivator. However, if the learning task is too difficult the player may be discouraged to continue playing and quit the game altogether. This can also happen if the challenges are too easy, which can lead to the player feeling bored. To put it differently, experiencing the learning task as either too challenging or not challenging enough can lead to motivational loss and a negative impact on learning.

Constructing learning tasks on an appropriate level is not an easy thing to do, especially since the initial project plan targeted the ages 11- to 13-year-olds (Swedish school grades 5-7). Not only is there a huge difference in what students know between these ages, there is also variation in ability and interests within the same age. One solution to tackle this problem was to limit the game's target audience to grade six (i.e., 12-year-olds). To gain better insight to what aspects of data literacy that are challenging for 12-year-olds, one group member visited a school in Eslöv to assist with data collection (a screening test) and making observations in three sixth grade classes. The main findings from the visit is that there are indeed large differences between individuals and classes, both in knowledge levels as well as in the general level of focus and interest. However, along with previous observations the exercises in the test were found to be manageable by the students, suggesting a reasonable level of difficulty. In our prototype, the introductory learning tasks were inspired by the exercises in the screening test, hence, it was useful to see students solve similar tasks and listen to their reasoning. For the first learning task, the player is instructed to help the biologist find the correct bar chart (among four alternatives) that correspond to his collected data about the amount of different animal species in Ecoville. The data is presented in a table next to the four different bar charts and the task is to identify and click on the correct graph. This first learning task is deliberately made to be less difficult than the following,

in order to spark interest and to increase the students' motivation. The purpose of the first part of the task (1a) in relation to data literacy is to learn how data in a table can be visually represented, and what makes a representation correct or incorrect (see Figure 1). In the second part of the task (1b), when students have identified the correct graph, they are presented with another graph that represents the corresponding data from the year before, and are then asked to reflect on whether the number of animals have increased, decreased or remained the same since the previous year (as shown in Figure 2). The purpose is to introduce students to the larger mystery by having them reach the conclusion that the number of animals are decreasing in Ecoville. Evidently, this is not good and something must be causing this.

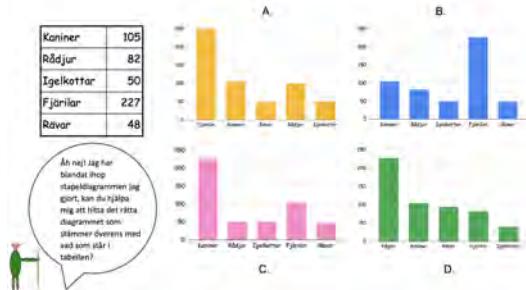


Figure 1. Learning task 1a in Prototype 1

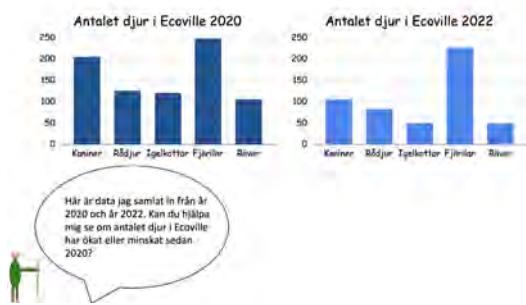


Figure 2. Learning task 1b in Prototype 1

The story

Regarding the story, choices were made to include two short introductory subtasks with the purpose to introduce the player to both data literacy and the mystery to be solved in the game. How to introduce the mystery was one of the first major challenges in the process of creating this digital learning game. The challenge as such consisted of how to introduce sufficient instructions that allow the player to proceed with the game. This relates both to information about the game's controls (i.e., *how to play the game*) and the game's objective (i.e., what the player needs to know to finish the game). Since a game is a rule-based system (Laine & Lindberg, 2020), the player needs to be made

aware of what the goals are, both regarding the main goal and subgoals throughout the game. Having clearly defined goals are, thus, one of the fundamental components of a game. Put simply, it is important that the player knows what to strive for (or, as Norman (2013) would put it, provide a good conceptual model by making the affordances, constraints and mappings visible). At the same time, it is important not to cognitively overload the player with too much information all at once due to the risk of the player ignoring the instructions. Considering the psychology of the game player (specifically, the limitations on memory and attention) is emphasised by Norman (2013) as a vital step to make the game more understandable and easier to use. Furthermore, not having clear goals or bombarding the player with too much information could dampen the motivational pull of the game (Laine & Lindberg, 2020). Having considered these potential issues, our solution consists in portioning out the information on what the game is about (i.e., what the main goal is) in the first couple of scenes. This decision was based on a design principle proposed by Laine and Lindberg (2020), where storytelling is used as an instruction tool. How this can be achieved in a clear and effective manner has been the primary focus in the first version of the prototype (Prototype 1).

In short, the introduction to the game world and characters starts with the player travelling to the town Ecoville to toss a coin into the town's famous wishing well. On the way, the player happens upon the Biologist, who is working on counting the animal population in and around Ecoville. The player is then pulled into the game's first learning task which involves helping the Biologist who has collected data but mixed up four bar charts visualising it. The player must choose the bar chart (identifying one of four alternatives) that best illustrates the data in the table stating the number of animals the Biologist has counted. When this first part of the learning task is completed, the second part of the learning task pops up. In the second part, the Biologist asks the player to reflect on whether the animal population in Ecoville has increased or decreased compared to a previous year. The correct answer is that the number of animals has significantly decreased in Ecoville. This is illustrated through a comparison of two bar charts with similar representations from 2020 and 2022 respectively, where a clear decrease of animals can be noticed in the latter diagram. This results in the Biologist wondering why the number of animals has decreased and inferring that it must have something to do with the recent groundwater shortage that Ecoville is experiencing. Why Ecoville has a groundwater shortage is a mystery and something that, the Biologist concludes, needs to be explored further. Perhaps, the Biologist asks, the player would like to take on the quest of solving what has happened to the groundwater?

Even if there already existed a narrative for the game before the authors of this report became part of this project, the decision to keep the plotline (i.e., solving what happened to the groundwater) was based on the idea of providing what

Laine and Lindberg (2020) refer to as an *epic meaning*. An epic meaning, such as saving the environment in Ecoville, relates to the final goal of the game. Epic meaning has linkages to the *storytelling and fantasy* motivator (which is described in subsection *New developments in gamified learning*). It is, according to Laine and Lindberg (2020), often employed as part of the storytelling as a way to make the player feel connected to something greater than themselves.

Implementation and technical aspects

The design choices made during the creation and implementation of the prototypes have been mainly focused on conceptual decisions. This led to the initial prototype (Prototype 1) not having a lot of effort put into usability aspects. As a result the only interaction design theories that were utilised in the implementation of Prototype 1 was in the form of design patterns, and Norman's design principles.

There are many design aspects of games that are seen as common practice and that are natural for most who have interacted with games. A few examples of this in Prototype 1 are that progression is made from the left side of the screen to the right, and that the dialogue-text is displayed in speech bubbles. Other design patterns used in Prototype 1 is the use of the left button on a computer mouse to interact with the game.

The learning tasks were the main interaction between player and game, and in that interaction some more thought was put into Norman's design principles. Examples of this are the feedback with changing colour tints of the answers, indicating if the correct answer was chosen or not. The answer turns green if it is the correct one, and red if it is incorrect. Other than this the only interaction that has had some design related consideration is the method of progressing the game state forward. In the prototype this is applied with the user pressing the left mouse button while hovering the mouse over a circular button with a arrow pointing to the right, with the idea of using the arrow as a signifier for the affordance of progressing the game state.

The design aspects and the consideration of design principles have purposefully been a lesser focus during the first part of the project (Prototype 1) and will have a larger role in next two prototypes.

Development and evaluation of Prototype 1

With the conceptual decisions above being made the first prototype was developed. The first step was to create a lo-fi version of the prototype to be the base for a hi-fi prototype. This was achieved in Google Sheets, as it fit the needs of a lo-fi prototype needed for the project. The lo-fi prototype defined the dialogue between the player's avatar and the Biologist, as well as the narrative text to be displayed in the game. The lo-fi prototype also included a version of the first learning task. This was then translated

into a hi-fi prototype using GDevelop. A description and evaluation of GDevelop is presented in *Appendix A*.

This first implemented prototype was then evaluated within the project-group and with the help of the project supervisors to gain an understanding of what could be used going forward and what changes could be made. In parallel with the brainstorming sessions and the creation of the lo- and hi-fi prototypes, a more thorough review of the existing literature on teachable agents was made. The insights gained from this was that the non-playing character called the Biologist was not a pure teachable agent, and had a poorly defined role in the game. It was unclear what the Biologist's role was in the eyes of the player. The dual role, of both acting as an in-game guide and a teachable agent, was deemed to be problematic following findings in Silvervarg and Måansson (2018) that the effect of a teachable agent depended on how the teachable agent's role was introduced and framed.

Following recommendations from our supervisors, the use of the cat as the player's avatar was discarded. Using animal-based characters is more suitable for games targeting a younger age-group. In addition, a recommended practice is to not mix both human and animal characters in the same game since this could potentially lead to confusion among the players.

Furthermore, it was established that the feedback in the initial state of the learning task was poor (partially due to a decision *not* to focus on feedback at this stage in the process). Feedback is a key feature of a digital learning resource. It offers support to the learning process and can function as a game motivator. At the time of evaluation, the only type of feedback that had been implemented was instant feedback meant to make the game system more intuitive when playing. It also contributes to the player experiencing a sense of control since the game reacts to the player's inputs, following Norman's (2013) design principles presented on page 6. As described above, the feedback in our prototype indicated that something in the game could be pressed. If the mouse hovered over, for example, a button in the game, its sprite lit up. This provided information regarding whether the player's choice in a learning task was the correct or incorrect answer. If the player chooses the incorrect answer the sprite lights a red colour and if the answer is correct the sprite lights a green colour. Our ambition was, however, to expand the function of feedback to be more informative. How this was achieved are presented in subsection 3.5 (*Prototype 3*).

3.4 Prototype 2

Revisit and update of the characters

In light of the evaluation of the first prototype, new ideas were defined and a new lo-fi prototype (Prototype 2) was created. This included adding a second non-playing character, so that the two roles that the Biologist initially filled would be separated into two non-playing characters. This decision left the story with three characters rather than

two. The new characters would be a non-player character with teachable agent-features (the, so called, *peer agent*), a non-player character that will act as a guide, and the player.

One complication with this decision pertains to how both non-player characters should be introduced to not repeat the same mistake where the roles of the non-player characters are ill-defined. That is, how can the roles of the non-player characters be introduced and framed to elicit the desired effects among the game players (see subsection *A new take on the story* for a description of how the characters are introduced). Moreover, since the player now had to keep track of multiple characters in the game a decision was made to change the perspective of the player. By removing the cat as the player's avatar and, instead, making the student playing the game become the player character *her- or himself* the hope is to minimise confusion. Another reason is to deepen the experience of immersion and flow. In removing the player's avatar we, hopefully, removed the extra step of the player having to build a relation to and identify with his or her avatar.

A decision was made to keep the Biologist as one of the non-player characters. In this prototype, however, the Biologist is only used as an in-game guide. The other non-player character has the role of a peer agent and is called *Holly*.

Using the non-player character Holly as a form of peer agent could act as a socio-motivational factor. It is possible that Holly's presence during the game's learning tasks could increase the sense of support, which is in line with findings demonstrated by Sjödén and colleagues (2011). The reason for having Holly as a peer agent was that it fits the game's narrative better, as opposed to having Holly be a teachable agent in the more classical sense. Holly differs from a "classical" teachable agent in that a teachable agent is always tested on the subject matter by a third party. In contrast, Holly's relation to the player will be more similar to the kind of relation the player has with his or her classmates. In other words, Holly will basically interact with the player as a peer. The idea is then to have Holly take on the role of a tutee and the player as the tutor in the data literacy-tasks. In situations relating to solving the mystery these roles may, however, alternate. In these cases, Holly will be the driving force and provide suggestions when, for example, something seems off and should be investigated further.

It would be interesting from a research perspective to investigate if a character with teachable agent-features like Holly, can elicit similar benefits (like, the protégé effect and the ego-protective buffer effect) as a genuine teachable agent. There is, to our knowledge, not much research on this topic. Our digital learning resource could here potentially be used as a research tool to explore this further. In section 6, future plans on how to develop Holly as a character and peer agent are presented.

A new take on the story

The changes to the plotline are as follows: the first scene where the player learns that he or she is travelling to

a town called Ecoville to toss a coin into the wishing well is now changed to the player learning that s/he, *together with a friend called Holly*, are travelling to Ecoville. On the way, the player and *Holly* happen upon the Biologist. The player and *Holly* are then pulled into the first part of the initial learning task. It is, however, revealed that Holly does not understand data literacy and expresses a desire for the player to teach Holly how to solve the first task. The second part of the learning task is the same as it was in the first prototype. The Biologist asks the player and *Holly* to take a stand on whether the animal population has increased or decreased since last year. Once again, Holly expresses a desire to be taught to understand how to interpret the data. With it being the introductory task, the idea, alongside introducing data literacy, is to strengthen the social bond between the player and the Biologist (he asks for Holly's help) as well as the social bond between the player and Holly (Holly asks for the player's help). When the second part of the learning task is completed the Biologist wonders why the number of animals has decreased and infers that it must have something to do with the recent groundwater shortage that the town Ecoville is experiencing. However, it is now *Holly*, instead of the Biologist, that stresses that the mystery of what has happened to Ecoville's groundwater needs to be investigated. Holly then turns to the player and asks if the player would be interested in helping her solve this mystery. The Biologist, who thinks this sounds like a great idea, proposes that the player and Holly should visit the café in Ecoville to search for more clues. The question of whether we have successfully introduced the plotline of the game was something that was evaluated, see section 4 (*User evaluation*).

Having Holly turn around and ask the player to take a stand on whether s/he would like to help solve the mystery (or not) is one way to enhance the player's perception of control. Other potential ideas of how the player can feel that they are influencing the game events is by choosing between dialogue options, routes to take, etc. Even if these choices lead to the same result, regardless of what choice is made, it can give the player a perceived freedom of choice. By making the game more interactive in this way, it will, hopefully, strengthen the game's motivational impact and positively affect the learning outcomes. When considering user-centered design principles, Norman (2013) highlights the importance of striking the right balance between automation and having the game player perceive a feeling of control. Taking away too much control can run the risk, according to Norman (2013), of the player adopting a more passive role.

3.5 Prototype 3

In this iteration, a third version of the lo-fi prototype was created which expanded on the ideas made in Prototype 2. The main goal of the third version was to create a testable prototype that could be evaluated in a classroom and with the intended target audience of Swedish 6th graders (see section 4 (*User evaluation*) for more details). Prior to the

evaluation, the third lo-fi prototype was translated into a playable beta-version using GDevelop. The focus of the third prototype has been on further developing functional aspects in the game that could support the user experience and the learning process (cf. subsections *Feedback* and *Other features*). In addition, two more learning tasks were added to expand the prototype in length, and updates on the existing learning task was made.

Feedback

One update in the existing learning task (and added in the new ones) relates to feedback, with the main objective for the feedback to become more informative and not only signal whether the chosen answer was correct or incorrect. With this update a feedback-textbox will automatically appear upon completing a learning task, which is done when the player selects an answer among the alternatives presented in the task. The reason for having the feedback-text appear immediately and cover a bigger portion of the screen was to increase the possibility of the player engaging with it. Research also claim potential benefits to having the feedback-text appear immediately after the player has selected a response. For example, Shute (2008) concluded that the value of immediate feedback is that it could lead to more efficient learning since it can help with real-time error correction.³

In the case that the player has chosen an incorrect alternative in a learning task, the textbox that appears will now have CCF-content. As discussed in *Background* (section 2), feedback that is critical and constructive is in line with the view that previous literature holds on what makes feedback effectively enhance learning (e.g., Black & Wiliam, 1998; Hattie & Timperley, 2007; Tärning et al., 2020). Following recommendations from e.g., Shute (2008), all CCF-content directly relates to the learning task (is contextualised) and not the player. Feedback (usually praise) that draws attention to the self and away from the task has been shown to be less effective than feedback that solely focuses on the task (e.g., Black & Wiliam, 1998; Hattie & Timperley, 2007). Something to consider when formulating the CCF-text is how much information the feedback message should contain. The feedback guidelines proposed by Shute (2008) to enhance learning state that feedback should be simple and specific. Specific in that it links to the performance and the learning goals of the task. Feedback should only generate enough information (e.g., clues or hints) to aid students in solving the task without giving too much away. If feedback is not presented in manageable units, it may invoke cognitive overload. This could, in turn, hamper learning (Shute, 2008). Formulating feedback suitable for everyone is challenging since some students may require more support than others. The focus in our prototype has been on trying to keep the CCF-text as

³The debate on feedback timing is, however, far from settled due to the fact that research on immediate versus delayed feedback has revealed inconsistent findings. See Shute (2008) for a discussion on the topic.

simple and to the point as possible while still following the template of the feedback message having one critical and one constructive part.

An example of how the CCF-text is written in Prototype 3 can be found in the first learning task. Here, the player is asked to figure out the bar chart that best represents the data in a table. As shown in Figure 3, one CCF-text the player receives when making an erroneous judgement is: “The pink bar chart and the table do not belong together.” This is followed by: “Find information to solve this task by looking closely at the height of each bar.”

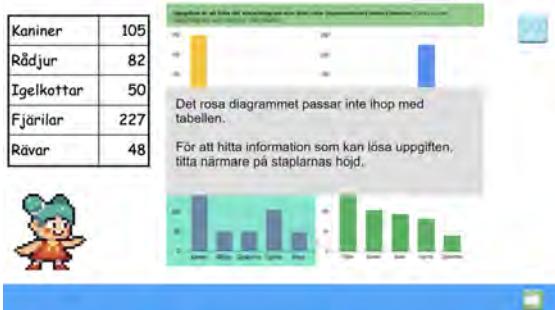


Figure 3. Example of CCF-text in learning task 1a in Prototype 3.

As mentioned in the literature review on feedback (section 2.2, e.g., Tärning et al., 2020; Silvervarg et al., 2021), previous studies indicate that students often choose to not engage with feedback – no matter how informative it is. According to Hattie and Timperley (2007), this can be ascribed to students perceiving it as evaluative punishment. Tärning et al. (2020) and Silvervarg et al. (2021) demonstrated that one way to mitigate the problem of feedback neglect is to make in-game characters direct their gaze and point toward something worth taking note of. This framing strategy was adopted into Prototype 3 by having the peer agent Holly use visual signals (specifically, point and gaze) toward the CCF-text (see Figure 3). It should be clarified that Tärning et al. (2020) and Silvervarg et al. (2021) both studied visually signalling via a teachable agent. This differs from our prototype in the respect that Holly is a peer agent, not a teachable agent in the classical sense. If students have a greater inclination to read CCF-text when Holly provided visual cues is something that will be discussed in section 4 (*User evaluation*).

Furthermore, a time constraint was implemented that makes it impossible to continue the game until all the informational text has been presented on the screen. The rationale was that this could prompt players to read the text and, in particular, the CCF-text. There is, however, a balance between the amount of time information can be frozen onto the screen before the player sees it as an annoying feature. As a way to make this less noticeable, the dialogue-text in the speech balloons unfolds word-by-word. It is only when all the text in a speech balloon has appeared on the screen that the player can continue onto the next scene. The

amount of text determines the time constraint which was a 0.05 second delay between each letter, creating the unfolding effect. Hence, having the words roll out provides a kind of invisible time constraint while ensuring the player that the game progresses as intended. See section 4 (*User evaluation*) for a discussion on the effect of the time constraint.

Additional improvements as to feedback were in the form of usability in that buttons and other objects were augmented with signifiers so that potential interaction would be clearer. One example of this is the use of the forward button that progresses the game state forward. When this forward button is active and can be clicked, it changes colour from grey to green. In the game, this interaction is first introduced with a tutorial pop-up that appears the first time it is possible for the player to perform the action. This method of having a tutorial pop-up is used for most interactive elements that can be clicked in the prototype. This includes, for example, the information button during learning tasks (for more details, see subsection *Other features*), the speech bubble over Holly that opens into dialogue during learning tasks, and the magnifying glass that symbolises that something can be inspected more closely (e.g., a graph or the newspaper). One exception to this is, however, when the player chooses an answer in the learning tasks. When the player is in the process of choosing an answer, hovering over one of the alternatives changes the tint of it, indicating that interaction is possible.

Learning tasks

In Prototype 3, some additional updates on the first learning task have been made. These changes have primarily consisted in the revision of text. Another change, suggested by our supervisors, was to not have the Biologist deliver the feedback-message in the first learning task. The reason for this was that it seemed a little strange for the Biologist to identify if the answer was correct or not since the learning task is about helping the Biologist find the correct bar chart. In the updated version of the prototype (i.e., Prototype 3), the feedback-message appears in a text-box. The updated version of the first learning task is shown in Figure 4 and 5.



Figure 4. Learning task 1a in Prototype 3

One considerable change made for Prototype 3 is the addi-



Figure 5. Learning task 1b in Prototype 3

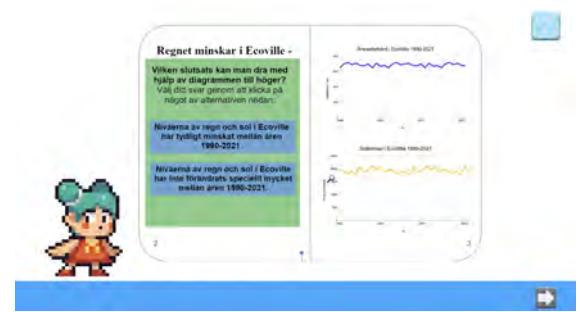


Figure 6. Learning task 2a in Prototype 3

tion of two new learning tasks that the player is instructed to solve. The second learning task is located by the player when they, together with Holly, enter the city centre and are about to visit the café (a suggestion from the Biologist) when they find a newspaper. When reading the newspaper, an article is visible with the heading “Less rain in Ecoville - lowest annual precipitation in 2 years!” to which Holly suggests the possibility that this might be the reason for the empty wishing well. For the first part of the second learning task (see Figure 6), students are being presented with two different line charts, one representing the annual precipitation in Ecoville between 1990-2021 and another line chart representing the annual amount of sunshine in Ecoville between 1990-2021. To reflect over what conclusions that can be drawn from the two graphs, the student is asked to choose between two alternatives. The first option states that the levels of precipitation and sunlight have clearly decreased during the last 30 years, while the second option states that over the last 30 years, the levels of precipitation and sunlight have not varied that much. As the lines in the graphs increase and decrease over time (with no visible trend), the learning goal is to get students to reach the conclusion that the levels of precipitation and sunlight have not clearly decreased over the last 30 years. This, in turn, relates to the second part of the learning task in which students get to decide if the heading in the article is misleading and the opportunity to change it to a more accurate one (see Figure 7). The purpose of the second learning task is for the students to reach the conclusion that because the levels of sunlight and precipitation have remained relatively stable, they cannot be the reason for the empty wishing well, as Holly suggested. Together with the observation that the original article heading was misleading and not a good reflection of the data presented in the graphs.

When Holly and the player continue to read the newspaper they come across a short article (with a corresponding pie chart) on the event of the one year anniversary of the soda factory *Läksa* in Ecoville (see Figure 8). This is linked to the third learning task in the final version of the prototype (Prototype 3). In the article, the local reporter have made a survey asking 100 people in the town for their preferred beverage: juice, lemonade or soda (by *Läksa*). The data is represented in a pie chart. Holly suggests that



Figure 7. Learning task 2b in Prototype 3

the article together with the graph implies that everyone in the entire world has *Läksa* (the soda) as their favourite beverage. The task for the player is to choose their response to Holly’s statement about what conclusions can be drawn from the data presented in the graph, by choosing one of three alternatives as their response. Either Holly is right in her statement about the inquiry and the player agrees (two of the three alternatives are versions of this option), or she has reached an incorrect conclusion and the player does not agree with Holly. The third learning task requires understanding of the relation between sample and population together with why Holly’s statement needs to be corrected. The goal is to make students reflect on the distinction between sample and population as well as to introduce *Läksa* (the soda factory) to the game and the mystery as a whole, as *Läksa* will have an increasingly important role for the unraveling of the mystery.



Figure 8. Learning task 3 in Prototype 3

Other features

Additional changes to the dialogue between the Biologist and Holly were made in Prototype 3 following suggestions from the supervisors as well as from the teachers we interviewed (see subsection 4.3 *Teacher interview* for further details). Changes included breaking up the text into smaller chunks to make it easier to read, and that consistency in language and terminology was used throughout the prototype.

An information button was added to the learning tasks and placed in the upper right corner on the screen. When clicking on the information button a text-box appears. In this text-box one can find general information on e.g., how to read different graphs. For example, the information in the text-box during the first learning task is a description on how to read and interpret a bar chart while, in the second and third learning task, one finds general information about line charts and pie charts respectively. In this way, the information button functions as a supportive tool that can be called upon if the player is in need of some guidance. However, no additional information is presented in the text-box which would be required for solving the task, i.e., the content is general information and the learning tasks can be solved without reading them. When confronted with a new learning task it can be difficult to know where to begin — especially if the task is perceived as challenging. The information button is meant to provide cues on how to interpret the graphs to make the task more manageable. It, thus, provides some direction which can help the player achieve learning goals. Fitts' law was taken into consideration during the implementation of this information button. The information button was placed in the periphery (i.e., the upper right corner) on the screen as it is not a vital button to progress the game state forward. When the information button is clicked, the text-box with information appears on screen. When the player wants to close the text-box, the exit button (in the form of a X-icon) is located near the information button that opened it. The information button is in the form of a question mark icon, as shown in, for example, Figure 8. Following Norman (2013), these icons were chosen as a result of their conceptual models. The question mark is conventionally used to indicate that extra information can be provided that offers help to the player. The X-icon is used to indicate that clicking it will close a window.

In an attempt to make Prototype 3 partially adaptive, the player is asked to enter their name at the beginning of the game, before the playing session starts. This input is later used at several occasions in dialogue-text that allows for the characters to address the player by his/her name. The idea is that this feature strengthens the social relationship between the player and Holly during conversation. Having Holly directly refer to the player by their name could strengthen the experience of Holly as a peer, as well as increasing the player's motivation to help and the feeling of

working together with her (Holly) during the learning tasks.

4 User evaluation

4.1 Research goals

The goal of the study was to evaluate the final prototype in order to identify strengths and weaknesses when used with the intended target audience. Importantly, we wanted to do this in a situation reflecting an ordinary learning environment in which the educational game is meant to be used in (i.e., a school), in order to increase ecological validity.

As a first stage, an online-interview was conducted with two middle school teachers. The teachers offered a pedagogical perspective by vetting all the text (including dialogue- and CCF-text) and the learning tasks to assess if the content were appropriate for the age group. Further details on this will be given in the subsection 4.2 (*Teacher interview*). The input from the teachers was implemented into the third version of the prototype, which was used in the evaluation with the targeted age group (described in subsection 4.4 *Study*).

4.2 Teacher interview

Before the user evaluation with the students from the targeted age group, the prototype (in a lo-fi format) was given to two middle-school teachers for them to look through. The teachers were not related to the class in which the user evaluation was later conducted. The purpose for involving the teachers was to get their input on the general tone of the language in the digital learning resource (especially in regards to the dialogue- and CFF-text), the level of difficulty of the learning tasks, as well as any other opinions or reflections they may have. Before reading through the material, the teachers were given instructions and questions about the prototype (see *Appendix B*) to keep in mind when reviewing it, which was later discussed during an online-interview. The 'instructions' and 'questions' included a short description of data literacy along with questions about the prototype for additional contextual information. During the interview with the teachers, the first part focused on the prototype. The second part of the interview concerned questions on their previous experience with using digital learning resources as well as how they tackle data literacy in the classroom. To clarify, the two teachers were first sent the prototype together with 'instructions' and 'questions' for them to look through. An online interview was then conducted where their feedback was discussed in combination with questions both in regards to the prototype and more general ones about the teachers view on data literacy and educational games (see *Appendix C*).

Working with the teachers, we got a chance to get valuable information which was used to make further adjustments to Prototype 3 used in the evaluation. The feedback from the interview mainly concerned specific word choices that were either inconsistent or potentially difficult for the students to understand. One suggestion from the teachers that was

brought up was the possibility of having some form of collective revision of some of the key concepts in the game before students play it, as this might be useful to ensure that no concept (specifically, concerning data literacy terminology) is unfamiliar to the player once the game is started. For example, the teachers did not think the students would understand which axis in a graph one is referring to when using the terminology *the vertical axis* and *the horizontal axis* if this is not explained to the students beforehand. Other potentially difficult words in the prototype were discussed regarding the dialogue-text and the use of expert words. Input from the teachers suggested that using difficult words or concepts can be beneficial for the students, yet it is important to ensure that the students understand their meaning. The teachers proposed that the character named *the Biologist* should explain what a biologist does since the students may not be familiar with this profession.

The aim of the teacher interview was to get a better understanding on how digital learning resources and data literacy are used in practice. It offered an unique perspective on what teachers may look for when choosing an educational game to include in their teaching plan. One suggestion from the teachers was the idea of embedding the educational game in the ongoing classroom activities. The teachers highlighted the need for educational games to be integrated into the curriculum and played while students are taught the same subject in the classroom. This way, a teacher can ensure that the students possess a basic understanding of the concepts that they need to achieve the learning goals covered in the educational game.

When answering the question on what, in their eyes, makes an educational game be perceived as a good educational tool, the teachers replied that the educational game should offer dialogue audio. This was seen as such an important feature that one of the teachers stated that she would not choose an educational game if it did not have audio. Furthermore, the teachers saw the need for a feature in the educational game that kept track of how many tries a student had made in a learning task before they succeeded.

When it comes to data literacy the teachers recognised that data literacy as a concept was new to them. After familiarising themselves with the concept through our description of data literacy, the teachers concluded that focus in the classrooms has primarily been on the notion of source criticism as well as to teach more concrete abilities like how to read graphs and tables of various kinds. This contrasts to other abilities that data literacy encompasses, for example the ability to analyse the message or narrative in relation to the empirical data visualised in the graphs. The kinds of graphs which, according to the teachers, will be the first that students learn to interpret are bar- and pie charts.

4.3 Formulating research questions

The aim of the evaluation was to get feedback and insights on our prototype from the targeted audience. Since the prototype is still under development we wanted to explore how

the educational game worked in general terms, primarily focusing on the content and composition. This includes investigating the functional design and flow of the prototype. If the prototype does not possess functions that make it easier to perform tasks and leads to a smooth flow in gameplay, it can negatively affect understanding the objective of the learning game. This may, in turn, distract the players from understanding what they are supposed to do.

Thus, our primary research questions for the user evaluation were: (i) How do students in the targeted age group understand the premise (e.g., the main goal, characters and social relationships) of the game? (ii) How did the students perform and understand the learning tasks? Specifically, we wanted to know if the difficulty level was at a reasonable skill-level, given the students current knowledge of data literacy.

4.4 Study

Participants

The study was carried out in a public school in the south of Sweden. A total of $N = 25$ participants (14 girls and 11 boys) from two sixth-grade classes (student ages 11 to 12) participated in the study. All students had middle class socioeconomic background.

Learning activities

The version of the prototype that was evaluated includes three different kinds of learning tasks. The goal when creating the learning activities was to not only build on students' ability to read charts and graphs but also involve them in other data literacy skills, such as developing an understanding of how to analyse and draw conclusions from data. The first learning task is related to mathematics and the knowledge of the visualisation of data using bar charts. The second learning task targets the ability to see trends in line charts (how something has changed over time) and use that knowledge to draw conclusions and value arguments. The third learning task is related to understanding sample versus population and what inferences can be drawn from visually represented data when combined with a descriptive text on how the data was collected.

Procedure and methods for data collection

The user evaluation began with all students receiving a verbal introduction in the classroom where the project group members introduced themselves and explained the purpose of their visit. The students were tested in pairs (except for one case where a group consisted of three participants) in a group room not far from their home classroom. How the students were paired up was decided by their teacher. This decision was based on how the students were seated in their home classroom.

Before giving more detailed instructions on the task, all

participants were asked to give their consent for sound recording during the evaluation. These sound recordings were later deleted and only used to compare with notes taken from each of the sessions. The students were seated in front of a laptop with a computer mouse. They were told that the prototype was still under development and that they were only going to play the first introductory scenes of the game. No information on the game's narrative was given to the students beforehand since this was one of the things we wanted to evaluate. To make the students more comfortable, it was emphasised that the prototype was under assessment — not the students themselves. Before they started the game, the students were encouraged to speak freely with each other about what they were experiencing and thinking during their own gameplay, i.e., the students were prompted to follow a think-aloud protocol.

Two members of the project group were present in the room to lead and guide the evaluation procedure, observe and take notes as well as assist with potential interaction-related or technical issues. It was decided that the project group members would primarily stay quiet but at the same time be available for any specific questions that might arise from the students. The project group members refrained from assisting the students in solving any of the game's learning tasks and only interacted with the participants to remind them to keep talking. At one occasion, a student pair were encouraged to talk with each other, as they had fallen silent during gameplay. Each session (i.e., playing the prototype from start to finish) lasted approximately 15 minutes.

After completing the interactive game play session with a think-aloud protocol, all students took part in a 5-minute semi-structured interview probing the student's experience of the game. The questions targeted the narrative (i.e., what, according to the students, the main goal of the game was), which characters the students met in the game and how they perceived them, and what the students thought of the learning tasks. The first two question topics related to the first research question and the latter to the second research question.

When the student pairs had completed the playing session and the interview, they were thanked for their participation before returning to their home classroom. Due to limited time, only the teacher were given a debriefing that was later passed on to the students.

The reason for the students being tested in pairs was to provide a more natural and 'safe' setting for the students to share what they were doing and thinking. Concurrent think-aloud methods can be used to access information on a participant's cognitive processes (Ericsson & Simon, 1980, 1998). This method of collecting data has been employed widely in various fields, among them in education research (e.g., Branch, 2000; Reinhart et al., 2022).

The interviews were used to explore the findings from the think-aloud playing sessions in more detail. It offered an additional and complementary way to investigate the research questions. The interviews allowed the students to describe

their experience of the prototype in their own words to see if it aligned with the observations from the project group members. In addition, conducting a retrospective interview allowed for minimal interference during the playing session with think-aloud protocol. This was to minimise disturbing the students' gameplay experience which, in turn, resulted in a more accurate evaluation of the prototype.

4.5 Results

Learning tasks

The findings from the evaluation study indicated that the students' overall impression of the prototype was positive. All participants seemed to find it interesting (which was elaborated during the interview) and were seemingly motivated to complete all the scenes and learning tasks in the prototype. The time duration for completing the prototype was on average 15 minutes, which was in line with prior expectations. The learning tasks had different results in regards to student performance. The first of three tasks was deliberately made to be easier to complete as a way to boost the students' confidence and motivation for the following tasks. This is reflected in the performance (see Figure 9), with the first part (1a) being solved on the first try by 92% of the students (11 of 12 pairs) and the second part (1b) by 100% of the students (12 of 12 pairs). The second learning task was found to be more difficult as the first part (2a) was answered correctly on the first try by 58% of the students (7 of 12 pairs), and the second part (2b) by 75% of the students (9 of 12 pairs). The third and final learning task (3) was answered correctly on the first try by 33% of the students (4 of 12 pairs).

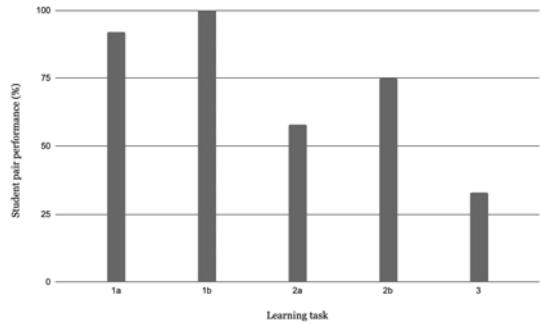


Figure 9. Student performance in the learning tasks with the percentage of student pairs answering correctly on the first try

This naturally gives cause to reflect on the level of difficulty for the task. It was, however, observed that in most cases students chose the wrong alternative simply because they did not read all the required text in the task or even all of the alternative answers. For example, one student pair stated in the interview that they did not bother to read the last (and correct) alternative in the third learning task since they were certain that one of the other alternatives was the

correct answer. This is another challenge for digital learning resources (as well as learning resources in general) that will be further explored in the *Discussion* (section 5). When asked about the learning tasks in the interview, all student pairs commented that the difficulty level was reasonable. The second and the third learning tasks were described as being more difficult by the participants. This is in line with the overall performance level on these learning tasks, as they had a lower success rate. One student pair stated that they appreciated that the learning tasks went from relatively easy to becoming more difficult as the game progressed. Three student pairs commented that the learning tasks required the participants to put in effort when trying to solve them. For instance, one participant, when talking about the learning tasks, said that "You really had to read it multiple times to make sure what it said and compare with the images [the graphs] and ask yourself: Is it really so?" (translation by the authors).

As for the feedback, it was observed that a majority of the student pairs did not read the CCF-text when confronted with it in a learning task. When the participants did engage with the CCF-text, it appeared to provide the student with some help in how to solve the task in question. For instance, during the third learning task the CCF-text redirected players back to the newspaper and encouraged them to read the article more carefully. The students that did as the feedback suggested, more quickly solved the task successfully.

The narrative

When asked about the narrative, 10 out of 12 student pairs said that the main goal of the game was to solve a mystery about what had happened to the town's water supply. Many of these student pairs made connections between the mystery and how it related to what they learned from the learning tasks. For example, student pairs concluded that the disappearance of water was the cause for the decrease of animals in Ecoville. Three student pairs made positive remarks on the game having an overarching mystery. Having a goal to aspire to and not only being faced with different tasks makes, according to these student pairs, the experience of playing an educational game more enjoyable. All students (12 of 12 pairs) commented during the interview that they thought playing the game was either "fun" or "kind of fun". In addition, many students gave a positive answer to playing the game again when it is a final product. For example, one student added during the interview: "I would like to see what happens". Another student started to speculate on the cause of the groundwater shortage in Ecoville.

Two student pairs did not explicitly link the mystery to a shortage of water. Instead, they mentioned that the educational game was about solving different tasks that related to reading graphs about how the animals in the town had decreased, how the annual amount of rain and sunshine had changed over the years, and favourite beverages. For

one of these two student pairs, the group project members observed that they did not take the time to read all the dialogue-text in between the learning tasks. Instead, this student pair tried to move through the game as quickly as the time constraint let them, only stopping and taking their time to read the text on the screen when solving the learning tasks.

Characters

All of the student pairs could correctly identify the two characters (Holly and the Biologist) that they had met when playing the game. When asked about the characteristics of Holly she was often described as 'curious', 'cool', 'kind', 'helpful', 'interested', 'adventurous', and 'someone who enjoys solving mysteries'. One student commented during the second learning task that "Can you [Holly] do something by yourself for once?" (translation by the authors), indicating that it was clear to the student that Holly's character was dependent on the player for solving the learning tasks. This was, however, the only negative comment in regards to Holly during the user evaluation. During the interviews as a whole it was clear that the students perceived Holly as a joyful character that sometimes is not very clever and needed their help. However, no student indicated that they felt a responsibility for making sure Holly gains understanding and new skills. When asked about possible improvements it was mentioned several times that students would prefer to be the one controlling Holly's movements (such as using the right and left arrow on the keyboard) in contrast to the current way which is clicking on the arrow button and "watching" her walk.

The Biologist was described by the students as 'kind', 'wanting to help' and 'interested in nature', among other things. Several student pairs mentioned that the Biologist was interested in solving the mystery.

Other remarks

No participant made any explicit comments on the time constraint during gameplay. One student pair tried to move on to the next scene before they had read the dialogue on the screen. They quickly realised that this was not possible (due to the time constraint), and this resulted in one of the students concluding: "We will have to read first." (translation by the authors). Another student pair (which was brought up in subsection *The narrative*) did not bother to read the text on the screen and, instead, continued to click on the arrow multiple times until the time constraint had run out and it was possible to move to the next scene.

A more general comment about the game that three of the student pairs remarked on was the amount of dialogue-text the participants had to read. These student pairs felt that there existed an unbalance between the amount of dialogue-text and what the player could do in the game. Another student pair noted, however, that they thought the amount

of dialogue-text was adequate.

5 Discussion

5.1 Task performance

In relation to the learning activities, the students' performance varied with decreasing accuracy for the second and third task. This might be explained by the increased level of difficulty and the more complex skills that the two final tasks are demanding. Both the second and third task requires more than the ability to simply interpret a graph (which would be more of a pure mathematical skill) as the success relies on the ability to reason and value arguments. This is central to the understanding of data literacy. The fact that some participants struggled with this further strengthens the notion that students need support in exploring and developing this skill of analysing and value arguments based on underlying data. Moreover, this skill has usually not been included in the curriculum for sixth graders according to the teacher interview. In the interviews with the student pairs, it became evident that during the third learning task (targeting the relation between sample and population) the students generally did not read all of the text or alternatives. Instead they chose their answer after reading only the first or the first and second alternatives. This could suggest that students are not used to these more "soft" questions where the essential aspects are in the details or in the specific phrasing when more than one alternative can be true, i.e., the students must compare two or more options in order to determine what holds to be the most true.

5.2 Feedback

The evaluation showed that students often avoided or neglected the feedback given in the learning tasks. This aligns with previous findings in the literature (e.g., Tärning et al., 2020). The aspiration was that Holly would mitigate this problem through the use of visual signalling since this has been shown, by Tärning et al. (2020) and Silvervarg et al. (2021) to have an effect on whether students read CCF or not. This seems, however, not to have been the case. One possible reason for this could be attributed to the students not feeling responsible for Holly, which would be a desirable characteristic and something that a teachable agent in the classical sense often elicits. Not feeling this responsibility might be due to the fact that it is never stated that Holly will be tested on any of her knowledge. However, it might just as likely depend on the students not having formed a strong enough social bond with Holly in the limited prototype that was evaluated. These two possible explanations do not contradict each other, rather, both can be factors at play here. Strengthening the relationship between the player and Holly should be addressed in future developments of the educational game. The characteristics the students associated with Holly were overall positive. During the interviews, the students describe Holly as helpful while at the same time be-

ing someone that needed help from the player. That the students viewed Holly this way could indicate that they perceive her as a peer, where they can play to each other's strengths and together solve the mystery. Moreover, several students expressed a preference for a future improvement to include the ability to control Holly's movements on the screen (e.g., using the right- and left arrow keys for making her walk to the right or left, respectively). This could possibly result in strengthening of the player's sense of control in the game. It could, however, lead to some confusion as the player is not to identify themselves with Holly's character. The idea of Holly is, again, to make the students regard her as a peer. Before incorporating such a function, prior consideration and additional research must be done.

It is worth noting that feedback neglect is a recurrent problem within education research. Literature on feedback has only begun to scratch the surface when studying ways that could engage students in more productive feedback behaviour. As a result, more research is needed to look at different strategies that can be used to influence feedback neglect. The implementation of a time constraint did not seem to have an impact on the students processing CCF. Furthermore, the evaluation revealed inconsistent findings regarding the time constraint whether it can compel students to read the text that was on screen (e.g., the dialogue-text in between learning tasks). For one pair it appeared to encourage the students to read, while it did not seem to have this effect on another student pair. The results of this study does not allow for any final conclusions. Investigating if a time constraint can be used as a form of nudging strategy would be interesting to explore in the future, and could possibly be done within the framework of this educational game.

5.3 The mystery

Another feature that was questioned during the user evaluation was the amount of text in the prototype, both related to the dialogue with Holly and the learning tasks. It appeared that the students did understand the mystery that is to be solved in the game, which suggests that enough background and clues were provided, and picked up by the participants. The students also expressed an appreciation for including a mystery in the game. This seems to be consistent with Laine and Lindberg (2020) and Laschke and Hassenzahl (2011), who conclude that storytelling can get the player more involved by working as a motivational factor and contributing to immersing the player in the game world.

In some cases the mystery was not entirely clear to the students, this can mainly be linked to them not reading all the text and therefore certain information went unattended. Balancing the amount of text without loosing the players motivation to read is a dilemma. Given the result from the user evaluation, it appears that in some cases the text needs to be shortened or edited so as not to overwhelm the player. A different approach is to have an introductory text in the beginning of the game that explains what the overall goal is. In this way, the dialogue between, for example, Holly and

the Biologist could be cut shorter. Our initial decision for the game was, however, to introduce the mystery slowly as a way to build up the storyline. The reason for this was, as previously mentioned, to not overwhelm the players with too much information at once as well as to make the game feel more exciting and intriguing. The amount of text is a constant trade-off between providing enough information and making sure that the text is read by the player. An additional aspect in which Ecoville as a learning resource differs from games that are played purely for entertainment is in its relation to education. The point of the game and the foundation which it is built upon is to make students learn something through playing it as (a possible) part of the teaching, this implies that more effort might be required from students. It should also be noted that the beginning of a game is often more text-heavy than the subsequent parts — especially if the game is framed by a narrative or backstory. Hopefully, the impression of, too much text, will be fine tuned as the prototype continues to be developed and the game expands in length.

Another possible alternative to cutting down the dialogue would be to add audio, with different voices for the characters that will read back the dialogue-text. In the teacher interview, audio was emphasised as a necessary feature to make the educational game more inclusive (e.g., for students with reading difficulties). Despite audio being commonly used as a supportive tool, it is not necessarily beneficial for learning. Previous literature on this topic has reported contradictory findings (e.g., Diao & Sweller, 2007; Moreno & Mayer, 2002). Findings from Knoop-van Campen, Segers, and Verhoeven (2020) suggest that redundant audio might have a negative effect on learning in the long run. The participants in this study were university students with dyslexia and their typically-reading peers. What Knoop-van Campen et al. (2020) found was that adding audio as reading support had a negative effect on the quality of knowledge retention across both groups. Another study by the same authors examined multimedia learning in 11-year-old children (with and without dyslexia). The results in Knoop-van Campen, Segers, and Verhoeven (2018) indicated that the use of audio support could be beneficial for children with dyslexia since it leads to more efficient study time (and can, thus, optimise learning). This does, however, not necessarily apply to children with typical-reading abilities. Even though the retention of knowledge remained similar across different conditions, the children with dyslexia spent more time when pictorial information was presented with only written text. This was in comparison to the two other conditions that either included audio-only or combined audio with written text (Knoop-van Campen et al., 2018). Since combining different modalities (e.g., combining audio with written text and pictures) has revealed inconsistent results, some caution should be taken before implementing audio into the educational game. This is, however, an avenue that could be explored further in the future.

5.4 Embedded learning

Another aspect that was highlighted during the teacher interview was the importance of embedding the educational game into the curriculum, and aligning it with other related learning activities in the classroom. To embed educational games is one of the key findings in Kapp (2012) meta-analysis that reviews research on how games can become better educational tools. Before playing, it could be beneficial if the teacher gives an introduction to the educational game and some basics of data literacy, as it is helpful if the students already possess some understanding of the concepts that will be covered in the game. It is, for example, easier to pick up on details and nuances when one already has prior knowledge on the subject in question. Thus, this is more a matter of strategy for improved learning rather than a reflection of a possible weakness in the learning resource in itself. Examples of basic concepts in relation to data literacy and our prototype could, for instance, consist in knowing what the x-axis and y-axis on a graph is as well as having some knowledge on how to read different kinds of graphs (e.g., bar- and pie charts). Possessing these abilities and having been introduced to the game beforehand, opens up for the students to focus more on what they actually should be learning from the game. Additionally, teaching the same subject by using different forms of media i.e., combining an educational game with more traditional teaching methods, could possibly lead to students getting a deeper knowledge and understanding of the learning objectives.

5.5 Summary

To summarise, the overall impression of the prototype was positive. The majority of the students understood that the mystery related to the disappearance of the groundwater in Ecoville. Some students, however, made remarks on the amount of text in the prototype. The balance between sufficient information and minimised text for the player to read is a difficult task that needs continuous work.

The second and third learning task was perceived as more difficult by the students which was also reflected in their performance. One explanation for this might be that they require more complex data literacy skills. Having the learning tasks be progressively more challenging was intended as a way to promote learning and keep the students motivated. At the same time, the learning tasks does not appear to have been too challenging. Feedback neglect is a common problem and a challenge for our learning resource as well, future work could be done related to time constraints to investigate this further. In addition, the social bond between the game player and Holly needs to be strengthened. Having a pedagogical agent such as Holly may result in students being more inclined to engage with feedback.

Lastly, from the interview with the teachers it was suggested that the game needs to be integrated with current curricula and other teaching activities for optimal use. This could help students reach the learning objectives covered in

the game.

5.6 Limitations of the study

The present study was limited in regards to the conclusions that can be drawn due to the prototype still being in an early development stage. Given the current length of the game it was not possible for any final conclusions to be drawn about the social relationship between the player and the character of Holly. This might benefit from further exploration as the idea of Holly as a social-motivational factor and peer character to the player is central to the future developments of the game.

Another limitation is that the students' behaviour when evaluating the prototype was only recorded through the use of concurrent think-aloud protocols and the retrospective interviews. It might be beneficial to combine these with an additional quantitative method, such as eye-tracking or using a software that gathers data on where and when the participants click on the screen. This was, however, outside the scope for the current evaluation but might be interesting to investigate further in the future. It could, for example, be beneficial if one wants to examine how many of the students engaged with the feedback in the learning tasks or if the implemented time constraint has any effect.

6 Future work

We have identified some ideas that would be worth exploring in this project. Adaptive difficulty is one such example as it would, hopefully, further improve the learning outcomes. A challenge for all pedagogical tools is the individual differences between students, which sometimes can be extremely varying. Making Ecoville adaptive would be desirable in order to make sure that all students have the same opportunity to learn at a level that suits them. Not only the learning tasks could be made adaptive. Following findings from Kirkegaard (2016) and Tärning et al. (2018), it would also be interesting to look into how to adapt peer agent Holly in relation to students' level of self-efficacy. However, making Ecoville adaptive would require a more complex implementation in addition to being time-consuming. At this point in the project, adaptivity has not been explored in implementation, and it is unclear to what degree it can be implemented using GDevelop.

Another future step would be to continue developing the prototype to further align with the design guidelines in section 2.4, among others Fitts' law and the usage of signifiers, and based on the results from the user evaluation. Revisions to the text could be made as well as to look further into strategies that could mitigate feedback neglect. Even if no conclusions could be drawn about how students perceived Holly, there were some indications that the students did not feel any responsibility towards the character. This was despite the students acknowledging that Holly wanted their help and depended on them to help her with the learning tasks. Feeling responsible for Holly would be desirable since this

could result in Holly being able to elicit the desired protégé effect and the ego-protective buffer effect.

One possible solution that could be explored in future developments of the prototype is the additional reinforcement of Holly's role. Specifically, when introducing Holly, a reinforcement (and reminders) of her desire to learn can be included in the dialogue. The idea to reinforce a pedagogical agent's role as a way to invoke a sought after behaviour is based on findings from Silvervarg and Måansson (2018). Furthermore, one way to strengthen the relationship between the game player and Holly could be to investigate the inclusion of socially oriented off-task topics. Gulz et al. (2011) showed that this could solidify the social bond between a player and a pedagogical agent (e.g., a teachable agent). This could, in turn, further increase the sought after protégé effect and the ego-protective buffer effect. Including off-task topics in the dialogue also improves learning and the player's experience of the game (Gulz et al., 2011). For example, using humour in storytelling has been shown to benefit learning (Laine & Lindberg, 2020).

In the interview, the teachers requested having a function in the educational game which allowed them to get information on how many tries a student has made in a learning task before solving it. This would be interesting to explore in the future. Such a function could be beneficial in order to track the student's progress in the game. Furthermore, it could give the teachers information and clues as to what concepts need to be revised or attended to more in the classroom. It can, thus, be used to embed the educational game into the curriculum as well as a way to ensure that learning occurs.

Finally, in the initial description of the project, it was suggested that a possible use of *The Mystery of Ecoville* would be as a research tool for mapping and exploring how students can be supported in their progress toward data literacy. This is also a possible future application that could be explored in more detail.

References

- Almeida, C., Kalinowski, M., & Feijó, B. (2021). A systematic mapping of negative effects of gamification in education/learning systems. In *2021 47th euromicro conference on software engineering and advanced applications (seaa)* (p. 17-24). (<https://doi.org/10.1109/SEAA53835.2021.00011>)
- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, 71, 550-585. (<https://doi.org/10.1016/j.chb.2016.08.049>)
- Bhargava, R., Kadouaki, R., Bhargava, E., Castro, G., & D'Ignazio, C. (2016). Data murals: Using the arts to build data literacy. *The Journal of Community Informatics*, 12(3), 197-216. (<https://doi.org/10.15353/joci.v12i3.3285>)
- Black, P., & Wiliam, D. (1998). Assessment and classroom

- learning. *Assessment in Education: principles, policy & practice*, 5(1), 7–74. (<https://doi.org/10.1080/0969595980050102>)
- Branch, J. L. (2000). Investigating the information-seeking processes of adolescents: The value of using think alouds and think afters. *Library & information science research*, 22(4), 371–392. ([https://doi.org/10.1016/S0740-8188\(00\)00051-7](https://doi.org/10.1016/S0740-8188(00)00051-7))
- Budiu, R. (2022). *Fitts's law and its applications in ux*. <https://www.nngroup.com/articles/fitts-law/>.
- Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *J Sci Educ Technol*, 18, 334–352. (<https://doi.org/10.1007/s10956-009-9180-4>)
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gameness: Defining "gamification". In *Proceedings of the 15th international academic mindtrek conference: Envisioning future media environments* (p. 9–15). New York, NY, USA: Association for Computing Machinery. (<https://doi.org/10.1145/2181037.2181040>)
- Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and instruction*, 17(1), 78–88. (<https://doi.org/10.1016/j.learninstruc.2006.11.007>)
- D'Ignazio, C., & Bhargava, E. (2016). Databasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, 12(3), 83-107. (<https://doi.org/10.15353/joci.v12i3.3285>)
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological review*, 87(3), 215. (<https://doi.org/10.1037/0033-295X.87.3.215>)
- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5(3), 178–186. (https://doi.org/10.1207/s15327884mca0503_3)
- Gulz, A., Haake, M., & Silvervarg, A. (2011). Extending a teachable agent with a social conversation module – effects on student experiences and learning. In *Artificial intelligence in education* (pp. 106–114). Berlin, Heidelberg: Springer Berlin Heidelberg. (https://doi.org/10.1007/978-3-642-21869-9_16)
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? – a literature review of empirical studies on gamification. In *2014 47th hawaii international conference on system sciences* (p. 3025-3034). (<https://doi.org/10.1109/HICSS.2014.377>)
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81–112. (<https://doi.org/10.3102/003465430298487>)
- Huotari, K., & Hamari, J. (2012). Defining gamification: A service marketing perspective. In *Proceeding of the 16th international academic mindtrek conference* (p. 17–22). New York, NY, USA: Association for Computing Machinery. (<https://doi.org/10.1145/2393132.2393137>)
- Inverarity, C., Tarrant, D. D., Forrest, E., & Greenwood, P. (2022). Towards benchmarking data literacy. In *Companion proceedings of the web conference 2022* (p. 408–416). New York, NY, USA: Association for Computing Machinery. (<https://doi.org/10.1145/3487553.3524695>)
- Kapp, K. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Wiley. Retrieved from <https://books.google.se/books?id=M2Rb9ZtFxccC>
- Kirkegaard, C. (2016). *Adding challenge to a teachable agent in a virtual learning environment* (Doctoral dissertation). Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-126940>
- Knoop-van Campen, C. A., Segers, E., & Verhoeven, L. (2018). The modality and redundancy effects in multimedia learning in children with dyslexia. *Dyslexia*, 24(2), 140–155. (<https://doi.org/10.1002/dys.1585>)
- Knoop-van Campen, C. A., Segers, E., & Verhoeven, L. (2020). Effects of audio support on multimedia learning processes and outcomes in students with dyslexia. *Computers & Education*, 150, 103858. (<https://doi.org/10.1016/j.compedu.2020.103858>)
- Laine, T. H., & Lindberg, R. S. N. (2020). Designing engaging games for education: A systematic literature review on game motivators and design principles. *IEEE Transactions on Learning Technologies*, 13(4), 804-821. (<https://doi.org/10.1109/TLT.2020.3018503>)
- Laschke, M., & Hassenzahl, M. (2011, 01). Mayor or patron? the difference between a badge and a meaningful story. *gamification-research.org*.
- Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of educational psychology*, 94(3), 598. (<https://doi.org/10.1037/0022-0663.94.3.598>)
- Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. *Computers in Human Behavior*, 71, 450-454. (<https://doi.org/10.1016/j.chb.2016.11.062>)
- Norman, D. A. (2013). *The design of everyday things*. London, England: MIT Press.
- Oceans of Data Institute. (2015). *Building global interest in data literacy: A dialogue*. http://oceanoftodata.org/sites/oceanoftodata.org/files/ODI%20Data%20Literacy%20Report_0.pdf.
- OECD. (2019). *Future of education and skills 2030: Conceptual learning framework*. https://www.oecd.org/education/2030-project/teaching-and-learning-learning-learning-compass-2030/OECD_Learning_Compass_2030_concept_note.pdf.
- Oliveira, W., Hamari, J., Shi, L., Toda, A. M., Rodrigues, L.,

- Palomino, P. T., & Isotani, S. (2022). Tailored gamification in education: A literature review and future agenda. *Education and Information Technologies*, 1–34. (<https://doi.org/10.1007/s10639-022-11122-4>)
- Pangrazio, L., & Sefton-Green, J. (2020). The social utility of ‘data literacy’. *Learning, Media and Technology*, 45(2), 208–220. (<https://doi.org/10.1080/17439884.2020.1707223>)
- Preece, J., Rogers, Y., & Sharp, H. (2015). *Interaction design, beyond human computer interaction*. NJ: John Wiley and Sons Ltd.
- Reinhart, A., Evans, C., Luby, A., Orellana, J., Meyer, M., Wieczorek, J., ... Nugent, R. (2022). Think-aloud interviews: A tool for exploring student statistical reasoning. *Journal of Statistics and Data Science Education*, 1–14. (<https://doi.org/10.1080/26939169.2022.2063209>)
- Rideout, V. (2016). Measuring time spent with media: the common sense census of media use by us 8- to 18-year-olds. *Journal of Children and Media*, 10(1), 138-144. (<https://doi.org/10.1080/17482798.2016.1129808>)
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report. Retrieved from <https://dalspace.library.dal.ca/handle/10222/64578>
- Ryan, & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67. (<https://doi.org/10.1006/ceps.1999.1020>)
- Schaaf, R., & Mohan, N. (2014). Making school a game worth playing: Digital games in the classroom. , 53. (<https://dx.doi.org/10.4135/9781483378534>)
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153–189. (<https://doi.org/10.3102/0034654307313795>)
- Silvervarg, A., & Måansson, K. (2018). How do you introduce an agent? the effect of introduction type on how a teachable agent is experienced by students. In *Proceedings of the 18th international conference on intelligent virtual agents* (p. 29–34). New York, NY, USA: Association for Computing Machinery. (<https://doi.org/10.1145/3267851.3267861>)
- Silvervarg, A., Wolf, R., Blair, K. P., Haake, M., & Gulz, A. (2021). How teachable agents influence students’ responses to critical constructive feedback. *Journal of Research on Technology in Education*, 53(1), 67-88. (<https://doi.org/10.1080/15391523.2020.1784812>)
- Sjödén, B., Tärning, B., Pareto, L., & Gulz, A. (2011). Transferring teaching to testing – an unexplored aspect of teachable agents. In G. Biswas, S. Bull (Eds.), 337-344. (https://doi.org/10.1007/978-3-642-21869-9_44)
- Skolverket. (2011, revised 2019). *Läroplan för grundskolan*, förskoleklassen och fritidshemmet. <https://www.skolverket.se/getFile?file=9718>.
- Skolverket. (2022). *Läroplan för grundskolan, förskoleklassen och fritidshemmet*. <https://www.skolverket.se/getFile?file=9718>.
- Soegaard, M. (2020). *Hick’s law: Making the choice easier for users*. <https://www.interaction-design.org/literature/article/hick-s-law-making-the-choice-easier-for-users>.
- Tärning, B., Lee, Y. J., Andersson, R., Måansson, K., Gulz, A., & Haake, M. (2020). Assessing the black box of feedback neglect in a digital educational game for elementary school. *Journal of the Learning Sciences*, 29(4-5), 511–549. (<https://doi.org/10.1080/10508406.2020.1770092>)
- Tärning, B., Silvervarg, A., & Gulz, H. M. (2018). Instructing a teachable agent with low or high self-efficacy – does similarity attract? *International Journal of Artificial Intelligence in Education volume*, 89–121. (<https://doi.org/10.1007/s40593-018-0167-2>)
- Van Audenhove, L., Van den Broeck, W., & Mariën, I. (2020). Data literacy and education: Introduction and the challenges for our field. *Journal of Media Literacy Education*, 12(3), 1-5. (<https://doi.org/10.23860/JMLE-2020-12-3-1>)
- Williams, S., Deahl, E., Rubel, L., & Lim, V. (2014). City digits: Local lotto: Developing youth data literacy by investigating the lottery. *Journal of Digital Media Literacy*.
- Wirtz, B. (2022). *The power of experience: The wonders of video game immersion*. <https://www.gamedesigning.org/learn/game-immersion/>.
- Wolff, A., Gooch, D., Cavero Montaner, J., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12, 9-26. (<https://doi.org/10.15353/joci.v12i3.3275>)
- Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children’s inquiries from complex data. *International Journal of Human-Computer Studies*, 129, 41-54. (<https://doi.org/10.1016/j.ijhcs.2019.03.006>)
- Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review*, 30, 1-23. (<https://doi.org/10.1016/j.edurev.2020.100326>)

Appendix A

Evaluation of GDevelop

During the project the implementation of the hi-fi prototypes has been done in the Integrated Development Environment (IDE) GDevelop 5. The objective of this IDE is to simplify game development with drag and drop elements and the possibility to create code with the help of a graphic user interface, rather than writing the code from scratch.

On top of the main project of developing The Mystery of Ecoville, one other part of the project has been evaluating GDevelop as a platform to implement the learning resource in this project as well as learning resources in general. The evaluation has been ongoing throughout the process and has led to some key discoveries and thoughts. This appendix aims to lay out the positive and negative aspects that have been identified and to discuss the viability of GDevelop in this environment.

GDevelop uses a simple interface to add objects to a game window, called scene, an example of this can be seen in Figure 10 below.

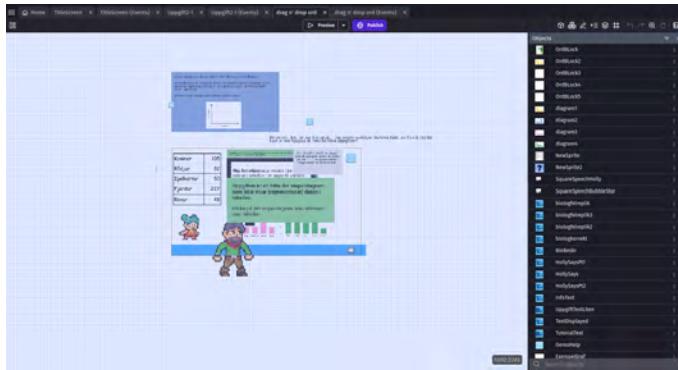


Figure 10. Example a scene from the developers perspective.

These objects can then be given certain attributes and behaviours, such as “platformer character”, “platform”, and “pathfinding”, to name a few examples. When the objects exist the developer can add define various “events” tied to these objects. An event can be the interaction between hardware and software, I.E. having pressed an arrow key can move a given object move in a certain direction. These events can also define the interaction between objects. Examples of this is if two objects collide they can light up or change animation. Creating and modifying these events and interactions is done by defining conditions and actions, see Figure 11.

These interactions follow simple rules of programming and can easily be understood with some programming knowledge.

Positive aspects of GDevelop

Initially there are quite a few positive aspects of GDevelop that could interest a developer creating a game. Alongside

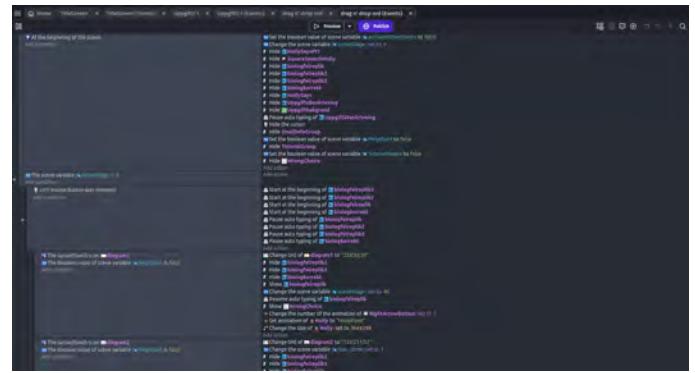


Figure 11. Example of the event page for a scene.

the beginner friendly nature of GDevelop there is also a large amount of tutorials to be found online, on top of traditional documentation. The main aspects found that are in GDevelop’s favour are listed below.

Asset library

The graphical objects or sprites used in game scenes can either be drawn in GDevelop with piskel (a pixel-art application that is integrated into GDevelop), imported from an external source or imported from the asset store. The store has an array of different types of sprites that can be used under differing licenses such as a Creative Commons CC0 public domain license, Attribution 4.0 International (CC BY 4.0) (which requires attribution), as well as assets to which a license can be bought.

Expansive documentation

To further help beginners and more experienced developers alike, there are very many tutorials and guides available. These range from simple “get started” videos to advanced guides on how to create animations for your sprites.

Overview of events

The handling of events can be quite intricate when writing code for interactive games. In GDevelop all events are accessible under one tab, making the handling of events more simple and easier to track. This tab also provides an overview of what conditions need to be met for certain interactions to take place.

Negative aspects of GDevelop

Revision control

If multiple developers are going to work on the same project individually there needs to be some sort of version control of the code to handle conflicts when sharing new and updated code between each other so as not to erase

progress. During development of the prototypes GIT was used for this purpose and the result was subpar. During the project issues arose with the way that GDevelop stores information. All objects and their interactions, variables and behaviours are stored in a JSON-object. This means that changes can't be tracked in a traditional way. This led to confusion when attempting to work simultaneously on the project as a result of not gaining understanding of the underlying code. No attempt was made to understand how to avoid these issues, as it would have taken up too much time and been counterproductive to the simplicity that GDevelop was supposed to provide. In the end GIT was used as a way to have version control and the possibility to share the project between multiple developers, but the work was not carried out simultaneously in the latter stages of the project to avoid issues regarding loss of updates.

Don't repeat yourself (DRY)

To save time during development and even during maintenance, it's necessary to follow the principle of "don't repeat yourself" (DRY). DRY is a method of work where the developer repeats as little code as possible. This is to avoid two main problems when updating or revising code. Firstly, if code is repeated, and the repeated code needs revision, then each instance of the code must be updated, rather than updating a single function or method. Secondly it can lead to an increased risk of error considering there are more opportunities for mistakes, and handling these mistakes is also more of a hassle as mentioned earlier. To combat this, the algorithm should be turned into a function that can be called upon with a single line of code, where needed. If the algorithm now needs to be changed this only needs to be done in one location, saving time and increasing readability. This can offhandedly be likened to referring to a section in a report instead of writing the section's content repeatedly.

Game art

This might not necessarily be an issue with GDevelop itself as there are a couple of ways of acquiring assets to use as sprites. Even though there are many different sprites available the assortment is limited. This means that if the game is gonna have a desired theme or atmosphere, pixel art is going to need to be drawn from scratch. It is a bit unfair to criticise GDevelop for this seeing as regardless of what development tool is used new art and models will be necessary.

Appendix B

Material sent to teachers before the teacher interview

Instruktioner:

Vi håller på att utveckla en digital lärresurs med målet att stärka elevernas förmåga för dataläskunnighet. Just nu finns en lo-fi prototyp (den ni får ta del av) som håller på att implementeras för att bli ett fungerande lärospel. Vi skulle därför vilja ha er hjälp med att utvärdera några aspekter av prototypen, framförallt i relation till språket som används, både i dialogen men också i den återkoppling som ges, och dataläskunnighets-uppgifterna.

Dataläskunnighet handlar inte bara källkritik - även om det ingår i begreppet. Det handlar om att kunna analysera och argumentera med (empirisk) data, till exempel om det finns goda skäl för att tro på det som någon säger eller påstår i förhållande till informationen (datan) som finns. I det ingår att kunna skilja på åsikt och fakta. Men dataläskunnighet handlar även om att lära sig tolka och dra slutsatser från det som står i ett diagram. Att vara dataläskunnig är en viktig förmåga för ett demokratiskt samhälle och knyter därför an till flera ämnen i skolan som matematik, samhällskunskap och naturvetenskap. I förlängningen är dataläskunnighet också relaterat till att ha förståelse för vetenskaplig metod.

Här nedan finns några frågor som vi är intresserade av att höra era åsikter om. Ni får väldigt gärna ha dessa i bakhuvudet medan ni utvärderar lo-fi prototypen. Vi vill naturligtvis även höra era åsikter om sådant som inte direkt frågas om. All feedback är bra feedback, så var gärna så kritiska som möjligt. Tack för din medverkan!

Frågor om prototypen:

I relation till dialogen/språket i spelet:

- Är den generella tonen ok?
- Används några ord som kan vara svåra eller främmande för åldersgruppen?
- Finns det några formuleringar som är problematiska eller otydliga?
- Finns det några ställen där det är alldelvis för mycket text på samma gång?

I relation till uppgifterna:

- Är svårighetsgraden ok? Är något uppenbart för lätt eller för svårt?
- Förstår man vad uppgifterna går ut på? Vad som ska göras och hur det relaterar till berättelsen (det större mysteriet) i spelet?

I relation till feedback på uppgifterna (Kritisk Konstruktiv Feedback):

- Ligger feedbacken på en bra nivå när det kommer till att förstå vad som blev fel och hur man kan korrigera det?

- Men även när det kommer till hur feedback-texten presenteras rent språkmässigt?

Appendix C

Questions about learning resources and data literacy for the teacher interview

Generella frågor om lärresurser:

- Hur ser ni på användandet av digitala lärresurser i skolan?
- Är det något ni själva använder er av?
- Vad tycker ni utmärker en bra digital lärresurs?

Generella frågor om dataläskunnighet:

- Hur hanteras dataläskunnighet i undervisningen idag?
- Finns det något som ni uppfattar inom dataläskunnighetsområdet som era elever generellt har extra svårt för?
- Finns det något annat som ni ser att det finns behov av när det kommer till dataläskunnighet?
- I relation till dataläskunnighet och källkritik, har ni som lärare märkt någon skillnad på hur elever resonerar i takt med att digitalisering/sociala medier/ökat informationsflöde?

Släpp en robot fri

Daniel Carlström Schad

da8454sh-s@student.lu.se

20 januari 2023

Kognitionsvetenskapen studerar de processer och mekanismer som underbygger mentala fenomen i såväl biologiska som artificiella system. Ett centralt problem för förkroppsrigade kognitiva system är planering av rörelse för navigering i komplexa miljöer. Det sätt som detta har implementerats i robotiken har hitintills varit väldigt olika hur navigering implementeras i biologiska kognitiva agenter. För att möjliggöra naturliga rörelser i rummet skapades en kontrollsysteem för rörelser i rummet för en sådan robot, *Epi*. Plattformen består av en tvåhjulad bas som kontrolleras genom differentiell styrning. En Lidarsensor mäter avståndet till hinder i omgivningen. Styrningsalgoritmer har implementerats för kontroll av plattformen genom *Ikaros*, vilket möjliggör integration med befintliga styrprogram för *Epi*. Genom en modulär konstruktion kan ett stort antal olika beteenden uppnås. Modular för simultan kartläggning och lokalisering har också implementerats, vilket möjliggör interaktion i ett allocentriskt koordinatsystem. Rörelseplanering implementeras genom en biologisk inspirerad modell av hippocampus funktion. Utifrån sannolikheten för ett hinder vid varje given position kan den totala konnektiviteten beräknas och detta kan användas för att skapa rörelseplaner genom iterativ beräkning av den totala förväntade konnektiviteten, *SR*. Detta möjliggör ett mellanting av modellbaserad och modelfri navigering.

1 Introduktion

Kognitionsvetenskapen innefattar flera olika discipliner som alla bidrar till vår förståelse av kognition, vare sig det rör sig om mänskor, djur eller robotar. De processer och mekanismer som underbygger mentala funktioner som till exempel rationalitet, minne, planeringsförmåga och kommunikation kan åstadkommas på många olika sätt och jämförelser mellan de olika systemen ger ytterligare information om de kognitiva systemens natur. Traditionellt sett har dessa funktioner studerats genom empiriska studier på mänskor och till viss del på djur. Utifrån dessa empiriska data har man sedan konstruerat modeller som på artificiell väg han lösa samma problem. Artificiell intelligens bidrar till kognitionsvetenskapen genom att (1) skapa teoretiska och matematiska modeller för hur kognition fungerar i mänskor och andra djur, samt (2) skapa nya modeller som ger oss insikt i förutsättningarna som krävs för en viss kognitiv funktion. Dessa två uppgifter bidrar till förståelsen av vad kognition innehåller då de säger oss något om den underliggande mekanismen som ger upphov till mentala fenomen, och *vad* det är som bidrar till dessa. Alla kognitiva fenomen är dock inte begränsade till en rent mental domän, så att det är oändligt

att ta hänsyn till den kognitiva agentes omgivning för att fullt ut förstå dess inre processer och beteenden. Det är just detta som avses med *embodied cognition* som under senare år har blivit en central del i det kognitionsvetenskapliga forskningsområdet. Den huvudsakliga tesen är att kognition inte endast sker i ett abstrakt rum, utan att den alltid är kopplad till omvärlden, ibland kanske till och med på ett sådant sätt att denna koppling är vad som utgör kognitionen (Shapiro & Spaulding, 2021). Man måste alltså ta hänsyn inte bara till den interna processen i sin modell av kognition, utan även den externa miljön. Nästan alla våra kognitiva processer vid ett givet tillfälle handlar om vår interaktion med omvärlden mer eller mindre direkt. Det är rimligt att de artificiella modellerna skall återspegla detta. Ett sätt att modellera förkroppsrigade och externalistiska system är att skapa artificiella system som interagerar med den externa världen: robotar.

Robotar som kognitiva agenter måste kunna interagera med omgivningen. Detta är ett komplext problem, och det har visat sig vara svårt att skapa agenter som kan ta hänsyn till den stora mängd information som omger agenter i världen (Pylyshyn, 1987). De mest framgångsrika formerna av robotar som idag tillverkas är ofta av en sådan typ att de löser detta problem genom att minska mängden information som roboten behöver förhålla sig till. Detta kan inkludera robotar som inte behöver representera någon information alls, som en gräsklippare som inte behöver representera hinder utan istället vänder när den träffar dem. Denna typ av kognitiv agent kan replikera många beteenden som man tidigare ansåg krävde en intern modell av omvärlden (Brooks, 1991). På den andra sidan av robotiken finner vi robotar som i allt större utsträckning skall efterlikna mänskor, till tanke och utseende. Dessa har en mycket stor mängd information att förhålla sig till och det är därför attraktivt att skapa representationella modeller, kanske för att vi upplever detta som sättet vi berarbetar informationen själva. Nackdelen med många av de symbolbaserade och representationella systemen har dock stora begränsningar vad gäller både mängden och typen av den information den kan bearbeta. Generellt sett behövs det ett välordnat problem med tydliga begränsningar för att ett system som SOAR eller LT (Newell, Shaw & Simon, 1958; Rosenbloom, Laird, Newell & McCarl, 1991; Simon, 1974) ska kunna nå en tillräcklig kompetens. I navigeringsfallet yttrar detta sig som en oönskvärd söktid för för att planera rutten genom en uttömande global sökning av alla möjliga vägar mellan den nuvarande positionen och målet. Alternativet till detta är att låta roboten vandra runt och genom operant betingning lära den vilka vägar som leder till snabba

belöningar. Dessa borde då vara de närmaste. Detta var länge en modell för hur navigering fungerade i djur, och är grunden för modern modellfri *reinforcement learning* (Dayan, 1993; Eshel m. fl., 2015; Gardner, Schoenbaum & Gershman, 2018; Schultz, 1998). Det finns dock vissa problem med en sådan modell. I en klassisk studie visade Tolman att råttor lär sig strukturen i en labyrinth då de utforskar miljön utan belöning (Tolman, 1948) då samma råttor sedan skulle navigera till en belöning kunde de göra detta nästan ommedelbart, vilket indikerade att de endast behövde lära sig *var* belöningen fanns inte *hur* de skulle ta sig dit. Tolman föreslog att detta var då de hade en intern *cognitive map* som representerade miljön. Detta har vidareutvecklats till *grid-cell* systemet som har lagts fram som en hypotes för navigering i vissa djur (Behrens m. fl., 2018; Bush, Barry, Manson & Burgess, 2015; Moser, Kropff & Moser, 2008; Moser, Moser & McNaughton, 2017). Vad denna kognitiva karta möjliggör är latent inlärning i avsaknad av stimuli. När dessa stimuli sedan blir kända kan de kombineras med den förexisterande kartan för målmedvetet beteende.

Roboten Epi är en humanoidrobot skapad vid avdelningen för kognitionsvetenskap vid Lunds Universitet inriktat på studiet av kognitiva funktioner hos robotar (Johansson, Tjøstheim & Balkenius, 2020). Epi finns i en mindre modell med bara huvud och axlar, samt en större som även har en torso och armar. Roboten har skapats för att ha en barnlikt utseende, som korresponderar med robotens kompetensnivå (Johansson m. fl., 2020). Epi kan genomföra stillasittande interaktionsstudier i vilka man till exempel kan undersöka ögonfärger som ger uttryck för emotioner eller personlighetstyper (Oka m. fl., 2019), eller mentala tillstånd (Lindberg m. fl., 2017). Epi har dock ingen förmåga att röra sig i rummet och kan därför inte interagera på ett situerat sätt genom positionering och rörelse, som är centrala delar av interaktion (Mondada, 2014). Vidare är navigering och målmedveten rörelse i rummet centrala delar av en kognitiv agents förkroppsligade interaktion med omgivningen. För att möjliggöra denna typ av interaktion på ett plausibelt sätt kommer denna studie att implementera ett system för navigering i Epi baserat på en neurologiskt inspirerad modell av navigering i däggdjur.

2 Teori

I denna sektion kommer de teoretiska ramverk som underbygger implementeringen att presenteras. Först presenteras den fysiska modeller för kontroll av motorerna. Därefter kommer det teoretiska ramverk som används för lokalisering och kartläggning att presenteras. Navigering kräver att Epi kan detektera de hinder som finns i omvälden och kräver dessutom en representation av den egna positionen i förhållande till dessa hinder. Detta kräver att Epi kan lokalisera sig själv i en komplex miljö samtidigt som roboten även måste konstruera den kognitiva karta som denna lokalisering sker i i realtid. För att lösa detta implementerades en version av simultan lokalisering och kartläggning. Slutligen presenteras navigeringsmodellen, som bygger på studier av hippocampus funktion i navigering hos djur. Detta leder till en matematisk modell som kan användas för navigering i Epi.

Motorkontroll

För att implementera motoriserad rörelse i rummet hos Epi används en cirkulär tvåhjulig plattform med 35,5 cm hjulbas och med två stödhjul som kan rotera fritt, se figur 1. Detta möjliggör rotation genom differentiell hastighet på de två drivande hjulen. Genom användning av en pulsgivare kan man kontrollera motorerna med mycket stor precision och bestämma en specifik position för hjulets läge. Detta medför att man kan bestämma exakt hur mycket höger respektive vänster hjul skall förflytta sig vid ett givet tillfälle. Motorn styrs med med strömtillförsel, så att en större ström leder till större hastighet på motorn. Genom att modulera strömtillförseln så att den korresponderar med pulsgivarens läsning kan vi skapa ett slutet kontrollsysteem som justerar strömmen till motorn för att behålla en given position oavsett vilka krafter som påverkar hjulets läge. Detta implementeras i en separat kontrollmodul (ODrive 3.6 56V).

Genom kontroll av styrspänningen på en fälteffekttransistor kan man kontrollera strömmen som passerar genom motorn. Den implementerar följande modell för styrning för att överlämna en önskad position till en hastighetsvektor (ODrive, 2021):

$$V_{target} = P_{gain}(P_{target} - P_{feedback}) + V_{feedforward}.$$

Utifrån V_{target} kan man räkna ut vilken ström som efftersöks och således vilken spänning som skall appliceras över fälteffekttransistorn:

$$\begin{aligned} V_{error} &= V_{target} - V_{feedback} \\ V_{\Sigma,t} &= V_{\Sigma,t-1} + V_{error} \cdot V_{\Sigma,gain} \\ I_{target} &= V_{error} \cdot V_{gain} + V_{\Sigma,t} + I_{feedforward} \\ I_{error} &= I_{target} - I_{feedback} \\ I_{\Sigma,t} &= I_{\Sigma,t-1} + I_{error} \cdot I_{\Sigma,gain} \\ U_{target} &= I_{error} \cdot I_{gain} + I_{\Sigma,t} + U_{feedforward} \end{aligned}$$

Detta ger oss ett slutet system som kan uppdateras med hög hastighet (8kHz i ODrive, 2021) och ger god kontrollförmåga.

Kinematiska modeller

Styrning av kontrollkortet sker genom att skicka kommandon till varje axel med det antal rotationer sedan start som det slutna kontrollsystelet som ekvationssystemet ovan skall upprätthålla. För att kunna styra robotten på ett tillfredsställande sätt måste vi konvertera detta till position i rummets euklidiska dimensioner. För att kunna navigera i en komplex miljö måste vi också kunna modellera robotens position som funktion av hjulens rörelse. Genom att integrera hastigheten över tid kan vi modellera position över tid. Vid varje givet tillfälle kan vi anta att robotten rör sig på en cirkulär bana. Om vi känner till hastigheten för de två hjulen kan vi beräkna rörelsehastigheten v och vinkelhastigheten ω som:

$$\begin{aligned} v &= \omega \cdot r \\ v_r &= \omega(r + \frac{l}{2}) \\ v_l &= \omega(r - \frac{l}{2}) \end{aligned}$$



Figur 1. Epi med hjulbasen för differentiell styrning.

där l är avståndet mellan hjulen. Lösning av ekvationsystemet ger:

$$\begin{aligned}\omega &= \frac{v_r - v_l}{l} \\ v &= \frac{v_r + v_l}{2} \\ r &= \frac{l}{2} \frac{v_r + v_l}{v_r - v_l}\end{aligned}$$

Om vi nu beräknar rörelsen kring cirkelbågen definierad av radie r med hastighet v och vinkelhastighet ω över en tid Δt erhåller vi (Thrun, Burgard & Fox, 2005):

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} + \begin{bmatrix} -r \sin \theta + r \sin(\theta + \omega \Delta t) \\ r \cos \theta - r \cos(\theta + \omega \Delta t) \\ \omega \Delta t \end{bmatrix} \quad (2.1)$$

Detta ger oss en modell över hur robotens position förändras i relation till dess utgångstillstånd. Genom att välja Δt av lämplig storlek kan vi approximera den sanna

positionen av roboten över tid. Notera dock att vi kan utgå ifrån odometri då integralen av hastighet över tid är avståndet som pulsgivaren anger.

Simultan lokalisering och kartläggning

För att kunna navigera i en komplex miljö bör Epi kunna kartlägga sin miljö i relation till sin egen position. För att göra detta måste Epi (1) kunna identifiera sin egen position, något som kräver en omgivning att lokalisera sig i; samt (2) identifiera avstånden till objekt i omgivningen, något som kräver en korrekt uppskattning av den egna positionen. Algoritmer som löser detta problem går under namnet SLAM [Simultaneous Localization and Mapping]. Jag har valt att implementera en version av FastSLAM 2.0 som delar upp omvälden i ett rutnät och för varje ruta i nätet uppskattar sannolikheten att denna ruta har ett hinder. Denna algoritm implementerar en version av ett partikelfilter (Hahnel, Burgard, Fox & Thrun, 2003; Thrun m. fl., 2005) som estimerar den posteriota distributionen av sensorvärdet genom att simulera ett antal olika alternativa utvecklingar av modellen över tid och selektion av de mest sannolika. Osäkerheten i positionsestimering som erhålls vid beräkning med ekvation 2.1 ovan kan användas för att uppskatta vilka positioner som är sannolika genom att integrera en modell av sensorernas brus i beräkningen av framtida position. Genom att matcha lidar-data i förhållande till en estimerad karta kan vi uppskatta hur sannolik en viss position för roboten är. Detta löser felackumuleringen som odometriskt modeller är känsliga för. Vi kan tolka detta problem probabilistiskt och ser då att den huvudsakliga uppgiften är att estimerat:

$$bel(x_{1:t}, m_t) = p(x_{1:t}, m_t | z_{1:t}, u_{1:t}) \quad (2.2)$$

Alltså den posteriota distributionen av möjliga positioner och kartor givet sensoriska inputs och rörelser. Eftersom karten och positionen är konditionellt oberoende kan vi faktorisera denna sannolikhet till (Grisetti, Stachniss & Burgard, 2005):

$$p(x_{1:t}, m_t | z_{1:t}, u_{1:t}) = p(m_t | x_{1:t}, z_{1:t}) p(x_{1:t} | z_{1:t}, u_{1:t}) \quad (2.3)$$

Låt oss skapa en ny tillståndsvariabel $x_{1:t} = (x_{1:t}, m_t)$ som representerar det totala tillståndet. För att estimerat $p(x_{1:t} | z_{1:t}, u_{1:t})$ kan vi applicera Bayes sats och anta en Markov-process med konditionellt oberoende över tid (Thrun m. fl., 2005):

$$\begin{aligned}p(x_{1:t} | z_{1:t}, u_{1:t}) &= \eta p(z_t | x_{1:t}, z_{1:t-1}, u_{1:t}) p(x_{1:t} | z_{1:t-1}, u_{1:t}) \\ &= \eta p(z_t | x_t) p(x_{1:t} | z_{1:t-1}, u_{1:t}) \\ &= \eta p(z_t | x_t) p(x_t | x_{1:t-1}, z_{1:t-1}, u_{1:t}) p(x_{1:t-1} | z_{1:t-1}, u_{1:t}) \\ &= \eta p(z_t | x_t) p(x_t | x_{t-1}, u_t) p(x_{1:t-1} | z_{1:t-1}, u_{1:t-1}).\end{aligned} \quad (2.4)$$

Notera att $p(x_{1:t-1} | z_{1:t-1}, u_{1:t-1})$ kan faktoriseras enligt samma metod. Vi kan nu estimerat den posteriota distributionen genom rekursion (Thrun m. fl., 2005). Svag induktion ger:

Bevis. För partiklar $x_{t-1} \in \mathbf{Y}_{t-1}$ fördelade enligt $bel(x_{1:t-1}) = p(x_{1:t-1} | z_{1:t-1}, u_{1:t-1})$ kan vi skapa partiklar $x_t \in \mathbf{Y}_t$ sådana att

$$x_t = p(x_t | x_{t-1}, u_t) p(x_{1:t-1} | z_{1:t-1}, u_{1:t-1})$$

med en sannolikhet proportionerlig mot vikten

$$\begin{aligned} w_t &= \frac{\eta \text{bel}(x_{1:t})}{p(x_t|x_{t-1}, u_t) \text{bel}(x_{1:t-1})} \\ &= \frac{\eta p(z_t|x_t) p(x_t|x_{t-1}, u_t) p(x_{1:t-1}|z_{1:t-1}, u_{1:t-1})}{p(x_t|x_{t-1}, u_t) p(x_{1:t-1}|z_{1:t-1}, u_{1:t-1})} \\ &= \eta p(z_t|x_t). \end{aligned}$$

Detta ger en posterior distribution enligt distributionen av vikter w_t :

$$\text{bel}(x_{1:t}) = \eta w_t p(x_t|x_{t-1}, u_t) \text{bel}(x_{1:t-1})$$

Utgångsfallet $\text{bel}(x_1)$ är trivialt att visa för partiklar dragna ur den priora distributionen. \square

Vi kan approximera denna metod numeriskt genom att begränsa antalet partiklar $y \in \mathbf{Y}$ och approximera sannolikhetsdistributionerna i beviset ovan genom att generera värden för x_t från en modell av transitionsdistributionen $p(x_t|x_{t-1}, u_t)$. Utifrån dessa kan vi sedan beräkna sannolikheten för de lidar-mätningar som gjorts, $p(z_t|x_t)$. Genom att generera nya partiklar slumpmässigt dragna efter dessa vikter kommer distributionen av alla partiklar approximera den posteriota uppskattningen av positionen. Vi kan nu beräkna den posteriota distributionen av kartor $p(m_t|x_{1:t}, z_{1:t})$. Markov-antagandet låter oss förenkla detta till $p(m_t|x_t, z_t)$. Då antalet partiklar närmar sig ∞ kommer distributionen att sammanfalla med den sanna distributionen. Med ett lämpligt val av antal partiklar kan vi approximera detta med ett finkt antal partiklar.

En modifierad version av FastSLAM-algoritmen återges i algoritm 1 (Grisetti m. fl., 2005; Thrun m. fl., 2005). Genom att integrera information från den posteriota distributionen av sannolika positioner och matcha detta mot de observerade värdena från andra sensorer kan vi bättre uppskatta den sanna posteriota distributionen. Vi kan på detta sätt approximera $\text{bel}(x_t) = p(x_t|x_{t-1}, u_t) \text{bel}(x_{t-1})$. Vi kan nu beräkna vikten w_t genom att beräkna sannolikheten för lidar-mätningarna givet kartan och den nya positionen x_t genom att jämföra z_t med vilka mätningar z_t^* som hade erhållits utifrån kartan m_{t-1} och x_t . Vi kan sedan uppdatera kartan utifrån positionen och de nya mätningarna. Vi kan minska tidskomplexiteten av algoritmen genom att endast genomföra omsampling av distributionen då vikterna indikerar att partiklar har avvikit från den sanna distributionen (Grisetti m. fl., 2005).

Navigation

Det finns många olika sätt att implementera navigering. Ett klassiskt alternativ är att låta robotten genomföra en uttömmande sökning av de olika vägar som är möjliga för att nå en destination. En sådan lösning är ofördelaktig då har hög tidskomplexitet. Det finns ett stort antal algoritmer som utgår från en *global modell* av omvärlden och som sedan genomför en uttömmande sökning av denna modell. Ett annat alternativ är att implementera en subsumption-arkitektur i stil med de som presenterats av Brooks (1991). Här försöker man undvika behovet av en modell av omvärlden och förlitar sig på lokala inlärda handlingsmönster som kan ge upphov till intelligenta beslut. I det första fallet erhåller vi ett system med god planeringsförmåga och anpassningsbarhet till föränderliga

miljöer på bekostnad av större beräkningskostnad *online*. I det modellfria paradigmet undviker vi online-beräkning på bekostnad av flexibilitet och inlärningsförmåga.

I detta arbete implementeras en alternativ modell baserad på *successor representationen* [SR] av omvärldens Markov-modell. Denna modell låter oss modellera den totala konnektiviteten mellan noder, vilket möjliggör navigation genom lokal sökning. Agenten behöver endast lokalisera den nästkommande nod i transitionsmatrisen som representerar omvärlden som har högst total konnektivitet med målet. Genom rekursion kan leder detta till en slags *gradient ascent* som leder till målet. Inlärning av dessa konnektiviteter kan åstadkommas analytiskt om man känner till systemets totala dynamik, eller genom *Reinforcement Learning* [RL] för en agent som inte har tillgång till denna dynamik. Successor representation förslogs först av Dayan (1993) som ett sätt att förbättra *temporal difference learning*, en modellfri metod för RL. Temporal differences learning uppskattar prediktionsfelet δ_t för belöning vid ett givet tillstånd s_t och använder detta för att uppdatera värdefunktionen

$$\delta_t = r_t + \gamma \hat{V}(s_{t+1}) - \hat{V}(s_t) \quad (2.5)$$

Eftersom detta är en iterativ metod för att uppskatta långtgående värdeförändringar kommer den att vara inflexibel då den kräver iteration över alla noder för att lära sig en förändring i värdefunktionen Gershman (2018). Temporal difference learning har kopplats till dopaminergera system i hjärnan (Momennejad m. fl., 2017; Russsek, Momennejad, Botvinick, Gershman & Daw, 2017; Starkweather, Babayan, Uchida & Gershman, 2017). Man kan till exempel se att dopaminresponsen i ventrala tegmentum hos Schimpanser som lär sig signaler för en belöning korrelerar starkt med skillnaden mellan förväntad och faktisk belöning än med belöningen i sig (Schultz, 1998). GABA:ergisk inhibition av dopamin i VTA bidrar till en minskad belöning då denna är förväntad (Eshel m. fl., 2015). Temporal difference learning är alltså en neurologiskt plausibel modell för inlärning av belöning i däggdjur.

För att lösa värdeberoendet av temporal differences learning kan vi faktorisera de två delarna av värdefunktionen som vi är intresserade av. Vi kan modellera framtida belöning som produkten av sannolikheten att vi berfinner oss i ett visst tillstånd och belöningen som vi erhåller i det tillståndet. Då finner vi:

$$V = M \vec{v} \quad (2.6)$$

där V är värdet vid varje givet tillstånd, M är successorrepresentationen och \vec{v} är en vektor med varje tillstånds omedelbara värde. Denna faktorisering löser värdeberoendet av temporal differences learning och gör att agenten snabbt kan anpassa sig till en föränderlig värdefunktion. Vi erhåller då enelffunktion (Gershman, 2018)

$$\delta_t(s') = \mathbb{I}[s_t = s'] + \gamma \hat{M}(s_{t+1}, s') - \hat{M}(s_t, s') \quad (2.7)$$

som för att agenten iterativt kan lära sig konnektiviteten mellan olika tillstånd. Det har föreslagits att de dopaminergera reward prediction error-mekanismer som observerats i VTA istället relaterar till denna uppdateringsmekanism (Gardner m. fl., 2018). Dopaminergera neuron i VTA visar förändrad aktivitet vid *oväntade fenomen*, även om dessa

inte är direkt relaterade till värdefunktionen. Man föreslår därför att den huvudsakliga funktionen är att uppdatera prediktiva system i stort. Detta skulle också möjliggöra latent inlärning, allstår inlärning utan belöning. Det visades resan av Tolman att råttor som fick utforska en miljö fritt sedan snabbare lärde sig att navigera till belöning då denna introducerades till den tidigare neutrala miljön än de som fick belöning från början (Tolman, 1948). Detta visar att de lärde sig om miljöns struktur före det att belöningen introducerades. Man kan alltså se successor represenstationen som en partiell modell av omgivningen som reducerar komplexiteten i en fullständig modell genom att representera den totala konnektiviteten i systemet snarare än alla möjliga vägar, samtidigt som man erhåller vissa av de fullständiga modellernas fördelar genom faktorisering av värde och framtida position. Om vi känner till dynamiken i systemet kan vi beräkna SR analytiskt:

$$M = \sum_{k=1}^{\infty} \gamma^k T^k = \frac{1}{I - \gamma T} \quad (2.8)$$

där I är identitetsmatrisen och T är transitionsmatrisen som representerar systemets dynamik.

I studier på människor har man konstaterat att de omedelbart kan uppdatera sina handlingsmönster då de erhåller ny information om värdet av olika tillstånd, men att de inte i lika stor utsträckning kan uppdatera sin uppskattning av vilka framtids tillstånd som är sannolika (Gershman, 2018; Momennejad m. fl., 2017). Detta talar för att M uppdateras genom iteration över möjliga tillstånd och sedan kan kombineras med värde för att ge optimala handlingsalternativ. Vidare evidens för att denna representation är neurologiskt rimlig är att många av de olika aktiveringsmönster som erhålls genom SR motsvarar de som observeras i hippocampus och entorhinala grid-cell system (Behrens m. fl., 2018; Gershman, 2018; Whittington, McCaffery, Bakermans & Behrens, 2022). En möjlig förklaring till detta är att grid-cell systemet representerar egenvektorerna för successor representationen (Whittington m. fl., 2022). Detta skulle ge en naturlig förklaring till de olika cell-typer som har observerats i mediala entorhinalcortex [mEC], och i hippocampus. Det har även föreslagits neurologiskt plausibla mekanismer för hur en representation av detta slag kan skapas och uppdateras över tid (Vértes & Sahani, 2019). Här föreslår man att nya observationer som avviker från förväntningar uppdaterar den lokala dynamiken i representationen. Men denna information måste spridas över den totala dynamiken i systemet, något som åstakoms genom ekvation 2.7 ovan. För att göra detta kan vi simulera ett stort antal slumpmässig transitioner i tillståndsrymden och uppdatera matrisen i linje med prediktionsfelet. Detta är precis vad som har föreslagits under sömn, och speglar även hippocampal återuppspelning (Momennejad, 2020; Vértes & Sahani, 2019).

Med detta i åtanke kan i nu trivialt lösa navigationsproblemet. Givet att vi har ett mål och inga andra mer eller mindre föredelaktiga positioner kan vi bestämma vår värdevektor i ett grid-system med diskreta positioner $s \in G$ och ett mål i s_Π till

$$\vec{v}_i = \begin{cases} 1 & \text{om } i = \Pi \\ 0 & \text{annars} \end{cases} \quad (2.9)$$

Detta är inget annat än att ta den kolumn i M som motsvarar s_t . Vi erhåller då

$$V_\Pi = M \cdot \vec{v}. \quad (2.10)$$

SR förutbestämmer omvärldens struktur åt oss, så att vi inte behöver göra komplicerade beräkningar i realtid. Notera att vi kan generera mer komplicerade navigeringsbeende genom att till exempel tillskriva alla positionen vi vill undvika ett negativt värde. Detta kommer leda till att agenten tar den närmsta vägen som undviker en vid plats. Eftersom den totala konnektiviteten representeras rekursivt i lokala tillstånd kan vi finna den väg med störst konnektivitet till målpositionen genom att utvärdera de tillstånd $s' \in S' = \{s : T(s_t, s) > 0\}$ som är näbara i ett tidssteg från det nuvarande tillståndet s_t

$$\Pi(s_{t+1}) = \underset{s'}{\operatorname{argmax}} V_\Pi(s'). \quad (2.11)$$

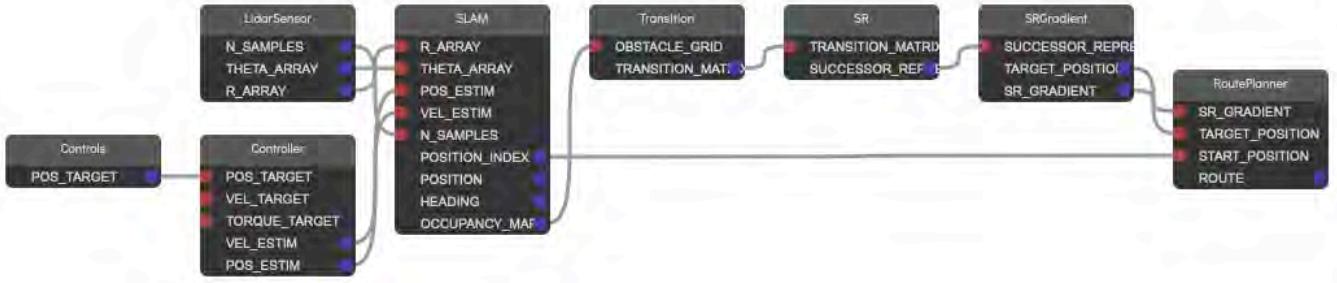
Denna navigeringsalgoritm ger en inkrementell tidskomplexitet på $O(nS')$ i fallet med endast ett tillstånd av värde. Denna metod fungerar i såväl Euklidiska rum som andra topologier med eller utan metrik. Alla rum som går att representera med grafteori kan omvandlas till denna typ av navigeringsproblem (Baram, Muller, Whittington & Behrens, 2018). Baram m. fl. (2018) fann även att denna metod kan användas för navigering i komplexa tvådimensionella kartor med och utan maskhål som förvränger metriken. Den lokala lösningen överensstämde med den globalt optimala, förutom då flera maskhål placerades bredvid varandra, då detta ökade konnektiviteten. I en sådan situation tog den simulerade agenten en suboptimal väg. Detta kan jämföras med hur människor navigerar i städer, där konnektiviteten inte är identisk över alla transitioner. Vissa möjliga färdvägar har högre konnektivitet än andra. Till exempel har tunnelbanesystem gjort att det ibland är lättare att ta sig till en relativt avlägsen plats än till en fysiskt närmre plats som inte är ansluten till tunnelbanesystemet.

3 Implemtentering

Arbetet med att implementera styrning av roboten var till att börja med starkt fokuserat på konstruktion av den erforderliga hårdvaran för elektrisk styrning av motorerna. Då det inte fanns någon programvara i C++ för kontroll av ODrive-kontrollmodulen behövde detta implementeras. För att kunna integrera plattformen med resten av roboten implementerades styrsystemet i Ikaros (Balkenius, Johansson & Tjøstheim, 2020). Ikaros är ett modulärt system för neurologisk simulering och kontroll av robotar skrivet i C++. Tidigare styrsystem för Epi är implementerade i Ikaros, och detta möjliggör smidig kontroll av roboten med ett web-interface.

Hårdvara

För att styra plattformen användes två motorer av typen ODrive D6374. Dessa motorer roterar i 150 rpm/V och verkar vid upp till 48V och 70 A. De körs vid 12 V och har konfigurerats till 5 A maximal ström för att minska de krafter som verkar på robotens ledar vid acceleration. Till dessa motorer har kopplats pulsgivare av typen CUI AMT102-V, som ger 8192 pulser per rotation. Mo-



Figur 2. Moduler i Ikaros för navigering.

torer och pulsgivare kopplas till en kontrollenhet ODrive v3.6 som använder programvaruversion 0.5.5. Uppdatering av programvaruversion kräver en uppdatering av USB-protokollets registerkoder. För att absorbera energi vid inbromsning används en effektresistor med $2.7\ \Omega$ resistans och 50 W maximal effekt. Kommunikation mellan kontrollenheten och styrdatorn sker via seriell kommunikation över USB 2.0. Kontrollkortet kan kontrolleras genom ASCI-kommandon över en seriell enhet för UNIX-baserade operativssystem.

Roboten är utrustad med hjul med en diameter på 200 mm som är placerade längs oberoende hjulaxlar på 355 mm avstånd, se figur 1. Hjulen är kopplade till den motsvarande motorn med en tandrem och har en utväxling på 15:92. Antalet varv motorn behöver snurra för att hjulet skall röra sig 1m är alltså 10,186.

Implementering i Ikaros

Kontroll i Ikaros sker genom sammankoppling av fält i olika moduler. Genom att sammankoppla moduler med olika funktionalitet kan roboten kontrolleras på ett lämpligt sätt. Ikaros kan köras i realtid med en fast uppdateringsfrekvens. Vid kontroll av Epi används en uppdateringsfrekvens på 40 Hz. Figur 2 visar de moduler som används för navigering. Modulerna är, i informationsbearbetningsordning:

- Controls:** Kontrollmodul för manuell styrning av roboten,
- Controller:** Kommunikation med motorernas kontrollkort,
- LidarSensor:** Lidar sensorn,
- SLAM:** Simultan lokalisering och kartläggning med FastSLAM,
- Transition:** beräkning av T för den estimerade Markovprocessen,
- SR:** Beräkning av M_{Π} analytiskt,
- SRGradient:** Beräkning av V_{Π} enligt ekvation 2.10 ovan samt
- RoutePlanner:** Iterativ beräkning av rutten till s_{Π} enligt ekvation 2.11 ovan.

I det följande beskrivs de olika modulernas funktion. I figur 3 visas användargränssnittet för styrning av roboten.

Kontrollsysteem

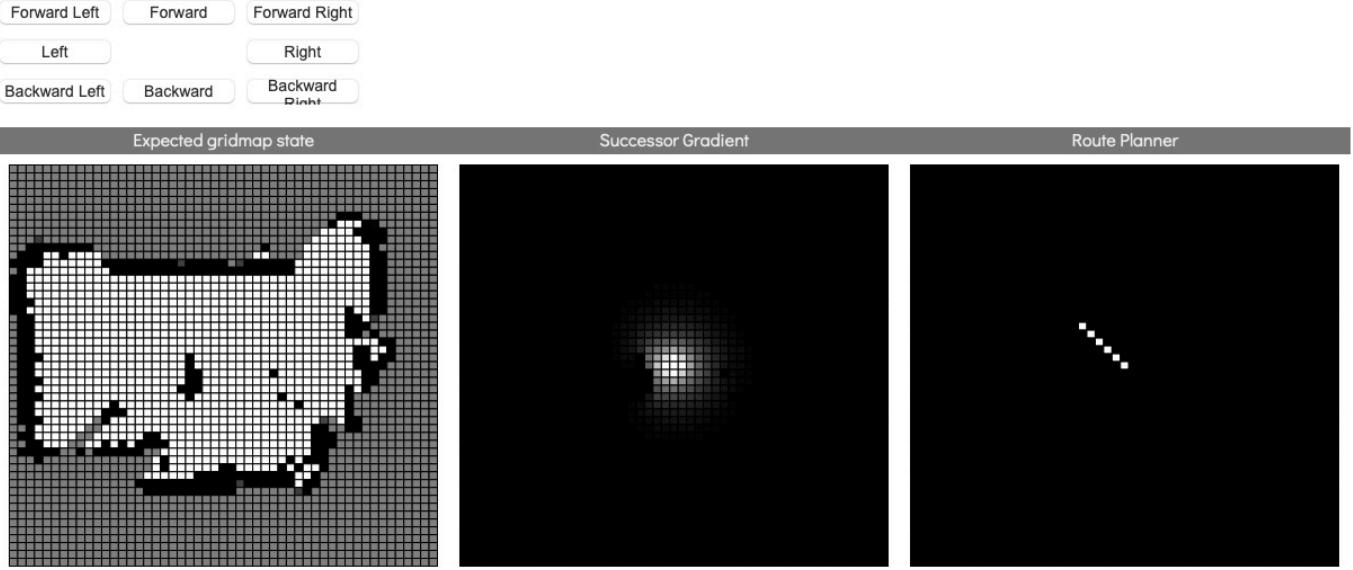
För att styra kontrollkortet existerar endast en drivrutin skriven i Python. Interaktion med kontrollkortet sker genom seriell kommunikation över USB. Kontrollenheten har konfigurerats med registrervärden avpassade för de motorer och pulsgivare som har installerats. Konfiguration av dessa parametrar sker via Python-verktyget som medföljer den officiella drivrutinen (ODrive, 2022). Dessa värden ger en god avvägning av motorprestanda och hastighet. Om roboten rör sig för snabbt finns det en risk att den uppfattas som aggressiv och detta kan också ha en negativ inverkan på många av de strukturella elementen i konstruktionen. Ararna på Epi är monterade på ett sådant sätt att det inte är lämpligt att accelerera för snabbt, då trögheten hos armens massa gör att den upplever skjukvrafter vid stora accelerationer. Konfiguration återges i appendix A.

Manuell styrning av roboten har implementerats genom ett antal knappar i ikaros web-interface som gör att man kan flytta roboten framåt och bakåt samt rotera den med- och motsols. Genom kombinationer av dessa funktioner kan man även förflytta den snett i sidled. Dessa har placerats i en separat modul.

Kontrollmodulen för motorerna accepterar ett antal parametrar, bland annat vissa konstanter för omräkningen mellan meter och radianer hjulrotation. Dessutom finns alternativ för input-filtrering för att minska initial acceleration vid nya kommandon. Modulen accepterar positions-, hastighets- och vridmomentsinput i meter, meter/s respektive Nm för de två hjulen och omvandlar dessa till kontrollkortets interna enheter. Modulen läser också antalet varv som har registrerats av pulsgivarna och returnerar den kumulativa sträcka som hjulen har roterat framåt i meter, samt den momentana hastigheten för de två hjulen.

Lidar-sensor

Navigation kräver att Epi känner sin omgivning och kan beräkna vilka rörelsemöjligheter som denna omgivning erfordrar. För detta ändamål används en lidar-sensor av modell Slamtec RPLIDAR A3. Denna sensor använder trigonometrisk kongruens för att uppskatta avståndet till ett föremål genom att sända en laserstråle mot föremålet i fråga och mäta infallspositionen för reflektionsstrålen på en digital kamerasensor. Eftersom avstånden inuti lidarsensorn är kända kan avståndet till den yttre reflektionspunkten kan uppskattas genom triangulationen med linsen och kamerasensorn. Genom att rotera lidar-sensorn i horisontalplanet kan vi mäta avståndet till föremål i



Figur 3. Användargränssnitt för kontroll av Epi. Bilden till vänster visar den förväntade kartan hos den bästa partikeln i algoritm 1. Vita områden är mycket osannolikt hinder, svarta är mycket sannolikt det och de grå har en okänd sannolikhet. Den mittersta bilden visar det förväntade värdet för varje position enligt ekvation 2.10 ovan. Den högra bilden visar den närmsta rutten enligt ekvation 2.11 ovan.

flera riktningar. Modulen i Ikaros använder sig av den SDK som tillhandahålls av tillverkaren (Slamtec, 2022). Eftersom lidar-sensorn inte roterar tillräckligt snabbt för att kunna mäta i realtid med en uppdateringsfrekvens på 40 Hz, körs lidar-sensorn i en separat tråd med en intern uppdateringsfrekvens på 7.5 Hz. Den minskade uppdateringsfrekvensen påverkar inte datakvalitén nämnvärt. Lidar-sensorn publicerar fält med vinkel och avstånd till hinder kring roboten. Lidar-sensorn har 25 m räckvidd och kan ta upp till 16000 mätvärden per rotation. Lidar-sensorn returnerar olika antal mätpunkter för varje mätning. Modulen publicerar ett fält vardera med vinklar och radier i ett polärt koordinatsystem med utgångspunkt i sensorns framåtriktning och med stigande vinkel med sols, samt antalet mätpunkter som har observerats vid den senaste mätningen. Denna monterades så att den är placerad normalt mot och 160 mm framför hjulaxeln. Det vertikala avståndet från marken är 349.98 mm.

En framtida möjlighet är att använda sig av de stereoskopiska kamerorna för att identifiera hinder kring roboten. Detta utesluter dock inte lidar-sensorn, då denna har möjlighet till heltäckande uppskattning av hindernas position, inte bara i den riktning som roboten riktar huvudet. Lidar är överlag mycket mer tillförlitlig i sin uppskattning av andra föremål i rummet, men bidrar inte med identifieringsförmåga för de objekt som den upptäcker. Vidare kan lidar-sensorn endast detektera hinder på samma höjd som sensorn.

Simultan lokalisering och kartläggning

Modulen för simultan lokalisering och kartläggning tar mätdata från lidar-sensorn och avståndsmätningarna från kontrollmodulen och använder dessa för att uppdatera positionen och kartan. Detta sker genom metoden som redovisades i algoritm 1 ovan. Mer konkret implementeras algoritmen genom att ett antal partikelobjekt skapas. Dessa objekt innehåller en karta, en position och en vikt, samt ett antal konstanter och andra parametrar.

Då varje partikel allokerar en egen karta i *heap*-minnet kan minnesförbrukningen bli relativt stor. För 4 Byte flyttal och en karta på $256 \cdot 256$ punkter med 100 partiklar kommer minnesförbrukningen för partiklarna vara 26 MB. Detta kan inte anses vara särskilt begränsande för en modern dator. Alla partiklar initieras med en startposition mitt i kartan och en initial karta konstruerad från lidar-sensorns första läsning. Därefter beräknas för varje partikel som i algoritm 1 ovan. I det nedanstående redogörs för de funktioner som där används.

Algorithm 1 FastSLAM($\mathcal{P}_{t-1}, u_t, z_t$)

Input: Array of N particles $\mathcal{P}_{t-1}^k = \langle x_{t-1}^k, w_{t-1}^k, \mathbf{m}_{t-1}^k \rangle$
Input: Control commands u_t
Input: Sensor measurements z_t
Output: Array of N particles $\mathcal{P}_t^k = \langle x_t^k, w_t^k, \mathbf{m}_t^k \rangle$

```

1:  $\bar{\mathcal{P}} = \emptyset$ 
2: for  $\langle x_{t-1}^k, w_{t-1}^k, \mathbf{m}_{t-1}^k \rangle$  in  $\mathcal{P}_{t-1}$  do
3:    $x_t^k = \text{sample\_motion\_model}(u_t, x_{t-1}^k)$ 
4:    $w_t^k = \text{measurement\_model}(z_t, x_t^k, \mathbf{m}_{t-1}^k)$ 
5:    $\mathbf{m}_t^k = \text{update\_map}(z_t, x_t^k, \mathbf{m}_{t-1}^k)$ 
6:    $\bar{\mathcal{P}}_t^k = \langle x_t^k, w_t^k, \mathbf{m}_t^k \rangle$ 
7: end for
8: if  $\sum_{\mathcal{P}_t} w_t^k < \frac{N}{2}$  then
9:   for  $k = 1$  to  $N$  do
10:    draw  $i$  with probability  $\propto x_t^i$ 
11:     $\mathcal{P}_t^k = \bar{\mathcal{P}}_t^k$ 
12: end for
13: else
14:    $\mathcal{P}_t = \bar{\mathcal{P}}_t$ 
15: end if
16: return  $\mathcal{P}_t$ 

```

För att dra slumpmässiga positioner från sannolikhetsdistributionen efter en observerad rörelse används algoritmen 2. Funktionen **sample_motion_model** beräknar

den nuvarande positionen med hjälp av odometriförändringen i ekvation 2.1 ovan givet

$$\begin{aligned}\hat{d}_l &= \mathcal{S}(d_l, \alpha d_l^2 + \beta d_r^2) \\ \hat{d}_r &= \mathcal{S}(d_r, \alpha d_l^2 + \beta d_r^2)\end{aligned}$$

där $\mathcal{S}(\mu, \sigma^2)$ är ett tal draget ur normalfordelningen med medelvärde μ och varians σ^2 . Vi kan på detta sätt approximera $bel(x_t) = p(x_t|x_{t-1}, u_t) bel(x_{t-1})$. Denna algoritmen utgår ifrån skillnaderna i erlagd sträcka i en riktning hos de två hjulen. Detta kan enkelt beräknas som $\Delta d = d_t - d_{t-1}$. Algoritmen följer en modifierad version av (Thrun m. fl., 2005).

Algorithm 2 sample_motion_model(x_{t-1}, u_t)

Input: Previous position $x_{t-1} = (x, y, \theta)^T$
Input: Movement commands $u_t = (\Delta d_l, \Delta d_r)$
Output: Current position $x_t = (x', y', \theta')^T$

```

1:  $\hat{d}_l = \mathcal{S}(d_l, \alpha d_l^2 + \beta d_r^2)$ 
2:  $\hat{d}_r = \mathcal{S}(d_r, \alpha d_l^2 + \beta d_r^2)$ 
3: if  $|\hat{d}_l - \hat{d}_r| < \varepsilon$  then
4:    $x' = \frac{\hat{d}_l + \hat{d}_r}{2} * \cos \theta$ 
5:    $y' = \frac{\hat{d}_l + \hat{d}_r}{2} * \sin \theta$ 
6: else if  $|\hat{d}_l + \hat{d}_r| < \varepsilon$  then
7:    $\Delta\theta = \frac{\hat{d}_r - \hat{d}_l}{l}$ 
8:    $\theta' = \theta + \Delta\theta$ 
9: else
10:    $r = \frac{l}{2} \frac{\hat{d}_r + \hat{d}_l}{\hat{d}_r - \hat{d}_l}$ 
11:    $\Delta\theta = \frac{\hat{d}_r - \hat{d}_l}{l}$ 
12:    $\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} + \begin{bmatrix} -r \sin \theta + r \sin(\theta + \Delta\theta) \\ r \cos \theta - r \cos(\theta + \Delta\theta) \\ \Delta\theta \end{bmatrix}$ 
13: end if
14: return  $x_t$ 

```

I algoritmen motsvarar l Epis hjulbas, α och β är empiriskt fastställda konstanter för felvariansen i kontrollkortets positionsestimering. I den faktiska beräkningen använder u_t den observerade positionen snarare än kommandon till kontrollkortet då detta kör asynkront och det inte går att garantera att hela kommandot har utförts före det att nästa stickprov dras ur rörelsediistributionen, något som gör att den observerade positionen är mer tillförlitlig.

Nästa steg i estimationsprocessen är att avgöra hur sannolik den nya positionen är givet den befintliga kartan och den omvärld som observeras av lidar-sensorn. Detta åstadkoms genom att jämföra det faktiska mätvärdet med avståndet i kartan. Givet ett uppmätt avstånd $z_i \in Z_t$ och ett predicerat avstånd z_i^* kan vi beräkna sannolikheten givet en modell av variansen i lidar-sensorns mätningar. I enlighet med Thrun m. fl. (2005) implementerades en sådan modell. Vi kan modellera felet i laser-sensorn genom att estimera fyra olika felkällor:

$$\begin{aligned}p(z_i|z_i^*) &= z_{hit} \cdot p_{hit}(z_i|z_i^*) + \\ &\quad z_{short} \cdot p_{short}(z_i|z_i^*) + \\ &\quad z_{max} \cdot p_{max}(z_i|z_i^*) + \\ &\quad z_{random} \cdot p_{random}(z_i|z_i^*)\end{aligned}\tag{3.1}$$

där $z_{hit} + z_{short} + z_{max} + z_{random} = 1$ är vikter som bestämmer sannolikheterna för de olika typer av fel som modelleras. Den första felkälla som modelleras är variansen i avläst avstånd, vilket kan modelleras som en normalfordelning $\mathcal{N}(z_i)$ med medelvärde z_i^* och standardavvikelse σ_{hit} .

$$p_{hit}(z_i|z_i^*) = \begin{cases} \frac{\mathcal{N}(z_i)}{\int_{z_{min}}^{z_{max}} \mathcal{N}(z_i) dz_i} & \text{om } z_{min} \leq z_i \leq z_{max} \\ 0 & \text{annars} \end{cases}\tag{3.2}$$

där nämnaren normaliseras distributionen nära maxima, z_{max} , eller minimala, z_{min} , avstånd.

Ett annat fel som förekommer med laser-sensorer är avstånd betydligt kortare än det sanna värdet på grund av oväntade händelser eller objekt, till exempel människor eller andra dynamiska objekt i omgivningen. Dessa är inte en del av den fasta kartan utan bör hanteras separat då deras positioner inte är permanenta Thrun m. fl. (2005). Vi kan modellera dessa enligt:

$$p_{short}(z_i|z_i^*) = \begin{cases} \frac{\lambda_{short} e^{\lambda_{short} z_i}}{1 - e^{-\lambda_{short} z_i^*}} & \text{om } z_{min} \leq z_i \leq z_i^* \\ 0 & \text{annars} \end{cases}\tag{3.3}$$

Där konstanten λ_{short} anger hur snabbt exponentialfunktionen minskar.

Ytterligare ett fel som förekommer är att lidar-sensorn ger maximala läsnignar på grund av interferens eller andra störningar. Detta kan hanteras genom att bestämma en fast sannolikhet för en maximal läsning (Thrun m. fl., 2005):

$$p_{max}(z_i|z_i^*) = \begin{cases} 1 & \text{om } z_i = z_{max} \\ 0 & \text{annars} \end{cases}\tag{3.4}$$

Sist av allt kan vi anta att det finns fel som vi inte har hanterat explicit. Vi modellerar dessa som slumpräckiga fel:

$$p_{random}(z_i|z_i^*) = \begin{cases} \frac{1}{z_{max}} & \text{om } z_{min} \leq z_i \leq z_{max} \\ 0 & \text{annars} \end{cases}\tag{3.5}$$

För att bestämma vilket det predicerade avståndet z_i^* bör vara kan vi använda ray-tracing för att estimera vilken ruta i kartan som är den första med ett hinder längs laserstrålen z_i :s bana. Implementeringen av denna algoritm återges i 3 nedan.

Algorithm 3 measurement_model(z_t, x_t, m_{t-1})

Input: Array z_t of N sensor measurements $z_i = (r_i, \theta_i)$

Input: Current estimated position x_t

Input: Previous estimated map m_{t-1}

Output: Probability of all measurements p

```

1:  $q = 1$ 
2: for  $i = 1$  to  $N$  do
3:    $z_i^* = \text{ray\_cast}(\theta_i, x_t, m_{t-1})$ 
4:    $q = q \cdot p(z_i|z_i^*)$ 
5: end for
6: return  $q$ 

```

Denna sannolikhet q är sannolikheten för att samtliga individuella laserstrålar mäter korrekt. Resultaten används som vikt i FastSLAM-algoritmen och representerar

$p(z_t|x_t)$ -termen i ekvation 2.4 ovan och möjliggör beräkning av partikelns positionssannolikhet givet kartan. Vi kan nu beräkna den uppdatera kartan.

För att uppdatera kartan måste vi först ha en modell för vilken information en sensormärning ger om en position i kartan. Denna modell återges i en modifierad algoritm från (Thrun m. fl., 2005) nedan.

Algorithm 4 update_map(z_t, x_t, m_{t-1})

Input: Array z_t of N sensor measurements $z_i = (r_i, \theta_i)$
Input: Current estimated position $x_t = (x, y, \vartheta)$
Input: Previous estimated map m_{t-1}
Output: Current estimated map m_t

- 1: **for** cell m_{t-1}^i in m_{t-1} **do**
- 2: Let x_i, y_i be the center of mass of m_{t-1}^i
- 3: $r = \sqrt{(x_i - x)^2 + (y_i - y)^2}$
- 4: $\phi = \text{atan2}(y_i - y, x_i - x) - \vartheta$
- 5: $k = \text{argmin}_j |\phi - \theta_j|$
- 6: **if** $r > \min(z_{\max}, z_t^k + \frac{\alpha}{2})$ **or** $|\phi - \theta_j| > \frac{\beta}{2}$ **then**
- 7: $m_t^i = m_{t-1}^i$
- 8: **else if** $z_t^k < z_{\max}$ **and** $|r - z_t^k| < \frac{\alpha}{2}$ **then**
- 9: $m_t^i = m_{t-1}^i + l_{\text{occ}} - l_0$
- 10: **else if** $r \leq z_t^k$ **then**
- 11: $m_t^i = m_{t-1}^i + l_{\text{free}} - l_0$
- 12: **end if**
- 13: **end for**
- 14: **return** m_t

Algoritmen identifierar den laserstråle som bäst passar med cellen i fråga. Om cellen befinner sig bortom maximala räckvidden, bakom ett hinder med tjocklek α eller utanför sensorkonen med vinkelspann β uppdateras inte cellen. Om den befinner sig inom sensorns räckvidd och innanför $\frac{\alpha}{2}$ från sensorns mätvärde antas den innehålla ett hinder. Om avståndet till cellen är mindre än mätvärdet är antages den vara fri. Vi kan nu uppdatera cellens värde med den motsvarande log-oddsen $l_{\text{free}}, l_{\text{occ}}, l_0$.

Modulen returnerar kartan med störst vikt, omvandlad till sannolikheter från log-odds, samt denna partikels position och riktning i meter/radianer samt index för den cell som roboten nu berfinner sig i.

Transition matrix

För att omvandla kartan med sannolikheter erhållen från SLAM-algoritmen till en representation av statistiska sannolikheter för konnektiviteten mellan celler kan vi anta att sannolikheten för en transition mellan s_t och målet $s' \in S'$ där S' är de grannar som s_t kan nå är

$$p_{s_t, s'} = \frac{m_t^{s'}}{\sum_{s \in S'} m_t^s}. \quad (3.6)$$

Detta ger en total sannolikhet för transitioner på 1 och en transitionssannolikhet för varje cell proportionerlig mot de andra målens förväntade sannolikhet för hinder. Moduler publicerar en matris av alla transitioner i kartan och är alltså av storlek n^2 .

Successor representation

Istället för att iterativt skapa successor representationen [SR] genom RL beräknar vi den analytiskt. I framtiden

hade det varit fördelaktigt att skapa en modell som använder RL för att estimera SR. Att beräkna SR analytiskt har en betydande tidskomplexitet. Givet en karta med storlek $n \cdot n$ är tidskomplexiteten $\approx O(n^{2.37})$ då vi måste invertera en matris med storlek n^2 . En iterativ metod har också fördelar då förändringarna i stora delar av SR-matrisen kommer vara minimala på grund av diskonteringsfaktorn γ som minskar påverkan av långa vägar mellan två noder till korta vägars fördel. Modulen publicerar matrisen med SR. För att använda denna modul för navigering behöver vi dock, såsom har påpekats i teoridelen, endast välja en kolumn i matrisen. Modulen *SRGradient* läser vilken cell i web-interfaceet som har klickats på och publicerar en matris med storleken $n \cdot n$ som representerar den totala konnektiviteten mellan varje cell och målcellen.

Navigering

Slutligen kan vi nu beräkna rutten till målet för att göra detta implementeras följande algoritm:

Algorithm 5 route_planner($V_\Pi, x_{\text{target}}, x_{\text{start}}$)

Input: Value over policy Π with target in x_{target}
Input: Start position x_{start}
Input: Target position x_{target}
Output: Route to x_{target} \mathbb{S}

- 1: $\mathbb{S} = (\{x_{\text{start}}\}, <)$
- 2: **while** $x_{\text{target}} \notin \mathbb{S}$ **do**
- 3: $\mathbb{S} = \mathbb{S} \cup \text{argmax}_{s' \in S'} V_\Pi(s')$
- 4: **end while**
- 5: **return** \mathbb{S}

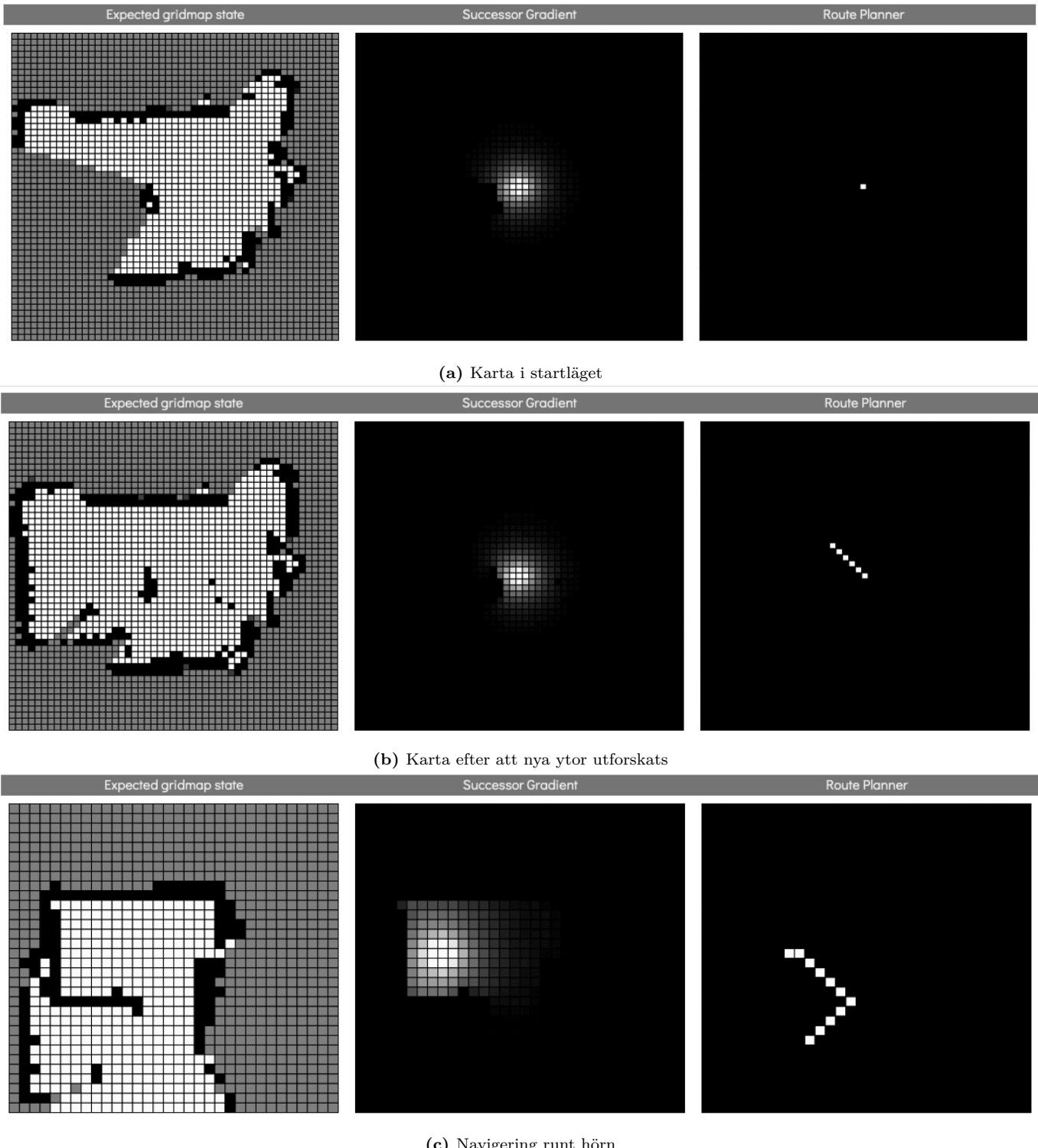
Notera att vi endast behöver utvärdera vilken av grannarna s' till den nuvarande positionen $\max(\mathbb{S})$ för att bestämma vilken väg som är bäst. Genom förberäkning av konnektiviteten, förslagsvis genom RL, kan vi snabbt hitta en väg till målet. Algoritmen kan även navigera kring hörn, se figur 5c, eller i komplexa labyrinter, se figur 4.



Figur 4. Navigering i en komplex miljö. Successor representationen kan navigera i alla graf-lika miljöer och kan därför hitta optimala vägar även i mycket komplexa miljöer.

4 Resultat

Navigeringsalgoritmen som implementeras i möjliggör planering av rörelser i komplexa miljöer som labyrinter eller inomhusmiljöer med flera rum. Detta kan åstadkommas genom förberäkning av konnektiviteten i rummet genom RL, eller analytiskt som i detta fallet. Denna förberäkning omvandlar ett globalt sökproblem till ett lokalt problem och möjliggör således navigering genom uppstigning med gradientmetoden. Genom att inkludera



Figur 5. Några exempel på navigering och kartläggning med Epi. Figur 5a visar kartläggningen före det att robotten förflyttat sig, så att ytorna bakom hindret i mitten av rummet är okända. I figur 5b har robotten kört runt hindrets hörn och kan därför kartlägga även dessa ytor. I figur 5c visas den planerade rutten för att ta sig runt ett hinder i mitten av rummet.

en SLAM-algoritm i navigeringsalgoritmen kan navigation ske i komplexa miljöer där sensorerna inte har omedelbar tillgång till hela kartan. Exempel på hur SLAM möjliggör kartläggning av mer komplexa miljöer se figur 5. För att denna typ av kartläggning inte skall leda till katastrofala fel måste felackumuleringen i odometri begränsas. För att göra detta simulerar vi flera olika partiklar med slumpmässiga fel och viktat dessa enligt hur bra de passar med kartan som den ser ut för tillfället. Detta kan dock leda till utarming av partiklar om variansen är för stor (Thrun m. fl., 2005) Denna utarmining leder om man har för få partiklar till ett ackumulerande fel. För att minska detta används adaptiv omsampling (Grisetti m. fl., 2005) för att minska risken att bra partiklar selekteras bort slumpmässigt. Grisetti m. fl. (2005) föreslår dock att man även kan använda lidar-mätningar och kartan för att genomföra lokaliseringen. Detta skulle leda till ännu med samlad posterior sannolikhet för positionen och skulle göra FastSLAM mindre benägen att degenerera, vilket leder till att färre partiklar kan användas. Implementering av detta skulle vara en betydande fördel.

Efterhand som Epi utforskar nya delar av kartan uppdateras dessa med information om var hinder finns. Specifikation av kartans storlek och upplösning avgör hur stor del av lidar-datan som tas med i kartläggningen. I figur 5c ser man att den nedersta väggen inte får plats inom kartan och därför inte tas med i beräkningen. Det hade varit fördelaktigt om man använde en egocentrisk referensram, som skulle möjliggöra större rörelser, på bekostnad av att hela kartan inte kan hållas i minnet samtidigt. För att möjliggöra större navigeringsområde kan man med fördel förse Epi med batterier och en mobil dator. Detta skulle göra att Epi inte är begränsad av kablage i sin navigering.

För att ytterligare utveckla navigeringsmodeller i Epi bör beräkningen av SR implementeras på ett iterativt sätt. Detta innebär att en sömnfas bör implementeras för Epi, då en stor del av den iterativa bearbetningen sker under sömn i de djurmodeller som studeras (Vértes & Sahani, 2019). Det hade också varit fördelaktigt om Epi kunde använda sig av de andra sensorer som roboten är utrustad med för att förbättra detektionen av hinder. Lidar sensorn är monterad 350 mm från marken och kommer därför endast att detektera de hinder som befinner sig på denna höjd. För att lösa detta hade man kunnat montera sensorn med en viss vinkel relativt hjulaxeln och omvandla den tvådimensionella kartan till en tredimensionell. Då skulle mätningar vid olika vinklar relativt vertikalplanet ske under sensorns rotationsbana. Detta är dock en relativt komplex lösning och skulle försämra antalet mätningar i varje voxel och på så sätt försämra datakvalitén. Ett annat alternativ är att lägga till ytterligare en lidar-sensor som mäter i vertikalplanet. Detta lägger dock till ytterligare kostnad och komplexitet och ökar beräkningskomplexiteten till $O(n^3)$, vilket är ofördelaktigt. Det mest rimliga är att inkludera information från de stereoskopiska kameror som redan finns monterade och som skulle möjliggöra inferens dels om de objekt som är hinder samt möjlighet att identifiera hinder som inte detekteras av lidar-sensorn.

Referenser

- Balkenius, C., Johansson, B. & Tjøstheim, T. A. (2020). Ikaros: A framework for controlling robots with system-level brain models. *International Journal of Advanced Robotic Systems*, 17(3), 1729881420925002. doi: 10.1177/1729881420925002
- Baram, A. B., Muller, T. H., Whittington, J. C. & Behrens, T. E. (2018). Intuitive planning: global navigation through cognitive maps based on grid-like codes. *bioRxiv*, 421461. doi: 10.1101/421461
- Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L. & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, 100(2), 490–509. Hämtad från <http://doi.org/10.1016/j.neuron.2018.10.002> doi: 10.1016/j.neuron.2018.10.002
- Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139–159.
- Bush, D., Barry, C., Manson, D. & Burgess, N. (2015). Using Grid Cells for Navigation. *Neuron*, 87(3), 507–520. doi: 10.1016/j.neuron.2015.07.006
- Dayan, P. (1993). Improving generalization for temporal difference learning: The successor representation. *Neural computation*, 5(4), 613–624.
- Eshel, N., Bukwich, M., Rao, V., Hemmeler, V., Tian, J. & Uchida, N. (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*, 525(7568), 243–246.
- Gardner, M. P., Schoenbaum, G. & Gershman, S. J. (2018). Rethinking dopamine as generalized prediction error. *Proceedings of the Royal Society B*, 285(1891), 20181645.
- Gershman, S. J. (2018). The successor representation: its computational logic and neural substrates. *Journal of Neuroscience*, 38(33), 7193–7200.
- Grisetti, G., Stachniss, C. & Burgard, W. (2005). Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling. I *Proceedings of the 2005 ieee international conference on robotics and automation* (s. 2432–2437).
- Hahnel, D., Burgard, W., Fox, D. & Thrun, S. (2003). An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. I *Proceedings 2003 ieee/rsj international conference on intelligent robots and systems (iros 2003)(cat. no. 03ch37453)* (vol. 1, s. 206–211).
- Johansson, B., Tjøstheim, T. A. & Balkenius, C. (2020). Epi: An open humanoid platform for developmental robotics. *International Journal of Advanced Robotic Systems*, 17(2), 1729881420911498. doi: 10.1177/1729881420911498
- Lindberg, M., Sandberg, H., Liljenberg, M., Eriksson, M., Johansson, B. & Balkenius, C. (2017, 8). The expression of mental states in a humanoid robot. I *International conference on intelligent virtual agents* (s. 247–250). Springer. doi: 10.1007/978-3-319-67401-8_32
- Momennejad, I. (2020). Learning structures: predictive representations, replay, and generalization. *Current*

- Opinion in Behavioral Sciences*, 32, 155–166.
- Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D. & Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nature Human Behaviour*, 1(9), 680–692. doi: 10.1038/s41562-017-0180-8
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137–156. doi: 10.1016/j.pragma.2014.04.004
- Moser, E. I., Kropff, E. & Moser, M.-B. (2008). Place Cells, Grid Cells, and the Brain's Spatial Representation System. *Neuroscience*, 31(1), 69–89. doi: 10.1146/annurev.neuro.31.061307.090723
- Moser, E. I., Moser, M.-B. & McNaughton, B. L. (2017). Spatial representation in the hippocampal formation: a history. *Nature Neuroscience*, 20(11), 1448–1464. doi: 10.1038/nn.4653
- Newell, A., Shaw, J. C. & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, 65(3), 151–166. doi: 10.1037/h0048495
- ODrive. (2021). *Control Structure and Tuning*. Hämtad 2022-10-30, från <https://docs.odriverobotics.com/v/0.5.5/control.html#position-control-loop>
- ODrive. (2022). *ODrive*. Hämtad från <https://github.com/odriverobotics/ODrive>
- Oka, N., Koda, T., Obaid, M., Nakanishi, H., Mubin, O., Tanaka, K., ... Johansson, B. (2019). Communicating Emotional State and Personality with Eye-color and Light Intensity. *Proceedings of the 7th International Conference on Human-Agent Interaction*, 214–216. doi: 10.1145/3349537.3352769
- Pylyshyn, Z. W. (1987). The robot's dilemma: The frame problem in artificial intelligence.
- Rosenbloom, P. S., Laird, J. E., Newell, A. & McCarl, R. (1991). A preliminary analysis of the soar architecture as a basis for general intelligence. *Artificial Intelligence*, 47(1-3), 289–325.
- Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J. & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology*, 13(9), e1005768.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, 80(1), 1–27.
- Shapiro, L. & Spaulding, S. (2021). *Embodied Cognition* (Winter 2021 utgåvan). Metaphysics Research Lab, Stanford University. Hämtad från <https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/>
- Simon, H. A. (1974). The structure of ill structured problems. *Artificial Intelligence*, 4(3-4), 181–201. doi: 10.1016/0004-3702(73)90011-8
- Slamtec. (2022). *Slamtec RPLIDAR Public SDK for C++*. Hämtad från https://github.com/slamtec/rplidar_sdk
- Starkweather, C. K., Babayan, B. M., Uchida, N. & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature neuroscience*, 20(4), 581–589.
- Thrun, S., Burgard, W. & Fox, D. (2005). *Probabilistic Robotics*. Cambridge, MA: MIT Press.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, 55(4), 189.
- Vértes, E. & Sahani, M. (2019). A neurally plausible model learns successor representations in partially observable environments. *Advances in Neural Information Processing Systems*, 32.
- Whittington, J. C. R., McCaffery, D., Bakermans, J. J. W. & Behrens, T. E. J. (2022). How to build a cognitive map. *Nature Neuroscience*, 25(10), 1257–1272. doi: 10.1038/s41593-022-01153-y

A Konfigurationskod

All källkod finns tillgänglig på GitHub (<https://github.com/dod12/ikaros>).

```
import time

import odrive
import odrive.enums as enums
from odrive import fibre

device = odrive.find_any()
axes = [device.axis0, device.axis1]

if device is not None:
    print("Found ODrive", device.serial_number)
else:
    raise IOError("Could not find ODrive")

control_mode = "none" # Control mode for filtering target position. Must be one of [None,
→ "bandwidth", "trap"].
tick_rate = 2

# Setup global settings
device.config.enable_brake_resistor = True
device.config.brake_resistance = 2.7
device.config.dc_max_negative_current = -0.001

# Setup per-axis settings
for axis in axes:
    axis.motor.config.current_lim = 5
    axis.controller.config.vel_limit = 10
    axis.motor.config.calibration_current = 10
    axis.motor.config.pole_pairs = 7
    axis.motor.config.torque_constant = 8.27 / 150
    axis.motor.config.motor_type = enums.MOTOR_TYPE_HIGH_CURRENT
    axis.encoder.config.cpr = 8192
    # axis.encoder.config.direction = 1 # Invert control direction for left encoder position.
    if control_mode == "trap":
        axis.trap_traj.config.vel_limit = 5
        axis.trap_traj.config.accel_limit = 1
        axis.trap_traj.config.decel_limit = 1
        axis.controller.config.inertia = 0
        axis.controller.config.input_mode = enums.INPUT_MODE_TRAP_TRAJ
    elif control_mode == "bandwidth":
        axis.controller.config.input_filter_bandwidth = tick_rate / 2 # Set to half the position
        → update frequency.
        axis.controller.config.input_mode = enums.INPUT_MODE_POS_FILTER
    else:
        axis.controller.config.input_mode = enums.INPUT_MODE_PASSTHROUGH
    axis.config.startup_motor_calibration = True
    axis.config.startup_encoder_index_search = True
    axis.config.startup_encoder_offset_calibration = True
    axis.config.startup_closed_loop_control = True

for axis in axes:
    axis.requested_state = enums.AXIS_STATE_IDLE

try:
    device.save_configuration()
except fibre.libfibre.ObjectLostError:
    print("Rebooting...")
    pass
```


Ansiktsigenkänning och åldersestimering på en Epi-robot

William Do

wi6126do-s@student.lu.se

Stella Johannsen

st4670jo-s@student.lu.se

Johannes Rosenfrost

jo7118he-s@student.lu.se

17 januari 2023

This report covers two project under the courses FoU Courses at Lund University. The task was to give Epi, a humanoid robot, the skills of recognizing some form of object. Facial recognition and age estimation was chosen because it would be interesting to see how a robot could interact with humans, and if they could estimate ages. With machine learning, a basic model containing 20 pictures of people was created that could detect human faces. Later, this model expanded to include over 600 pictures for better precision, and a class to detect other robots of the same type. For age estimation, we conducted an experiment with two different models and humans where we compared their age estimation capabilities. During the implementation of the interaction between humans and robots, Epi had different interactions depending on whether it detected a human or a robot, and when Epi detected age categories. In the future, it is important that the model should be implemented with data augmentation for even better precision. This would solve the brightness and color problem during identification, and should also help Epi's distinguish humans from robots better. The results from the age estimation experiment showed that humans were in general better than both ML-models in predicting ages from pictures. In the future, it can be a good idea to train the age estimation on the dataset, so that the results might improve, or find a dataset with higher resolution. This report will describe why we made the choice with facial recognition, how we programmed the model, recorded the movements, difficulties during the projects, and how we implemented age estimation and conducted the experiment regarding age estimation.

1 Introduktion

1.1 Bakgrund

Detta projekt genomfördes under kurserna Neuromodellering, kognitiv robotik och agenter och Kognition, lärande och avancerade teknologi vid Lunds universitet. Vårt mål i den första kurserna var att få Epiroboten som utvecklats av Lund Universistets Cognitive Robotics Lab (hädanefter robotlabbet) att känna igen olika objekt och enkla egenskaper hos objekten. För att kategorisera objekten och dess egenskaper behövdes en maskininlärningsmodell som skulle tränas i Keras för att sedan importeras i Ikarossystemet som styr Epi. Vårt mål i den andra kurserna var att utveckla den tidigare

modellen vi hade för att kunna åldersestimera mäniskor. Vi testade först att implementera åldersestimering på vår nuvarande modell, men denna var inte effektiv. Istället använde vi oss av två andra modeller för att utföra åldersestimeringen eftersom dessa var bättre på detta. Vi hade några tekniska krav som vi kommer beskrivas senare, och några egna satta mål. Vi tog tidigt i delkurserna och bestämde vilka objekt och vilka egenskaper som Epi skulle känna igen och reagera på. Vi valde ansikten som objekt, och tilldelade sociala interaktioner som egenskaper hos Epi. Detta utvecklades i den andra delkursen till att åldersestimera mäniskor.

1.2 Ansiktsigenkänning

Varför valde vi just ansikten? Ansiktsigenkänning är en viktig egenskap för mäniskor. Med hjälp av denna funktion kan vi identifiera andra mäniskor, men även läsa av ansiktsuttryck, identifiera ålder, kön och så vidare. Detta i sin tur leder till en mer utvecklad social interaktion, där mäniskan kan bland annat identifiera intention, känslor, intresse hos andra mäniskor. Ansiktsigenkänning tillåter oss också att identifiera personer vi tidigare pratat med för en fortsatt interaktion, men även familjemedlemmar, vänner, respektive och andra mäniskor vi stöter på dagligen.

Denna funktion utvecklas nästan direkt som mäniskor föds. Nyfödda barn visar en preferens för objekt som är ansiktsformade jämfört med objekt som inte är det, och denna färdighet utvecklas desto äldre vi blir. Ansiktsigenkänning är alltså en viktig funktion för oss mäniskor, eftersom det är en så stor del av vår interaktion och sociala kognition (Simion & Giorgio, 2015).

1.3 Åldersestimering

Åldrande är en stokastisk process som påverkar variationer i ansiktsform och textur. För att kunna utföra en åldersestimering behövs faktorer skapas för att kategorisera åldrante. Det finns alltså ett klassificerings- eller regressionsproblem när ålder ska uppskattas i ett ansikte. Mäniskor tränas på att se ålderstecken sedan födslen och för att ta fram modeller som hanterar en sådan klassificering behövs först dessa ålderstecken extraherats. Detta sker vanligtvis genom observation av ansiktsformen samt texturen på ansiktet (Angulu, Tapamo & Adewumi, 2018). Därefter kan en algoritm användas för att utföra den faktiska åldersestimeringen.

1.4 Epi



Figur 1. Stora Epi till vänster, lilla Epi till höger, bilden är hämtad från <http://www.epi-robot.org/>

Epi är en humanoid robot som utvecklas av robotlabbet. Det finns två olika modeller av Epi: en större och en mindre (se Figur 1). Den större har en hel överkropp och rörliga armar och den mindre är en byst som enbart har huvud, nacke och axlar. Vi har arbetat med den lilla versionen av Epi, eftersom vi har flyttat den fram och tillbaka. Den stora är svårare att förflytta och har periodvis varit upptagen i ett annat projekt. Epi har många likheter med en mänsklig: den kan tex se, tala, och röra sig, men det finns vissa begränsningar. Den första begränsningen ligger i Epis rörelsefrihet. Epis huvud har bara två rörelsefrihetsgrader; den kan vrida sig upp och ner, samt 90 grader åt sidan. Epi kan till exempel inte lägga huvudet på sned, då den inte har den rörelsefriheten. Den andra begränsningen gäller Epis ögon. Dessa består utav två stycken Raspberry Pi Zero med Camera Module V2 installerade. Dessa stödjer 1080p30, 720p60 och 640x480p90 videoströmmar. Kameran vi fick tillgång till använder sig av 640x480 formatet. Utöver detta konverterar kameran bilderna till ett h.264 video format. Det går att ändra på inställningarna till ögonen, både format och upplösning. Epis kontrollsysteem är baserad på neurofysiologisk data och är implementerad med hjälp av Ikaros. (Johansson et al. 2020).

1.5 Ikaros

Ikaros är ett öppet ramverk för realtidskontrollering av robotar och hjärnmodulering på systemnivå. Lund University Cognitive Science (LUCS) började utvecklingen av Ikaros-systemet år 2001 och idag används systemet av flera universitet och forskningsinstitut runt om i världen (Ikarosproject.org, u. å.).

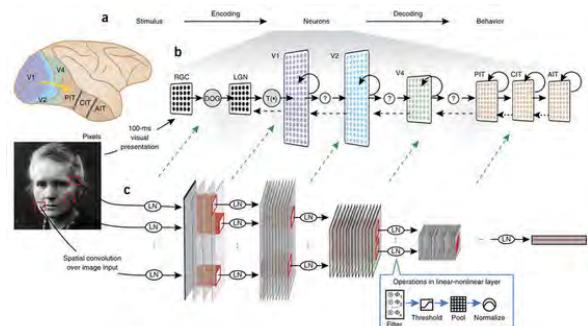
Ikaros-systemet är uppbyggt av flera moduler. En modul har en input för att ta emot inkommande data och en output för utgående data. I själva modulen processeras inputdataen innan den skickas ut till outputen. Flera moduler kan kop-

pas samman med varandra och skapa ett nätverk av moduler som utgör en modell i Ikaros. Ikaros-ramverket innehåller flera färdiga standardmoduler exempelvis: Input/Output-moduler för att hantera filer och data in och ut från systemet, hjärnmodells-moduler för olika hjärnmodeller samt robotmoduler för att hantera olika robotsystem, motorer och andra hårdvaror i robotar. Utöver detta så använder man Ikaros plattformen för att köra Epi (Balkenius, Morén & Johansson, 2007)

I detta projekt har främst robotmodulerna för inspelning av rörelser och ljud med tillhörande grafiska gränssnitt används.

1.6 Robotlabbet och hjärnmodellering

Robotlabbet arbetar inom två spår, varav det ena är hjärnmodellering. Som tidigare nämnt är Ikaros en plattform som Epi använder sig av för att köras. Ikaros är också en modellering av hjärnan, där hjärnans processer och funktioner ska efterliknas. Ikaros består, som tidigare nämnts, av olika moduler. Dessa gör specifika saker och efterliknar specifika delar av hjärnan. Modulerna kan samarbeta med flera andra, precis som olika regioner i hjärnan gör. Vi har arbetat utifrån detta spår (hjärnmodellering) eftersom vår modell innehåller ett Convolved Neural Network (CNN). Detta nätverk liknar hjärnans syncentrum. Det innehåller snarlika komponenter som gör specifika saker för att skicka denna information vidare i processen. CNN-modeller liknar med andra ord hjärnans syncentrum. Hur dessa liknar varandra kan man se i Figur 2).



Figur 2. Liknelser mellan hjärnans synbearbetning och hur ett CNN nätverk bearbetar bilder. Bilden är hämtad från <https://medium.com/nerd-for-tech/understanding-convolutional-neural-network-cnn-9f5ec8a308ac>

När det gäller den visuella informationen så bearbetas den först från informationen våra ögon inhämtar. Detta gäller en maskin eller en robot. Den inhämtar information från en kamera som sedan bearbetas. När det gäller människor skickas information genom olika delar i hjärnan för att sedan nå occipitalloben och i förlängning den primära synbarken och sekundära synbarken (V1-V5) (Michael Gazzaniga, 2019). I synbarken bearbetas den "enklaste"informationen först och komplex visuell information bearbetas senare i den visuella processen. Samma process finns

även i en CNN-modell där enklare information bearbetas tidigt för att sedan bearbeta mer komplex information. Även om CNN har många liknelser till människans visuella bearbetning är den inte en perfekt likhet. För vårt ändamål så bedömer vi att CNN liknar människans visuella bearbetning tillräckligt (Xu & Vaziri-Pashkam, 2021).

1.7 Robotlabbet och social interaktion

Som tidigare nämnt arbetar robotlabbet i två spår. Det andra spåret är social interaktion med robotar. En social robot är en robot som antingen interagerar med andra robotar eller med mänskor. Interaktionen ska efterlikna den som mänskor har mellan sig i en social kontext, det vill säga Human-Human Interaction (HHI). I situationer med robotar ersätter man en mänsklig med en social robot, och har då en Human-Robot Interaction (HRI). Detta kan göras av många olika anledningar, bland annat för att bättre förstå HHI eller för att undersöka hur både mänskor eller robotar interagerar med varandra. För vår del undersökte vi detta spår utifrån ansiktsigenkänning; vilken funktion har ansiktsigenkänning för en robot som har den egenskapen? Här gällde det att undersöka vilket beteende Epi skulle visa upp när den kände igen mänskor, och vilket beteende Epi skulle ha när den kände igen andra Epis. Epi är inte autonom och kan inte utföra rörelser eller andra interaktioner själv. Vi undersökte vilka kommunikativa egenskaper som Epi skulle uppvisa och spelade som tidigare nämnt in dessa via inspelningsplattformen. De sekvenser vi spelade in representerar hur mänskor hälsar på varandra, och en skillnad gjordes på robotar och mänskor. Vi antog att Epi hade en mer vänlig inställning till andra Epis, då dessa kan ses som dess vänner och en mer neutral inställning till mänskor då dessa är främningar. Detta ligger i linje för hur mänskor hälsar på varandra (Fox & Gambino, 2021; Goodrich & Schultz, 2007).

1.8 RestNET50

ResNET50 är en CNN-modell som är förtränsad på ImageNet-dataset och kan klassificera bilder i 1000 olika objektkategorier (MathWorks, 2022). Som ett resultat har CNN-modellen lärt sig funktionsrepresentationer för en mängd olika bilder. Modellen kan implementeras med hjälp av Keras api (Keras, 2022).

1.9 SSR-Net

SSR-Net eller Soft Stagewise Regression Network är förtränsad på IMDB-WIKI som innehåller över 500 000 bilder med ålders labels. SSR-Net är en CNN-modell som byggs för att vara kompakt och inte för tekniskt krävande i motsvarighet till tidigare modeller. SSR-Net är inspirerad av dessa tidigare modeller och använder sig av en multiklass klassificering i flera steg för att sedan genomföra regression på värdena utifrån klassificeringen. Det finns problem med multiklass klassificering enbart i ålderestimering eftersom ålder är en kontinuerlig variabel. Varje åldersklass har en hög korrelation istället för att vara separata klasser, det vill säga det är

inte en separat klass mellan 25 och 26 på grund av den höga korrelationen. Det är fördelaktigt att genomföra regression på värdena utifrån multiklass klassificeringen, för att kunna genomföra en ålderestimering mer effektivt. Vilket SSR-net gör. (Yang, Huang, Lin, Hsieh & Chuang, 2018)

1.10 WideRes

WideRes är även denna en CNN-modell. Till skillnad från andra CNN-modeller så tittar denna modellen inte djupt i de olika lagrena som skapas i CNN-modellerna utan tittar brett istället, därav namnet WideRes. WideRes har färre lager men arbetar dubbelt så snabbt. Detta innebär att WideRes inte är lika krävande av datorn. WideResNet-arkitekturen (WideRes) är också förtränsad på IMDB-WIKI. Modellen har tagits fram för att tränas snabbare än tidigare modeller. Den tränas snabbare genom att använda sig av "residual blocks" där man kan lägga till fler lager per block, och bredda lagrena genom att lägga till flera plan av egenskaper. Det vill säga att varje lager kan upptäcka flera egenskaper. Sist kan man använda sig av blocken för att öka filteringskapaciteten i varje lager så att varje lager kan filtrera relevanta egenskaper. (Zagoruyko & Komodakis, 2016)

2 Metod och Implementering

Projektet har delats upp i 2 olika studier. I den första studien var syftet att lära en robot att känna igen något föremål, något som senare valdes till ansikten och ansiktsidentifiering. I den andra studien var fokuset på ålderestimering med ansikten och en jämförelse av olika maskininlärningsmodeller för detta. Arbetsprocessen har inneburit flera val och beslut som påverkat resultatet av projektet. Detta kapitel beskriver samt motiverar metoder och implementeringar som beslutats och utförts under processens gång. Kapitlet är indelat i två delkapitlar, en för varje studie som utförts.

2.1 Studie 1

2.1.1 Syfte

Syftet med den första studien var att lära en robot att känna igen föremål och dess egenskaper med hjälp av maskininlärning. Maskininlärning var något som arbetsgruppen hade begränsade erfarenheter av. Detta innebär att under arbetets gång var det viktigt att ta reda på mer om maskininlärning samtidigt som direkta arbetsuppgifter utfördes. Initialt fanns många idéer om i vilken riktning gruppen skulle ta projektet. För att Epi skulle kunna ha nytta av denna erfarenhet i framtiden bestämdes det att projektet skulle ha något att göra med mänskor. Med en brainstorming togs flera intressanta ämnen fram.

Ansikten valdes som den riktning projektet skulle ta eftersom att detta kändes mest naturligt ur ett mänskligt perspektiv. Ansiktsigenkänning är något som Epi kan ha nytta av i framtiden, och något som var intressant för arbetsgruppen att undersöka och implementera själva. Det är en teknik som används i samhället, som till exempel identifiering i

mobiltelefoner. Det fanns även underlag för att fortsätta utveckla denna riktning, med ansiktsuttryck och integrering.

2.1.2 Teknisk implementering och utförande

Den tekniska implementeringen av detta projekt består av flera deluppgifter som kommer att beskrivas i detta kapitel: Datainsamling, val av maskininlärningsmodell, träning av modellen och förvandla klassificerare till detektor. Arbetet utfördes i iterationer, där de flesta av de nämnda deluppgifterna genomfördes upprepade gånger. Projektet är helt skrivet i Python och använder Tensorflow- och OpenCV-biblioteken samt Keras api.

Det första steget i processen var att samla in dataset som motsvarar de objektklasser vi valt att detektera: mänskliga ansikten och Epirobotar. Dataset är en kollektion av bilder som vi använde för att träna vår modell för att känna igen olika objekt. Bilderna på mänskliga ansikten samlades in från datasetet *CelebFaces Attributes Dataset (CelebA)* från databasen *Know Your Data* (KnowYourData, 2022; Tensorflow, 2022). Att hitta bilder på Epirobotar var inte lika lätt, eftersom det fanns ett begränsat antal tillgängliga online. I de tidiga iterationerna använde vi så många bilder vi kunde hitta, och därefter utökade vi datasetet med bilder från robotlabbet.

Nästa steg var att välja en maskininlärningsmodell för att träna på datamängden vi samlade in för ansikts- och objektklassificering. Modellen vi valde var ResNET50. Det finns många olika modeller vi kunde ha använt, men vi valde att använda ResNET50-modellen eftersom det var lätt att implementera med hjälp av Keras api och även på grund av att det finns en uppsjö av resurser som använder ResNET50-modellen som vi baserade vår lösning på, vilket gjorde att vi sparade mycket tid och kunde lägga mer fokus på andra delar av projektet.

Efter att vi valt ResNET50-modellen var det dags att träna modellen med vår datamängd, vilket skapade en klassificerare för ansikts- och objektkänning. En klassificerare är en maskininlärningsmodell som tar emot en bild som indata, och som utdata ger en label för en av de möjliga klasserna som i vårt fall är *human* eller *epi*.

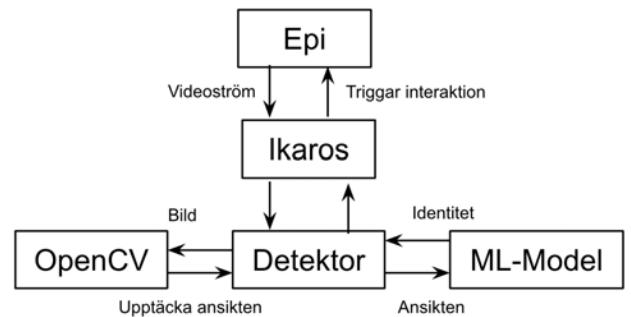
Fördelen med att använda en förtränad modell som ResNET50 är möjligheten att utnyttja *transfer learning* för att bygga en anpassad klassificerare, vilket i vårt fall skulle klassificera bilder av mänskliga ansikten och Epirobotar. Transfer learning går ut på att ta funktioner som en modell lärt sig av ett problem, för att sedan utnyttja dem på ett nytt liknande problem (Keras, 2020). Eftersom ResNET50-modellen är tränad på att klassificera objekt i 1000 olika objektkategorier, fann vi att den var lämplig att användas för att detektera mänskliga ansikten och Epirobotar. Transfer learning är det viktigaste konceptet i den tekniska delen av projektet eftersom det är väldigt svårt att bygga en ansiktsklassificerare från grunden och det skulle kräva enorma tidsresurser och expertkunskap inom CNN och maskininlärning för att kunna åstadkomma, vilket inte ligger inom ramen för detta projekt.

I den första iterationen tränades modellen med 20 bilder av mänskligansikten. Programmet kunde känna igen ansikten rakt framifrån med hjälp av datorns webbkamera, men på grund av en begränsad dataset saknade den noggrannhet. Därefter utökade vi datasetet och samtidigt beslutades det att ännu en objektklass skulle läggas till för att känna igen andra epirobotar genom att lägga till bilder på Epi, och resultaten blev bättre.

Vi implementerade även en detektor som använder OpenCV-biblioteksmetoder för kameraåtkomst och ansiktsdetektion så att Epis kamera ska kunna upptäcka ansikten i realtid (OpenCV, 2022). Från detektorn skickas en bild på ett ansikte till modellen som returnerar om ansiktet är mänskligt eller tillhör en Epi-robot. Programmet visar även flödet från Epis kamera så att resultaten kan övervakas manuellt. Detektorprogrammet triggar även relevanta rörelser och ljud med Epi.

Den sista delen av arbetsprocessen innefattade att spela in rörelser och ljud för att Epi ska kunna reagera när den känner igen någonting. Interaktionen valdes som en nickning eftersom att detta är ett mänskligt sätt att hälsa utan att använda händerna. Från början var tanken egentligen att lära Epi att hälsa med en handrörelse, men roboten med armar fanns inte tillgänglig så här långt in i projektet. En nick uppåt innebar att en robot identifierats, medan en nick nedåt var för en människa. Det valdes även att låta Epis ögon ändra färg beroende på vilket objekt den kände igen, där blå färg indikerade att en robot hittades, medan grön innebar människa.

En enkel schematisk bild över den tekniska implementeringen kan ses i figur 3



Figur 3. Schematisk bild över den tekniska implementationen

2.2 Studie 2

2.2.1 Syfte

Syftet med den andra studien var att utforska modeller för ålderestimering som kan fungera på Epi och sedan jämföra huruvida dessa modeller presterade jämfört med mänskor. Genom att testa olika modeller och se vilka som genererade en korrekt ålderestimering kan slutsatser dras kring vilken typ av modell, klassificeringsmodell, regressionsmodell eller hybridmodell, som är bäst lämpad för detta ändamål.

Ett annat mål med studien är att jämföra resultatet mellan de bäst lämpade modellerna och mäniskor. En sådan jämförelse kan resultera i slutsatser kring om modeller kan implementeras på ett sätt som gör de bättre på ålderestimering än en mäniska.

Utifrån syftet utformades en frågeställning:

Kan en maskininlärningsmodell vara bättre än mäniskor på att estimera den sanna åldern på en mäniska?

Fyra hypoteser formulerades därefter:

- H1: Alla modeller är bättre än mäniskor.
- H2: Enstaka eller vissa modeller är bättre än mäniskor
- H3 (Null): Inga modeller är bättre än mäniskor.
- H4: Mäniskor inom vissa åldrar eller visst kön är bättre än andra på att känna igen mäniskor.

2.2.2 Teknisk implementering

Bilderna som användes för att utföra denna studie kommer från UTKFace-datasetet som är ett dataset med över 20 000 ansikten på personer mellan 0 till 116 år. Samtliga bilder i datasetet är tagna från internet och bilderna är justerade samt beskruva så att främst ansiktet är synligt. Varje bild i datasetet har information om ålder, kön och etnicitet på personen på bilden. (Song & Zhang, u. å.)

Ett försök gjordes med den tidigare ResNet50-modellen som användes i studie 1, där modellen fick tränas på UTKFace och ett mindre antal tester utfördes. Testerna visade dock att modellen var ganska dålig och vi valde därför att leta efter andra modeller. Vi valde att leta efter regressionsmodeller som var förträpnade på ålderestimering istället för klassificeringsmodeller för att bättre passa in med vårt experiment.

Modeller och implementeringar som användes i Studie 2 är SSR-Net skapad av Yang m. fl. (2018) och implementation av Chernov (2019) samt WideResNet50 av Zagoruyko och Komodakis (2016) och implementation av Zhang (2018) har använts. Dessa modeller valdes för denna studie främst på grund av att själva modellerna är förträpnade för ålderestimering och är effektiva samt snabba på att beräkna fram en ålder. Själva modellerna och dataseten fanns tillgängliga lokalt och kördes i en bärbar dator. Båda modellimplementationerna har en funktion för att med hjälp av en webbkamera eller en videoström ge en fluktuerande ålder i realtid. Denna funktion används för att koppla ihop med Epis ögon för demostrativt syfte. En annan funktion för att göra det möjligt att skicka enstaka bilder istället för en videoström har implementerats för båda implementationerna. Funktionen används i experimentet för att skicka in bilder från UTKFace. Tillskillnad från Studie 1 utfördes ingen träning, utöver den modellskaparna/utgivarna gjort sedan tidigare, eftersom modellerna redan är förträpnade för ålderestimering. UTKFace-datasetet kommer därför endast användas för att testa och jämföra modellerna med mäniskor. Om en mäniska registreras hos någon av våra två modeller SSR, eller

WideRes så börjar denna skriva ut ålder i realtid, modellen skriver ut åldern kontinuerligt, den kan till exempel fluktuera mellan 20-40 år på någon som är 30.

För att skapa interaktion med de nya ålderestimeringsfunktionerna har ett antal ljudklipp med en röst som säger olika åldersspann: 0-15 år, 16-30 år mfl. spelades in och implementerades i Epi. Tanken är att Epi ska säga åldern på personen den ser framför sig. Likt studie 1 skapades ett pythonprogram för att koppla ihop maskininlärningsmodellerna, Epis kameror samt interaktionerna från Epi med varandra.

2.2.3 Utförande

För att besvara frågeställningen sattes ett experiment upp, där testpersoner, samt de två modellerna, skulle genomföra en ålderestimering på ett antal bilder tagna från testdatasetet.

Datasetet som användes var UTKFace där sann ålder till varje bild fanns tillgängligt. För varje testperson skapades ett testset bestående av 30 slumpmässiga bilder från datasetet. Enda manipulationen av datasetet är att vi lagt bilderna i egna ålderskategorier. För att få ett heltäckande resultat valde vi att dela upp datasetet genom att kategorisera bilderna i olika åldersspann, med fem år mellan varje grupp. Grupperna som skapades gick från 1- 5 till 96- 100år. Vi valde att dela upp åldrarna i kategorier för att inte missa några datapunkter. När vi genomförde experimentet så tog vi ut värdet av den sanna åldern, detta gjordes även i resultat, eftersom vi hade tillräckligt med datapunkter för sann ålder. Vi tog och genomförde ålderskategoriseringen för att välja ut vilka bilder som fanns med i testsetten. Utöver åldersuppdeleningen av bilderna utfördes urvalet av bilder för ett testset slumpmässigt. Detta genomfördes genom att först slumpa ifrån vilken åldersgrupp bilden skulle tas från för att sedan slumpa vilken individuell bild från den åldersgruppen som skulle presenteras för testpersonen. Denna procedur utfördes för varje bild som lades in i ett testset.

Experimentet utfördes både via enkät och intervju. Först fick deltagarna呈现出 sin ålder och sitt kön, inga andra personuppgifter samlades in. Denna datan används enbart för jämförelse för att se om det finns någon skillnad mellan kön och ålder hos testdeltagare. Därefter presenterades 30 olika bilder på ansikten varpå deltararen fick estimera åldern av personerna på bilderna. Deltagarna uppmanades att ge en ålder i heltalet som sedan noterades.

Alla bilder, som testades på deltagarna (mäniskor) användes också till att testa modellerna som också fick utföra en ålderestimering. Modellerna skickar ut ett heltalet som utdata för sin ålderestimering på en bild som noterades.

Resultatet sammanställdes och datan från testdeltagare jämfördes med de olika maskininlärningsmodellerna genom att bland annat titta på jämförelseåldern det vill säga felmarginen eller den absoluta differensen mellan den estimerade åldern och den faktiska åldern.

Totalt utfördes tester på 34 personer mellan 17-71 år. Medelåldern på testdeltagarna var 31,2 år och medianen på 23 år. Totalt användes och estimerades 1049 bilder.

3 Resultat

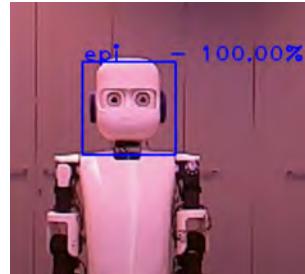
3.1 Studie 1

3.1.1 Igenkänning

Om en mänsklig identitet identifieras i Epi:s synfält registrerar Epi det och det skapas där med en grön ram kring mänskligans ansikte vilket kan ses i figur 4. Det visas även en text som specificerar vad för objekt Epi ser, här det för en mänsklig står "human". Till höger om texten dyker den procentuella sannolikheten att det är just det ansiktet som specificeras i texten, som uppträder i rutan. Procentsatsen säger alltstå hur säker modellen är att ansiktet är från en mänsklig eller en Epi-robot. Då en annan Epirobot noteras i Epi:s synfält skapas istället en blå ruta runt ansiktet med texten "epi" samt sannolikheten för att det är en Epi, vilket demonstreras i figur 5.



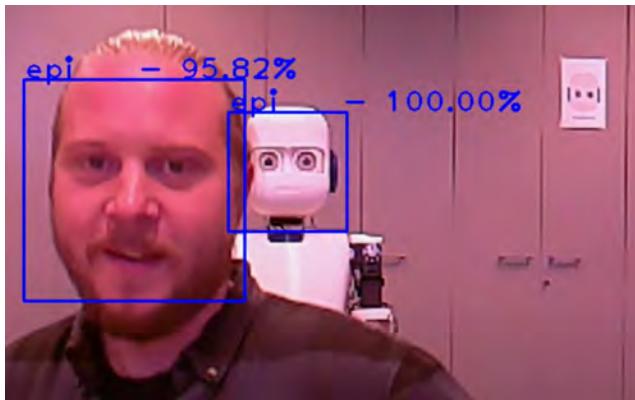
Figur 4. Igenkänning av en mänsklig i videoströmmen



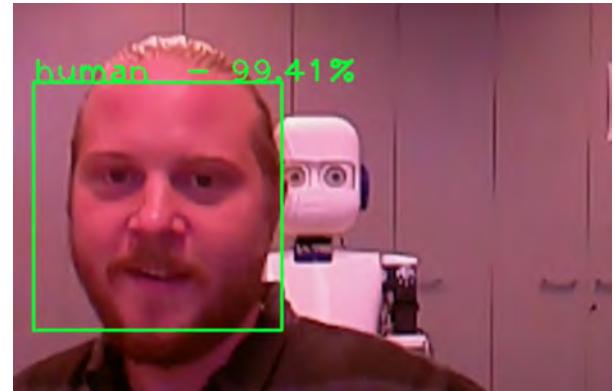
Figur 5. Igenkänning av en Epi-robot i videoströmmen

Resultatet för igenkänningsförmågan hos Epi har dock varierat utifrån vissa variabler, som exempelvis ansiktsform på testpersonen, kön på testpersonen och typ av ljus i videoströmmen.

I figur 6 demonteras ett resultat som påträffats vid upprepande tillfällen, där Epi misstalar en mänsklig för en annan Epirobot. I bilden kan vi urskilja att procentsatsen för sannolikheten att mänskligan är en Epirobot är hög. Figur 7 visar samma videoström ett antal millisekunder senare där Epi känner igen att ansiktet tillhör en mänsklig.



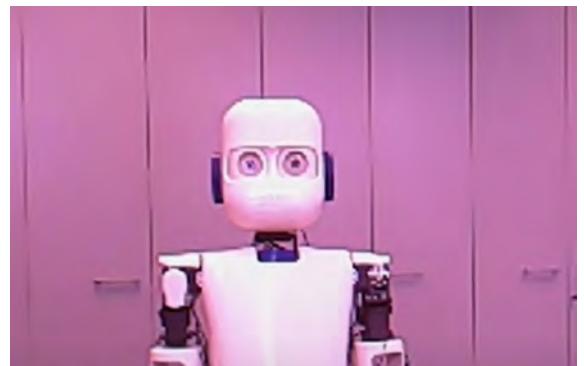
Figur 6. Epi misstalar en mänsklig för en annan Epi-robot



Figur 7. Epi känner igen att det är en mänsklig i videoströmmen

Ett annat resultat som uppmärksammats under testningen av ansiktsigenkänning är att Epi har enklare att känna igen ansiktet på kvinnor jämfört med män. Detta har framförallt noterats då Epi ofta känner igen kvinnor i profil, men sällan känner igen män i profil.

Under hela projektets gång har ett annat resultat observerats, nämligen att ljuset i videoströmmen har stor betydelse för huruvida Epi kommer känna igen ett ansikt. Detta resultatet har i de flesta fall förekommit då Epi ska känna igen en annan Epirobot i en ljus miljö, vilket visas i figur 8.



Figur 8. Epi känner inte igen Epi-roboten

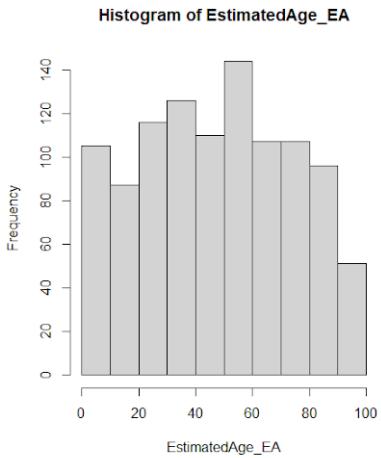
3.1.2 Interaktion

Epi kan även interagera med den mänskliga eller Epiroboten som förekommer i synfältet. Interaktionen sker både med rörelser och ljud. När mänskligan registreras av Epi förändras robotens ögonfärg till gröna. Efter det nickar roboten nedåt och säger "Hello". Under den tiden roboten pratar färgas munnen gröna. Då Epi uppfattar en annan Epirobot ändras istället munnen och ögonfärgen till blåa. Efter det nickar Epi uppåt och hälsar med "Hello".

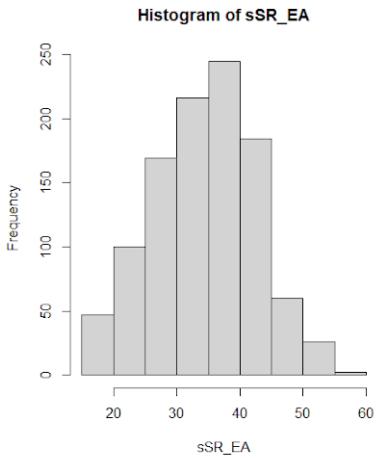
3.2 Studie 2

3.2.1 Åldersestimering

Resultatet har visat sig vara överraskande och vi har bekräftat vår nollhypotes. För att sammanställa resultatet på ett begripligt sätt så presenteras inte den fullständiga statistiska analysen. Istället kommer fokus vara på resultatet utifrån de statistiska analyserna, som genomfördes via R. Nedan följer de grafer, histogram och annan information om den statistiska analysen, för att sedan diskutera figurerna närmare.



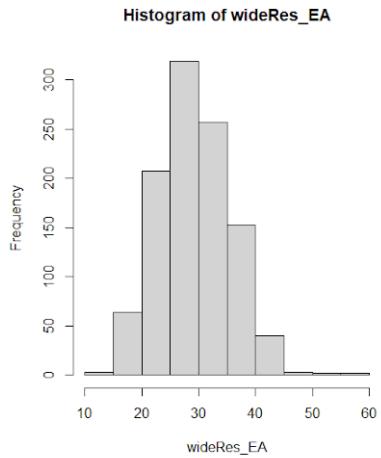
Figur 9. Histogram över testdeltagares gissningar



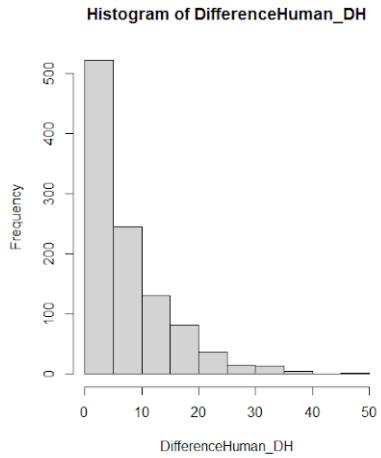
Figur 10. Histogram över SSRs gissningar

3.2.2 Tidig analys

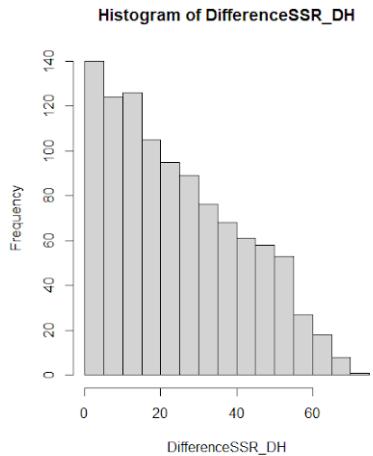
När vi undersökte datan insåg vi att modellerna är sämre än vad vi tidigare har trott. Vi började analysera datan genom att undersöka histogrammen för den estimerade åldern. I figur 9 kan vi se hur den estimerade åldern av testdeltagarna har en bred spridning i hela åldersspannet. Medan i figur 10 och i



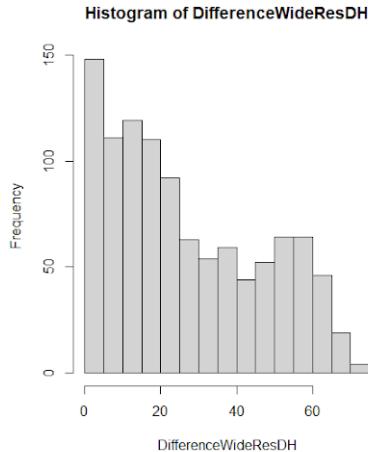
Figur 11. Histogram över WideRes gissningar



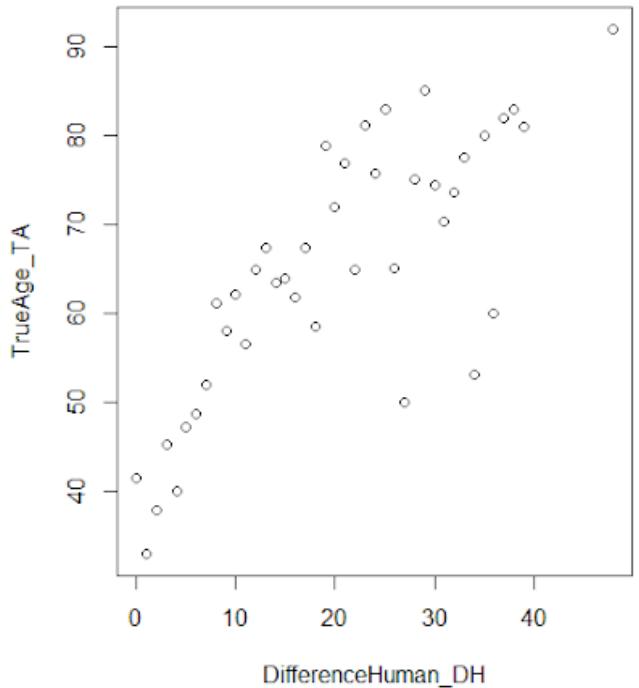
Figur 12. Histogram över människors jämförelseålder



Figur 13. Histogram över SSRs jämförelseålder



Figur 14. Histogram över WideRes jämförelseålder



Figur 15. Människors genomsnitt

14 så ser vi att både SSR och WideRes tenderar att estimera åldern mellan 20-40 år. Vi kontrollerade hur pass modellerna och människornas estimerade ålder korrelerade med den sanna åldern. För människor så korrelerade gissningarna med 93%, vilket innebär att den estimerade åldern överensstämmer väldigt bra med den sanna åldern. För modellerna var samma korrelation 55% och 56%, det vill säga de hade en lägre korrelation mellan estimerad ålder och sann ålder.

När vi hade undersökt den bakomliggande dataen så kontrollerade vi även hur det såg ut med skillnaden mellan den sanna åldern och den estimerade åldern häданefter, jämförelseåldern (differensen), det vill säga; om deltagaren gissade 55, och den sanna åldern är 54 så är jämförelseåldern 1. I figurerna 12,13,14, kan vi se hur människor och båda modellerna ligger i jämförelseåldern. Vi undersökte också standardavvikelsen för människor och modellerna i jämförelseåldern. För människor var avvikelsen på 7,39, för SSR 17,2 och WideRes 19,7.

När vi undersökte snittet för människors jämförelseålder kontra den sanna åldern, det vill säga den absoluta snitt-jämförelseåldern är: 7.314 år. Detta betyder att människor är i snitt 7,3 år ifrån den sanna åldern. Denna jämförelsen tog bara hänsyn till den sanna åldern, se figurerna 9,16 och 17 för snitt-jämförelseåldern.

Vad som var intressant i jämförelseåldern är att människor är väldigt duktiga på att estimera åldern. Jämförelseåldern ligger nästan helt och hållt inom 0-10 år, medan modellerna har en mycket större spridning i jämförelseåldern.

3.2.3 Människor

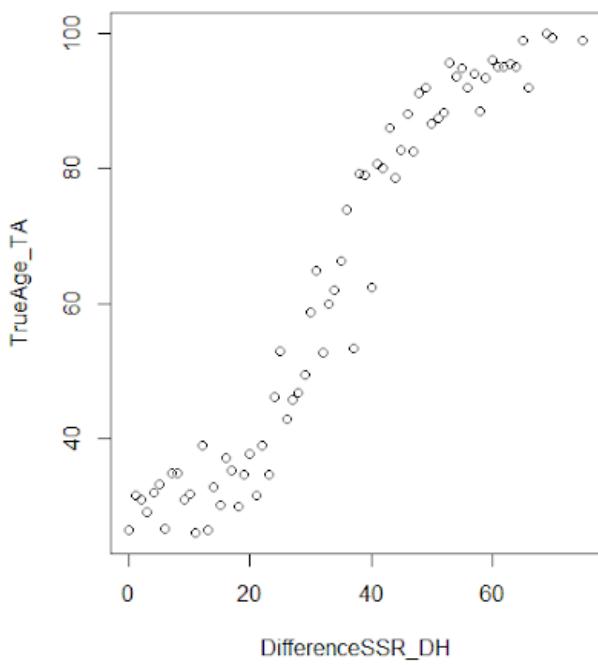
I figur 18 kan vi se att människor gissade i totala snitt 7,7 år (grön). Lägsta snittet, det vill säga hur pass nära människor var på korrekt svar utan att svara korrekt, är 0,35 (blå) och snittet längst ifrån sann ålder var 15,1(röd). I denna snittanalys kontrollerades vad snittet är, givet vad den estimerade åldern var samt vad den sanna åldern är. På sådant

sätt kunde vi få ut de tre olika värdena, där vi kan se lägre, genomsnitt och högre snittet. Enkelt uttryckt så snittade människor sin jämförelseålder med 7,7 år.

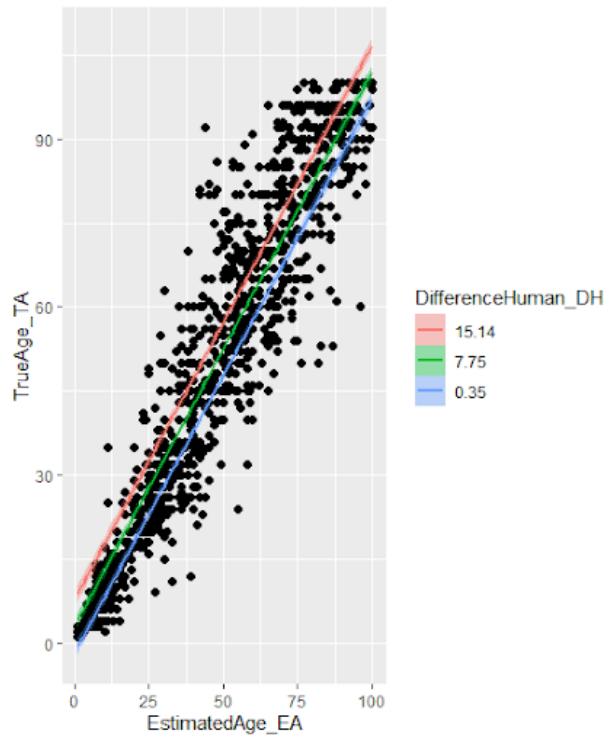
När människor inte fick rätt så var vi väldigt nära, och om människor gissade fel så var detta i snitt inte högre än 15,1 år. Människorna gissade över hela åldersspannet, som tidigare nämnt. I figur 19 kan vi se hur människor gissade i snitt. Totala snittet ligger på 48,49 år och undre snittet på 21,97 år och högre på 75 år, med andra ord så gissar vi kanske högre på yngre människor, lägre på äldre människor och precis mittemellan totalt sett.

Vi undersökte om kön på deltagarna påverkade resultatet med bland annat ett t-test vilket det inte gjorde (p-värde var 0.801). Deltagarnas ålder var inte heller en faktor för hur pass ofta de gissade rätt eller fel och hur jämförelseåldern ser ut. Detta återkommer senare i diskussion. Testdeltagarna fick totalt 88 korrekta gissningar av 1049. Testdeltagarnas gissningar var i snitt 7,3 år ifrån korrekt svar. Människor har en sannolikhet på cirka 8,3% att gissa helt rätt, om vi inte tar hänsyn till den estimerade åldern och deras jämförelseålder. I modellen som vi körde tog vi hänsyn till dessa värden eftersom detta kunde eventuellt påverka resultatet. I den statistiska analysen som vi utfört med de relevanta variablene har har alla människor en sannolikhet på ca 7,7% att gissa helt rätt.

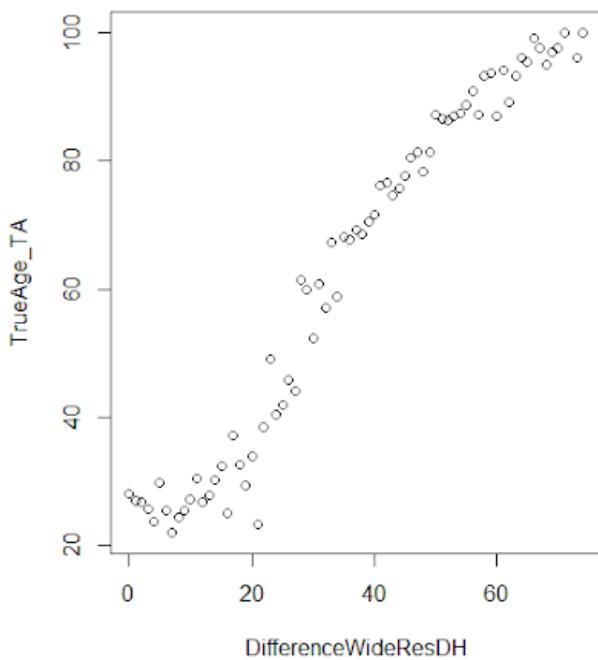
Vi undersökte även om deltagarnas ålder påverkade jämfö-



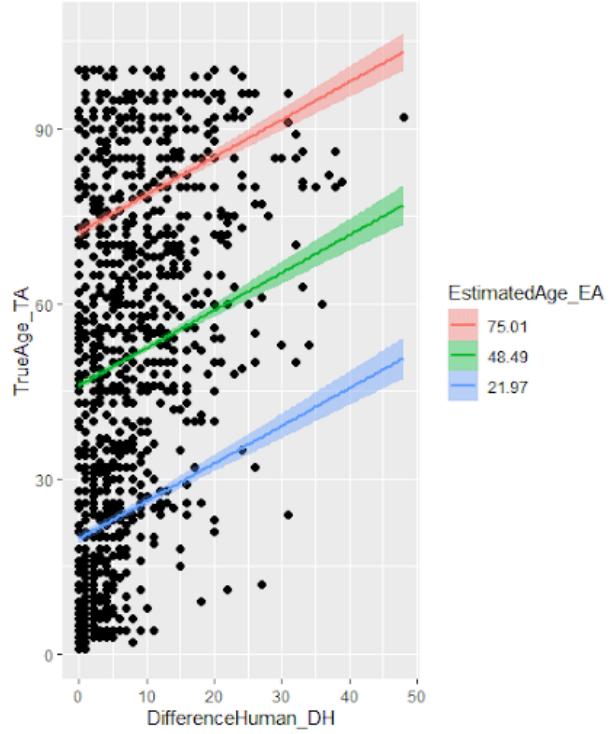
Figur 16. SSRs genomsnitt



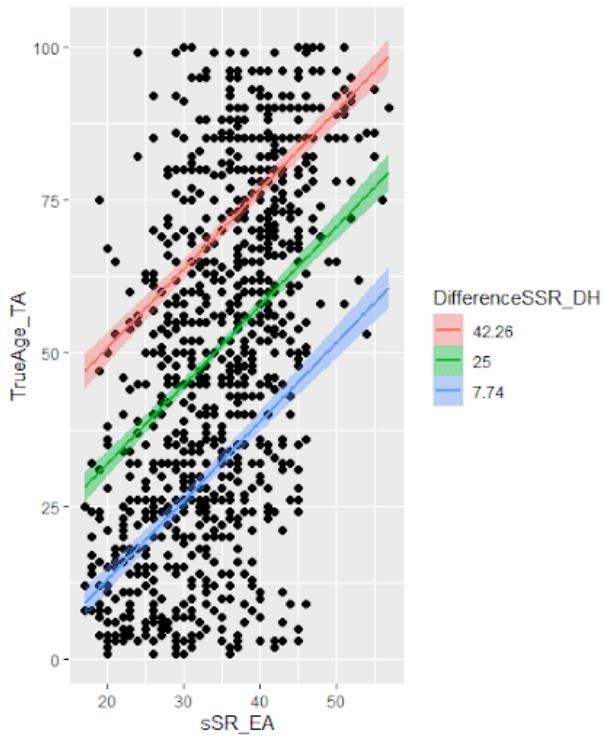
Figur 18. Bild för människors jämföresålder, i förhållande till estimeradålder och sann ålder



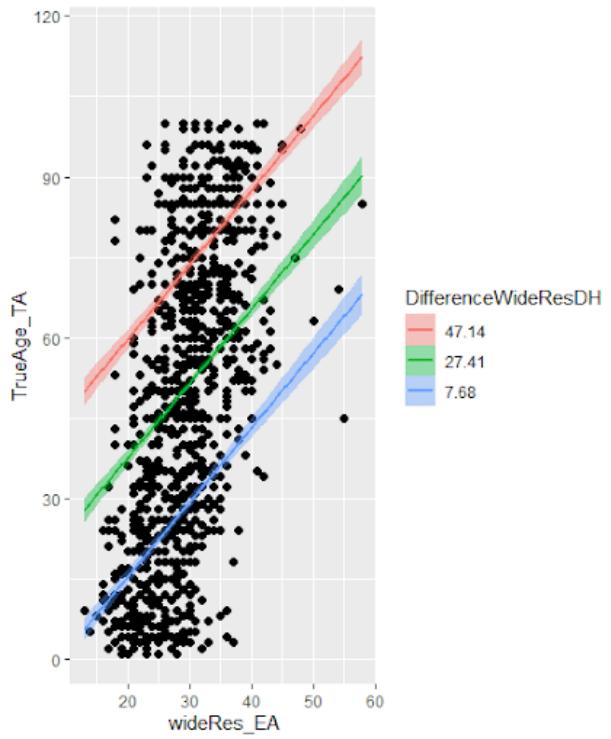
Figur 17. WideRes genomsnitt



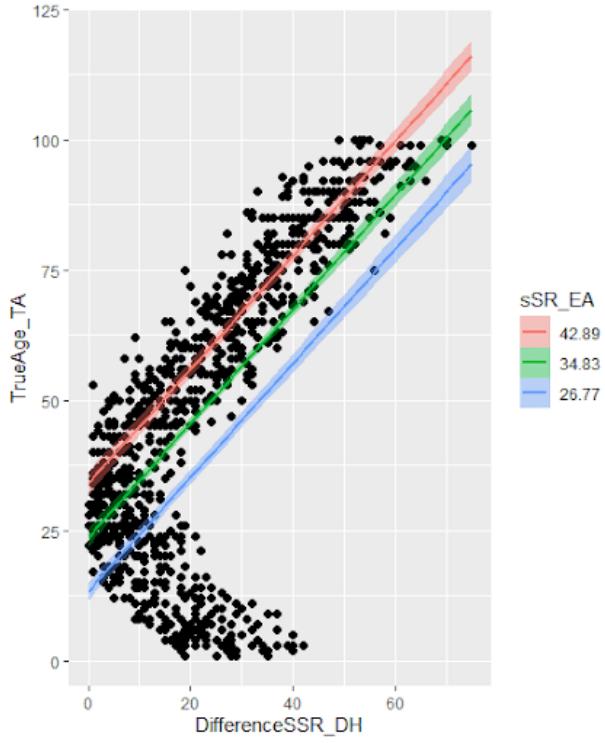
Figur 19. Bild för människors estimerade ålder, i förhållande till jämförelseålder ålder och sann ålder



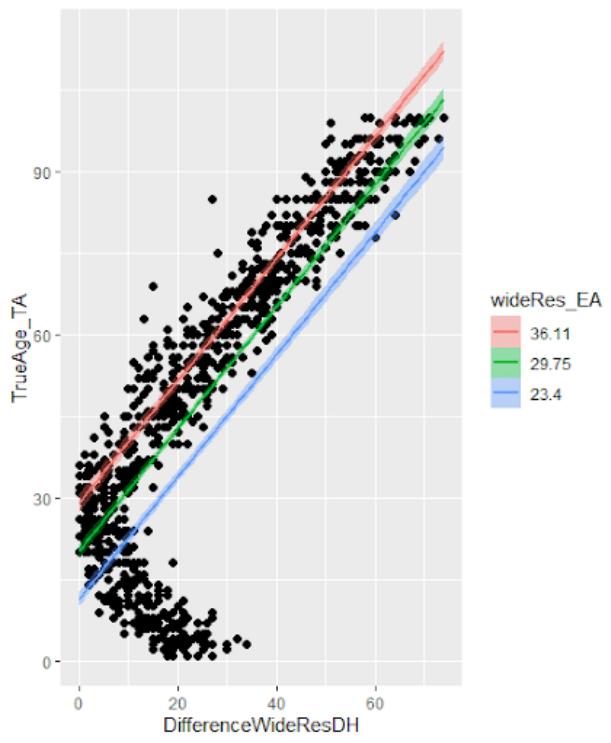
Figur 20. Bild för SSRs jämförelseålder, i förhållande till estimerad ålder och sann ålder



Figur 22. Bild för WideRes jämförelseålder, i förhållande till estimerad ålder och sann ålder



Figur 21. Bild för SSRs estimerade ålder, i förhållande till jämförelseålder och sann ålder



Figur 23. Bild för WideRes estimerad ålder, i förhållande till jämförelseålder och sann ålder

reelseåldern och den estimerade åldern. Men på grund av att majoriteten av deltagarna var under 35 år så gav inte detta några relevanta resultat. Vi jämförde dock hur den yngsta deltagaren, den äldsta och en deltagare runt 30 skilde sig emellan. Jämförelseåldern samt deras estimerade ålder skilde sig inte markant, vilket innebär att ålder på testdeltagare inte påverkade vad de gissade.

3.2.4 SSR och WideRes

När vi undersökte jämförelseåldern för modellerna såg vi att dessa är betydligt högre än för mäniskor. SSR:s absoluta jämförelseålder är i snitt 23,6 år och för WideRes är denna på 25,8 år. Detta betyder att båda modellerna gissade 23,6 och 25,8 år från den sanna åldern. När vi undersökte detta med samma förhållande som för mäniskorna, det vill säga hur ofta modellerna har rätt och vad deras gissningar var, så ser vi att båda modellerna oftast gissar att åldern är mellan 20-40 år. I figur20 kan vi se att SSR gissar mycket runt 20-40 år, oavsett vad den sanna åldern är. Snittet är också betydligt högre än för mäniskor. SSR snittade 25 år för jämförelseåldern, 7,7 år ifrån rätt när de inte fick rätt och 42,2 år längst ifrån rätt svar, se figur 21. SSR gissar också som lägst i snitt 26,7 år, genomsnitt 34,8 och högst 42,8 år, se figur 20 igen. För WideRes så ligger siffrorna på liknande nivå se figur22 och 23 . SSR fick totalt 4 rätt och WideRes 8 rätt, SSR:s gissningar var i snitt 23,6 år ifrån sann ålder och WideRes var 25,8 år. I de bayesianska modellerna som vi genomförde har SSR cirka 0,3% att gissa rätt och WideRes cirka 0,7% att gissa rätt.

3.2.5 Statiska modeller

En undersökning kring hur sann ålder, estimerad ålder, jämförelseålder och rätt eller fel påverkades genom olika linjära regressioner utfördes också. Vi undersökte hur dessa modeller såg ut i ”anpassade värden” (fitted values), det vill säga; vi tränade en ny modell på värdena vi hade fått från de bayesianska modellerna. Posteriorisannolikhet genomfördes för att se om modellerna kunde tolka resultatet bra eller inte, det vill säga om sannolikheten överensstämmer med den observerade datan, vilket den gjorde till stor del. De modeller som var mest relevanta återfinns i appendixen.

4 Diskussion

4.1 Studie 1

Som resultatet visar är vårt program inte helt felfritt. I många fall har programmet detekterat och klassifierat mänskliga och Epi:s ansikten korrekt. I vissa fall klassificerade programmet och modellen ett mänskligt ansikte som en Epirobot. Det finns flera anledningar till att detta kunde ha hänt. En anledning är på grund av likheterna mellan egenskaperna hos mänskliga ansikten och Epiansikten. En annan anledning kan vara att datamängden vi samlade in på Epibilder och mänskliga ansikten inte är tillräckligt stor och

varierad. Ju fler bilder som används för att träna modellen desto bättre kan den klassificera objekt i rätta kategorier.

Vi kunde ha samlat in fler bilder manuellt men det finns ett annat sätt att göra det med hjälp av dataförstärkning. Dataförstärkning eller *data augmentation* går ut på att artificiellt öka en datamängdsstorlek genom att tillämpa vissa transformationer på originalbilderna. Till exempel horisontell och vertikal reflektion, rotation eller förstoring, färginversion och förändringar i ljusstyrkan. Detta skulle avsevärt öka storleken på vårt dataset och följdaktligen nätverkets generaliseringsförstående av objektklassificering. Det diskuterades mycket kring användningen av dataförstärkning men till slut hade vi inte tillräckligt med tid för att implementera det. Detta är något som kan göras i nästa kurs.

Det är värt att nämna att det även finns ett sätt att förbättra klassificeraren genom att använda metoden för finjustering. Finjustering eller *fine tuning* består av att frysas upp hela modellen vi erhållit efter att ha utfört *transfer learning* och sedan träna om modellen på den nya datan vi samlat in med dataförstärkning med en mycket låg inlärningshastighet. Detta kan potentiellt uppnå meningsfulla förbättringar av vår klassificerare genom att stegvis anpassa de förträna funktionerna till den nya datan (Keras, 2020).

En annan felkälla som kan påverka detekteringen är användningen av OpenCV:s ansiktsdetektering. OpenCV:s ansiktsdetektering är främst avsedd för att upptäcka mänskliga ansiktet men Epi ansikte har visat sig vara tillräckligt likt i många fall för att den ändå ska registreras som ett ansikte. En annan teknisk lösning kan användas för att förbättra detekteringen av Epi:s ansikten.

4.1.1 Avgränsningar

En avgränsning vi har fått göra i projektet är att inte arbeta direkt med Ikaros, plattformen Epi körs från, som vi tidigare tänkt. Ikaros är uppbyggd med olika moduler som gör specifika saker, det finns exempelvis en motorcontrol som efterliknar just motoriken. Det finns även flera visuella moduler som samarbetar för att efterlikna den visuella processen som mäniskor har. Detsamma gäller för deep learning och igenkänning. En modul för att implementera maskinlärningsmodeller med Ikaros är under utveckling, med eftersom denna modul inte var färdig har vi i denna delkurs behövt hitta andra lösningar för att genomföra den sociala interaktionen. I Ikaros används en grafisk inspelningsplattform för att utföra olika handlingar. Denna inspelningsplattform körs i webbläsaren och kommunicerar tillbaka till Ikaros och bygger på att en bestämd sekvens spelas in för att sedan skickas till Ikaros som den specifika sekvensen. För att utföra interaktionen har vi i detektor-programmet implementerat funktioner så att en (<http://>)request skickas till inspelningsplattformen som Ikaros sedan utför. Denna lösning medför dock att Ikaros-systemet som Epi är uppbyggd av inte har någon större vetskap om att den ser ett ansikte. Ett problem som tillkom med denna lösning är att Epi inte kan initiera två på varandra följande rörelser om Epi registrerar två mäniskor i rad. Detta problem medför att Epi behöver registrera ett annat objekt, i

denna fallet en Epi-robot, mellan roboten kan hälsa på den andra människan. Epi kan alltså inte hälsa på flera männskor samtidigt eller i rad, utan behöver se ett annat stimuli för att utföra en ny handling.

4.2 Studie 2

Resultatet visar att våra modeller är långt ifrån optimerade. Varför modellerna hade så hög jämförelseålder skulle bland annat kunna bero på den låga upplösningen på bilderna i datasetet. Vi använde oss av ett dataset som hade lågupplösta bilder, eftersom datasetet var tillräckligt stort (över 20.000 bilder) och tog mindre plats. Detta var motiverat för enkelheten i att dela upp datasetet i flera testsets. De kunde inte vara för tekniskt krävande. Båda modeller var förträna de på ett annat dataset och vi genomförde inte en träning på UTKFace. Detta kan påverka hur pass bra modellerna var. Eftersom vi inte tränade så fungerade hela datasetet som en validering istället, det är bra att ha ett stort valideringsdataset. Problemet med att vi inte tränade innebär att modellerna inte hade någon tidigare erfarenhet av bilderna. Det finns vissa problem med detta, till exempel upplösningen på bilderna, ljussättningen i bilderna, ansiktsuttryck hos personerna i datasetet och så vidare. Vi genomförde inte träning av modellerna eftersom dessa var redan förträna de på ålderestimering. Baserat på artiklarna kopplade till modellerna så verkade dessa prestera ganska bra. Modellerna gav också bra resultat när vi själva testade modellerna. Vi hade kunde testat fler bilder och se till att de bilder vi testade hade ett bättre åldersspann, istället för att ta slumpröviga bilder. Vad som var intressant var den låga distributionen i den estimerade åldern för modellerna, de gissade förväntningsvärt ofta på 20-40 år. Detta kan bero på den låga bildupplösningen på en del bilder vilket kan göra det svårt för modellerna att upptäcka rynkor och andra ålderstecken från ansiktet.

När det kommer till testdeltagarna och de resultaten vi fick från dessa så var det överraskande hur duktiga testdeltagarna var på att ålderestimeringa. Människorna fick ett förväntningsvärt bra snittvärde och lyckades bättre än vad vi trodde från början. När vi testade oss själva fick vi intrycket av att vi var längre ifrån den sanna åldern. Detta var heller ingenting som vi kontrollerade och hade vi gjort det, så hade resultatet inte varit lika förväntningsvärt. Vi fann det dock intressant och roligt att männskor var så pass duktiga som de var. En del testdeltagare uttryckte att de använde sig av en eller flera referensansikten, dvs ansikten på personer man vet åldern på exempelvis en familjemedlem, som man sedan jämför med för komma fram till en specifik ålder. Testdeltagarna uttryckte också man tittade bland annat på tänder, rynkor och hår för att ålderestimeringa. En del testdeltagare upplevde svårigheter att ålderestimeringa på vissa bilder då dessa var väldigt suddiga och lågupplösta vilket kan ha påverkat resultatet.

4.2.1 Avgränsningar

Vi valde att genomföra testet utan Epi på grund av tidsbeschränkningen, men även för att vi ändrade fokus i den andra

delkursen. Tidigare låg fokus på ansiktsigenkänning och att förbättra denna modellen. Medan i den andra delkursen så låg fokus på att jämföra ålderestimeringsmodeller. Detta gjorde att vi inte arbeta med Epi utan skapade ett experiment som effektivast kunde jämföra olika modeller. Vi valde att begränsa oss till modeller som är snabba eftersom att ålderestimeringa kan vara krävande, och för att vid en eventuell implementation i Epi ska kunna köras utan större problem. De valda modellerna i Studie 2 är därför relativt enkla, vilket kan förklara modellernas mindre bra resultat. För att få ett bättre resultat kunde andra mer komplexa modeller användas.

Bilderna från UTKFace var relativt lågupplösta och kan ha påverkat resultatet. Andra datasetet med högre bildkvalitet kan användas istället vilket kan ha förbättrat resultatet för modellerna.

Vi lyckades ha 34 testdeltagare som tittade totalt på 1051 bilder. Vi var väldigt nöjda med detta, men en avgränsning var att hitta deltagare. Eftersom de flesta vi känner och studerar med är jämnåriga, så hade vi många yngre männskor. Detta kan ha gett ett mindre tillförlitligt resultat för alla männskor. Det hade varit bättre om vi hade sett till att åldersgrupperingen i testdeltagarna var bättre spridd. Vi kunde också genomfört experimentet på flera deltagare, för att förbättra resultatet.

4.3 Förslag för fortsatt arbete

För att vidareutveckla studie 1 diskuterade vi olika idéer, bland annat en utökning av objektigenkänningen. Förutom mänskliga ansikten och Epirobotar kan objektigenkänningen utökas med andra föremål som exempelvis mobiltelefon, telefon och mugg. Utökningen av objekt kommer göra så att Epi får större kunskap kring sin omgivning och möjliggör exempelvis att en rörlig Epi kan upptäcka hinder eller farliga föremål.

En annan tanke som diskuteras är att utveckla och använda två olika parallella modeller. En modell för ansiktsigenkänning som kan känna igen olika personer och en annan modell för objektigenkänning för att identifiera specifika objekt.

Gällande Epirobotens sociala interaktion, det vill säga hur Epi ska reagera när den ser något, kan även denna utvecklas ytterligare när objektigenkänningen har utökats med fler objekt och ansikten.

En annan punkt som diskuteras är att utveckla "data augmentation". Detta innebär att man utvecklar ett program som med exempelvis endast en originalbild skapar många olika snarlika bilder/kopior genom att förvränga, ändra ljusstyrka, spegla, rotera originalbilden med mera. Genom att utveckla detta kan maskininlärningsmodellen som användes i Studie 1 förbättras även med ett mindre antal träningsbilder.

När det gäller ålderestimering kan man utveckla projektet vidare med använda andra mer komplexa maskininlärningsmodeller eller att ha ett dataset med bättre upplösning på bilderna, detta skulle troligen ge bättre resultat. Detsamma

gäller testdeltagarna, genom att ha en bättre fördelad åldersgruppering kan man undersöka huruvida ålder är en faktor för åldersestimering. Sist skulle man kunna implementera modellerna på Epi och genomföra testet med Epi.

En annan tanke som diskuteras är att testa flera modeller, eller att testa på olika datasets. Om man testar olika datasets per modell, kan man också få resultat för vilka dataset som är bättre och sämre. Genom att undersöka olika datasets, kan man kontrollera vilken modell och vilket dataset som har effektivast åldersestimering och jämföra datan med varandra.

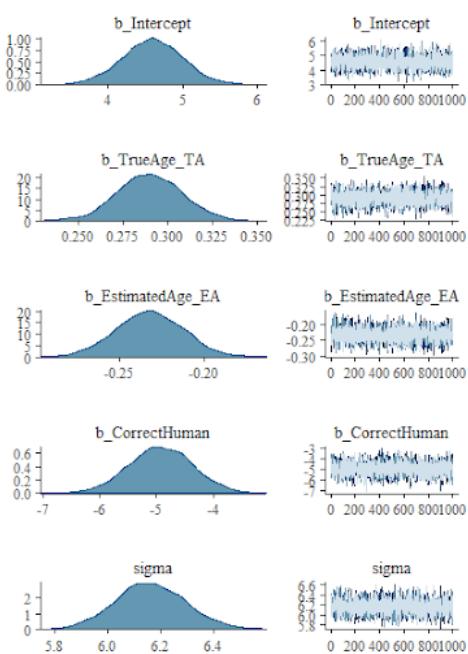
Stort tack till:

Vi vill tacka Birger Johansson som var vår handledare under båda delkurserna och gav oss väldigt givande råd under kursernas gång. Vi vill även tacka Jacob Persson och Adla Jebara som var med i den tidigare delkursen och var med under Studie 1 för det hårda arbetet och goda samarbetet. Sist vill vi också tacka våra testdeltagare som hjälpte oss utföra vår undersökning i Studie 2.

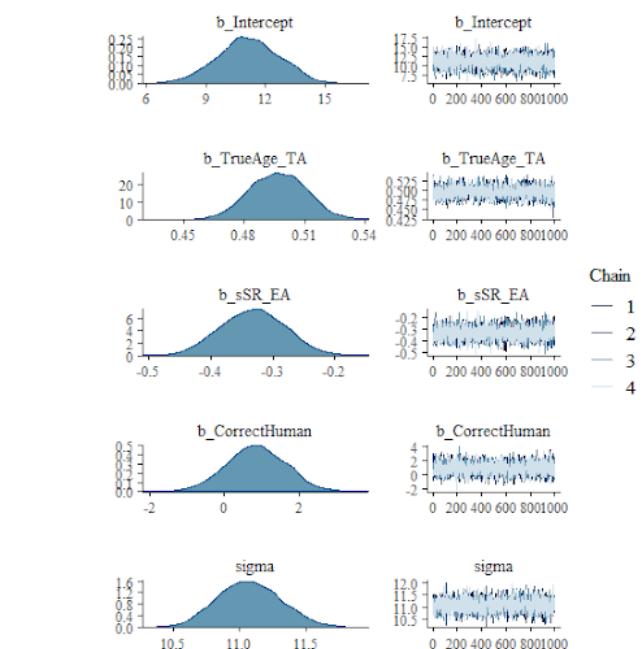
Referenser

- Angulu, R., Tapamo, J. R. & Adewumi, A. O. (2018). Age estimation via face images: a survey. *EURASIP Journal on Image and Video Processing*, 2018(1), 1–35.
- Balkenius, C., Morén, J. & Johansson, B. (2007). System-level cognitive modeling with ikaros. *LUCS: Lund University Cognitive Studies*, 133.
- Chernov, P. (2019). *Age and gender estimation. open-source projects overview. simple project from scratch.* Hämtad från <https://medium.com/@pavelchernov/age-and-gender-estimation-open-source-projects-overview-simple-project-from-scratch-69581831297e>
- Fox, J. & Gambino, A. (2021). Relationship development with humanoid social robots: Applying interpersonal theories to human–robot interaction. *Cyberpsychology, Behavior and Social Networking*, 24(5). Hämtad från <https://doi.org/10.1089/cyber.2020.0181>
- Goodrich, M. A. & Schultz, A. C. (2007). Human–robot interaction: A survey. *Foundations and Trends in Human–Computer Interaction*, 1(3), 203–275. Hämtad från <https://doi.org/10.1561/1100000005>
- Ikarosproject.org. (u. å.). *About ikaros.* Hämtad från <http://www.ikaros-project.org/about/>
- Keras. (2020). *Transfer learning and fine-tuning.* Hämtad från https://keras.io/guides/transfer_learning/
- Keras. (2022). *Resnet and resnetv2.* Hämtad från <https://keras.io/api/applications/resnet/#resnet50-function>
- KnowYourData. (2022). *celeb-a.* Hämtad från https://knowyourdata-tfds.withgoogle.com/#tab=STATS&dataset=celeb_a
- MathWorks. (2022). *Resnet50.* Hämtad från <https://se.mathworks.com/help/deeplearning/ref/resnet50.html>
- Michael Gazzaniga, G. R. M., Richard B. Ivry. (2019). Cognitive neuroscience: The biology of the mind.
- OpenCV. (2022). *Getting started with videos.* Hämtad från https://docs.opencv.org/4.x/dd/d43/tutorial_py_video_display.html
- Simion, F. & Giorgio, E. D. (2015). Face perception and processing in early infancy: inborn predispositions and developmental changes. *Frontiers in Psychology* 6, 6. Hämtad från <https://DOI:10.3389/fpsyg.2015.00969>
- Song, Y. & Zhang, Z. (u. å.). *Utkface.* Hämtad från <https://www.kaggle.com/datasets/jangedoo/utkface-new>
- Tensorflow. (2022). *Celebfaces attributes dataset.* Hämtad från https://www.tensorflow.org/datasets/catalog/celeb_a
- Xu, Y. & Vaziri-Pashkam, M. (2021). Limits to visual representational correspondence between convolutional neural networks and the human brain. *Nature Communications*, 12(1). Hämtad från <https://DOI:10.1038/s41467-021-22244-7>
- Yang, T.-Y., Huang, Y.-H., Lin, Y.-Y., Hsiu, P.-C. & Chuang, Y.-Y. (2018, 07). Ssr-net: A compact soft stagewise regression network for age estimation. I (s. 1078–1084). doi: 10.24963/ijcai.2018/150
- Zagoruyko, S. & Komodakis, N. (2016). *Wide residual networks.* arXiv. Hämtad från <https://arxiv.org/abs/1605.07146> doi: 10.48550/ARXIV.1605.07146
- Zhang, C. (2018). *Easy real time gender age prediction from webcam video with keras.* Hämtad från <https://www.dlogy.com/blog/easy-real-time-gender-age-prediction-from-webcam-video-with-keras/>

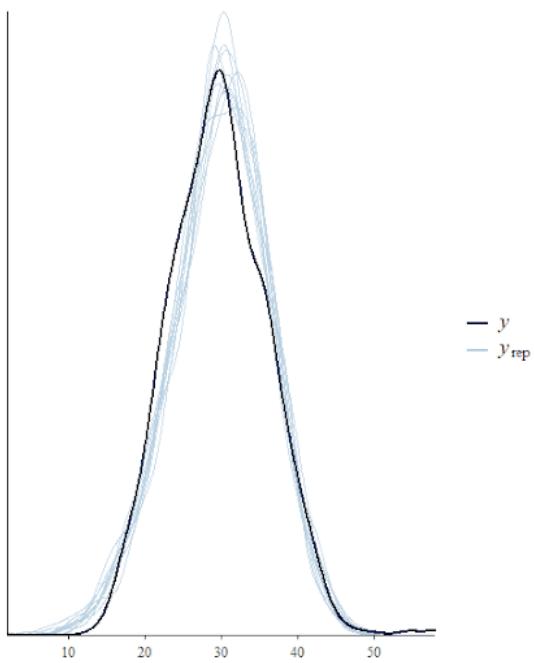
5 Apendix



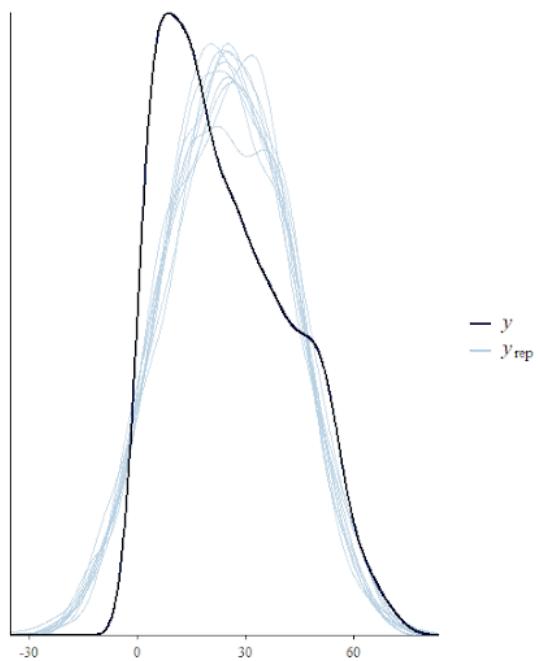
Figur 24. Människors modell där jämförelseåldern ses som funktion av sann ålder, estimerad ålder, korrekt svar



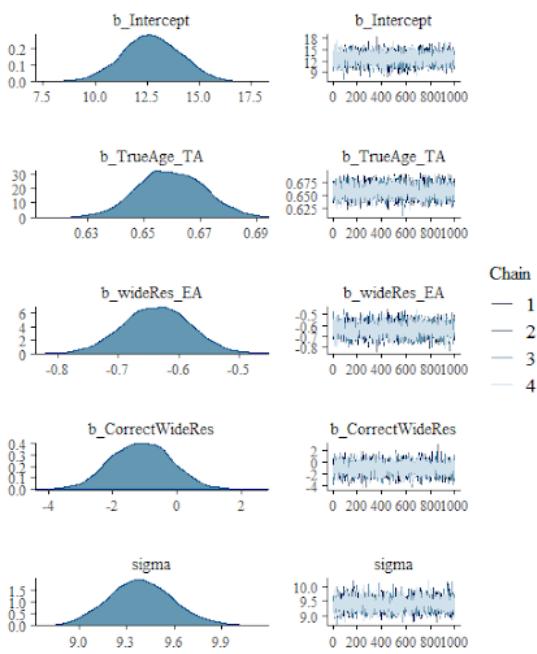
Figur 26. SSRs modell där jämförelseåldern ses som funktion av sannålder, estimerad ålder, korrekt svar



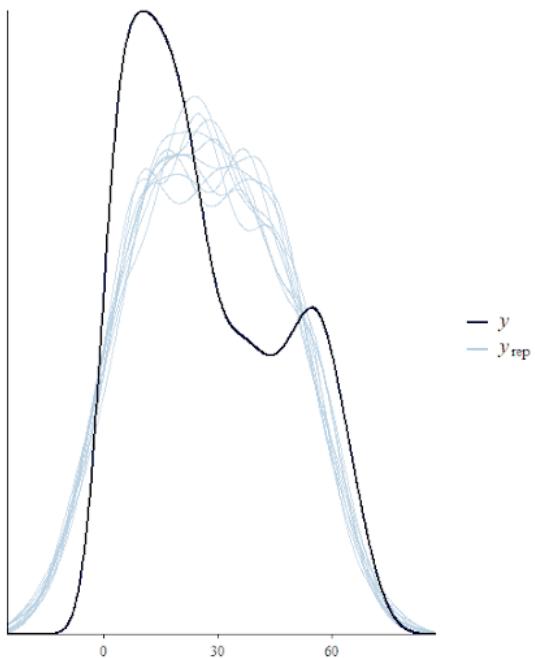
Figur 25. Posteriorisannolikhets jämförelse av människans modell



Figur 27. Posteriorisannolikhets jämförelse av SSRs modell



Figur 28. WideRes modell där jämförelseåldern ses som funktion av sannalder,estimeradålder,korreksvar



Figur 29. Posteriorisannolikets jämförelse av WideRes modell

Social Distraction in Virtual Reality

Victoria García

vi2454al-s@student.lu.se

Daniel Harrysson

da4082ha-s@student.lu.se

Emil Johansson

em3837jo-s@student.lu.se

Louise Pham

lo6823ph-s@student.lu.se

The use of virtual reality has expanded in the last few years and the use of virtual lectures, inside a interactive virtual environment, is now possible. This study aims to explore how people behave in a virtual classroom during a lecture. What will be specifically studied is how social distractions can affect the people taking a virtual lecture, and whether these types of lectures influence knowledge retention and behaviour. The experimental platform was developed from a pre-exisiting VR classroom. The participants took part of a yoga lecture, individually, together with agents that are animated to move like real people that have different behaviours. The different behaviours of the agents were either erratic, curious or idle. The result of this study was that there was no difference in knowledge retention regardless of if the agents were acting erratically or focusing on the lecture. However, when the agents were erratic, the participant looked around more as well as were more prone to interact with the implemented features of screenshotting and laser pointing.

1 Introduction

The use of virtual reality (VR) and its popularity has been on the rise for the last couple of years (Alsop, 2022). Mainly due to its interactive and immersive use in the video game industry (Clement, 2023). However, VR's integration into other media has now become more prevalent and stretched beyond video games. One reason for this is the technology company Meta's large scale investment into development of the technology (Rapoza, 2022). More and more ways of interacting with one another in a fully immersive virtual space have been released, like VR-chat and Meta's Horizon Worlds (*VRChat*, n.d.) (*Meta Horizon Workrooms*, n.d.), leading to these interactions not just being used for entertainment purposes but ideas of broader uses such as virtual meeting and education developed to be held in a virtual space (*Meta Horizon Workrooms*, n.d.). With VR's increasing popularity, more research has come out exploring how a fully immersive virtual space compares to similar interactions in real life and if VR is a valid alternative to its real world counterparts. Several researchers, such as Lund's Education Technology Group (Tärning, Ternblad, Haake, Gulz, & Nirme, 2022), have started exploring the medium's capabilities as an educational tool. Despite this, as far as we are aware is that little focus has been put towards understanding social interaction

in VR and if using such a medium can be more distracting than conventional educational methods. Is it a viable option to have social interactivity alternatives in VR or does it elicit too much distraction? In this study we aim, through a virtual classroom with different behavioural conditions, to see if social distraction can indeed be seen in VR. More specifically if unfocused behaviour from other perceived individuals in a virtual space elicits a similar type of behaviour in one self.

Our main research questions are:

- Is an individual affected by a social distraction in a VR environment?
- What effect does social distraction have on an individual's ability to retain information in a VR environment?
- How does group behaviour manifest in a VR environment?

2 Background and Theory

The following sections discusses relevant previous research. Building the foundation for our experiment and the structure behind it.

2.1 Previous Research on distraction in VR

Tärning and colleagues (Tärning et al., 2022) explored VR in an educational setting to see if distractors that are usually present in today's classrooms, specifically media on screens, have a similar effect in VR. They found that children aged fourteen and fifteen, don't engage with virtual screens in VR, contrary to previous research made in a real life setting. Although their findings are inconclusive, they argue that this does not necessarily mean accepting the null hypothesis, i.e such distractions in VR do not have the same effect. The authors' explanation as to why this could be is that the task stimuli and the distractor stimuli do not accurately mirror the real-life classroom setting it is trying to emulate. The effectiveness of distractors themselves (the screens) might have less to do with the content shown and more so to do with competing high-cognitive goals and emotions "a friend is receiving (task-irrelevant) information that might be both useful and fun for you to take part of." From comments of the

participants, it can be interpreted that the social presence in the experiment was lacking.

2.2 Social Influence and Group Behaviour

When we as people interact we tend to change or behaviour in slight ways (Gass 2015). We intentionally and unintentionally emit influence over others which changes the interaction and depending on how it is done it can be categorised in different ways. For the purposes of this study social influence is of interest due to this study intending for the participant to be unaware of any sort of influence happening. Specifically, it can be said that "Social influence involves intentional and unintentional efforts to change another person's beliefs, attitudes, or behaviour." (Gass, 2015). Due to the similarities of social influence to other forms of psychological influence a distinction needs to be made for clarification of its use within this research context. Whilst certain forms of influence like persuasion generally is made with the overt intention of changing a target as well as that said target needs to some degree be aware of this (Gass, 2015). Social Influence has more to do with peripheral processing of one's surroundings therefore the target might not be aware of the change in behaviour or attitude (Gass, 2015). An example of social influence is group behaviour, conforming to the norms of the individuals around one self (Gass, 2015). To our knowledge no extensive research has been done on the topic of how it pertains to VR.

2.3 Presence in Virtual Reality

By changing different variables within a virtual space one's experience can be perceived as more realistic; this has been defined as presence by VR researchers. According to Lee (2016) there exists three versions of presence:

Physical Presence: A psychological state of experiencing a virtual object as if it were real.

Social Presence: A psychological state of experiencing virtual individuals (agents or avatars) as if they were real. This is mainly going to be discussed in this study.

Self Presence: A psychological state of acquiring self ownership over one's virtual self as if it were one's real self (Lee, 2006).

Some of the variables for increasing perceived presence are:

Egocentric reference frame (ERF) (Klatzky, 1998) refers to the mental model of one's surroundings which is structured through a first person point of view, and is then centered around the body's position within space, which goes together with the body's inner spatial situated model.

By creating a virtual environment that takes into account one model of reality it leads to immersion a feeling of being fully present within said world. Although this does not mean it has to be hyper realistic it means speaking to how we see and refer to our position in said space being presented (Alce et al., 2022).

2.4 Social Presence and Avatars in Virtual Reality

We can also see in for example in the work of Fox (2015) that the use of computer controlled agents or human controlled avatars, produces different interaction and engagements from users. To specify, the study said that "When people perceived their virtual interactants to be human-controlled avatars, social influence was significantly stronger than when they thought they were interacting with computer-controlled agents." (Fox et al., 2015). If the environment the user inhabits is constructed in a way where the interactants they meet are perceived to be human-like or controlled by humans, it will lead to higher engagement from the users (Fox et al., 2015). The interactants themselves do not have to look like humans. However, anthropomorphism as well as generally having characteristics that makes the user perceive the interactants as human controlled, can help with engagement as well as social presence. An example of seeming human controlled can be the way the agents move. Previous research has shown that high movement realism creates an increased perceived social presence (Greenwald, Wang, Funk, & Maes, 2017).

3 Method

This section introduces the method in how the test is to be conducted as well as the components included in the test. This acts as a foundation for how all tests are to be carried out. To address the research questions, test will be conducted and conclusions will be drawn from observations of the test and interviews.

3.1 Participants

The focus population for this study is students at Lund University. Students can be assumed to have some previous experience which vital since if the participants had no previous knowledge about lectures, they might have had a harder time understanding what they are supposed to do in the VR environment or how the lecture functions as a basic concept. Participants with no or a small amount of previous experience with virtual reality are preferred. No further limitations were chosen for the participants.

Factors such as age and gender are gathered through a pre-test survey since interesting data might emerge from these data-points. However, there is no limit to how many participants can be of a specific gender or age as long as they are a student of Lund University. A total of ten participants took part in the study (60% Male). Of the ten participants, nobody (0%) had any extensive experience with yoga prior to the experiment.

3.2 Test Procedure

At the start of the test, the participant is briefed by the moderator. The purpose of the moderator and other roles is later described in section 3.7. The moderator explains to the participant what will be included in the test and any inform-

ation that the participant might need to complete the test. The moderator introduces any other people that might be in the room and explains their roles. After this, the moderator will have the participant sign the pre-test survey, and a consent form to gather useful data-points and to make sure that the participant is consenting to research being made from their behaviour.

The experiment is presented as a research of memory retention and experiencing a lecture inside a VR environment rather than a research surrounding distractions. The participant will be informed that there are other participants in the same VR environment, as well as that the lecturer is pre-recorded, and that the test will proceed when the participant has been informed of and accepted the consent form.

Once the consent form has been filled out, the test will begin. First, the participant gets to choose a color for their VR avatar. They are shown which buttons to press in order to activate a laser pointer or a screenshot, and are told that the screenshots will be able to be accessed at the end of the whole lecture. The participant will watch through the first part of the lecture, while each virtual agent will exhibit one of the five behaviours chosen by technical engineer. The technical engineer's role will be explained in later sections. The behaviour the agents display are either erratic, neutral, waiting (as in idle) and curious. These are explained more in detail in section 4.3.1.

When the participant is done with the first lecture, a memory questionnaire is presented where the participant is tested to see if they remember key parts of the lecture. After this, the next lecture will start playing and the participant once again goes through the lecture and then gets presented with another memory questionnaire. This memory questionnaire is also to see if the participant remembers key parts from the second lecture. After this, a social presence survey is presented and the participant gets time to fill this one in as well. When the last survey is filled in, the participant is done with the test. Each test is planned to have a maximum time of 30 minutes.

Finally, the participant is informed that the other participants in the VR environment were also pre-recorded and that the test is conducted in order to research distractions inside a VR environment. They are also informed that no screenshots were actually taken.

A pilot test was also performed to see how the lecture worked without the presence of agents and only a lecture and a lecturer, but following the above procedure.

3.3 Test Environment and Equipment

The tests were conducted at the Cognition & Philosophy VR Lab at Lund University. The equipment used for the test was a VR-headset, Varjo VR 3, and a computer with enough performance to run Unity, a Steam Valve controller and another computer or mobile device to fill in the surveys during the tests.

3.4 Platform Description

The platform used for this experiment consists of a classroom, agents, a lecture and a lecturer. The lecturer, the animations of the lecturer, audio-tracks, classroom and presentation was already included at the start of this project, as this platform has been used for other experiments. The laser pointer as well as the illusion of taking a screenshot was implemented during this project.

3.4.1 Classroom

The classroom consists of a screen in the middle and two different readable posters on the side walls. One of the posters is of a map of Brazil, and the other one is an image of the periodic table. The posters are there for immersion purposes, as this resembles how a classroom could look.

3.4.2 Virtual Agents

The two animated agents consist of a VR headset as a representation of a head, as well as a VR controller as a representation of one hand. Each set of controller headset has its own unique color. These are placed next to each side of the participant, one at the beginning of the experiment and the other one will appear at the start of the lecture. The animation is based on recordings of headset and controller movement made while listening to the lecture in the VR environment, and was recorded using the software Brekel.

3.4.3 Lecture and Lecturer

The lecturer holds a presentation of the different types of Yoga that exist in the world and presents information surrounding each type. This presentation is split up in to two parts of four minutes each, totalling to eight minutes. The lecturer directs his gaze towards both the participant and the agents.

As previously mentioned, both the presentation and lecturer were already implemented at the start of this project, this includes the animations for the body and eye contact of the lecturer, the voice of the lecturer, as well as the contents of the lecture which is a text and bullet point presentation with no pictures.

3.4.4 Laser Pointer

A laser pointer was implemented during the course of this project. Naturally, the laser pointer is red. The laser pointer is activated manually by the user, by using the trigger button on the controller to turn it on and off.

Through Unity it was possible to detect the users button presses and count how many times the laser was turned on and off. This was logged in order to be analyzed.

3.4.5 Screenshot

In order to simulate a screenshot being taken, a point-light-object with a halo effect is used for the agents. This light-object is scripted so as to allow the light to only flicker for a short time in order to simulate a camera taking a picture with flash. The light-objects are attached to the headset of the agents and do not make a sound. The technical engineer is also responsible for the agents use of this.

The users flash is a point-light-object that is also attached to the users headset, however the radius of this light is expanded and much stronger than the agents. When used, it creates a flash in the whole field of view of the user. It also makes a short camera shutter sound. However, it does not actually take a screenshot and only simulates it. The user can activate it by pressing the B button on the controller.

The number of screenshots taken were also logged for further analysis.

3.5 Surveys and Memory Questionnaires

After each part of the lecture the participant was handed a memory questionnaire divided into parts A and B with eleven questions each. These questionnaires, seen in appendix A and B respectively, measure how well the participant retained the information presented in the lecture, with the help of different questions that test their knowledge of the subject. The point of this is to see if the increased erratic behaviour in the virtual agents distracts the participants enough for the participant to not remember key points in the lecture. These questions are multi-choice questions, short written text questions and likert scale choices.

Since the lecture is in the same order no matter if the participant starts with the erratic virtual agents, the curious ones or the neutral ones, the surveys will also be in the same order. Before leaving, the participants also filled out a social presence, seen in appendix C, survey to see if the participant thought that the virtual agents were actual people or not. The different questions were either answered with a likert scale or free form text answers.

In order to avoid the participant answering that the virtual agents were not real out of fear of being fooled, a manipulation check was also included in the survey phrased in such a manner as to lead answers. This was presented as a free form text question at the end of the survey.

"Since this is a new form of experiencing a lecture some people feel that they are not actually interacting with other people. How did you experience the presence of the others in the task?"

The survey as a whole is constructed this way, having the questions framed in a neutral way so the participants won't answer a specific way out of fear of being wrong.

3.6 Interview

After the test was finished a debriefing was performed in order to tell the participants the actual purpose of the experiment, as well as asking them questions about their experience

of the experiment as a whole.

3.7 Roles

When conducting a test, it is very important to make sure that everything works when it is time for testing. If problems occur during a test, it could affect results and waste precious time and resources. To combat this, roles have been introduced to make sure that everyone involved in conducting the test is aware of what they are supposed to do.

3.7.1 Moderator

The moderator's purpose is to welcome the participants, making sure they are aware of what they are going to be testing and that the participant is completing the surveys and questionnaires in the previously decided way. It is very important the tests are conducted the same way each time, to make sure that each participant has the same conditions and that the results are not affected by some participants conducting the experiment in different ways. The moderator is also responsible for presenting the consent form, answering any questions that the participant may have, providing support to the participant if needed and also presenting all the surveys that the participant will have to fill in.

3.7.2 Observer

The observer's duty is to study and document the participants actions during the experiment. This is done through visual observation both through the screen which shows what the participant sees, but also in real life observations. There can be multiple observers at once.

3.7.3 Technical Engineer

The technical engineer is responsible for the program running as planned during the testing. If there is a decision that recordings should be made as well, then the technical engineer will also be responsible for this.

3.8 Data collection

When planning the test cases and making a test plan, a set of data collection points was decided. These data points were as follows.

- **A:** Surveys and memory questionnaire results - By gathering results from the surveys and questionnaires presented in 3.5, there can be an indicator if the distractions caused by the surrounding agents have a significant impact on the amount of correct answers the participants fill in on the survey.
- **B:** Observing loss of focus - Recordings will be made from the perspective of the participants in the virtual reality environment. These recordings will be used to declare an area of focus, which is the area where the lecture notes are presented and the lecturer is stationed.

If the view is moved outside of this area, a "loss of focus" is declared. Both the amount of times loss of focus has been observed and the percentage of time the participant spends while having loss of focus will be measured. For example. If the lecture is 20 seconds and the participant looks outside of the area of focus for 10 seconds. The results would be 50% of the lecture was spent not looking within the area of focus. The participant also looked twice to the left of the area of focus, resulting in them having 2 counts of loss of focus.

- **C:** Recorded button presses - There are two buttons on the controller that the participant uses during the test, one is for the laser pointer as explained in 3.4.4 and the other is for the screenshot function as explained in 3.4.5. When the participant presses one of these buttons, it gets recorded and is stored in a log file that can be accessed after the test. These logs will be used to determine the amount of clicks made during a test per participant. The laser pointer is measured in the amount of times it is turned on or off, while the screenshot function is measured per amount of times it is activated.
- **D:** Observations of behaviour during tests - While the tests are conducted, observations will be made on whether the participants change their behaviour according to the agents. Examples of this could be that someone refuses to use their laser pointer because the agents uses theirs, or that someone shines their laser pointer at an agent because the agent used their laser pointer at them. These observations can also be made while watching the recordings after the tests are done.
- **E:** Comments and answers on oral questions - During the test, observers will listen for eventual comments made by the participant, these comments are then written down if any are made. After the test has been conducted a small interview is made as described in 3.6. Any answers that are of interest will also be written down by the observers.

4 Implementation of the Prototype

This section will go into further detail surrounding the different tools used in order to implement the prototype. A VR-ready computer was also used.

4.1 Hardware

The following subsection consists of the different equipment used for the experiment as well as for the recording of animations.

4.1.1 Varjo VR 3 Headset

The headset which is used for this project, both during the development of the prototype as well as the experiments themselves, is called Varjo VR 3. (*VR headset - varjo VR-3: Highest Resolution Virtual Reality*, 2022)

4.1.2 Steam Valve Index Controllers

The hand controllers used for tracking in this project, both during the development of the prototype as well as the experiments themselves, are called Steam Valve Index Controllers (*Controllers - Valve Index*, n.d.). The trigger button is programmed for the laser-pointer to turn on and off, and the B buttons is programmed to take a screenshot.

4.2 Software

The following subsection consists of the different programs and software used for the experiment as well as for the recording of animations.

4.2.1 Brekel

Brekel is an affordable tool used to capture and record the motion of VR controllers, headsets and trackers (*Brekel*, 2022). It supports all OpenVR/SteamVR based devices, such as HTC Vive and Steam's own controllers.

The tool is used to record different animation stages that are later on programmed to represent the virtual agents. As of now, all the recordings have taken place while the lecture has been running simultaneously. The reason being to make the animations as plausible and realistic as possible, and not just random movements happening at random moments.

Brekel is fairly simple and straight forward to use, which also has its disadvantages. Since the recordings are supposed to match the lectures held inside the project, it is crucial that the recordings are synchronized with the lecture, as well as the laser pointer and screenshot functionalities that the agents are intended to use. There was no way of recording button presses on the controller during the recordings. However, a solution to this was a mixture of recording the amount of button presses made, through a Unity script, as well as manually pressing buttons on the keyboard in order to activate the agents screenshots and laser pointers.

4.2.2 Unity

The development environment used in this project was Unity. It is a game engine used for real time 3D development and is a popular environment for VR development, (*Unity*, n.d.). The programming language for the scripts is C#. For the asset creation, such as the laser pointer and screenshot function, two of Unity's main functionalities, were used:

- **GameObjects** - A basic object that does not do much by itself, but rather serves as a container for components.
- **Light component** - A component used to handle and control light related functionalities in Unity.

Unlike Brekel, Unity was not nearly as simple nor straight forward to first time users. There were more things that had to be taken into account. One such thing was what application programming interface (API) the project used. The project was built using OpenVR's API. When trying

to incorporate SteamVR Unity Plugin to the project, the plugin conflicted with the headset's Varjo system, making it very hard to advance the development. The programs were switching VR views with each other and started to distort the view inside the headset as well as in Unity. The periphery was warped, and sometimes the whole view would turn upside down. Another API that several forums were advising about was OpenXR, which SteamVR did not yet support.

4.3 Design process and Development

This subsection discusses the implementations that were made, their design choices and solutions to problems that arose during the development phase of the prototype.

4.3.1 Agents and Behaviours

The agents are meant to resemble real people, so as to create social presence around the participant. The placement of these agents was chosen so as to not be in the way of the participant, but still noticeable in the peripheral view of the participant. That way it is unavoidable for the participant to not turn their head in order to detect what is going on around them.

The agents behave differently depending on what part of the lecture is currently being presented. These behaviours were made into lecture long animations, so as to not disturb the social presence by being noticeably split into parts. The first part of the lecture has one set of animations, and the second another set.

In total, five separate animations were recorded and applied to the two agents;

- **E:** Erratic - The animation is excessively moving their head and controller, and is constantly using the interactive functions while erratically shifting their focus from the participant, the other agent, the lecturer and the classroom.
- **N:** Neutral - The animation does not use any of the implemented tools. Instead, they are focused on the lecture and not paying any attention to the other entities in the room.
- **C1:** Curious 1 - The animation is mostly observing the classroom and focuses on the lecture. The functions are rarely used, but when they are it is made to look as if the agent is trying out the functionalities of the controller.
- **C2:** Curious 2 - The animation is acting in a reactive manner, and hence uses the implemented functions more often, not only in regards to the actions and movements of the other animation it is being paired with but also in regards to the content on the lecture slides.
- **W:** Waiting - The animation looked idly around in the classroom without trying out the implemented functions.

Each recording is around four minutes long in order to match the length of one lecture. Depending on the different experiment condition (whether it was a neutral start or an erratic start) the agents could have two different animation patterns. For the sake of efficiency in the development process animation **E** was used for both agents before and in between lectures. However, the agents' animations started with different parts of the animations in order to avoid mimicking each other frame by frame.

When choosing the movements for the recordings we included typical human like movements such as sighing motions, straightening of backs, head scratching and waving.

Importing these animations with their respective meshes to the project (headset and VR-controller) and programming the functions onto them was harder than expected. We encountered a problem where the animation of the controllers would stop working for some animations but work for others when trying to only use one set of meshes. After a long time trouble shooting and debugging, we realised that the name of the controllers had different ID:s in almost all animations. For reasons unknown, the controller ID changed each time the controller went into an inactive mode where it tries to save battery.

Another problem we encountered was that the agents kept changing positions in the classroom based on what animation was currently being played. This happened even though we had a parent object surrounding the child object in Unity, with the animation controller attached to the parent object. Instead of being fixed to our chosen starting points for each agent, the child object still behaved as though it was an anchor object and moved based on its position that it was recorded in, in the real world. In order to test whether or not our solutions actually worked, we had to re-implement the 3D model with the new animations each time.

We suspect that Brekel is behind both of these problems. Our solution was to perform all recordings of the animations in one session, *before* the controller could go inactive, and in the same positions in order to avoid the cuts and jumps in the animation as much as possible. The results turned out much better, with the controller problem completely solved and the problem with repositioning and cut animations kept to a minimum.

4.3.2 Animations

As previously mentioned, both agents, left and right, visually consisted of a headset and a controller. Three animations were applied to each agent. The left agent had the animation patterns **W-E-W-N** and **W-N-W-E**, and the right agent had **W-C2-W-C1** and **W-C1-W-C2**. The four different animation stages corresponds to the lecture's different stages: before first lecture → first lecture → between lectures → second lecture. For each lecture, the agents appeared in different colours. The left agent appeared as blue during the first lecture and green during the second lecture, while the right agent appeared as pink during the first and

yellow during the second. This was made in order to distinguish the agents from each other, but also as an attempt to convince the participant that they encountered four different people during the whole test session. The participant was also given the choice to choose from one of the four different colours in the beginning of the session to further convince them of being in a virtual environment with four other different participants, whom have also chosen a color.

4.3.3 Laser Pointer Iterations

The different versions of the laser pointer were completely different implementations and were not iterations of each other.

The first pointer consisted simply of a light component attached to a GameObject. This laser pointer did not look like a laser itself, but rather like a dot on the wall. It was noticeable and could have been used for the experiment, however it was not possible to program it, as it was a light-component and not a light-object which made it impossible to create a working script for it in order to turn on or off on command. This laser pointer was also not as disruptive as the following iterations, as it was not a stream of light that interfered with the field of view of the participant.

The second iteration of the laser pointer consisted of a 3D-object of a long cylinder with a small circumference. The idea was to create a material for this 3D-object that was gradually transparent the further away the laser pointer got. However, this proved to be more challenging than anticipated and did not look like a light source.

For the final iteration, a line-renderer was used by gradually making the pointer transparent until the furthest point of the laser pointer. It had all the previously mentioned qualities we wanted it to have, such as looking like a light source and gradual decrease in strength. The reason for the laser pointer being red is simply because most laser pointers used today are that same color. This laser pointer can be seen in figure 1.

In order to adjust the laser pointer to the controller in the VR environment, one person had to sit in the VR environment while giving verbal feedback to another person in Unity, for whether to move the light-object in a certain direction in order for the laser-pointer to look like it is coming from the participants hand controller. As previously mentioned, this 3D model with its animations had to be re-imported several times, which meant that a lot of time was used doing the same placement calibration for the laser pointer.

4.3.4 Screenshot Iterations

Although the agents' screenshot animations were somewhat quickly implemented, it was not possible to use the same approach for the users screenshot animation. The agents screenshot light-objects used a halo effect, but this effect can only be seen if placed far enough from the light-object. Which meant that when placed on the users virtual headset, it was not visible for the user.



Figure 1. The final laser pointer iteration pointing at the presentation, both from the users and the agents' perspective. A floating headset, representing one of the agents can also be seen, as well as the lecturer.

Originally what was wanted was a white flashing light on the users actual field of view, as a white filter rather than a point-light. However, due to how the project was originally created, it was not possible to add this in an acceptable manner.

What was finally settled on, through a lot of trial and error, was setting the strength of the light on a very high setting, as well as making the radius of the light large. That way it looked like a white filer over the field of view, as long as the user was looking somewhat forward. Both the users and the avatars screenshot function can be seen in figure 2

Luckily, the script for the flicker, which randomises the intensity of the light-object, could be re-used for the users' light flicker.

4.4 Controllers and Application Flow

Each agent has their own animation controllers with their respective scripts that are controlled by the technical engineer. The animation controllers contains the animations mentioned in the previous section. The script enables the technical engineer to change the animation state by pressing keys on the keyboard.

When the application is started by pressing play in unity, the right agent will already be spawned and in its waiting animation. Pressing "1" will spawn the left agent in its waiting animation, and after some seconds the teacher appears to start the first lecture. This decision was made to mimic the possible technical difficulties, latency and different session start up times that tend to occur in collaborative sessions.

The keys "E" and "R" are then pressed as soon as the teacher spawns to start animation **E - erratic** and animation **C2 - curious 2** respectively, or the keys "N" and "M" to



Figure 2. A faint light of the avatars screenshot function can be seen. The users perspective is slightly lighter from using their own screenshot. A floating headset, representing one of the avatars can also be seen, as well as the "teacher".

start animation **N - Neutral** and animation **C1 - curious**. 1. When the lecture is over and shortly before the test user puts on the VR headset again, the key "W" is pressed to remove both agents from the application to reappear again in their waiting animations after some time and in different colours. When the key "2" is pressed, the second lecture starts. Again, either "E" and "R" or "N" and "M" is pressed to start the chosen animations.

While recording with Brekel we had the lecture running simultaneously to make the animations look as natural as possible. During the recording, we were therefore able to use the different functions. Therefore, we originally had the desire to automate the laser pointer and screenshot functions. However, as mentioned in section 4.2.1 the button clicks were not recorded. We tried using the finger tracking motion in order to determine when the different functions were being used, but these did not show up in Unity. As an alternative, we considered randomising the interval and amount of when the functions should be used, but after some discussion back and forth we decided that making the behaviour seem as natural as possible was more important. By randomising the usage of the functions we believed that the agents would exhibit non-human like behaviour, especially in conjunction with our highly intentional animation states. Therefore, we instead opted to manually control when the agents would use the functions. "A" and "S" controlled the screenshot and laser pointer for the left agent, and "K" and "L" was used for the right agent.

The keys "B", "Y", "P" and "G" are can be pressed to change the participant's VR controller in the respective colours blue, yellow, pink and green.

To ease any possible confusion that may arise from all the different keys, a detailed manuscript of when to start the animations **E, N, C1, C2, W** at each stage was written.

5 Results

This section will present and evaluate the test results gathered from quantitative, as well as qualitative, data.

5.1 Analysis

We used a paired sample t-test to investigate any significant differences between the two conditions. This was done utilising r studios Version 2022.12.0+353 (TeamPosit, 2022) assuming a cof inter of 95% and alpha of 0.05.

5.2 Quantitative evaluation

The paired samples t-test used to compare any differences in result in the questionnaire between the condition showed that the erratic condition ($M= 3.8$, $SD= 1.4$), when the erratic agent animation was present, and the neutral condition ($M= 3.8$, $SD= 2.39$), when the neutral agent animation was present, had no significant difference between them $t = 0.81$, $p>0.05$. However, it can be seen that certain participants scored higher than average in the neutral condition. Specifically, one participant scored 8 out of 11 points in the neutral condition, compared to the mean of 3.8 points out of 11. Results also indicated that all participants performed worse on memory questionnaire B than A, with memory questionnaire A having a mean score of 4.36 out of 11 compared to B's 3.55 out of 11.

When looking at the percentage of the total lecture time the participants were not looking at the main focus of the lecture, as described in 3.8 part B, the paired samples t-test revealed that the erratic condition ($M= 0.0839$, $SD=8.25$) and the neutral condition ($M= 0.0419$, $SD= 4.83$) had a significant difference between them $t = 2.8$, $p<0.05$. The erratic condition showed to have a much higher percentage of time spend not looking at the main focus of the lecture, and had the participant instead looking at other parts of the classroom or at the other avatars.

The amount of times the participant looked away from the main focus of the lecture as described in 3.8 part B was also significant $t = 3.09$, $p<0.05$, between the erratic condition ($M= 6.9$, $SD= 5.95$) and the neutral condition ($M= 4.7$, $SD= 6.00$). The results show that the participants had a loss of focus in the lecture a lot more times during the erratic condition compared to the neutral condition.

The laser pointer had a mean usage of 15.8 times per test, a median of 8.5 uses and a standard deviation of 24.96, which shows that the amount of uses varied a lot between different participants. One participant had an anomaly where they clicked the laser pointer button 84 times, which could explain the high standard deviation. The screenshot function had a mean of 4.8 uses per test and a standard deviation of 5.07.

5.3 Qualitative evaluation

In the social presence survey and the post-test interview, the participants expressed the following points.

5.3.1 Virtual Agents

The majority of participants (80%) expressed that the virtual agents were perceived as human until they were told that they were pre-recorded. They also expressed the feeling that they did not feel isolated in the virtual environment and that the presence of the "others" was felt at all times. One participant told us the agents had a robotic feel to them. These participants also tried to interact repeatedly with the agents during the lecture and by getting no response correctly assumed that they were in fact not other participants.

5.3.2 The Lecture

The virtual environment was according to the participant not seen as equivalent to a regular in person lecture. Neither the environment or the lecture itself. This seems to be because of the medium of VR it self. Such as the tools being used during the lecture and the audio quality of the lecturer speaking. The headset used caused some discomfort for one participant that according to them "took them out of the experience of being present in a lecture hall". Another participant expressed while reflecting on their experience that there are elements that causes engagement such as the movement of the lecturer being life like to how a lecture would move in real life.

5.3.3 Observed behaviour

While observing both in real time but also the recorded material, a few behaviours were noted. A general behaviour that was spotted was that if the avatar shone their laser pointer at the participant, there was a big chance that they would shine theirs back, resulting in loss of focus for the participant.

Another behaviour spotted was on the contrary side of the spectrum. Some participant told outright and showed that they refused to use their screenshot or laser pointer due to the avatars being "annoying" or "disturbing".

6 Discussion

6.1 Social Distractions

When looking at the results, we can see that the participants seemed to be more distracted during the erratic state rather than the neutral state. This would make sense, since one would think that a lot of noise around the participant would make them more distracted. This is also supported by both the percentage of time the participant looked away from the focus point of the lecture, the loss of focus, but also by the amount of times the participant looked away, the counts of loss of focus. These results show that the erratic state has a significant effect on the results compared to the neutral state as shown in 5.1. One would think that the results of the surveys would also be affected, but the analysis indicated that the different conditions does not affect how much one retains from the lecture. It is very interesting that the results show up this way. In one way, they show that people pay less at-

tention to the lecture when surrounded by erratic behaviour, but at the same time the results on the information gathered is the same even if the participant is more distracted. This could indicate that the medium of VR causes us to retain information in a different manner than how one would be affected in a real world setting, similar to the results that Tärning and colleagues (2022) discussed in their study that there could be something within VR that makes our reaction and processing of what would usually be in real life would be seen as distracting example media screens in a classroom. However due to the small sample size this is hard to answer definitively. Further investigation with a comparison of a real life equivalent is needed in order to see if such an over all effect is present in VR. There was one participant who only partook in the VR lecture without any distractions, and this participant scored higher than the others. However a larger sample size would be needed for any conclusions to be made. It should also be noted that the spread of participants scores in the neutral condition is higher than the erratic with certain individual scoring several points higher than the average, although this does not infer any causal relationship.

When looking at observed behaviour, we could see that people who used the laser pointer a lot were also prone to having a loss of focus both in percentage of time but also in counts of loss of focus. This showed especially through a few participants who used their laser pointer far more than the rest, and thus also had a high loss of focus percentage and counts of loss of focus. There was one exception where a participant had a high number of laser pointer usages, a low loss of focus count and percentage while having a low score in the survey. This could indicate that the laser pointer itself might be a distraction, but there is not enough evidence to come to a complete conclusion in this case.

Another observed behaviour was that some participants would ignore the avatars and strictly focus on the lecture. An expected outcome of this would be a higher score, but this was not seen in the results either as those strictly focusing on the lecture scoring the same as the other participants on average. This could be a result of the questions in the memory questionnaire being too hard, lowering the average score to a score where participants guessing have an equal chance to those who listened to the lecture. This would have to be tested with a larger control group that completed the lecture without distraction, something we did not have the resources to complete in this test.

Apart from the knowledge retention, the result might indicate that social influence in VR could affect one similarly to how it works in real life. Indicated by the behaviour by participants during the experiment using the laser pointer and screenshot functions more in one of the conditions and through this looking away from the point of focus during the lecture. There seems to be no noticeable difference in the behavioural change displayed by the participants and how Gass (2015) describes how social influence from the ones around you unintentionally changes attitudes and behaviours. Indi-

viduals around oneself displays a certain type of behaviour and therefore the individual themselves are more likely to mirror said behaviour to fit in the group. Although as seen here and in previous literature one type of behaviour can also lead to reactionary behaviour of nonconformity, such as in the individual perceiving the agents as distracting and therefore not wanting to participate. Changes in behaviour in regards to agent behaviour can also be connected to the work of Fox (2015) as well as Greenwald, Wang, Funk, & Maes (2017). Which says that perceiving agents as humans will lead to higher engagement. If we like Fox says engage with Agents in games and collaborative exercises as if they were human when we perceive them as such. One could imagine that the psychological process present during real life interaction could be present in VR. In other words It could be that the process of peripherally processing information to assimilate ones behaviour could work the same in a virtual space, as long as they are perceived as human. As said before this is simply speculation in regards to previous research and with the small sample size as well as comments mentioned in 5.3.2 indicating that the differences in the medium of VR and real life lectures all of this will require further investigation.

6.2 Virtual Agents

The participants view of the agents is in line with previous research indicating that perceiving virtual agents as human is likely as long as they are presented as such and meets a threshold of credibility (Fox et al., 2015). This could be attributed to the movements of the agents which as a result of the motion capture displaying all the same range of motions that the participant themselves can perform as described previously by Greenwald and Colleagues (2017). As such they do not need to look like realistic human as in this study but simple having an anthropomorphic model could be argued to be enough. The models for the agents in this study follow the suggestions of Fox (2015). Utilising human characteristics as in the shape of the model and possessing the same degree of freedoms as the participant, making it more believable to be human (Fox, 2015). As Argued by Greenwald and colleagues (2017) the likely hood of them being perceived as human is amplified on top of the models features by the motion agent being that of how humans move. Smooth and acting according to the same levels of freedom as the participant.

6.3 Implementation

There are several factors in the implementation, specifically surrounding the trade-off between making the agents as life-like, action wise, as possible, versus making them as easy to use as possible. As can be seen throughout the whole of section 4.3, there were many keys that the technical engineer had to keep track of. Additionally, controlling the agents' functions manually might have created a more human-like behaviour, but instead introduced an increased risk of human error. Even while following the written manuscript, one

technical engineer might perhaps use the functions more frequently or less than another, or they might start some animations too late or too early. Since the behaviour and outcome of a session is so heavily reliant upon the technical engineer, it feels almost inevitable that having different technical engineers during the different sessions produces different outcomes. Inconsistencies like these could cause liabilities in the data.

Despite this, however, given the time limitations and resources offered this was the best workaround we could find. Moreover, considering this project's main research questions, their focal point being social influence and group behaviour, we decided that having realistic agents were of higher importance.

6.4 Methodological discussion

An idea to survey the participants before the test was discussed, in order to see if they had any difficulties focusing on lectures in general, in a non-virtual environment. This information could then be used to see if they perform worse than people who filled in that they do not have trouble focusing on lectures. But an issue occurred where there was no way to confirm that the people actually had problems focusing in a distraction filled environment. Without the possibility to confirm this, the data would be useless since someone might feel like they have a problem concentrating in distraction filled environment while they don't, or vice versa. Thus this part was removed and a decision was made to include all students from Lund University.

This is not to say that a pre-requisite like this could not be useful, however, due to time constraints, it would have worked against the study rather than for it. It is possible that the factor of perceived distractibility could have some co-relation to distractibility in a VR environment. Whether this is true or not needs to be further researched.

7 Conclusions

Throughout this study, it can be clearly observed that individuals tested on are, to different degrees, affected by social distractions in VR. Information retention was however not observed to be decreased or increased by the VR environment. All participants made a choice intentionally or unintentionally to behave in a certain way due to the avatars. Whether it was to mimic, react to or simply ignore the avatars. However, to definitively answer the questions posed in this study a larger sample size is needed. Currently we can not claim that simply because a lecture is structured for a virtual environment it will cause distractions to be less potent. This can simply be a side effect of having a small sample sized caused by the time frame of this study. As mentioned in our results there are indications of the effect of the differing conditions on behaviour which should in theory affect ones performance as seen in real life distraction studies (Tärning et al., 2022). We therefore suggest that future research into this topic should on top of having more

participants also have a larger control group without agent influence and a comparative real life test group which can give a stronger indication of differences between the mediums if such a difference as discussed earlier would exist.

Another part that would need to be tested is the memory questionnaire itself. An issue that occurred in the test is that the memory questionnaire might have been too hard for the participants even without the distractions in the VR environment. This would have to be tested with a control group that could verify that the memory questionnaire should be easy enough to complete with good enough grades without the distractions.

A point worth mentioning is that we received a comment questioning the use of VR as a medium. The participant stated that it seemed unnecessary to use VR the way it was used in this project, since we did not utilise any of the strengths of VR to actually promote knowledge retention. It could be interesting to see if a higher level of immersion would affect the social influence in any way.

References

- Alce, G., Eriksson, J., Wallergård, M., Engqvist, J., Lyons, L., & Hallberg, J. (2022). *Virtual reality - en introduktion*. Studentlitteratur AB.
- Alsop, T. (2022, Jan). *Vr headset unit sales worldwide 2024*. Retrieved from <https://www.statista.com/statistics/677096/vr-headsets-worldwide/>
- Brekel. (2022, Oct). Retrieved from <https://brekel.com/>
- Clement, J. (2023, Jan). *Global vr gaming content market size 2025*. Retrieved from <https://www.statista.com/statistics/499714/global-virtual-reality-gaming-sales-revenue>
- Controllers - valve index. (n.d.). Retrieved from <https://www.valvesoftware.com/en/index/controllers>
- Fox, J., Ahn, S. J.-G., Janssen, J., Yeykelis, L., Segovia, K., & Bailenson, J. (2015, 09). Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence. *Human-Computer Interaction*, 30, 401-432. doi: 10.1080/07370024.2014.921494
- Gass, R. H. (2015). Social influence, sociology of..
- Greenwald, S., Wang, Z., Funk, M., & Maes, P. (2017, 06). Investigating social presence and communication with embodied avatars in room-scale virtual reality. In (p. 75-90). doi: 10.1007/978-3-319-60633-0_7
- Lee, K. M. (2006, 01). Presence, Explicated. *Communication Theory*, 14(1), 27-50. Retrieved from <https://doi.org/10.1111/j.1468-2885.2004.tb00302.x> doi: 10.1111/j.1468-2885.2004.tb00302.x
- Meta horizon workrooms. (n.d.). Retrieved from <https://www.meta.com/se/work/workrooms/>
- Rapoza, K. (2022, October). *The metaverse is failing, but this is one investment that will not die*. Retrieved from <https://www.forbes.com/sites/kenrapoza/2022/10/21/the-metaverse-is-failing-but-this-is-one-investment-that-will-not-die/>
- ?sh=643755c91834
- Tärning, B., Ternblad, E., Haake, M., Gulz, A., & Nirme, J. (2022, January 21). Lessons learned from a study on distractions in virtual learning environments: Reliability, ecological validity, and an elusive social component. *Presence*, 28, 65-85. doi: 10.1162/PRES_a_00342
- TeamPosit. (2022, Dec). *Rstudio: Integrated development environment for r*. Retrieved from <http://www.posit.co/>
- Unity. (n.d.). Retrieved from <https://unity.com/Vrchat>.
- Vr headset - varjo vr-3: *Highest resolution virtual reality*. (2022, Oct). Retrieved from <https://varjo.com/products/vr-3/>

A - First yoga survey

Yoga A

* Required

1. Kod *

2. Vilken eller vilka religiösa traditioner utgör Yogas ursprung?

Check all that apply.

- Buddhism
- Sufism
- Hinduism
- Jainism

3. Vad är (huvud)syftet med Yoga? *

Mark only one oval.

- Befria sitt sinne från det världsliga
- Ökad rörlighet
- Gemenskap
- Bryta dualismen mellan kropp och sinne

4. Vad är Sanskrit för "världsaltert" ? *

Mark only one oval.

- Bhakti
- Moksha
- Asana
- Brahman

5. Vad innebär "Jnama yoga" ? *

Mark only one oval.

- Fysisk träning
- Teoretiskt kunskapsinhämtande
- Sprituell hängivelse
- Praktiskt handlande

6. Vilken av förgreningarna likställs ofta med "Yoga" i västerländsk kultur? *

Mark only one oval.

- Jnama Yoga
- Bhakti Yoga
- Laya Yoga
- Hatha Yoga

7. Vad innebär "Hatha Yoga" ? *

Mark only one oval.

- Den ursprungliga formen
- Meditation
- Utövandet av fysisk aktivitet enligt givna principer
- Teoretiskt kunskapsinhämtande

8. Vilken typ av Yogaformer är vanligast på Yogastudios i väst? *

Mark only one oval.

- Dynamisk Yoga
- Ashtanga Yoga
- Flödande Yoga
- Fysiskt ansträngande Yoga

9. Hur många former av dynamisk yoga tas upp i föreläsningen? *

Mark only one oval.

Två

Tre

Fyra

Fem

10. Vad strävar man efter inom "Hatha Yoga" ? *

Mark only one oval.

Att utföra serier av kroppsställningar strikt så som de beskrivs

Att uppnå ett flöde

Att anpassa Yoga till utövaren

Att uppnå fysisk och mental balans

11. Vad utmärker utförandet av Vinyasa Yoga (jämfört med Ashtanga Yoga)? *

Mark only one oval.

Större fysisk utmaning

Mer fokus på det mentala

Striktare hållning till originalformerna

Friare serier kroppsrörelser, och kan anpassas till utövares preferenser

This content is neither created nor endorsed by Google.

Google Forms

B - Second yoga survey

Yoga B

* Required

1. Kod *

2. Vilket år publicerades boken "Yoga Mala" ? *

Mark only one oval.

- 1954
- 1968
- 1982
- 2000

3. Vilka kroppsliga effekter har utförandet av Asanas?

Check all that apply.

- Sträcka ut muskler
- Ökad vakenhet
- Bygga styrka
- Bygga flexibilitet

4. Vad kallas medveten andning inom Yoga? *

Mark only one oval.

- Karma-andning
- Dhristi-andning
- Mula-andning
- Ujjayi-andning

5. Vilka muskler är involverade i den första kroppsålset? *

Mark only one oval.

- Inre magmuskler
- Bäckenbotten
- Benmuskulatur
- Halsmuskulatur

6. Vad kallas det andra kropplåset på Sanskrit / inom Yoga? *

Mark only one oval.

- Dhristi Bandha
- Uddiyana Bandha
- Laya Bandha
- Keta Bandha

7. Vilka muskler är involverade i det tredje kroppsålset? *

Mark only one oval.

- Inre magmuskler
- Bäckenbotten
- Benmuskulatur
- Halsmuskulatur

8. Vad kallas det tredje kropplåset på Sanskrit / inom Yoga? *

Mark only one oval.

- Sura Bandha
- Laya Bandha
- Karma Bandha
- Jalandhara bandha

9. Vilken är den tredje (och sista) principen bakom Asanas? *

Mark only one oval.

- Bhakti (blicken)
- Dhristi (blicken)
- Sura (andning)
- Ujjayi (andning)

10. Hur ska ögonen vara när man utför Asanas enligt Ashtanga Yoga? *

Mark only one oval.

- Avslappnad blick
- Stängda
- Transcendent blick
- Fokuserad blick

11. Vilka av följande är kända Asanas inom Ashtanga Yoga

Check all that apply.

- Solhälsning
- Krigarposition
- Sittande framåtböjningar
- Bryggor

This content is neither created nor endorsed by Google.

Google Forms

C - Social influence survey

Slutenkät

Svara på frågorna och motivera gärna

* Required

1. Kod: *

2. Hur kändes denna virtuella föreläsningen jämfört med en vanlig föreläsning? *

Mark only one oval.

Inte alls som en verlig föreläsning

1

2

3

4

5

6

7

Precis som en verlig föreläsning

3. Fanns det tidpunkter där det kändes som att den virtuella miljön var verlig? *
Motivera gärna.

4. Upplevde du att du var i det virtuella rummet med andra personer eller upplevde * du att du var där ensam?

Mark only one oval.

Ensam

1

2

3

4

5

6

7

Med andra personer

5. I detta test, fanns det någon tidpunkt då det kändes som att du faktiskt var i ett klassrum på en föreläsning? *

Mark only one oval.

Inte alls

1

2

3

4

5

Hela tiden

6. Kände du att det fanns något som distraherade dig under föreläsningarna? Om ja, motivera gärna. *

7. Vilken föreläsning var enklast att hänga med i enligt dig? *

Del 2

8. Fann du att det var stor skillnad på beteendet av deltagarna mellan de olika föreläsningarna? Om ja, motivera gärna *

Del 3

9. Eftersom detta är ett nytt sätt att uppleva en föreläsning, så känner vissa att de inte interagerar med verkliga människor. Hur upplevde du de andra deltagarna under föreläsningen? *
-
-

This content is neither created nor endorsed by Google.

Google Forms

To Be a Robot

Simulating Robot Perception in Virtual Reality

Mia Huovilainen

mi1102hu-s@student.lu.se

Vanessa Nguyen

va4274ph-s@student.lu.se

Alexander Wåhlander

al2183wa-s@student.lu.se

In this study, we investigated how social interaction is affected when constraining the perception of participants akin to the perception of the research robot EPI. By simulating these perceptual constraints in virtual reality, we investigated how social efficiency, head movements, and quality of enjoyment and engagement are affected by constrained perception in an experiment involving participants playing a cooperative game, a so-called tower game, with a confederate. We measured the duration of each turn of the game as well as total time taken to complete the task, the extent to which participants had their head rotated towards a certain zone, and ratings of participants' experience of the interaction. We used a between-subjects design, with a total of 19 participants between ages 21-30. The analysis indicated that the experiment group generally took longer to complete the task, but they showed a larger improvement over trials. However, no significant difference of head movements between the two groups was found, but we found a small difference in enjoyment of the interaction for the experiment group relative to the control group. A detailed discussion of the results are presented with regard to relevant theories and related works.

1 Introduction

Today robotic research is a popular subject; the field of social robotics in particular has received a lot of attention for very understandable reasons. The idea of interacting with artificial machines, that in some capacity may even resemble humans, such as androids, is undoubtedly an idea that has captured the imagination of the public (see Bartneck et al., 2020, chap. 2); one only has to look at the wealth of media being produced examining the subject (Wikipedia, n.d.). While many are interested in how it may feel to interact daily with a robot, far fewer have thought about how the interaction may be experienced from the point of view of the robot; research is certainly lacking regarding this subject.

This work - a collaboration between the Lund University Cognitive Science group and the Lund University Faculty of Engineering - is an attempt to simulate a social interaction from the point of view of a robot, given that a human cannot actually *be* a robot and understand its subjective phenomenological experience. In other words: to proximate an understanding of what it is like for a bot to be a bot (see Nagel, 1974) - assuming that robots are even capable of possessing subjective phenomenological experience (cf. Searle, 1980). We used virtual reality (VR) to simulate the limited

perceptual abilities of the robot EPI, which is a research robot developed by the Lund University Cognitive Science Robotic Group, designed to be used in experiments in developmental robotics (Johansson et al., 2020). It's designed to resemble a humanoid child while still remaining decidedly robotic, with facial features such as large eyes in order to inspire a childlike impression (Johansson et al., 2020). The robot is controlled using the Ikaros system¹, a system designed for biologically motivated robot control.

This work focuses on one particular aspect of these social interactions, namely perception, in order to find out if the perceptual abilities of a robot may hamper it socially. Henceforth, when we refer to perception, we refer to visual perception in particular; furthermore we take visual perception to include not only eyesight per se, but also supporting motor actions, in this case head movements. To investigate this, we have designed a virtual reality environment, in which an exploratory experiment has been conducted with two human players that cooperated to complete a so-called tower game, a cooperative block-building game (see Salter et al., 2015), which was played as a turn-taking game. One of the players had constrained perceptual abilities, approximating those of EPI, thus allowing for quantitative measurements of the constraints' effect on the cooperative interaction under investigation. The participants' subjective experience was measured through a questionnaire. Through this experimental method, we aim to answer how perceptual constraints, like those of the robot EPI, affect the efficiency and experience of social interaction, as well as look into whether head movements are employed as a strategy to overcome the constraints. As the work is highly exploratory - due to a limited amount of previous studies about the subject - two rounds of pilot experiments were conducted and are reported.

Despite being a novel avenue of research, we do hypothesise that the perceptual constraints implemented in the experimental conditions will affect the quality of the social interaction negatively, leading to the following formal hypotheses:

H1: The turns of constrained participants should take longer than the turns of the controls.

H2: Constrained participants will spend more of their time having their head rotated, as well as showing greater rotation.

¹<https://github.com/ikaros-project>

H3: Constrained participants will rate the quality of the social interaction as being lower than the controls.

As mentioned, we hope that answering these questions may grant important insights for the development of future robots, and also allow for a greater understanding of human-robot interaction (HRI) as well as human communication with limited perception.

The rest of the text will be structured as follows: the next section will be dedicated to a presentation of the theories and empirics underlying our work, which in turn is followed by a presentation of the methods utilised and the design rationale, followed by a presentation and analysis of our experimental results. Concluding the text is a section dedicated to discussing the results, the work as a whole, a reflection on the relevance of this project as well as suggestions for future work.

2 Related work

Communication and interaction

Mondada (2014) reasons that human communication is inherently multimodal, meaning that there is an array of different communicative resources that humans use in an interaction. These communicative resources can be both verbal (speech) and non-verbal (e.g. head movement, gaze behaviour, gestures, etc.), and the resources used depend on the contingencies of the current situation (Mondada, 2014). In an experiment, Froese et al. (2014a,b) limited the perception of participants to a one-dimensional environment through human-computer interaction. The participants were tasked with finding and discerning the avatar of a game partner from a number of distractor objects, with a vibrating device being their only form of feedback. Despite these restrictions, the participants were mostly successful at the task, demonstrating the ability of humans to conduct social interactions in environments both unfamiliar and limited in nature.

Efficiency in turn-taking communication

Turn-taking is generally considered to be one of the most fundamental ways to organise a face-to-face conversation. It demands participants to understand when someone has finished communicating in order to give a response. Hence this rapid flow of turn-taking is one of the building blocks of a conversation (Levinson, 2016). Considering its foundational element in communication, turn-taking games are a common way to study social interactions that require a cooperative environment; Cason et al. (2013) used a turn-taking game to investigate whether teaching is an important aspect in promoting efficiency-enhancing, history-dependent strategies, while Szymanik et al. (2013) investigated whether it's possible to predict participants' reaction time and reasoning strategies in a turn-taking game called Marble Drop.

Another important aspect of communication is its efficiency; Steinfeld et al. (2006) presented a few measurements for measuring effectiveness and efficiency. According

to them, this can be accomplished by determining the percentage of the task that was accomplished, while measuring efficiency can be performed by measuring the time required to complete a task (Steinfeld et al., 2006).

There are currently no related works, to the authors knowledge, that studies social efficiency in a turn-taking game. However, to combine the collaborative nature of "turn-taking" in a game with a measurement of efficiency, thus using the time taken to complete a game turn as a variable dependent on social efficiency, might be a relevant method for studying social interactions.

Visuomotor integration

Visuomotor integration is the coordination of neuronal activity between vision-related and motor-related parts of the brain (Shin et al., 2009). It is a skill that underlies e.g. observational learning, which can be seen as a fundamental human skill in which we learn simple and complex social behaviours by observing others (Ramsey et al., 2021), but it also is integrated in most daily activities that require perceptual, motor, and cognitive abilities (Pratviel et al., 2021). To investigate visuomotor coordination, a study conducted by Huang et al. (2022) found that visuomotor timing can be improved with sustained visual attention in a task of synchronising finger movements to a sequence of flashing visual stimuli. Combining VR to this subject, researchers in Miller et al. (2017) created a virtual environment with a full-body motion capture system and an eye-tracking system in order to study the relationship between the vision and motor processing of subjects involved in different tasks, for instance having participants doing a physical movement (e.g. leaning) in response to a visual stimuli. Additionally, Pratviel et al. (2021) conducted a principal component analysis with a hierarchical cluster analysis on different visuomotor tasks in VR, and saw that the two clusters with best performance were separated by opposite head movement strategies, i.e. one cluster had limited head movement and the other cluster had excessive head movement. They also reported that the cluster in which the characterising component was excessive head movements, did not qualify as a successful strategy since this cluster had low reaction time which was a measure for performance (Pratviel et al., 2021).

To connect these findings to our work, the tower game is certainly a task that requires visuomotor integration, and it seems to be affected by disrupted visual attention, however, the relationship between head movements and visuomotor performance seems unclear.

Human-robot interaction

One experimental technique that has been extensively used in HRI is the Wizard of Oz paradigm (WoZ) which lets a person (usually a researcher or confederate) remotely operate a robot to some degree (Riek, 2012). It has been used, for example, to investigate the first-person experience of being a robot (Thellman et al., 2017). They found that alter-

ing the participant's FOV to the robot's camera view and giving them the ability to shift the perspective freely by moving their head, led to participants reporting that their perspective shifted from their own physical self to the robot's; feeling as if they were in the position of the robot (Thellman et al., 2017). Another study used WoZ to investigate social interaction strategies in which the robot's inherent perceptual capabilities were included, however, in this study, it was an expert, and not the participants that controlled the robot (Sequeira et al., 2016). Looking at research on head movements in participants that interact through a mobile robotic telepresence, no significant relationship has been shown between the representation of head movement and social presence (Bamoallem et al., 2016). When looking at the influence of FOV on MTR in an interaction, it was shown that a restricted FOV on the horizontal plane actually promoted interaction and that the participant used the robot body to rotate towards people in the interaction instead (Kiselev et al., 2014).

Virtual reality

Virtual reality has been used as a tool for studying HRI in a virtual environment, in which it already has been validated for robots such as mobile operators (Inoue et al., 2005; Weistroffer et al., 2014) and non-humanoid social robots (Sadka et al., 2020). When in a virtual environment, the concepts of "immersion" and "social presence" are vital, the latter more important when engaging in a social interaction. According to Oh et al. (2018), social presence is enhanced by having a conversation partner represented by an avatar. It is further stated by Oh et al. (2018) that realistic behaviour, such as non-verbal communication, also enhances social presence; following this, we implemented an animated avatar in the form of a humanoid robot in our experimental procedure (see Figure 1).

Immersion is meanwhile described as "a property of the valid actions that are possible within the system" (Slater, 2009, p. 3551), with valid actions taken to be the actions a person within VR can take that result in changes to perception or changes to the environment. Slater refers to immersive virtual reality as possessing a set of valid actions that approximate the valid actions of a real-life equivalent of the situation (2009). Slater (2009) goes on to define two illusions that are important to consider in virtual reality: the place illusion (PI) and the plausibility illusion (Psi). PI is the sensation of *being there* in the environment, while Psi is the illusion of sensing that the situation you're in is real (Slater, 2009). He further explains that immersion and PI are defined through the *sensorimotor contingencies* that they support; sensorimotor contingencies being the structure of rules in which sensory changes made by various motor actions are governed (O'Regan & Noë, 2001). Immersion and PI are very similar, the difference being that immersion demarcates where PI can occur (Slater, 2009).

In VR, the field of view (FOV) governs the area visible to the user; it tends to differ somewhat from the normal human

field of view, which amounts to a combined horizontal FOV of 180° (Lin et al., 2002). In early virtual reality, the field of view tended to be set at 60° but is now most often set at over 100° in modern systems as it has been found to affect distance and movement perception (Hopper et al., 2019). High FOV has also been reported to increase immersion (Lin et al., 2002).

Slater (2009) points out that valid sensorimotor actions in VR can be of the same range as in physical reality, but still not be sufficient to support those sensorimotor contingencies in VR that approximate sensorimotor contingencies in reality. One example can be related to Alfano & Michel (1990) who investigated how visuomotor performance is affected by a constrained FOV, reporting that there were effects on eye-hand coordination, and the more restricted the FOV of the participants was, the more time it took to complete the task, and more errors were made. Slater (2009) writes that having a head-tracked head mounted display (HMD), allowing a full range of movement but with limited FOV, will lead individuals to try to look around and perceive the scene as a whole, but since the visual capture is limited, individuals will have to learn to perceive the environment in a different way. He further discusses that this could include greater and different movement patterns of the head in order to capture the same visual information as in reality (2009).

3 Material and design

EPI

Given the focus of the present work on the perceptual constraints of the robot EPI, only the perceptual specifications of the robot will be presented. The head of EPI is controlled by six servos, with the main movement being controlled by two, allowing for yaw and pitch. Each individual eye is controlled by a servo, allowing them to move sideways at 475 °/s, similar to human eye saccades; the eyes are however incapable of vertical movement (Johansson et al., 2020). Additionally, each eye is equipped with Adafruit's Zero Spy Camera for Raspberry Pi Zero (Adafruit Industries, n.d.; Johansson, in conversation), connected to a Raspberry Pi Zero running User space Video4Linux (UV4L) (Johansson et al., 2020). The cameras have a 70° focal angle (Adafruit Industries, n.d.), and support 1080p30 and 640x480p90 resolutions (Johansson et al., 2020).

Virtual environment

In this study, participants were tasked with cooperatively building a tower in VR. The virtual environment consists of a room with a table, two boxes where the building blocks are stored on each side of the table, and two models of the tower that is to be built, one representing the front of the tower and one representing the back. Otherwise, the environment is completely empty with plain walls (see Figure 3). The design of the room is kept simple in order to not distract the participants with irrelevant stimuli; this is inspired

by the concept of *Shared Diminished Reality*, a framework conceptualised by Vuarnesson et al. (2021), meant to isolate the effect of particular factors in an interaction by reducing the amount of sensory data in the environment. Reducing the available stimuli should also help avoid the breakdown of any of Slater's (2009) illusions; while PI is probably not much of an issue given the ease of repairing the illusion, Psi would be more of an issue with more available stimuli, as this would imply more factors would have to be considered in order to maintain the illusion. Fundamental concepts, such as cause and effect relationships between the actions of the participants and the building blocks are in place, which should be enough to maintain the Psi at a basic level; moreover, both the participant and the confederate were in the same physical room during the experiment, which allowed for natural verbal communication to take place. Considering these factors, the possibility of a breakdown of the Psi should not pose much of an issue for our study.

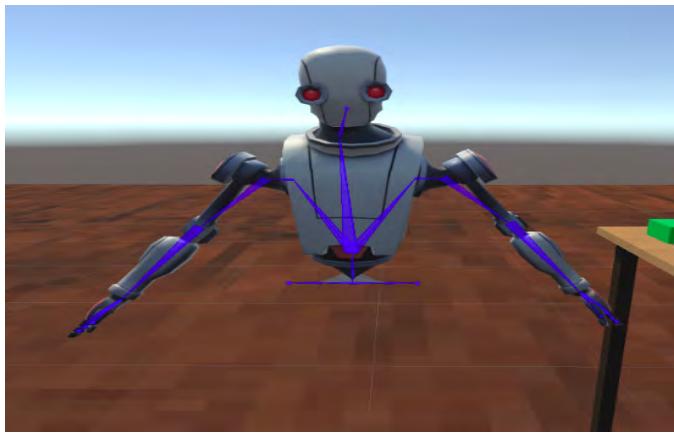


Figure 1. Upperbody of the humanoid robot

The avatars were animated using Unity's Animation Rigging tools (Unity Technologies, n.d.-a). Through inverse kinematics in the hand joints, the avatar's arms move in accordance with the hand controllers. Correspondingly the head of the avatar is mapped to the HMD so that the avatar's head movements align correctly with the player's.

As for the playing aspect of the experiment, the building blocks were made up of square, colourful blocks, inspired by those used in Salter et al. (2015). In order to make these blocks interactable and have physics (e.g. gravity), each one has a *Throwable* script (Valve Corporation, n.d.-a) and a *RigidBody* component (Unity Technologies, n.d.-c) attached. The *Throwable* script enables the player to pick up an object by hovering their hand above it and pressing the grab button. While the *RigidBody* component makes it possible for the blocks to collide with other objects, thus preventing them from falling through the table. To make the building process as effortless as possible, the blocks were designed to snap to a specific place. This means that if a correct building block

is selected, when placing it near a predetermined point in space it will snap onto a designated position. By doing this each block will be accurately placed and remain fixed despite accidental collisions. To achieve this, any block that possibly fits into the tower structure has a script attached to it, which has the function of creating a corresponding snap zone.

The constraints implemented is a between-subjects factor, meaning that half of the participants were constrained, and the others acted as a control group. To limit the FOV, a black canvas with a transparent rectangle in the middle was placed in front of the participant's virtual avatar. The canvas follows the movements of the HMD to consistently be positioned right in front of the camera. Thus, the vision of the participant is limited to this transparent square, while much of their peripheral vision is completely blocked. The FOV can be modified by adjusting the width of the transparent rectangle. As for limiting head movements to achieve two degrees of freedom, a fade function was implemented and set to execute when any rotations on the roll axis oversteps a set margin of 10 degrees in both directions. In other words, every time the participant tilts their head to either side the fade function will set the screen to completely black when overstepping the threshold.

A system overview of the VR environment for the experiment with the experimental group is shown in Figure 2; note that the FOV canvas and the fade effect are inactivated for the control group.

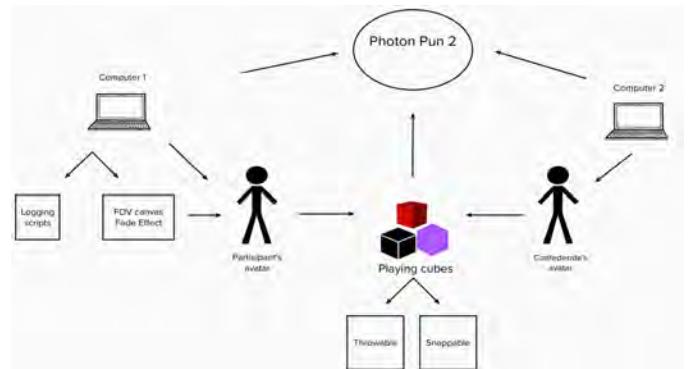


Figure 2. System overview of the VR environment

Virtual reality equipment

The virtual reality environment was constructed using the game engine and game development toolkit Unity (Unity Technologies, n.d.-d), along with its integrated XR Interaction Toolkit (Unity Technologies, n.d.-e). Third-party software for Unity was used to allow for both parties of the experiment's social interaction to be instantiated in the environment together; the software used was Photon Pun 2 (Exit Games, n.d.-a). It works by providing a game server that a client gets forwarded to whenever a room is created or joined (Exit Games, n.d.-b). The avatars instantiated for

both players were assets downloaded from the Unity Asset Store (Unity Technologies, 2019). The camera fade effect to simulate the absence of a degree of freedom was made by Castaño Estrella (n.d.), and the interior of the environment was made using an asset, developed by Unity Asset Store (n.d.), downloaded from the Unity Asset Store. The snapping function for the game objects was implemented with the help of the open-source toolkit VR Builder (MindPort, n.d.). The virtual reality system used was the Valve Index (Valve Corporation, n.d.-c), which ran on the SteamVR software (Valve Corporation, n.d.-b).

Video analysis

All instances of the experiment were recorded using the recording software incorporated in the program GeForce Experience (Nvidia Corporation, n.d.) and stored on an external hard drive. The recordings were analysed using the annotation software ELAN 6.4 (Max Planck Institute, 2022). Using ELAN, we measured each turn taken as well as how often the fade effect was activated during the game.

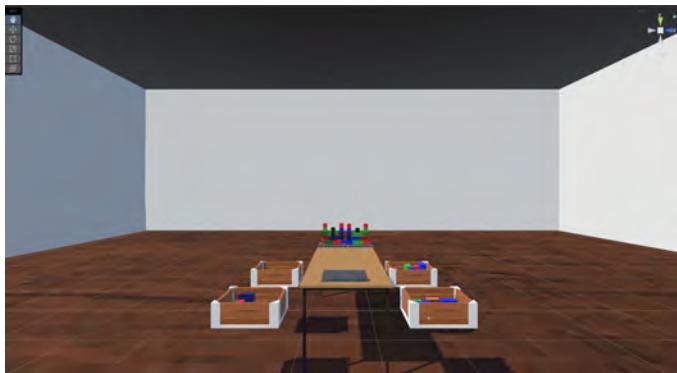


Figure 3. Virtual Environment

4 Method

Pilot participants

A total of 8 people participated in the pilot experiments (Mdnage = 23), 5 identified as male and 3 as female. None of the participants reported any deficiencies in colour vision. On average the participant's earlier experience with VR was low (Mdn = 2 on a Likert scale from 1 to 5, with 5 representing very much). Four participants were assigned to the experimental group, and four to the control group. The participants were recruited from Lund University's student population through personal networks. Participation was completely voluntary and the participants received no compensation except for Swedish fika.

Pilot procedure

The experiment took place at a VR laboratory at the Faculty of Engineering at Lund University. Two rounds of pilot studies were conducted to test the experiment design and to evaluate whether manipulating the perceptual abilities of the participants caused simulator sickness. During recruitment, participants were told of the general nature of the experiment, i.e. to study a social interaction in VR. In the VR laboratory they were only told to engage in an experiment and got to sign a consent form (see Appendix). They received instructions on the tower game and got to familiarise themselves with the virtual environment and how the hand controllers worked before carrying out the experiment.

The experimental group had their perception limited as described earlier, while the control group did not. After training, the participants and controls were placed in the earlier described virtual environment, and given the task of building a structure mirroring a model tower, out of the materials provided. This task was completed cooperatively with an unconstrained confederate, who was one of the researchers - this was to be able to better control for confounding variables. After the experiment, the participants got to answer two questionnaires: one asking about their experience of the interaction and whether they felt constrained in head rotation and field of view, the other one being the SSQ. The questionnaires are described in *Measurements*. They then were debriefed on the specifics of the project and were thanked for their participation.

During observation of the participants in the experiment condition from the first round of the pilot, the participants did not seem to understand the function of the fade effect. We speculated that this might be due to a lack of understanding of the causality between their rotation of the head and the screen fading to black. Initial analysis of the report from the interaction questionnaire reported that there was no difference between the participants in the experiment or control groups regarding their experience of being able to rotate their heads freely. The aim of the fade effect was for the participants to avoid rotating their heads on the roll axis, wherefore it was decided to give the participants instructions to move their head in such a way as to trigger the fade effect during practice, thus allowing for a causal relationship between the fade effect and movement of the head on the roll axis to be established. These instructions were based on the Socratic method, wherein the participants had to work out the mechanics of the fade effect themselves and vocalise them to the researchers. This change was incorporated into the second round of the pilot experiments, wherein it was observed that the participants seemed more aware of the fade effect and how to prevent it; this change was then incorporated into the experiment.

Participants

A total of 20 people participated in the experiment, 12 identified as male and 8 as female (Mdnage = 24, Range = 21-

30). None of the participants reported any deficiencies in colour vision. Regarding the participant's earlier experience with VR, they had low experience on average ($Mdn = 2$ on a Likert scale from 1 to 5, with 5 representing very much). Of these participants 10 were assigned to the experimental group and 10 to the control group. Of all 20 participants, only 19 were included in the data analysis due to incomplete data materials for one participant from the experimental group. The participants were recruited from Lund University's student population through personal networks. Participation was completely voluntary and the participants received no compensation except for Swedish fika.

Experimental procedure

The experiment took place at the same VR laboratory at Lund University. Similarly to the pilot procedure, participants were only told that they were to participate in a social experiment in VR during recruitment and were given the same consent form. The participants were assigned to either the experimental group or the control group based on the order of their experiment time slots: every other participant was assigned to the experimental group and vice versa. The experimental group had their perception limited in the way described above, while the perception of the controls was not limited in any way.

The participants were given simple instructions about their task, and were then given an opportunity to briefly practise. During the practice round, the participants were instructed to move their head in certain ways in order to trigger the fade effect, whereafter the researchers verified that the participants knew the cause of the fade effect and to avoid it. The participants were shown the model of the tower to be built and given additional instructions during the practice round, and were taught how to use the controllers for manipulating the building blocks.

The participants then engaged in the cooperative building task with the confederate, with whom they were free and encouraged to communicate verbally, in order to ease cooperation. After the conclusion of the building task, the participants were instructed to answer a questionnaire, before receiving a full debriefing.

5 Measurements

Social efficiency

The experiment was designed to investigate the impact on social efficiency by perceptual constraints akin to those of the robot EPI, i.e. $H1$. Our main measurement was the temporal difference (as in Steinfeld et al., 2006) between turns taken during the game by constrained participants and controls, as well as the overall time it takes to perform the task; to do this we measure each individual turn taken by a participant in the game. We take a turn to be the time from when the confederate has placed a block (i.e. the block snaps into place) to when a participant's block snaps into place. The

analysis was done manually using ELAN (Max Planck Institute, 2022). A separate script was also used in Unity to log the execution time of the whole experiment.

Head rotation

In order to investigate $H2$, whether the head movements of the constrained participants are more extreme in comparison to a non-constrained participant, as Slater (2009) might implicate, we examined how participants rotated their heads on the yaw axis; in particular whether they rotated their heads more in order to get a good view of the boxes with building blocks that were adjacent to them. We measured the participants' head rotations on the yaw axis using a C# script implemented to continuously register the HMD's rotation during the experiment. The head rotations were sampled with a frequency of 50 Hz using Unity's default frame rate (Unity Technologies, n.d.-b). The constraint for head rotation on the roll axis, the fade effect, was measured through the amount of times that the effect was enabled during the experiment, using ELAN (Max Planck Institute, 2022), to see if it would have any effect on the data.

Questionnaire

To measure the experience of the interaction in order to test $H3$, we employed a questionnaire with a 5-point Likert scale that was adapted from Herath et al. (2018), which measures co-presence and co-experience in human-robot interaction. Since the questionnaire is about how participants perceive a robot as a social agent, we had to modify the questionnaire to be appropriate in a human-human interaction. Irrelevant items were discarded and the chosen items were modified in order to see how participants perceived the interaction with the other player. Three additional items were added concerning the constraints on FOV and head rotation asking to what extent they felt constrained by the field of view, if they felt that they could move their head as they wished, and to what extent they felt annoyed by not being able to move their head as they wished. These items were on the same scale as the other items.

Changing the FOV may also be a risk factor for *simulator sickness*, a type of motion sickness that occurs in the absence of actual motion (American Association of Psychology, n.d.). It has been reported that a high FOV may increase the risk of simulator sickness (Lin et al., 2002), yet it has also been reported that low FOV causes bodily discomfort, dizziness, unsteadiness, and disorientation (Alfano & Michel, 1990); this necessitated consideration of whether our experimental procedure caused simulator sickness, which we measured using the Simulator Sickness Questionnaire (SSQ) developed by Kennedy et al. (1993) during the pilot. The items were translated by us into Swedish. See Appendix for both questionnaires.

6 Results

Statistical analysis

The statistical analysis will be presented in the order of each hypothesis. All analyses were conducted in R (4.1.1).

Hypothesis 1: The turns of constrained participants should take longer than the turns of the controls.

Three analyses were performed to investigate *H1*: (1) If a turn on average is longer in the experimental group in comparison to the control group, (2) if a trial effect can be observed in the participants and if there is a difference between the groups, and (3) whether there is a difference between the groups regarding how much time they take up of the entire interaction.

The turns of the participants in the experimental group ranged between 3.43 and 138.69 seconds per turn with a median of 18.59 seconds. The turns of the participants in the control group ranged between 4.03 and 46.60 seconds with a median of 12.53 seconds.

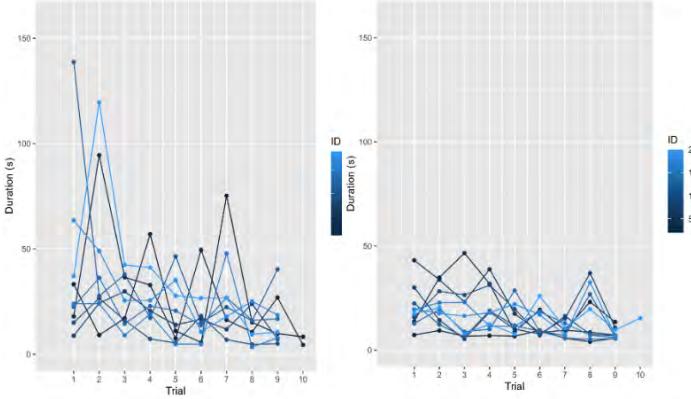


Figure 4. The duration of each turn for the different participants in each trial, divided between the experimental group (left) and the control group (right).

Thus the variance of time taken per turn seems to be much larger for participants in the experimental group compared to the control group. Analysis of the time taken per turn by the participants was done using Bayesian linear regression in which the response variable (time each turn) was log-transformed, employing a mixed model, assigning each individual their own intercept to control for individual differences. We found that the control group was circa 29% (95% CI [-50.84, 0.00]) faster than the experimental group per turn. The posterior distribution was 91% positive, indicating a high degree of confidence in the result, however with a large credible interval so the effect size is uncertain.

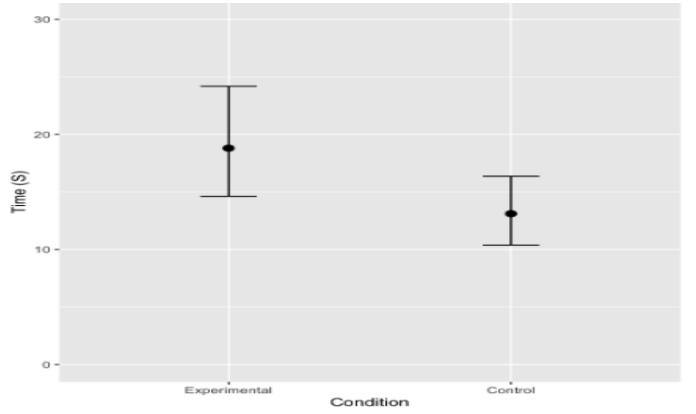


Figure 5. Time difference of an average turn between the groups.

We controlled whether the fade effect, given that it was two seconds of time in which the participants, was responsible for the longer turns of the experiment group. This was investigated for the experiment group, using Bayesian linear regression, through which we found that for each fade, time increases with 13.88% (CI 95% [9.42, 19.72]). This has an incremental effect that can increase the time if the amount of fades are large, but there are still some uncertainties in how large this effect is.

In addition, we analysed the improvement over time per group, that is, whether the participants became faster over time, and if they did, how much faster they became. This was tested with Bayesian linear regression, in which time was log-transformed, using mixed models, assigning an intercept for each individual to control for individual differences, with trial as a predictor. Three statistical tests were done; one to examine the overall improvement over time across the groups, and one to examine the improvement over time of each individual group. All trials except 3 were 9 turns long; 3 trials were 10 turns long due to participant errors, of these two of the trials were part of the experimental group and one was part of the control group. They were included in the analysis since we judged that these anomalies would not significantly affect the results, as the analyses were of time per trial.

Overall, there is an improvement over time of circa 10.42% (95% CI [-13.06, -6.76]), meaning participants become around 10.42% faster per subsequent turn. As can be seen in Figure 6, participants in the experimental group take almost double as long as on their first turn compared to the control group, but they rapidly improve; at their last trial they are as fast as the control group. The predicted time for the experiment group at turn 1 is 31.91 (CI 95%, [21.26, 47.30]) seconds and 11.23 (CI 95% [7.71, 16.57]) seconds at turn 9; the posterior distribution was 100% positive which indicates a very high degree of confidence in the results. The predicted time for the control group at turn 1 is 19.07 (CI 95% [14.92, 24.89]) seconds while it is 9.16 (CI 95% [7.04, 11.98]) seconds at turn 9; again the posterior distribution

was 100% positive, thus we can be confident in the results. An analysis per group indicates that the participants in the experimental group become 12.19% (95% CI [-16.47, -7.69]) faster per trial, while the participants in the control group become 8.61% (95% CI [-12.19, -4.88]), which amounts to a 4.3 percentage point difference.

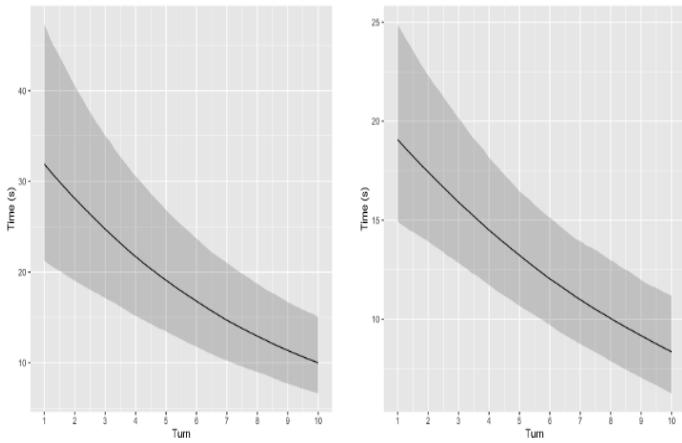


Figure 6. Predicted change in time per turn over time, experimental group (left) and control group (right). Note that they start at different points on the y-axis, but end at roughly the same place.

The last analysis examined how much of the absolute time per trial - as compared to the confederate - was taken up by the turns of each of the groups. The experimental group on average took up 59% of the total time per trial, while the control group on average took up 52% of the time. The analysis was done by first measuring the total amount of time each trial took using a logging script during the experiment, and then we measured the length of each individual turn through video analysis. The length of each individual turn per person was summed together, then divided by the total time of the experiment; this number was analysed using a Bayesian logistic regression model, using a beta distribution, in order to examine the differences between the groups. The analysis indicated that the experiment group took up 7.4% (95% CI [-2.00, 16.68]) more of the total time on average than the control group. The posterior distribution was 95% positive, indicating a high degree of confidence in the result, however we have to be cautious with the credible interval spanning over zero.

Given that the fade effect in and of itself lasted two seconds, we decided to investigate whether the triggering of the fade effect impacted the result. This was done using Bayesian logistic regression analysis on the percentage of the absolute time each participant in the experiment took up per trial, using a beta distribution. Looking at the participant with the lowest amount of fades (=3), the model predicts that this participant will take up 54.78% of the time (CI 95% [43.00, 65.33]), and the participant with the highest

amount of fades (=78) will take up 72.29% of the time (CI 95% [51.65, 88.10]). Looking at the contrasts, the models predict a difference of 16.91 percentage points (CI 95% [-8.42, 40.45]). When looking at the difference between a person with 3 fades and 4 fades, the models predicts a difference of 0.2% (CI 95% [-0.11, 0.67]), showing that the incremental effect of fade on the total amount of time in the interaction is small, but going up to extreme amounts of fades, the effect is larger. The posterior distribution showed that 90.9% is positive, however we have to be aware of the credible intervals being quite large so we cannot be too sure about the extent of the effect of the fade.

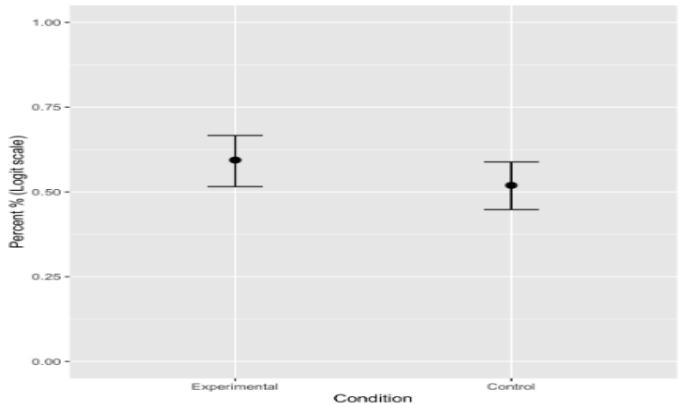


Figure 7. Difference in the percentage of time each group takes up of the total time in the interaction, on a logit scale.

The results from the analyses indicate support for *H1*, as there is marked time difference between the trials of the experimental group and the control group. Furthermore, both groups improve over subsequent trials, with the average length of turns approaching some 10 seconds; participants in the experimental group improved significantly faster than those in the control group, with their first turn on average being almost double that of the controls, while their last turn was on average as long as that of the controls. Lastly, the experimental group occupies on average 7.4% more of the time of the social interaction of the trial compared to the control group. However, we have to be aware of the effect of the fade effect, for participants with a large amount of fades this effect can be large.

Hypothesis 2: Constrained participants will spend more of their time having their head rotated, as well as showing greater rotation.

We decided to look into one area of interest, namely the boxes containing the building blocks, and determine the zones towards which the participants' heads were rotated on the yaw axis to the degree that they could see the boxes. Being in the zone means that a participant's head is rotated on the yaw axis to the degree that they can see one of the boxes adjacent to them. Since there are two boxes on either side of the

participant, the zone is split into two parts. Because we did not use eye-tracking, we cannot be sure when participants are actually looking at the boxes; instead we are interested in head rotations corresponding to viewing the boxes. The cut-off values for deciding when the participants were looking at the boxes were determined manually, by having one of the researchers instantiated inside the virtual environment, and another researcher viewing the inspector in Unity to determine the corresponding value on the yaw axis. Since the experiment group can see less because of the canvas, the different groups have different cut-offs for when the entire box is in their field of view. For the control group, to see the entire boxes, they had to rotate their head either 150 or 30 degrees. For the experiment group, this was 180 and 15 degrees respectively. A head rotation of 90 degrees corresponded to looking straight ahead.

An analysis was conducted, comparing both groups to both cut-offs, to see whether: (1) the experiment group rotated their head towards these zones in comparison to the control group and (2) whether the experiment group showcased greater head rotations in comparison to the control group.

The percentage of time that the head was registered as being above or below the threshold on the yaw axis, as earlier described, was calculated. Using the cut-offs for the control group, on average, the experiment group spent 23% of the total time having their head rotated inside the zones of interest (that were described earlier) on the y-axis, meanwhile the control group spent on average 24% of the time inside the zones. Using the cut-offs of the experimental group, the average was 6% for both the experimental group and the control group.

A Bayesian logistic regression was conducted, using a beta distribution for the thresholds for the control group, and zero-inflated beta distribution for more extreme cut-off.

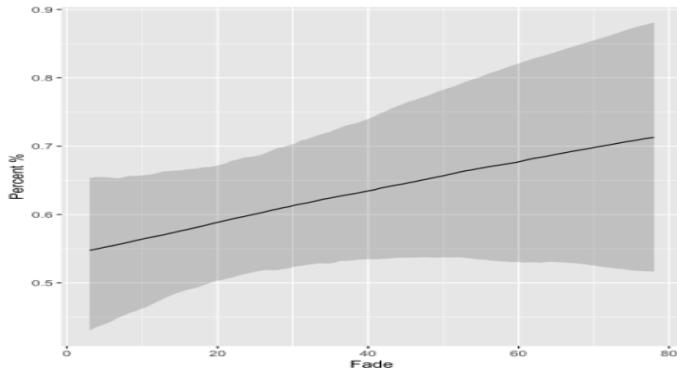


Figure 8. Percentage of time occupied by a participant in a interaction, as a function of number of times the fade effect is initiated during the experiment. Only modeled for the experimental group.

For the first prediction, in contrast to what we predicted, the statistical model shows that the control group rotated

their head 3.31% more relative to the experiment group (CI 95% [-6.42, 13.79]). However, this effect is quite small, and since the credible interval spans across zero, we cannot be sure of this effect; in addition, 75.7% of the posterior distribution was positive so there is uncertainty about the results. For the second prediction, using more extreme cut-offs, the result was also contrary to what we predicted. The control group rotated their head 0.8% more relative to the experiment group (CI 95% [-4.85, 6.36]). Since the credible interval is large and spans across zero, we cannot be sure of this effect, and even if it was meaningful, this effect is practically negligible. To summarise, H2 is not supported by our data.

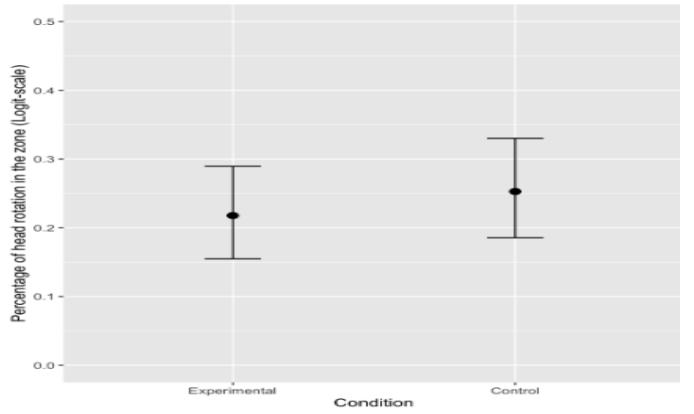


Figure 9. Percentage of time participants have rotated their head on the y-axis within a zone, with cut-offs of 150 and 30 degrees.

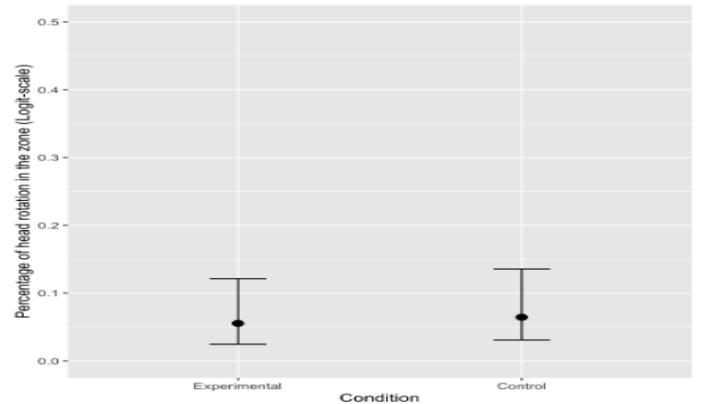


Figure 10. Percentage of time participants have rotated their head on the y-axis within a zone, with cut-offs of 180 and 15 degrees.

Hypothesis 3: Constrained participants will rate the quality of the social interaction as being lower than the controls.

To analyse the quality of the social interaction, participants were asked to rate their experience of the interaction using

the questionnaire detailed in *Measurements*. All items were then summed (items 2, 5, and 16-18 were reversed before they were summed). The sum of these points were following a log-normal distribution so this variable was log-transformed. The average value of the sum of points between the two groups was 64.44 for the experimental group and 70.90 for the control group. To investigate whether the ratings on the questionnaire differed between the two groups, a Bayesian linear regression model was used for the log-normalised sums. The analysis showed that the control group enjoyed the interaction more by 9.96% (CI 95% [-1.21, 21.50]) relative to the experimental group. The posterior distribution was 96.1% positive, indicating confidence in these results, however we have a large credible interval spanning across zero so there is still uncertainty of the effect size.

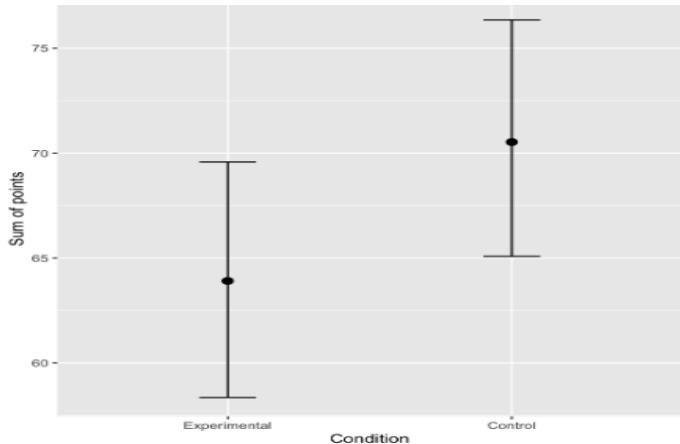


Figure 11. Difference in total score between the groups, predicted by the model. Note that the scale on the y-axis is zoomed in.

With these results, we can find some support for *H3*, in which the control group had a slightly higher average value relative to the experimental group. However, the difference is small with a large credible interval.

7 Discussion

Results

Hypothesis 1 was supported to some extent by the results of our analyses. The difference between the length of turns of the experiment and control groups was 29%, which is a fairly large effect given the median length of the turns. Furthermore, the experiment group took up 7.4% more of the overall time of each trial compared to the control group. These results are perhaps unsurprising, as the constraints were quite limiting, but it has some bearing on our research question; this result indicates that perceptual limitations approximating those of the robot EPI have a significant effect on the quality of a social interaction. Yet we should be careful to

draw too strong conclusions from these results, as the fade effect in and of itself might have an effect on the temporal aspects of the interaction. Triggering the fade effect effectively renders a participant blind for two seconds, wherefore one might reason that the fade effect leads to a de facto two second time penalty per trigger; one could even quite rightly reason that the difference between the groups can be explained by the time penalty imposed by the fade effect. We did however control for this, and found that the fade effect increased the length of a turn by 13.8% when triggered. We would however be amiss to not consider the huge variability between how many times the fade effect did trigger in the experiment group. One participant - an extreme outlier - managed to trigger the fade effect no less than 78 times due to poor compliance with instructions, thus racking up what amounted to a huge time penalty. This was not the only outlier in the data and as such it is at this point premature to put down the time difference between the groups to only the fade effect. However, the two second time penalty imposed by the fade effect probably did have some effect in the grand scheme of things - a 13.8% time increase adds up incrementally - some trials as a whole were probably lengthened by a few seconds due to the fade effect. Claiming that the fade effect in and of itself had no effect, especially in the case when the amount of fade effect went up to 78, would be quite preposterous given that this specific participant spent no less than 156 seconds completely blind. Overall, the 2 second time penalty imposed by the fade effect probably cannot explain the time difference between the groups by itself, but it probably did have some effect. It is of course possible that the fade effect might have affected the performance of the participants in other ways than by simply imposing a 2 second time penalty: for example, the fade effect might have disrupted visuomotor integration through a break in visual attention (see Huang et al., 2022), which could lead to the participants becoming temporarily disoriented.

Another finding was that time per turn decreased over trials quite sharply. The experiment group in particular improved significantly over the course of the experiment; at turn 1 they took on average 31 seconds, but at turn 9 they were down to 11 seconds. This becomes remarkable when contrasted with the control group who went from 19 seconds on average at turn 1 to 9 seconds at turn 9. The experiment group not only improved relatively more than the control group, but became almost as fast as the control group. This indicates that the participants in the experiment group seem to have been able to get used to their perceptual constraints rapidly, and managed to perform well despite their limited perception. This finding is not dissimilar to the findings of Froese et al. (2014b), who found that participants quickly became skilled at a social task in a computer mediated one-dimensional environment, thus showing humans to be very skilled at social tasks despite obvious perceptual limitations. Before we get to the task of hailing the human ability to overcome any and all obstacles, we should note that the variations between the length of individual turns in the experiment group in the present work were very wide.

Thus we cannot be entirely certain if the improvement over time of the experiment group only can be attributed to human adaptability, as it is possible that outliers in the data could have an effect on the results; this should be investigated through replications with a larger sample size than the current work.

Hypothesis 2 was not supported by our data. When looking at the task itself, it may not be strange since the task requires the participants to rotate their heads so they can see the boxes which contain the building blocks necessary for task completion. However, the experimental group did not seem to show more extreme head rotations in comparison to the control group. This is surprising considering they had a canvas limiting their FOV, requiring them to rotate their head more than the control group in order for them to see the entire box. These results are not in line with what was speculated by Slater (2009), who discussed that being in a VR environment with limited FOV could lead to more extreme head movements in order to capture the same visual information as in reality with a normal FOV. However, other studies showed that more head movements did not affect the experience of the interaction in mobile robotic telepresence (Bamoallem et al., 2016), and that head movements did not covary with reaction time in visuomotor tasks (Pratviel et al., 2021). Pratviel et al. (2021) found that good performance in visuomotor tasks could be achieved by both fixed head movement and by excessive head movement, but they saw that head movements did not covary with reaction time, in which they concluded that large head movements can not qualify as an effective strategy. We speculate the same may be true in our study, that excessive head movements may not be the strategy that the participants in the experimental group employed in order to combat their constraints.

We did not investigate what position in space participants' head was rotated towards most, but we encourage future studies too look more into different types of analyses of head movement to get a more clear view on how humans employ head movements in social interaction, in order for the field of HRI to gain a greater understanding on how to implement head movements in robots. According to our data however, it didn't seem like the perceptual constraints that we implemented affected the head movements in the experimental group.

It would be interesting to look into what kind of strategies humans employ when facing perceptual limitations. One initial anecdotal observation that we noticed during the trial of the participant with the most fades was that they moved their head and torso very fast and almost frantic in comparison to other participants in the experiment group, whose movements were more slow and careful. We cannot formally confirm these observations from our data, but we speculate that the risk of triggering the fade effect increases with speed of movement, and that one strategy employed is to move more slowly and carefully. This is something that can be investigated in future research.

Hypothesis 3 was supported by our data, although with

a small margin, with a difference of 9.96%. This is still a distinction nonetheless regarding the quality of the interaction between the two groups. However, with such small differences, further research is needed to confirm and support these results. The majority of the participants, regardless of group, expressed that they thought the experience was overall enjoyable. This could be explained either by their lack of experience in VR or the personal connection to the researchers. One of the constrained participants reported how they thought the experience was interesting and how it expanded their understanding of VR despite how annoying it was not being able to move their head freely. Another participant explained that since they were an acquaintance of the confederate, the task seemed more engaging and the interaction more enjoyable. Although a few constrained participants described their annoyance at the fade effect and how they lost immersion because of it, some may have thought of the constraint perception as only a part of the game. One of the constrained participants in particular reported how they thought the task was more engaging and fun because of the constrained perception, since the task became more challenging due to it. These reports suggest that there are more factors at play other than the perceptual constraints in regards to the quality of a social interaction in VR.

Work process

We started the project by familiarising ourselves with Unity and C# since none of us had extensive experience with these before. Research was done by reading documentations and following YouTube tutorials. In parallel, we also discussed what interaction phenomena we wanted to study and test by which resulted in the decision of focusing on joint attention. To gather inspiration, we did a literature study and searched for articles that were relevant to our research. This process continued in parallel with the implementation phase of our project since we continually made minor changes to what we wanted to investigate. To find articles, we searched the databases LubSearch and Google Scholar by using different combinations of keywords such as "human-robot interaction", "social interaction", "joint attention", "restricted perception", "social robotics", "tower game", "social presence", and "virtual reality". We read the abstract, and if it seemed relevant to our research, the article was saved in a shared folder that everyone had access to for further reading. Different interaction settings were discussed in the beginning but we decided quite quickly that we wanted to test this through some sort of game. The ideas we discussed were tic-tac-toe, some sort of "Disarm the bomb", and a tower game. The tower game was decided because of its cooperative nature, and it would not be too difficult for us to implement it. However, we noticed that it would be complicated to measure joint attention since we most likely would have to look at eye gaze through eye tracking, or find some system that annotates instances of joint attention. This would be more complicated in VR, and given our limited experience, we decided to look at social efficiency by measuring time and have a questionnaire that

measures some aspect of the social interaction.

After discussions with our supervisor, we decided upon implementing a technique for the experiment where one of the researchers will be a confederate interacting with the participant. This decision was made in order to achieve a more controlled test environment, consequently facilitating the process of collecting and analysing data. The same confederate was attendant in every session, except for two sessions because of sickness, in which one of the other researchers stepped in as a substitute. Because the confederate did not have strict rules to follow, we judged that it would not impose a risk for our results to have another researcher as a confederate for two sessions.

The implementation phase began with setting up our environment and the multiplayer functionality in Unity (see two previous sections for more on design rationale and implementation). The multiplayer functionality ran into some difficulties, due to the complexity of Unity's own multiplayer solution, Netcode for GameObjects, and our own lack of experience with Unity and C development. Due to the difficulties experienced, we made the decision to explore third-party multiplayer services and chose Photon Pun 2 (Exit Games, n.d.-a) as our multiplayer solution.

Furthermore, there were configuration difficulties regarding constructing the avatars. The first issue was when importing the avatar package from Didimo, since Unity in the computers used at Lund University was initially not rendered with the Universal Render Pipeline (URP), thus the avatars didn't instantiate correctly. The solution was to download the URP and reconfigure the Package Management. The next obstacle was during the animation process. Due to many YouTube tutorials using deprecated functions and objects (e.g. XR Rig) the mapping of the VR controllers and HMD wasn't accurate with the avatars. By rewriting the source code and continually testing we were able to find a solution at last. Since the software Plastic SCM was used for parallel programming, we experienced additional issues with merging, in which we received compilation errors and difficulties with merging the different versions of the project. After troubleshooting, we realised that many of these issues were grounded in packages and assets not being installed and imported correctly (or at all) on the specific computer. After installing and importing the assets and packages required on the computers used, everything worked smoothly except for the avatars from Didimo, which we couldn't get working. Our solution was to switch avatars to the Robot Kyle avatar provided by Unity.

After this, we had to implement the restricted perception of EPI to the player, which meant manipulation of head movements, the field of view, and camera resolution. To simulate only 2 degrees of freedom for rotation, we thought of disabling the rotation tracking for one axis (i.e. roll). This turned out to be hard to implement, and when looking for answers, we only saw recommendations against this type of solution because it can potentially cause simulator sickness. As an alternative solution, we decided to implement a fade-

out function when the HMD registers a roll outside a certain range (see previous sections for details on implementation).

We also tried to limit the FOV directly in the user interface but noticed that our change was overridden by the system. As a solution, we tried to implement this by more or less attaching "blinders" to the avatar. See the previous section for implementation details. The camera resolution, like the FOV, is hardcoded in the Valve Index, whereby a solution external to the HMD is necessary. We have not found a solution to this, and we decided to omit the effect since we reasoned that the camera resolution in EPI is probably high enough that decreasing the resolution for the participants probably won't make a great difference.

Before carrying out the planned experiment, we evaluated our design by means of a pilot experiment. Here we tested how participants reacted to our VR environment and the perceptual constraints in particular (and investigated whether the constraints lead to unbearable simulator sickness through the Simulator Sickness Questionnaire (Kennedy et al., 1993)). This also gave us an opportunity to try our post-test questionnaire. During the first round of the pilot experiment we observed that the constrained participants didn't fully understand the cause of the fade effect, as mentioned in , we decided to adjust our procedure which resulted in a better outcome (participants were able to establish the causal relationship between head rotation and the fade effect) in the second round.

The adjustment was incorporated into the experiments which were carried out during some two weeks in the middle of December. The experiments generally went smoothly with only some minor hiccups, for example the data of one participant was lost due to a software bug in Unity. Nonetheless, we completed data collection a few days before Christmas, and the data was analysed over the holiday season using ELAN and R.

We discussed whether the time for each turn could be automatically computed, rather than manually during video analysis, but decided against this since there was a risk that an automated script would not catch any errors that might occur. As an example, one participant put a block into place and thought it was correct, so the confederate continued on with their turn, but it was shown later that the block never snapped into place, thus this was a "fake turn". Had we used a program, it would have led to a systematic error in the measurement since it would measure wrong checkpoints. Since we managed to catch this mistake during video analysis, we decided to measure this "fake turn" as a real one. The statistical analysis required a lot of processing of the data and working with distributions that were new to us, so the process was complicated at times, but generally it went well without major complications.

After results were extracted, the only remaining task was to complete the report; the fruits of that work can be found in preceding sections of the present text.

Limitations

There were some notable observations as well as reported deviations within the test sessions worth mentioning as these may have an impact on the results. There were some technical issues regarding the synchronisation of the objects in the VR environment for four participants (from both groups), leading to some distractions, thus extra time was spent in confusion for these participants during the game.

Although all participants had normal colour vision, four of them were in need of glasses and had to remove them during the game for the HMD to fit which may have resulted in inferior vision in the VR environment for these participants. However, according to reports from these particular participants, their vision was sufficient to see and distinguish all objects in the game. In the experiment group, two participants were notably affected by the fade effect a higher number of times in comparison to others. There's a possibility that these participants had not remembered or misunderstood the given instructions, and hence didn't fully achieve the proper behaviour that was being tested in the experiments. Additionally, one participant in the test group had trouble regarding the use of the VR controllers which led to difficulties picking up objects during the game which took time. As we have mentioned before, some participant mistakes occurred, leading to some participants playing ten turns instead of nine. One participant began putting the first block into place, even though the confederate was supposed to start, because they began the game before we gave the instructions that the confederate was to start first. Another participant put a block into the snap zone, thinking they were done, but in the subsequent turn it turned out that this block never snapped into place. For future versions of this experiment, using the tower game, it would be beneficial to provide the participants feedback on when the block has snapped into place to avoid these kinds of situations. During one turn, one participant dropped their building block just before they put the block into the snap zone, but this was treated as a complete turn, and the tower was treated as complete with one block missing.

Another limitation with this study is that we have a small participant pool that is homogeneous regarding age and occupation (most were students or recent graduates). As we also have mentioned before, many of the participants were acquaintances with one of the researchers, so playing the game with someone of the researchers being in the room may have affected how enjoyable the interaction was, which may act as a confound. We have to be cautious about our results because of this, as well because this research question in general is still to a great extent unexplored.

As also previously discussed, the implementation of a constrained head movement through the usage of a fade effect may have contributed to the results regarding the time taken for the experimental group to complete the task.

Future work

The scope of this work is very limited but it lays a foundation for future works to expand upon and it introduces aspects that need further exploration. Additionally this experiment needs to be tested on a larger scale, e.g. with a larger number of participants in a variety of age groups, to validate the presented results and strengthen our analyses. One approach to further expand this field is to look at alternative ways to implement the restriction of head movements. Another proposal is to study other turn-taking interactions apart from game related ones, for example non-conversational interactions or conversational ones where each person is located separately, to further investigate how the quality of interaction would be affected when the persons in the interaction are not adjacent in real life. Other alternatives to measurements, for instance using eye-tracking to study joint attention, a motion tracking-based system to let participants manipulate a robot's body (as in Thellman et al., 2017) or other questionnaires regarding engagement and immersion to analyse participants' experiences, are additional suggestions for future works. Studies about perceptual abilities of robots in the field of HRI are needed to bring nuances and to expand the research field. The importance of taking into account the inherent perceptual limitations of the robot in HRI is strengthened by Sequeira et al. (2016), as they explained that in order to extract meaningful information that is potentially useful for the design of social robots, their perceptual and behavioural limitations have to be taken into consideration during the interaction studies.

8 Conclusion

The purpose of this study was to investigate the effects of constraining the perception of participants engaged in a social block building game in a way akin to the robot EPI; simulation of the perceptual constraints was done using virtual reality, wherein participants engaged in a game with a confederate. We found that the duration of constrained participant turns in the game were longer in comparison to controls, that the constrained participants took up more time of the overall game compared to controls, and that the constrained participants rapidly adapted to the constraints to the point that they became almost as skilled as the controls at the game. Furthermore, the control participants seemed to enjoy the experience slightly more than constrained participants, and reported feeling slightly more engaged in the social experience. However, our hypothesis that constrained participants should move their heads more erratically than controls was not supported. Further work is needed to establish whether these results might have any bearing on future robot design, as this work was exploratory in nature and would benefit from further refinements to the research method.

References

- Adafruit Industries. (n.d.). *Zero spy camera for raspberry pi zero*. Retrieved on 12 October 2022, from <https://www.adafruit.com/product/3508>
- Alfano, P., & Michel, G. (1990, 03). Restricting the field of view: Perceptual and performance effects. *Perceptual and motor skills*, 70, 35-45. doi: 10.2466/PMS.70.1.35-45
- American Association of Psychology. (n.d.). *Apa dictionary of psychology*. Retrieved on 23 October 2022, from <https://dictionary.apa.org/simulator-sickness>
- Bamoallem, B. S., Wodehouse, A. J., Mair, G. M., & Vasanthan, G. A. (2016). The impact of head movements on user involvement in mediated interaction. *Computers in Human Behavior*, 55, 424-431. doi: <https://doi.org/10.1016/j.chb.2015.09.016>
- Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijser, M., & Šabanović, S. (2020). *Human-robot interaction: An introduction*. Cambridge University Press. doi: 10.1017/9781108676649
- Cason, T., Lau, S.-H., & Mui, V.-L. (2013, 10). Learning, teaching, and turn taking in the repeated assignment game. *Economic Theory*, 54. doi: 10.1007/s00199-012-0718-y
- Castaño Estrella, D. (n.d.). *Camera fade*. Retrieved on 7 September 2022, from <https://assetstore.unity.com/packages/vfx/shaders/fullscreen-camera-effects/camera-fade-155626>
- Exit Games. (n.d.-a). *Photon engine*. Retrieved on 5 October 2022, from <https://www.photonengine.com/en-US/Photon>
- Exit Games. (n.d.-b). *Pun Basics Tutorial*. Retrieved on 24 October 2022, from <https://doc.photonengine.com/en-us/pun/v2/demos-and-tutorials/pun-basics-tutorial/intro>
- Froese, T., Iizuka, H., & Ikegami, T. (2014a, jan). Embodied social interaction constitutes social cognition in pairs of humans: A minimalist virtual reality experiment. *Scientific Reports*, 4(1). doi: 10.1038/srep03672
- Froese, T., Iizuka, H., & Ikegami, T. (2014b). Using minimal human-computer interfaces for studying the interactive development of social awareness. *Frontiers in Psychology*, 5. doi: 10.3389/fpsyg.2014.01061
- Herath, D., Jochum, E., & Vlachos, E. (2018, 12). An experimental study of embodied interaction and human perception of social presence for interactive robots in public settings. *IEEE Transactions on Cognitive and Developmental Systems*, 10, 1096 - 1105. doi: 10.1109/TCDS.2017.2787196
- Hopper, J. E., Finney, H., & Jones, J. A. (2019). Field of view and forward motion discrimination in virtual reality. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 1663-1666). doi: 10.1109/VR.2019.8797756
- Huang, Y., Zhong, S., Zhan, L., Sun, M., & Wu, X. (2022, 01). Sustained visual attention improves visuomotor timing. *Psychological Research*, 86, 1-8. doi: 10.1007/s00426-021-01629-9
- Inoue, K., Nonaka, S., Ujiie, Y., Takubo, T., & Arai, T. (2005). Comparison of human psychology for real and virtual mobile manipulators. In *Roman 2005. ieee international workshop on robot and human interactive communication, 2005*. (p. 73-78). doi: 10.1109/ROMAN.2005.1513759
- Johansson, B., Tjøstheim, T. A., & Balkenius, C. (2020). Epi: An open humanoid platform for developmental robotics. *International Journal of Advanced Robotic Systems*, 17(2), 1729881420911498. doi: 10.1177/1729881420911498
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), 203-220. doi: 10.1207/S15327108IJAP0303_3
- Kiselev, A., Kristoffersson, A., & Loutfi, A. (2014, 03). The effect of field of view on social interaction in mobile robotic telepresence systems. In (p. 214-215). doi: 10.1145/2559636.2559799
- Levinson, S. C. (2016). Turn-taking in human communication – origins and implications for language processing. *Trends in Cognitive Sciences*, 20(1), 6-14.
- Lin, J.-W., Duh, H., Parker, D., Abi-Rached, H., & Furness, T. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings ieee virtual reality 2002* (p. 164-171). doi: 10.1109/VR.2002.996519
- Max Planck Institute. (2022). *ELAN*. Retrieved from <https://archive.mpi.nl/tla/elan>
- Miller, H., Bugnariu, N., Patterson, R., Wijayasinghe, I., & Popa, D. (2017, August 10). Development of a novel visuomotor integration paradigm by integrating a virtual environment with mobile eye-tracking and motion-capture systems. In *2017 international conference on virtual rehabilitation, icvr 2017*. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICVR.2017.8007481
- MindPort. (n.d.). *VR Builder Manual*. Retrieved on 23 October 2022, from <https://www.mindport.co/company>
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137-156. doi: 10.1016/j.pragma.2014.04.004

- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83(October), 435–50. doi: 10.2307/2183914
- Nvidia Corporation. (n.d.). *GeForce Experience*. Retrieved on 7 December 2022, from <https://www.nvidia.com/sse/geforce/geforce-experience>
- Oh, C., Bailenson, J., & Welch, G. (2018, 10). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, 5. doi: 10.3389/frobt.2018.00114
- O'Regan, J., & Noë, A. (2001, 11). A sensorimotor account of vision and visual consciousness. *The Behavioral and brain sciences*, 24, 939-73; discussion 973. doi: 10.1017/S0140525X01000115
- Pratviel, Y., Deschodt-Arsac, V., Larrue, F., & Arsac, L. M. (2021, December). Reliability of the Dynavision task in virtual reality to explore visuomotor phenotypes. *Scientific Reports*, 11(1). doi: 10.1038/s41598-020-79885-9
- Ramsey, R., Kaplan, D. M., & Cross, E. S. (2021). Watch and learn: The cognitive neuroscience of learning from others' actions. *Trends in Neurosciences*, 44(6), 478-491. doi: <https://doi.org/10.1016/j.tins.2021.01.007>
- Riek, L. D. (2012). Wizard of oz studies in hri: A systematic review and new reporting guidelines. , 1(1). doi: 10.5898/JHRI.1.1.Riek
- Sadka, O., Giron, J., Friedman, D., Zuckerman, O., & Erel, H. (2020). Virtual-reality as a simulation tool for non-humanoid social robots. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3334480.3382893
- Salter, D. A., Tamrakar, A., Siddiquie, B., Amer, M. R., Divakaran, A., Lande, B., & Mehri, D. (2015). The tower game dataset: A multimodal dataset for analyzing social interaction predicates. In *2015 international conference on affective computing and intelligent interaction (aci)* (p. 656-662). doi: 10.1109/ACII.2015.7344639
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. doi: 10.1017/S0140525X00005756
- Sequeira, P., Alves-Oliveira, P., Ribeiro, T., Di Tullio, E., Petisca, S., Melo, F. S., ... Paiva, A. (2016). Discovering social interaction strategies for robots from restricted-perception wizard-of-oz studies. In *2016 11th acm/ieee international conference on human-robot interaction (hri)* (p. 197-204). doi: 10.1109/HRI.2016.7451752
- Shin, S., Crapse, T. B., Mayo, J. P., & Sommer, M. A. (2009). Visuomotor integration. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of neuroscience* (pp. 4354–4359). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-29678-2_6384
- Slater, M. (2009, 12). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364, 3549-57. doi: 10.1098/rstb.2009.0138
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Scholtz, A., & Goodrich, M. (2006). Common metrics for human-robot interaction. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/1121241.1121249
- Szymanik, J., Meijering, B., & Verbrugge, R. (2013). Using intrinsic complexity of turn-taking games to predict participants' reaction times. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35.
- Thellman, S., Lundberg, J., Arvola, M., & Ziemke, T. (2017). What is it like to be a bot? toward more immediate wizard-of-oz control in social human-robot interaction. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3125739.3132580
- Unity Asset Store. (n.d.). *POLYGON office building*. Retrieved on 16 September 2022, from <https://assetstore.unity.com/packages/3d/environments/urban/polygon-office-building-82282#publisher>
- Unity Technologies. (n.d.-a). *Animation Rigging*. Retrieved on 23 October 2022, from <https://docs.unity3d.com/Packages/com.unity.animation.rigging@1.0/manual/index.html>
- Unity Technologies. (n.d.-b). *MonoBehaviour.FixedUpdate()*. Retrieved on 5 January 2023, from <https://docs.unity3d.com/ScriptReference/MonoBehaviour.FixedUpdate.html>
- Unity Technologies. (n.d.-c). *Rigid Body*. Retrieved on 23 October 2022, from <https://docs.unity3d.com/ScriptReference/Rigidbody.html>
- Unity Technologies. (n.d.-d). *Unity*. Retrieved on 5 October 2022, from <https://unity.com/>
- Unity Technologies. (n.d.-e). *XR Interaction Toolkit*. Retrieved on 5 October 2022, from <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.1/manual/index.html>
- Unity Technologies. (2019). *Space Robot Kyle*. Retrieved on 28 October 2022, from <https://assetstore.unity.com/packages/3d/characters/robots/space-robot-kyle-4696>
- Valve Corporation. (n.d.-a). *Interaction System from the Lab*. Retrieved on 23 October 2022, from https://valvesoftware.github.io/steamvr_unity_plugin/articles/Interaction-System.html
- Valve Corporation. (n.d.-b). *SteamVR*. Retrieved on 5 October 2022, from <https://www.steamvr.com/en/>

Valve Corporation. (n.d.-c). *Valve Index*. Retrieved on 5 October 2022, from <https://www.valvesoftware.com/en/index>

Vuarnesson, L., Zamplaras, D., Laroche, J., Dumit, J., Lutes, C., Bachrach, A., & Garnier, F. (2021). Shared diminished reality: a new vr framework for the study of embodied intersubjectivity. *Frontiers in Virtual Reality*, 2, 646930.

Weistroffer, V., Paljic, A., Fuchs, P., Hugues, O., Chodacki, J.-P., Ligot, P., & Morais, A. (2014). Assessing the acceptability of human-robot co-presence on assembly lines: A comparison between actual situations and their virtual reality counterparts. In *The 23rd ieee international symposium on robot and human interactive communication* (p. 377-384). doi: 10.1109/ROMAN.2014.6926282

Wikipedia. (n.d.). *Category: Android (robot) films*. Retrieved on 6 October 2022, from [https://en.wikipedia.org/wiki/Category:Android_\(robot\)_films](https://en.wikipedia.org/wiki/Category:Android_(robot)_films)

Appendix

The experiment questionnaire reported with mean and median for each questionnaire item, divided between the experimental group (B) and the control group (K).

Questionnaire item	Mean for group K	Mdn for group K	Mean for group B	Mdn for group B
Age	24.5	24	24.33	24
How much experience do you have with Virtual Reality?	1.8	1	2	2
I felt involved in our interaction	4.7	5	4.22	4
I felt detached during our interaction	1.3	1	2.11	2
I did not feel that there was a sense of social distance between me and the other player	3.6	4	3	3
I felt that our interaction was stimulating	4.3	4.5	4.22	4
I felt that there was a sense of distance between me and the other player	2.2	2	2.77	3
I tried to create a sense of closeness between us	3.3	3	3	3

Questionnaire item	Mean for group K	Mdn for group K	Mean for group B	Mdn for group B
How well were you able to observe the body language of the other player?	3.5	4	3.11	3
How often did you make a sound (for example laugh and speak) in response to the other player?	3.9	4	3.77	4
How often did you smile in response to the other player?	3.5	3.5	3.11	3
How often did you want to or did you speak to the other player?	4.5	4.5	4.44	4
To what extent did you feel mentally immersed in the experience?	4.4	5	3.77	4
How involving was the interaction?	4.6	5	4	4
How engaging was the game?	3.8	4	3.77	4
How was the overall movement quality during the interaction?	3.4	3	3.55	3
Overall, how satisfying, or enjoyable was the interaction?	4.4	4.5	4.11	4

Questionnaire item	Mean for group K	Mdn for group K	Mean for group B	Mdn for group B
To what extent did you feel constrained by the field of view?	1.7	2	2.88	3
To what extent did you feel you could move your head as you wished?	4.5	5	2.33	2
To what extent did you feel annoyed by not being able to move your head as you wished?	1.3	1	3.55	4

Enkät om simulatorsjuka

I denna enkät kommer du bli ombedd om att svara på några frågor gällande hur du kände dig när du befann dig i VR-miljön. Detta för att undersöka huruvida du upplevde simulatorsjuka, vilket är en form av åksjuka. Det är viktigt att du svarar sanningenligt i enkäten. Fyll ut enkäten i din egen takt; den bör inte ta mer tid än 3 minuter. Dina svar kommer att vara anonyma.

* Required

1. Jag kände mig generellt obekvämt under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

2. Jag kände mig utmattad av VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

3. Jag fick huvudvärk av VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

4. Jag kände ansträngning i mina ögon under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

5. Jag saliverade mycket under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

6. Jag svettades mycket under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

7. Jag kände mig illamående under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

8. Jag hade svårt att koncentrera mig under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

9. Mitt huvud kändes tungt under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

10. Min syn kändes suddig under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

11. Jag kände mig yr när mina ögon var **öppna** under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

12. Jag kände mig yr när mina ögon var **stängda** under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

13. Det kändes som att världen snurrade under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

14. Min mage gjorde sig känd under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

15. Jag behövde rapa under VR-upplevelsen. *

Mark only one oval.

- Inte alls
- Lite
- Ganska mycket
- Väldigt mycket

16. Vill du tillägga något övrigt?

This content is neither created nor endorsed by Google.

Google Forms



LUNDS
UNIVERSITET

SAMTYCKE TILL DELTAGANDE I FORSKNINGSPROSJEKT

Mia Huovilainen, Vanessa Nguyen, Alexander Wåhlander

Du kommer att få delta i ett forskningsprojekt som genomförs av Mia Huovilainen, Vanessa Nguyen och Alexander Wåhlander, inom ramen för ett samarbete mellan mastersstudenter i Kognitionsvetenskap och från Lunds Tekniska Högskola vid Lunds Universitet. Syftet med forskningen är att studera en interaktion i Virtual Reality. Interaktionen kommer att skärmminspelas, vari datan kommer att analyseras.

Forskningsmaterialet kommer hållas konfidentiellt i den utsträckning som aktuell lag föreskriver och kommer inte presenteras på ett sätt som möjliggör identifikation. Dina svar och dina resultat kommer att behandlas så att inte obehöriga kan ta del av dem. Om du motsätter dig någon av användningarna av det inspelade materialet så kan avtalet anpassas.

Mitt deltagande är helt och hållet frivilligt och jag känner att jag fått tillfälle att få mina frågor kring studien besvarad och/eller vet vem jag ska vända mig till för att få svar på mina frågor. Jag är medveten om att jag när som helst under experimentets gång kan avbryta mitt deltagande utan att behöva förklara varför.

Med din underskrift nedan intygar du att du har läst och förstått informationen ovan, att du samtycker till inspelning av material samt ovan nämnda bruk av materialet.

Lund den ____ / ____ 2022

Underskrift

Namnförtydligande



LUNDS
UNIVERSITET

CONSENT FORM FOR PARTICIPATION IN RESEARCH PROJECT

Mia Huovilainen, Vanessa Nguyen, Alexander Wåhlander

You are going to participate in a research project conducted by Mia Huovilainen, Vanessa Nguyen, and Alexander Wåhlander, within the framework for a collaboration between master's students in Cognitive Science and from the Faculty of Engineering at Lund University. The aim with this research is to study an interaction in Virtual Reality. The interaction will be screen recorded in which the data will be analyzed.

The research material will be kept confidential to the extent prescribed by relevant law, and will not be presented in a way that will enable identification. Your answers and your results will be treated in a way such that unauthorized persons will not have access to them. If you oppose any type of use of the recorded material, the agreement can be modified.

My participation is entirely voluntary and I feel I've been given the opportunity to have my questions about the study answered and/or I know who to turn to have my questions answered. I am aware that I can at any time during the experiment revoke my participation without having to explain why.

With your signature below you affirm that you have read and understood the information above, and that you consent to recording of material as well as the above mentioned use of the material.

Lund ____ / ____ 2022

Signature

Name clarification

Under dina fötter - på Uppåkra - den arkeologiska metoden som ingång till källkritik och källtillit

Madeleine Loostrand
li6870lo-s@student.lu.se

Charlotta Ekelöw Nilsson
ch6315ni-s@student.lu.se

Anna Dahlström
an7360da-s@student.lu.se

I denna rapport redogörs för processen att ta fram ett lärspel för Uppåkra Arkeologiska center, numera Uppåkrastiftelsen, för årskurs 3 och 4. Resultatet är ett digitalt arkeologispel, som vi valt att döpa till "Under dina fötter", där eleven får genomföra en typ av "skattjakt" och bekanta sig med intressanta fynd från platsen. Spelet, som utspelar sig ovanför den ännu inte utgrävda järnåldersboplatsen, låter barnen i rollen som arkeolog få "gräva upp" fynd, lära sig om dätidens Uppåkra genom dem, samtidigt som spelet behandlar frågor om hur vi kan veta något om det förgångna utifrån spår och fragment från därriden. Fynden hanteras alltså som källmaterial, och blir ett konkret sätt för åldersgruppen att erfara hur fynd kan bärta på information, som kan tolkas på många sätt, men få välgrundade slutsatser genom källkritiska och vetenskapliga metoder. Spelet är förutom en sökning efter fynd ett slags frågespel, där spelaren får alternativ mellan vilka slutsatser om därriden som går att dra utifrån ett visst fynd. Spelaren bjuds därmed in till att reflektera över de epistemiskt och ontologiskt grundade frågorna: hur var därriden, och hur vet vi det? Under den konceptuella och tekniska utvecklingen av spelet har vi gjort flera besök på plats för att lära känna miljön, pedagogerna och verksamheten i stort. Vi har tagit stort intryck av dessa besök och möten vid utformningen av spelet. En grundläggande ambition vi haft har varit att förstärka upplevelsen av Uppåkra som en historiskt viktig plats, att med digitala medel låta barnen erfara att den till synes helt vanliga åkern i själva verket har gömda "skatter" under marken, och bjuda in dem att "göra saker" i och med miljön. Vår målsättning har varit att få en känsla av att spel och plats smälter samman, att spelet "händer" i verkligheten. Läsplattan får i spelet flera funktioner, och förutom att hantera text, bilder och "själva spelet" förvandlas den under spelets gång till olika arkeologiska redskap såsom metalldetektor, spade, skärslev och såll, och spelaren "arbetar" med rörelser för att få fram fynd ur jorden. Läsplattans kamera aktiverad under sökning och förvandlar den till ett fönster, så att marken under syns. Spelet utspelar sig därmed inte enbart "i" läsplattan, utan i samspelet mellan platsen och spelaren och dennes handlingar. Här har vi tagit utgångspunkt i principer inom teoribildningar som situerat och transaktionellt lärande, med fokus på att lärandet sker inom kontextuella villkor och uttrycks genom handlingar. Det är perspektiv som rimmar med Uppåkrastiftelsens pedagogiska vision, där barns deltagande och lust att lära är centrala fokus.

1 Introduktion

Det förgångna dör aldrig. Det är inte ens förgångenget. - William Faulkner

Arkeologiskt arbete - att utvinna kunskap ur tysta material - med hjälp av vetenskaplig metod, är ett intressant möte mellan det konkreta, nutida, och det förgångna. Frågorna tycks hoppa sig vid tanken på att stöta på ett ting som vilat i jorden i tusen år. Vad är det? Vem har haft det? Hur användes det, och till vad? Hur hamnade det här? Hur var det att leva då? Att dra slutsatser om hur därridens människor levde och hade det utifrån en liten keramikbit kan förefalla en mystisk och dunkel process. Vi tyckte därmed att det skulle vara intressant att låta barnen få agera som arkeologer för att avtäcka och demystifiera hur detta "kunskapande" går till. Inom arkeologin är källmaterialet hos äldre fynd ofta svårtolkat jämfört med källmaterial inom exempelvis samhällsvetenskapen. För att få ett "stumt" källmaterial att berätta om gångna tiders händelser som rör mänsklig verksamhet behöver arkeologer därmed arbeta fram logiska slutledningar, arbetshypoteser, bedöma deras sannolikhet, och falsifiera/verifiera hypoteser. Den källkritiska metoden är därmed en metod för att identifiera svaga punkter i logiska resonemang och förhindra förhastade slutsatser (Blomqvist, 1989). Hur visar man för ett barn den vetenskapliga process som gör att vi utifrån en liten oansenlig keramikbit kan veta saker om därriden? Kan man få barn i 9-10 årsåldern att förstå något så nyanserat som att det finns olika grader av sannolikhet i vetenskapliga berättelser, att spekulation kan förädlas till en välgrundad slutsats med vetenskaplig metod?

Uppåkra erbjuder en unik och autentisk kontext för att upptäcka att fynd är källor med potential att berätta om historien. Att se den ouppgrävda åkern, där historien ligger dold, och samtidigt hitta fynd, som tidsspår, är en spännande kontrast mellan gåta och spår. Fynden är konkreta, spänande, och lockar till både fantasi och försiktighet. Frågan som pockar: vad berättar dessa fynd om därridens Uppåkra och de människor som levde där? Följs nästan genast av "hur kan vi veta det?" och här uppstår en spänande kontext för att med arkeologen som guide, och fynden som utgångspunkt, ge konkreta exempel på hur man vetenskapligt kan arbeta fram hypoteser, hur de underbyggs och hur det går till. Genom att visa upp att många olika discipliner sammankommer för att pröva olika hypoteser får eleverna en insyn i hur arkeologisk kunskap "knackas fram" med den källkriti-

iska/vetenskapliga metoden - men att det trots det handlar om berättelser som har olika grader av sannolikhet. Vår utgångspunkt genom projektet har varit att barns källkritiska förmåga kan stötta om de får konkret erfarenhet av hur en arkeolog i samarbete med andra hanterar och värderar information och drar slutsatser i mötet med specifika fynd. Att de genom detta upptäcker att kunskap och vetande är en stegvis process, med gradskillnader, där vissa saker om det vi vet om dätidens Uppåkra är givna, annat är sannolikt, och det jämte det finns spekulationer som väntar på att verifieras eller avfärdas, och att det är så allt vetande ter sig. Att använda en yrkesgruppens verklighet och perspektiv som ingång till källkritik är dock inte något vanligt fenomen i skolans värld. I Skolforskningsinstitutets systematiska forskningsöversikt om undervisning i källkritik i samhällskunskap (2022), som kan sägas vara ett pedagogiskt stöd till lärare, beskrivs ett antal metoder för att undervisa om källkritik. Metoderna delas upp i kategorier där eleverna fått arbeta med (i) texter, (ii) multimedia eller (iii) diskussion som verktyg för att stärka den källkritiska förmågan. Exempel på situerad och social inlärning - såsom att lära sig källkritik genom att se hur det används inom en yrkesgrupp, under guidning av en företrädare för samma yrkesgrupp, finns ärenot inget exempel på. Det är intressant, eftersom källkritik är en metod, och många andra (ofta praktiska) metoder lärs ut genom lärlingsskap. När vi fick möjligheten att utveckla en lärresurs för Uppåkrastiftelsen för årskurs 3 och 4, och lärde känna den pedagogiska verksamheten och hur arkeologerna jobbade, såg vi en möjlighet att använda yrkeskunndet hos arkeologen som utgångspunkt för att i samspelet med eleverna "göra" källkritik. Genom att lotsa eleverna genom den arkeologiska processen med att utvinna kunskap ur "tysta" källor (fynd) låter vi elever på ett undersökande och delaktigt sätt få möjlighet att arbeta utifrån källkritiska frågeställningar.

2 Teori

2.1 Uppåkramodellen

Uppåkra arkeologiska center, numera Uppåkrastiftelsen, har i samarbete med forskare och pedagoger utvecklat "Uppåkramodellen", där man genom att lyfta delaktighet och sinnesaktiverande aktiviteter vill främja nyfikenhet och ge förutsättningar för lärande (Stiftelsen Uppåkra Arkeologiska Center, 2022). Widén & Dahlbeck utkom 2022 med en slutrapport kring pilotprojektet Uppåkramodellen och de nämner *undersökande* och *variation* av lärandemiljöer och arbetssätt som bidragande faktorer till det goda lärandet. "Arkeologiskolan", där femteklassare får gräva tillsammans med arkeologer och hitta fynd som sedan analyseras, är ett exempel på en pedagogisk aktivitet med hög grad av undersökande och delaktighet. Enligt Widéns & Dahlbecks (2022) slutrapport av pilotprojektet Uppåkramodellen väckte upplevelsen av arkeologiskolan en historiemedvetenhet hos eleverna, en känsla av att "*historien finns närvarande i jorden under dem*" (Widen & Dahlbeck, 2022:12). Efter besöket på

arkeologiskolan framkom det att eleverna reflekterade över tid som en del av historia. Eleverna reflekterade över att männskor bott i Uppåkra i så många år och funderade över vad de gjort och upplevt jämfört med vad de själva gör idag. Barnen uttryckte också enligt Widén & Dahlbeck vikten av att "*få leka och prova sig fram till en historisk arkeologisk kunskap*" (Widén & Dahlbeck, 2022:13) och menar på att barnen visade att de fått en förmåga att reflektera kring historiska fakta och att de också kunde använda fiktion för att "*leka sig fram till ett meningsfullt lärande*" (Widén & Dahlbeck, 2022:13). Widen & Dahlbeck beskriver också hur elevernas upplevelser på arkeologiskolan inte endast formas av interaktionen med pedagogen eller det fysiska grävandet utan även genom olika medier och uttrycksformer som bild, film och ljud. Vårt uppdrag var att ta fram en lärresurs för årskurs 3 och 4, som utifrån pedagogernas erfarenheter inte riktigt är mogna för delta i arkeologiskolan, som behandlar källkritik och på något sätt levandegör den fysiska utemiljön. Vi ville ta fasta på metodiken i Uppåkramodellen och införliva detta i vår lärresurs och samtidigt synliggöra den arkeologiska metoden. I vår lärresurs blev detta centrala teman. Vårt fokus blev att skapa nyfikenhet genom att förmedla och levandegöra kunskap på ett lustfyllt sätt, där barnen kroppsligt får utforska och upptäcka fynd på platsen, och vad fynden berättar om därriden. Spelet är däremed designat för att komplettera och förstärka miljön, samtidigt som pedagogerna på plats bidrar till spelet med pedagogisk för- och efterbearbetning.

2.2 Transaktionellt, socialt och förkroppsligat lärande

En viktig del av Uppåkramodellen och dess pedagogik är deltagande och socialt lärande. På samma sätt som exempelvis arkeologiskolan är ett exempel på sociokulturellt lärande så är även vår applikation det. Utifrån ett sociokulturellt perspektiv är våra handlingar och vårt tänkande ihopkopplat med det sociala sammanhanget, och de analytiska gränserna mellan individer, kognitiva processer och språk löses upp (t.ex. Goodwin, 2000, Hutchins, 1995, Säljö, 2000). Lärande kan förstås som ett resultat av delaktighet i en social och kulturell kontext, där deltagare ständigt söker efter gemensamma grunder för att upprätthålla och fortsätta de aktiviteter de deltar i. Att ta arkeologens roll, och i handling få engagera sig med platsen, kan utifrån det ses som ett slags tillblivande som inte handlar om att enbart ackumulera kunskap. Uppåkramodellens fokus på deltagande i en autentisk kontext kan särskilt motiveras utifrån ett transaktionellt perspektiv på lärande. I teoribildningen är relationen mellan elev, situation och miljö i ständig förändring, där mening uppstår i de handlingar som görs, och i genomlevandet av konsekvenserna av dem. (Östman & Öhman, 2010; Östman, 2003). Som ett resultat av att spelet sätter igång processer där barnen "gör" källkritik och arkeologi, förändras relationerna mellan dem själva, platsen och pedagogen (som då viktigt nog, också är arkeolog). Det transaktionella perspektivet förklarar också betydelsen av att deltagarna får gräva fram fynd på platsen eftersom "görandet" förvand-

lar betydelsen av platsen - den förvandlas från åker till en skattgömma. När spelarna får se fynden som bilder och information och illustrationer som kompletterar, förändras betydelsen av fynden - guldbägaren får en annan innebörd när spelaren ser den i händerna på en präst i templet. Att barnen får veta att man kan dra många sorters information och gissningar från fynd gör också att de får uppleva olika kvaliteter hos fyndet - det är inte bara en "sak" utan ett spår, en ledtråd, med många betydelseaspekter. Genom det handbegrifliga i att genom lärspelet lokalisera, gräva och sälla fram fynd, får barnen en känsla för jordens informationsbärande kvaliteter - de försiktiga rörelserna med skärsleven förstärker detta. Genom rörelserna "gör" eleverna principen att jorden är viktig som källmaterial. I det översta lagret kan spaden användas, eftersom jorden blandats runt av plogar under åren och jorden tappat sin informationsbärande kvalitet. Men de fynd som ligger längre ner behöver hanteras med försiktighet och mindre handrörelser eftersom jordlagret i sig bär på information om mänsklig aktivitet. Genom att föra in de fysiska elementen i spelet får spelarna också erfara kulturlagren och att man förflyttar sig längre bak i tid ju djupare ner i dem man kommer. Den viktigaste resuren, som varit vår centrala utgångspunkt vid designen av spelet är inte platsen och fynden - utan pedagogerna och deras erfarenhet som arkeologer. Deras perspektiv och samspel med eleverna är det kraftfullaste verktyget för att skifta barnens uppmärksamhet och "se" annorlunda på fynden. Spelet är bara en del av det som sker under barnens besök på Uppåkra, och tänkt fungera väl med övriga pedagogiska aktiviteter och väcka barnens frågor och nyfikenhet, som kan fångas upp senare. Det spelet kan erbjuda, som är svårt för pedagogerna, är att förvandla åkern till en utgrävningsplats och låta spelarna få egna upplevelser av att gräva fram och hitta fynd. Genom dessa handlingar förändras betydelsen av platsen, fynden, vad det innebär att veta. Vår förhoppning är att det engagerar och fördjupar de samtal som sker mellan barn och pedagog efteråt, när de fortsätter att diskutera vad fynden lär oss om därtiden. Att informationen om fynden, och hur man drar slutsatser om dem, väcker tankar om hur kunskap arbetas fram. Att det ger erfarenhet av den vetenskapliga metoden, och stöttar källtillit såväl som källkritisk förmåga. I den komplexa väven mellan att gräva upp fynd och resonera kring dem utifrån olika frågeställningar och påståenden i en miljö där utgrävningar sker, med arkeologer och pedagoger intill sig, bygger användaren kunskap och förmågor i en sociokulturell miljö, vilket är ett gott exempel på situerat lärande (Lave & Wegner, 1991., De Pietro, 2013).

2.3 Källkritik

Källkritik som begrepp har i Sverige sitt ursprung i en uppsättning metodregler som historiker använder för att avgöra om historiska påståenden är vetenskapligt användbara för ett visst syfte eller ej, samt för att förhindra förhastade slutsatser. Inom andra discipliner syftar processen framförallt till att bedöma informationens trovärdighet och användbarhet, i relation till källan (Skolfor-

skningsinstitutet, 2020). Källkritik är besläktad med andra former av informationsbedömning, såsom *digital källkritik* och *civic literacy*, och de kognitiva processerna involverade i ett "informationsrelaterat hantverkskunnande" har sannolikt gemensamma nämner. I Skolforskningsinstitutets systematiska forskningssammanställning *Kritiskt tänkande och källkritik - undervisning i samhällskunskap* (2022) ses källkritisk förmåga som en central del i *kritiskt tänkande*, med motiveringen att elever i en kritisk tankeprocess ofta utgår från någon typ av information (Skolforskningsinstitutet, 2020). I Guath & Nygrens (2022) studie av vuxnas förmåga till digital källkritik diskuteras kopplingen mellan digital källkritik och *civic literacy* som att en individ med hög informationslitteracitet tycks operera utifrån vissa *frågeställningar*. De ser stora likheter mellan Breiviks förslag att personer med god källkritisk förmåga ställer tre grundläggande frågor "[...]: how do you know that—what evidence do you have for that—who says so—how can **you** find out." (Breivik i Bawden, 2001, p-242-243) och modellen över civic literacy av McGrew et al. (2017, 2018): 1. *Sourcing: Who is behind the information?* 2. *Evidence: What is the evidence?* 3. *Comparing: What do other sources say?*

Källkritik kan sägas ha fler betydelser än de redan nämnda. Inom samhällsvetenskapen kan källkritik handla om att värdera information i relation till källan som förmedlar informationen, bedöma vilken trovärdighet en källa har, och insikten att vissa källor väger tyngre än andra. Det senare hänger ihop med *källtillit*, att vi känner förtroende för att information som kommer från en specifik källa är äkta, oberoende, icke-tendentios och relevant. Källtillit, att lita på att information som verkar trovärdig som kommer från vissa aktörer man bedömer som trovärdiga, bottnar många gånger i att vi har förtroende för att de informationsalstrande processer/metoder källan använder sig av är väl avvägda och genererar tillförlitlig information. Sundin och Heider (2018:61) lyfter fram betydelsen av källtillit:

"Att förhålla sig till information på nätet, eller för den delen utanför, handlar därför inte enbart om kritik, utan lika mycket om tillit och förtroende. Om vi alltid ska ifrågasätta alla institutioner, personer eller dokument omöjliggörs ett demokratiskt samtal. Att inte lita på någonting ligger i praktiken inte långt ifrån att inte ifrågasätta någonting. Även med en grundförståelse som säger att kunskapen alltid är i rörelse måste vi acceptera att vissa källor och institutioner rimligtvis är mer trovärdiga än andra. Källkritik är viktig, men den måste därför paras med vad vi kan kalla för en källtillit och såklart även med allmänbildning och kunskap, även kunskap om hur kunskap skapas och hur vetenskap fungerar."

Källkritisk förmåga är inte någon generell förmåga, utan typiskt sett utvecklas den källkritiska förmågan i takt med och i förhållande till specifika ämneskunskaper. Att ha god källkritisk förmåga inom ett område betyder alltså inte per automatik att förmågan är god inom ett annat. Den verkar

därmed vara kopplad till förmågan att känna igen mönster, upptäcka avvikeler och perceptuell närhet till ett ämne. Källkritisk förmåga uppstår och uttrycks i mötet med information och dess relation till tidigare inlärning. När vi bedömer en källas trovärdighet, är det dess informationsprocessande kvaliteter vi primärt utvärderar. Arkeologisk metod, som är ett källkritiskt förfarande, kan därför vara intressant att använda som utgångspunkt för att studera källkritik, eftersom den på ett så konkret sätt synliggör den vetenskapliga processen med att förädla hypoteser till teori och fakta. Distinktionen mellan fiktion (påståenden som inte kan falsifieras/verifieras) och hypoteser (som kan verifieras och bli teori eller falsifieras och förkastas) kan kopplas till de fynd som eleverna hittar under skattjakten, och i samtal om fynden tillsammans med arkeologen. Genom insyn i den arkeologiska metoden får barnen syn på hur hypoteser kan grundas genom empiri. Fundamentet till att känna källtillit - att lita på vissa källors information blir underbyggd när eleverna får se hur en kvalitativ metodik går till. De av Breivik och McGrew (2017, 2018) identifierade grundläggande källkritiska frågorna "Vilka är bevisen för den slutsatsen? Vad säger andra källor?" har varit utgångspunkter för oss då vi formulerat frågor och information i spelet.

3 Arbetsprocess

Utgångspunkten för projektet var att ta fram en lärresurs för Uppåkra arkeologiska center där spelaren får navigera i miljön, samtidigt som teman kring källkritik, källtillit och forskningsprocessens kollektiva natur behandlas. Projektets inledande fas hade följande iterativa fokusområden; (i) inläsning av material kring källkritik, Uppåkras historia och befintliga pedagogiska program (ii) generering av idéer för hur man kan förhöja upplevelsen av den arkeologiska platsen (åkern) med hjälp av teknik, (iii) bygga förståelse av uppdragsgivarens (Uppåkra arkeologiska center) befintliga pedagogiska program och hur fysiska och digitala lärmiljöer används idag, samt (iv) deras tidigare erfarenheter av digitala resurser, vad som funkat bra och mindre bra.

Av olika anledningar blev vårt första besök på plats i Uppåkra uppskjutet, och vår första kontakt blev via Zoom. I väntan på att se vilka avgränsningar de fysiska miljöerna kunde föra med sig, kom vi därmed att lägga relativt mycket tid på att generera idéer på hur projektet skulle kunna utformas och hur olika tekniska lösningar skulle kunna tillämpas för att göra "åkern" mer spänande - AR, VR, "UppåkraGo", audioguide med spelinslag, var några av alternativen vi skissade på i detta läge. Att göra ett platsbesök på Uppåkra och få se hur pedagogerna - som också är arkeologer - arbetar, samt hur de pedagogiska miljöer som byggts upp användes av dem, gjorde att projektet kom in i en tydligare avgränsningsfas, och vi kunde då pröva våra olika idéer mot förhållandena på plats. Vi kunde också väga in pedagogernas beskrivningar av tidigare erfarenheter av digitala resurser som utvecklats - det visade sig att Uppåkra sedan länge haft samarbeten med olika externa aktörer som utvecklat digit-

ala resurser specifikt för platsen. Några av dessa kan idag användas av besökare, såsom VR-miljöer kring templet och på åkern. För flera projekt hade appar tagits fram, men slutat användas när de inte uppdaterats och tekniska licenser gått ut. En AR-lösning, som vi inledningsvis hade haft som ett av våra huvudspår, hade prövats, men inte fallit särskilt väl ut då det var svårt att se visualiseringarna utomhus i solljus. Pedagogerna framhöll att det var svårt att använda "åkern" pedagogiskt - eleverna blev lätt distraherade när det inte fanns något som fångade deras intresse. Särskilt för yngre barn, i 9-10-årsåldern, var det utmanande att försöka förmedla information kring bosättningen och levandegöra platsen på de fria ytorna som verkar locka till spring och stoj. Samtidigt ville pedagogerna gärna visa upp platsen eftersom det är där det *riktiga* Uppåkra ligger, i bemärkelsen att fornlämningarna finns där under marken. Vi hade inledningsvis tänkt att en AR-visualisering skulle kunna göra åkern intressantare, men också förutsätt - precis som man tidigare erfarit på Uppåkra - att väderförhållanden kan göra det en mer eller mindre lyckad upplevelse. Ett annat av våra tidiga spår - att använda ljud för att berika upplevelsen - var att ta fram en audioguide med olika karaktärer som berättar en historia, skulle komma att överlappa ett befintligt pågående projekt med att ta fram en audioguide för besökare. Eftersom vi hade olika tidsramar insåg vi att det hade blivit svårt att koordinera arbetet inom ramen för kurser, så den idén övergavs.

Frågan hur ett lärspel som behandlar källkritik/källtillit skulle kunna fungera som ett komplement till existerande miljöer och fungera för åldersgruppen, föll på plats när vi fick en demonstration av hur pedagogerna använder *Underjorden* - en lärmiljö där ett utgrävningsschakt iscensätts. Här är det källarlokalen på det arkeologiska besökscentret som genomsyrat en transformation - väggarna har målats i olika bruna nyanser, likt de kulturlager som möter en arkeolog och vars färger vittnar om de olika mänskliga aktiviteter som ägt rum under tidens gång. På väggarna syns fyra skuggor efter de stolphål som fanns att se då man grävde ut under kulhuset. Högt upp, nära det svartmålade taket, finns någon ölburk som sticker fram ur väggen, en tiokrona, och längre ner på väggarna sticker flera kopior av äldre fynd som hittas på Uppåkra fram. Detta rum används idag för att visa upp hur arkeologer, tillsammans med många andra forskare och specialister, arbetar med fynd och jord som källmaterial. Fynden kanske "upptäcks" i jorden, men kunskapen om vad fynden berättar är en mödosam kollektiv process som involverar många discipliner, något som blir tydligt i utställningen *Underjorden* och den tillhörande presentationen given av pedagogerna. Varje fynd har på olika sätt kopplat in expertis från olika områden, allt från forensiker och kriminalare till osteologer och experter inom DNA-analys. Kopplingen till de av McGrew (2017, 2018) och Breivik (i Bawden, 2001) identifierade källkritiska frågorna som togs upp i teoriavsnittet vävs in skickligt av pedagogerna då de förklarar hur de "läser" berättelser som vissa fynd förmedlar.

För oss blev det på så vis tydligt att *Underjorden* var

en mycket bra resurs i det att den passar för målgruppen, kopplar ihop konkreta ting med historiska "berättelser" och visar metodiken för att få fram dessa berättelser. Här utgår pedagogen, som är arkeolog, från barnens frågor om fynden och genom att ställa källkritiska frågor som "hur vet vi det" illustreras hur processen med att ta fram arkeologisk kunskap gått till. Potentialen i att inviga barnen i den arkeologiska metoden och därigenom "göra källkritik" på ett förhållandevis lättillgängligt sätt för åldersgruppen blev därigenom tydlig för oss. Att ta fram ett lärspel som kompletterar aktiviteten i Underjorden och förankrar/bearbetar den blev vårt huvudfokus, samtidigt som vi ville levandegöra "åkern" och mottot "*det viskar under foterna*". Under den andra delkursen arbetade vi framförallt med att implementera dessa idéer i ett spel. Vi gjorde fyndanalyser där vi tittade på vilka fynd som var särskilt intressanta att ta med i spelet, utifrån vad de förmedlade för information om Uppåkra. Fynd som förmedlade något om dess religösa betydelse eller relation till omvärlden var särskilt intressanta, precis som fynd som visade på hur människor levde då och att vissa föremål, såsom kammar och spännen, är lika de vi använder idag. Vi gjorde ett par besök på Uppåkra där vi gjorde testningar av hur spelet fungerade på plats, vad som var lagom fördräjning innan fynd hittades, hur audiofeedback upplevdes och så vidare. Vi fick tillfälle att tillsammans med pedagogerna fundera på hur spelet tematiskt och visuellt skulle presenteras, och vilken roll spelaren skulle ha. Särskilt värdefullt var att gemensamt få göra en inventering kring vilka de redan upparbetade resurser var, och tips på hur de eventuellt skulle kunna införlivas i spelet. En sådan var 3D-skannade bilder av fynd, som vi försökte få in i spelet men tyvärr föll det på att de inte var färglagda och därmed såg mindre "äkta" ut än de foton vi sedan kom att utgå från. En annan resurs var de texter till fynd som författats av arkeologerna och historiska museet i Lund, som vi använt som utgångspunkt i spelet. Dessutom fanns en film med animeringar och illustrationer, som används i *Underjorden*, som vi kunde ta bilder ifrån för att levandegöra texterna och förmedla i bilder hur fynden användes förr. Dessa resurser har varit väldigt värdefulla för att få en enhetlighet och tematik i det visuella. Vi funderade också mycket på vilken grad av gamification vi skulle göra i spelet, och sökte svar i en del litteratur. Med stöd från det, bestämde vi att ha låg grad av gamification och incitamentsstruktur, och i stället jobba med att få tydlighet i spelets grundprocesser så att spelaren vet hur och vad som ska göras, och få allt att se prydligt och lättläst ut. Vi nöjde oss med att införa en progressbar, så att spelaren får erfara att den gör framsteg och rör sig mot slutmålet. När spelappen fått sitt innehåll, utseende och struktur fick vi möjlighet att testa den på kursare och lärare för projektkursen. Därigenom fick vi se att spelet fungerar som tänkt, och verkar engagera spelaren, men att det finns stor personlig variation i hur de olika rörelserna utförs utifrån instruktionerna, i form av olika takt och rörelsemönster. Att pedagoger finns på plats och med några enkla rörelser kan visa detta innan barnen sätter igång, är väldigt värdefullt. Vi fick också feedback på att vår progressbar, där man får

en ny stjärna för varje fynd, ibland kunde uppfattas som ett betyg på spelarens insats eller på fyndet den hittat och inte ett sätt att visa var i spelet spelaren befann sig. Även tankar kring möjligheter att gå bakåt i spelet togs upp, då spelaren kan vilja läsa frågor och svaralternativ på nytt efter att den fått svaret. Spelet väntar på att testas med målgruppen på plats i Uppåkra.

4 Lösningar

4.1 Konceptuella lösningar

Spelet är alltså en form av skattjakt/arkeologisk utgrävning där eleverna upptäcker fynd och sedan får information och frågor knutna till dessa. Aktiviteten sker på åkern, som ett sätt att dra en koppling till de fynd som faktiskt till stor del fortfarande finns under marken. Plattan används som en simulerad metalldetektor, och ger "utslag" i form av ljud som liknar de som används vid metalldetektering samt pulserande vibrationer då användaren söker efter fynd. Därefter ökar intervallet på ljud-och vibrationsåterkopplingen när användaren söker vidare, och när en slutlig hög nivå då användaren är framme vid en fyndplats. Användaren får sedan gräva upp skatten genom att använda plattan som en simulerad spade i luften och får då audiofeedback i form av ljud som associeras till grävande samt visuell feedback i form av en illustration av ett växande hål på skärmen. När fyndet efter ett visst antal spadtag grävts fram spelas en liten ljudslinga, och en bild på fyndet kommer upp, tillsammans med en *progressbar*. I nästa steg visas fyndet på en del av skärmen tillsammans med kort information om fyndet. En fråga med svarsalternativ introduceras. Frågorna och svarsalternativen är tänkta att skapa associationer mellan olika teman och Uppåkra (såsom att det är en religiösiktig plats, där handelsutbyte med omvärlden skett, och att vissa fynd vittnar om både skillnader och likheter mellan dätid och nutid). Frågor och svarsalternativ är också tänkta att visa upp fynden som källmaterial samt spegla den vetenskapliga processen (exempelvis att olika discipliner kan samverka för att ge kunskap om fynden).

Efter att ett svarsalternativ valts, får användaren (oberoende av svar) information om det/de rätta svaret/svaren samt lite mer information hur "kunskapandet" kring ett fynd gått till. Exempel på sådan information är vilka discipliner/metoder som används, och exempel på vad som går att underbygga vetenskapligt, och exempel på sådant som vi med moderna metoder fortfarande har svårt att fastställa. Förutom att spelaren får veta vilket/vilka svarsalternativ som är rätt och varför, samt varför andra svarsalternativ är fel eller orimliga har vi valt att inte ha förstärkande feedback i form av belöningar vid rätt svar eller "straff" för felsvar (exempelvis poängsamlande). Anledningen är dels att ett sådant spel kräver en tekniskt mer komplicerad lösning, men framförallt öppnar det upp möjligheten att frågor och svarsalternativ kan få en annan funktion - exempelvis kan flera alternativ passa in, och genom det kan alternativ användas som kontrasterande fall eller sida vid sida för att förmedla

information. Frågorna och alternativen har alltså ett epistemologiskt värde och blir ett sätt att förmedla information och visa upp exempel på hur fynden används som källmaterial, vad vi kan veta och hur, och vad som kan vara svårt att få kunskap om. Genom att inte vara av typen rätt eller fel uppmanar frågorna till att väcka tankar och insikter om hur kunskap kan byggas upp. Att frångå tävlingsmomentet hoppas vi leder till att spelarna, som gör skattjakten i par, dessutom blir mer resonerande än fokuserade på att ha rätt eller undvika fel, och tar längre tid på sig att läsa frågor och svarsalternativ och den efterföljande förklaringen. Efter att spelarna fått "facit" på frågan kan de välja mellan att söka efter nästa fynd och aktivera metalldetektorn, eller läsa mer om det fynd de hittat. Ovan beskrivna struktur kan sägas vara den ryggrad som utgör skattjakten, men med variationen att vissa fynd ligger djupare ner, eller är av annat material, och därmed får skrapas eller sällas fram. Några fördelar som vi sett med denna lösning är att det är ett exempel på en skalbar arkitektur; man kan lätt fortsätta att lägga in fler fynd, men också testa en mindre version av spelet på plats och sedan göra modifieringar eller bygga ut lösningen. En sådan breddning skulle kunna vara att föra in andra typer av information i spelet, särskilt under "läs mer-knappen", där man skulle kunna knyta in mer material, såsom kartor, ljud och bildmedia för att stödja läsförståelsen och förmedla visuell information. Uppåkra har mycket tillgängligt material som skulle kunna fogas in - särskilt intressant är kanske de intervjuer med experter som visar hur vi kan dra kunskap om keramik eller flinta, som skulle kunna klippas ner och läggas in i spelet eftersom dessa rimmar så väl med lärandemålet att visa den kollektiva, och multidisciplinära sidan av den vetenskapliga metoden.

4.2 Tekniska lösningar

Uppåkras tekniska resurs som var relevant för oss var tio stycken androidplattor (Huawei Mediapad T5) som kan användas vid skolbesök. Dessutom finns professionella bilder tagna på vissa fynd som kunde användas. Appen programmerades med hänsyn till detta och gruppens kunskaper och använder därför Java som programmeringsspråk och utvecklades i Android Studio. För att "hitta" fynden hade vi tänkt använda oss av en pedometer (steigräknare) i plattan. Dock visade det sig att den typ av platta som Uppåkra använder sig av inte stödjer den funktionaliteten. En annan idé var att använda oss av specifika koordinater för fyndplatserna, men vi ville se till att spelet var anpassningsbart och flexibelt, så den lösningen strök vi också. Därför fick vi tänka om, och använder oss istället av en timer. Timern slumpar en tid mellan 25-35 sekunder, och när tiden har gått ut så är fyndet "hittat". Detta är inte en helt optimal lösning då spelet inte registrerar om spelaren faktiskt rör på sig eller ej, men det är en alternativ lösning som stödjer en liknande spelupplevelse trots de tekniska begränsningarna. Återigen kommer pedagogernas roll vara viktig, för att se till att spelarna faktiskt går runt och söker efter fynd på åkern. När ett fynd hittas signalerar plattan genom ljud och vibrationer att användaren

är framme vid fyndet. Utöver detta upplever spelaren också den fysiska platsen genom att enhetens kamera används för att visa att man är i omvärlden och letar efter fynd. Plattan blir ett "fönster" mot marken under, vår applikation blir på så vis en form av förstärkt verklighet (*augmented reality, AR*). Vår förhoppning är att detta ska skapa en känsla av att spelaren upptäcker fynden och platsen Uppåkra simultant.

Gester implementerades också för att simulera användningen av olika arkeologiska verktyg såsom spade (grävning), skärslev (skrapning) och såll (skakning). De olika gesterna är inspirerade av hur verktygen används av arkeologen i verkligheten och registreras med hjälp av plattans accelerometer när man hittat till fynden. När ett fynd har hittats och tagits fram med hjälp av verktygen så ställs användaren inför en kort information om fyndet samt den tillhörande flervalsfrågan. Användaren väljer ett alternativ, och skickas sedan vidare i spelet till svarssidan. Som tidigare nämnts så finns ofta inget rätt eller fel, utan svaret består av en beskrivning av den arkeologiska processen och vad som går att veta och inte veta med hjälp av denna. Sedan har spelaren valet att läsa mer, fördjupad information om fyndet eller att fortsätta spelet i jakt på nästa fynd. Spelloopen innehåller i nuläget sex stycken fynd med tillhörande bild och information, och spelet avslutas genom att spelaren skickas till en outro-vy när alla fynd är hittade. Från outro-vyn har spelaren möjligheten att välja att spela spelet igen. Samtliga fynd och frågor bidrar till att skapa en förståelse för dätidens Uppåkra, samt med hjälp av gamification skapa glädje i att upptäcka och lära sig mer om platsen.

Vi har även valt att ha en genomgående visuell design som anspelar på de mönster, färger och material som användes i Sverige under järnåldern och som även överensstämmer med Uppåkras grafiska profil. Detta återigen för att förstärka stärka känslan av att "*det viskar under fötterna*" och anspela på den kultur som en gång fanns på platsen.

Uppåkra har under sin verksamhetstid haft flera tekniska lösningar som använts en kort tid, för att sedan överges. Anledningar har varit att lösningarna inte längre passar in med deras verksamhet och licenser som har utgått. Vi har därför velat underlätta och möjliggöra att applikationen kan få en längre hållbarhet och fortsätta att utvecklas i Uppåkras regi. Vi har sedan start tänkt på denna överlämning och utvecklat applikationen och kodens struktur på ett sådant sätt att det ska vara lätt för Uppåkra att sätta sig in i den och göra egna tillägg. Vi har även skapat en instruktionsfil som övergripande förklarar hur den tekniska lösningen fungerar och är uppbyggd, i hopp om att ytterligare underlätta för Uppåkra att modifiera/vidareutveckla vår lösning.

5 Reflektion och fortsatt utveckling

På det stora hela har projektet gått smidigt framåt och inte stött på några större problem. Vägen har förstås inte varit spikrak, utan inneburit ett sökande framåt, där olika lösningar och idéer prövats - men våra initiala mål har varit tydliga sedan vårt första besök på Uppåkra. Resan att

hitta ett uttryck och implementera dem i spelet har varit en mer laborativ process eftersom vi inte gjort ett spel tidigare eller hade full översikt över vilka resurser vi kunde använda (bilder, video, m.m.). Från början upplevde vi det som att vi hade två uppdragsgivare - dels universitetets ursprungliga beskrivning av projektet, och dels de önskemål som pedagogerna på Uppåkra haft, som dessutom har förändrats en del under projektets gång i takt med att spelet tagit form och vi demonstrerat olika versioner av det. Exempelvis kom mer tydliga önskemål upp om att i spelet visa hur kulturlager växer över tid, eftersom pedagogerna haft erfarenhet att det är något som elever har svårt att greppa utifrån en verbal beskrivning. Vi hade inte möjlighet att göra en full animering av denna process, och har i stället fått jobba med bilder - kanske är det något som kan få ta form i en framtid utveckling. Våra besök på Uppåkra har varit avgörande för hur spelet har utformats, och vi har inspirerats mycket av både platsen och pedagogerna. Platsen i sig, som upplevs som torftig, står i bjärt kontrast till dess betydelse. Vi har velat göra ett spel som förändrar upplevelsen av åkern. Inspirationen att utveckla lärspelet med utgångspunkt i den arkeologiska metoden som en ingång till källkritik kom ur mötena med pedagogerna. Platsen, dess begränsningar och pedagogernas önskemål fungerade på så vis som tydliga målsättningar för oss i vår konceptuella utveckling och har gjort det lättare att avgränsa vad som behövde hända i vårt lärspel - fokus på fynd och den arkeologiska metoden, och att "få göra något spännande" på åkern.

En viktig målsättning sedan start har varit att säkerställa att applikationen vi skapar har relevans och fortsätter vara tillgänglig för Uppåkra efter projektets slut. Uppåkra har, som tidigare nämntt haft problem med andra tekniska lösningar då de inte varit hållbara (av olika anledningar) över tid. Att hitta en lösning för styrningen av metalldetektorn var däremot inte helt lätt då vår utgångspunkt - att använda GPS för utplacerade fynd på en specifik del av "åkern", visade sig svår att implementera. Det berodde på att resolutionen på plats, som delvis är avhängig hur många master som finns i närheten, gjorde att det var väldigt svårt att sätta en beräkning som skulle generera en lagom "radie" för att hitta fynden. Skulle radien vara för liten - under 1 meter, skulle spelare uppleva att de gick till rätt plats men ändå inte hittade fyndet. Skulle radien vara för stor, skulle de bara behöva gå några steg innan de fick utslag för ett fynd - de skulle hitta saker hela tiden. Kortfattat hade GPS fungerat bättre i större fysiska miljöer/stadsmiljöer, men på Uppåkra är det mer praktiskt att använda sig av alternativa metoder. För att åtgärda alla dessa problem, användes en timer istället för att hitta fynden. Detta möjliggör att man med precision kan styra hur länge en spelare måste leta innan den hittar något - förvisso av slump, men vi tror inte det kommer vara uppenbart eftersom ljudfeedbacken har olika lägen som signalerar att man kommer närmre fyndet. En annan aspekt av styrningen gällde huruvida vi skulle implementera en vektorpil som indikerar riktning. När vi testade den idén på plats insåg vi då att det kanske skulle få spelare att springa in i häckar och liknande, och att enbart ljudfeedback tillåter

att naturliga begränsningar på plats implementeras i spelupplevelsen. Ytterligare ett plus med den nuvarande lösningen med enbart timer är att tekniken inte är platsberoende över huvud taget. Detta gör att pedagogerna på Uppåkra har en större frihet i att välja vilket område som de vill att användningen ska ske, och gör att pedagogerna kan välja att dela upp en grupp, om de exempelvis har två skolklasser samtidigt, på olika ytor. Användningen är på detta sätt flexibel.

Spelet väntar på att testas på den tänkta målgruppen och på plats på Uppåkra. Vår förhoppning är att spelet kommer att förvaltas av Uppåkra och fortsätta utvecklas i deras regi, där fler fynd kan läggas till. Något vi funderat över är att göra information och medier som finns i spelet tillgängliga efter spelet avslutas, eller om spelaren vill gå tillbaka och läsa om något fynd i efterhand. Frågor om hur materialet förs över till eleverna efteråt, och blir tillgängliga i skolmiljön, är också något att bygga vidare på, liksom hur man kan förankra spelet inför besöket genom lärarhandledningar eller genom att någon kommunikation som riktar sig till eleverna i klassen någon dag före ett besök för att förbereda dem och väcka nyfikenhet. Vi ser också att det är ett pedagogiskt intressant grepp att utgå från arkeologens villkor och metoder när barn och ungdomar arbetar med källkritik, och att lärare efterfrågar och söker stöd i hur undervisning om källkritik kan stötta. Att använda fynd och jord som källmaterial är ett alternativ jämfört med de mer traditionella metoderna där elever arbetar med texter, media och diskussioner. Det finns gott om stöd för att lärande kan stötta i en autentisk miljö med riktiga problem, och att platsen och de sociala relationerna är viktiga komponenter som ger riktning åt lärandet ifall eleven deltar och handlar i den.

6 Referenser

- Almeida, F., Bolaert, H., Dowdall, S. et al. *The Walk-About framework for contextual learning through mobile serious games*. Educ Inf Technol 20, 415–428 (2015).
- Bawden, D. (2001). *Information and digital literacies: a review of concepts*. J. Document. 57, 218–259. doi: 10.1108/EUM0000000007083
- Blomqvist, L. (1989) Arkeologisk metod [http://www.fnf.nu/_filer/bagerfeldt/06%200m%20Arkeologisk%20Metod%20\(1989\).pdf](http://www.fnf.nu/_filer/bagerfeldt/06%200m%20Arkeologisk%20Metod%20(1989).pdf)
- De Pietro, O. *Authentic and Situated Learning with the Use of an Adaptive Search Engine and a QR-Code in Mobile Mode*. International Journal of Digital Literacy and Digital Competence, vol. 4, no. 3, July 2013. Gale Academic OneFile,
- Goodwin, Charles (2000). Action and Embodiment within Situated Interaction. *Journal of Pragmatics* 32, 1489-1522.
- Guath, M. Nygren, T. (2022). *Civic online reasoning among adults: An empirical evaluation of a prescriptive*

theory and its correlates. Frontiers in Education 7:721731, doi: 10.3389/feduc.2022.721731 <https://www.hist.uu.se/nyheter/?id=19210&typ=artikel&lang=sv>

Hutchins, Edward (1995). *Cognition in the Wild.* Cambridge, MA: MIT Press.

Lave, J., Wenger, E. (1991). *Situated learning. Legitimate peripheral participation.* Cambridge, UK: Cambridge University Press.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., and Wineburg, S. (2018). *Can students evaluate online sources? Learning from assessments of civic online reasoning.* Theory Res. Soc. Educ. 46, 165–193, doi: 10.1080/00933104.2017.1416320

McGrew, S., Ortega, T., Breakstone, J., and Wineburg, S. (2017). *The challenge that's bigger than fake news: Civic reasoning in a social media environment.* Am. Educ. 41, 1–44. Available online at: https://www.aft.org/sites/default/files/ae_fall2017.pdf

Stiftelsen Uppåkra Arkeologiska Center. (2022, 27 okt) *Uppåkramodellen* <https://www.uppakra.se/uppakramodellen/#top>

Skolforskningsinstitutet (2022) *Kritiskt tänkande och källkritik - undervisning i samhällskunskap* <https://skolforskningsportalen.se/forskning/kritiskt-tankande-och-kallkritik-undervisning-i-samhallskunskap/>

Sundin, O., & Haider, J. (2018). Källkritik, självkritik och källtillit. I U. Carlsson (Red.), Medie- och informationsskunnighet i den digitala tidsåldern - en demokratifråga: Kartläggning, analyser, reflektioner (s. 59- 61). Nordicom.

Säljö, Roger (2000). Lärande i praktiken. Stockholm: Prisma.

Widén, P., Dahlbeck, P., *Rapport av pilotprojektet Uppåkramodellen eller barn gör historia på riktigt.* (2022) hämtad 2022-10-27, från: <https://www.uppakra.se/media/uploads/2022/09/Slutrapport-Uppa%CC%8Akramodellen.pdf>

Östman, Leif & Öhman, Johan (2010). A Transactional Approach to Learning. Paper presenterat på John Dewey Society, AERA Annual Meeting in Denver, Colorado, 30 april – 4 maj.

Östman, Leif (2003). Transaktion, sociokulturella praktiker och meningsskapande: ett forskningsprogram. I Östman, L. red. Erfarenhet och transaktion i handling: en rapport från projektet lärande i naturvetenskap och teknik. Uppsala: Pedagogisk forskning i Uppsala 147.

Arkimedes övertygelse:

Inlärning och övertygelse av verkliga fenomen i virtuella lärmiljöer

Daniel Zander

da8183za-s@student.lu.se,

Sonja Holmer

so8232ho-s@student.lu.se

Jacob Lundell

ja7516lu-s@student.lu.se

Joel Waye

jo5502wa-s@student.lu.se

Kaspian Jakobsson

ka1251ja-s@student.lu.se

7 februari 2023

Abstract

Virtual reality (VR) och dess användning för utbildningsändamål studeras flitigt. Trots detta finns det i nuläget bristande empiriskt stöd för teknologins pedagogiska verkningsgrad. Vi föreslår en konceptuell utvidgning där inte bara inlärning utan även övertygelse studeras i förhållande till en virtuell simulation av ett verkligt fenomen. Med förankring i diverse konceptuella ramverk ämnar vi utforma en visuellt och interaktionsmässigt optimerad virtuell miljö där deltagaren får utforska ett fysiskt fenomen: lyftkraften av objekt i vätska, vilket beskrivs av Arkimedes princip. Med VR-applikationen som verktyg vill vi uträtta en jämförande studie som besvarar tre huvudfrågor vad gäller applikationens möjlighet till att: (1) övertyga personer om giltigheten av Arkimedes princip, (2) stödja inlärning/förståelse av principen. Vi utforskar även (3) det potentiella sambandet mellan nivån av övertygelse och förståelse. Resultat från 13 deltagare indikerar att det finns ett positivt samband mellan deltagarnas förståelse av Arkimedes princip och övertygelsenivån. Studien diskuterar möjliga implikationer för liknande forskning, samt utvecklingen av olika virtuella lärmiljöer.

1 Introduktion

I takt med dagens framsteg inom virtual reality (VR), med hänsyn till tillgänglighet, grafik och beräkningskraft, görs diverse artefakter mer och mer kapabla till att efterlikna och ersätta verklig sensorisk input samt motorisk (och annan) interaktion med omvärlden. Detta låter på ett sätt även de mest fantasilösa av personer träda in i en alternativ verklighet. Hur ”övertygande” en sådan VR-miljö är, associeras ofta med användarens subjektiva upplevelse av *presence* (närvaro), dvs. känslan av att faktiskt ”befinna sig i” den virtuella miljön. Denna upplevelse av presence sägs i sinom tur vara avhängig VR-applikationens mätbara *immersionsnivå*, där immersionsnivån utgörs av systemets kapacitet till att stödja naturliga sensomotoriska kontingenser (resolution, synfält, tracking, frihetsgrader, etcetera) (O'Regan & Noë, 2001; Slater, 2009). Samtidigt kan frågan om VR och övertygelse vara vilseledande: ”[for] it seems impossible that people would

[ever really] believe the virtual world to be the real thing” (Slater, 2018, s. 432, vår betoning).

En relaterad om något annorlunda fråga är:

kan en virtuell demonstration av ett *verkligt* fenomen göra personer övertygade om giltigheten (alternativt sanningen, validiteten, etcetera) av ett sådant fenomen?

I vår pilotstudie ställs denna fråga inom en utbildnings-tillämpad kontext av VR. Passande nog har kunskapsförmedling av särskilt naturvetenskapliga koncept visat lovande resultat, vilket kan tänkas bero på teknologins relativt obegränsade interaktions- och manipulationsmöjligheter — i linje med teorier om VR och *förflyttningsaktivt* lärande (Xu, Kang & Yan, 2022). VR-teknologin möjliggör även ett framlyftande av vissa så kallade *kritiska* aspekter av, säg, ett fysiskt fenomen, varpå man kan styrka lärandesituationer inom viktiga ämnen som fysik och matematik (Bussey, Orgill & Crippen, 2013; Kullberg, Runesson Kempe & Marton, 2017). Mot denna bakgrund har vi valt att implementera en virtuell laboration av Arkimedes princip, vilken beskriver ett vätskedynamiskt fenomen, vars förståelse vilar på domänspecifik kunskap av koncept som densitet, vätska, massa och gravitation (Minogue & Borland, 2016). Vi ämnar på experimentell väg, med ett jämförande experiment, undersöka om en virtuell (interaktiv) demonstration av Arkimedes princip kan: (1) övertyga personer om dess sanning, giltighet eller validitet, samt (2) stödja inlärning/förståelse av principen. En tredje fråga berör det potentiella sambandet mellan de två, dvs. (3) är en större övertygelsenivå associerad med bättre förståelse av principen? Vi återkommer i vad som följer till begrepp som *immersion* och *presence*, då diverse teoretiska ansatser inom VR-assisterad inlärning introduceras. Dessa konceptuella ramverk är tänkta att motivera designen av både VR-applikationen och experimentet. Därefter presenteras metodologin och arbetsprocessen, med fokus på implementering av VR-miljö och konceptuell design av laboration, samt utformandet av ett övertygelse- och kunskapstest. Slutligen presenteras även resultatet av pilotstudien med påföljande diskussion.

2 Teoretisk bakgrund

Frågeställningen som med denna studie ska besvaras berör flera stora begrepp, så som *kunskap*, *inlärning*, och *virtuella miljöer*. Följande avsnitt ger en överblick över den forskning som gjorts inom relevanta områden. Relevansen för VR-applikationens design beskrivs mer ingående i efterföljande stycke, Metod och arbetsprocess.

2.1 Virtuella inlärningsmiljöer

VR-applikationen som deltagarna baddas in i är en *technology-enabled embodied learning environment* (TEELE). Detta nyintroducerade konceptuella ramverk presenteras i en reviewartikel (Xu m. fl., 2022) och har som syfte att stödja framtagandet av teknikstödda lärmiljöer. I begreppet teknikstödda lärmiljöer rymmer både virtuella och digitala miljöer. Immersion kan, enligt TEELE-ramverket, beskrivas utifrån tre olika dimensioner: sensorisk, fysisk, och kognitiv immersion.

Sensorisk immersion behandlar bland annat ljud- och bildkvaliteten, samt information som deltagaren erfär genom andra sensoriska kanaler. Hög auditiv och visuell upplösning ökar immersion, och även känslan av presence hos deltagaren. Haptisk feedback (i form av exempelvis kraft-feedback) kan ytterligare öka både immersion och inlärning (Magana & Balachandran, 2017; Xu m. fl., 2022, s. 112). Baud-Bovy och Balzarotti (2017) argumenterar till och med för att motoriska handlingar kan vara svåra att utföra utan haptisk feedback.

Ju mer aktivt en deltagare interagerar med den virtuella eller digitala miljön, desto högre blir nivån av fysisk immersion. Begreppet inkluderar alltså både mängden motoriska rörelser som deltagaren utför, samt hur aktiv och passiv hen tillåts vara i miljön. En studie visade exempelvis att studenter som tillåts interagera med relevanta objekt i sin lärmiljö även presterade bättre på efterföljande kunskapstest av fysikkoncept, än de studenter som enbart passivt observerade animationer relaterade till samma koncept på en skärm (Johnson-Glenberg & Megowan-Romanowicz, 2017). Anledningen till att aktivitet och fler *embodied interactions* med systemet även bidrar till högre fysisk immersion (och ger en bredare bas för inlärning) är att eleverna känner sig mer involverade och upptagna med ämnet i fråga. Inte bara kvantiteten spelar roll: kvaliteten av interaktionen och meningsfullheten är också en faktor att ta i beaktning. Helkroppsrörelser ökar risken för tekniska glitches, och noggranna avvägningar bör därför göras, skriver Xu m. fl. (2022).

Detta spelar väl in med nivån av den kognitiva och operationella kongruensen mellan rörelser och deras syfte. Dessa är båda del av dimensionen kognitiv immersion. Hög kongruens möjliggör konkretisering och externalisering av koncept och kan på så sätt avlasta individen kognitivt (Xu m. fl., 2022; Xu & Ke, 2014).

2.2 Presence och övertygelse

Den subjektiva känslan av *presence* uppstår som en funktion av de (mätbara) immersiva egenskaper i VR-applikationen som beskrivits ovan. Ökad känsla av presence i virtuella miljöer har även visats stödja inlärning (Dengel, 2020; Mikropoulos, 2006).

Det är inte bara applikationens inneboende egenskaper (vad gäller design och tekniska aspekter) som kan påverka känslan av presence (Jicol, 2022), utan även deltagarens emotionella tillstånd och dess förväntningar på applikationen är avgörande. Användaren av en virtuell applikation måste också själv hämma sin misstro gentemot miljön (*suspension of disbelief*). Hen måste alltså övertygas av att det som sker i miljön faktiskt är veridiskalt.

2.3 Illusion of place and plausibility

På grund av sin komplexitet och förekomst i ett flertal ramverk, har begreppet *presence* fått en utvidgad semantisk betydelse, vilket försvårar meningsfulla vetenskapliga diskussioner enligt Slater (2009). Slater presenterar därför ett annat ramverk enligt vilket graden till vilken människor agerar realistiskt i en virtuell miljö beror på både på hur stark illusionen av att faktiskt befinna sig *i miljön* är för individen, samt hur händelser i VR-applikationen ger intryck av att vara verkliga. Detta kan ske trots att individen är medveten om att miljön är virtuell. Den senare inkluderar bland annat förståelsen för att händelser i applikationen kan vara reaktioner på agentens egna handlingar; något som knyter an till en sensomotorisk teori om perception, där medveten perception resulterar från bemästrandet av de regelbundenheter som råder mellan agentens motoriska aktivitet och effekten av sensoriska stimuli (O'Regan & Noë, 2001). På så sätt ersätts *immersion* och *presence*-begreppen som förklaring till en agents beteende i en virtuell miljö av de två illusionerna, som kallas för *Place illusion* och *Plausibility illusion* (Slater, 2009). Båda är nödvändiga för att frambringa det beteende hos människor som även förekommer i verkligheten.

2.4 Designprinciper från CTML och inlärning

I designandet av en multimedial lärmiljö finns det flera vägleddande designprinciper som har sitt ursprung i *Cognitive theory of multimedia learning* (CTML) (Mayer, 2005). Dessa är förankrade i kognitionsvetenskapliga och inlärningsteoretiska modeller så som *cognitive load theory*. En lärmiljö ska, enligt CTML, vara utformad på så sätt att det minskar den externa kognitiva belastningen, som kan uppkomma i distraherande miljöer eller modalitetsintegrerande uppgifter.

Några av huvudprinciperna som Mayer (2005) presenterar är modalitetsprincipen, och signalering. Modalitetsprincipen säger att inlärning främjas av att informationen presenteras genom olika modaliteter, exempelvis visuellt och auditivt, jämfört med om informationen skulle presenteras enbart i en *modalitet*, eftersom kompletterande information kan presenteras simultant och integreras.

Med signalering menas uppmärksamhetsledande signaler

för att dra individens uppmärksamhet till relevanta delar av miljön. Detta minskar den externa kognitiva belastningen genom att individen själv inte måste söka efter dessa aspekter i lärmiljön. Signalering har även undersökts i virtuella miljöer och man hittade då att textannotering främjade inlärning på recall-nivå (Albus & Seufert, 2022), men saknade effekt på några av de andra nivåerna av inlärning, så som förståelse och transfer (Bloom, 1956). Dessa tre inlärningsnivåer står i hierarkisk relation till varandra, där recall-nivån, som den första av de tre, innefattar en ytlig kunskap om koncept och ord. Förståelse, som följer recall, innehåller att analoga problem kan lösas och förstås. Högst i hierarkin är det som Albus och Seufert (2022) kallar för transfer-nivån. Inlärning på denna nivå innebär att individen kan lösa helt okända problem inom samma domän, eftersom en sådan djup förståelse för koncepten uppnåtts (Albus & Seufert, 2022).

3 Metod och arbetsprocess

På uppdrag av Educational Technology Group (<https://lucs.lu.se/etg>) vid Lunds universitet fick vi i uppgift att implementera ett självvalt fenomen i olika lärmiljöer, för att sedan kunna jämföra hur väl de kan stödja en *övertygande* demonstration. Innan experimentstrukturen växte fram, diskuterades, både inom gruppen och med handledare, ett lämpligt fenomen som tar vara på VR-teknologins potential. Under denna inledande process snubblade vi över Arkimedes klassiska princip som beskriver den kraft som föremål upplever när de sänks ned i vätska. Vi insåg att visualiserings- och interaktionsmöjligheterna i VR kan underlätta vid åskådliggörandet av denna princip (t ex. med visuellt representerade kraftvektorer). Arkimedes princip bygger på flera delkoncept som i sig inte är triviala och därav ofta leder till konceptuella missförstånd (Minogue & Borland, 2016). Samtidigt har Arkimedes princip demonstrerats utan VR-teknologi sedan 200-talet f.Kr. och en motsvarande verlig laboration av principen kan därför stödja ett naturligt jämförande av inlärning och övertygelse. Efter beslutet om fenomenen inleddes genast ett antal parallellt löpande processer med att bygga HiFi-prototyp, designa experiment (och midFi-prototyp av laboration) samt utforma ett preliminärt kunskaps- och övertygelsetest. Nedan presenteras, i ovan angiven ordning, våra teoretiskt och pragmatiskt motiverade lösningar på samtliga uppgifter.

3.1 Design av experiment och laboration

För att besvara frågorna som nämnades i introduktionen har vi utfört ett jämförande experiment. Testgruppen fick utföra en interaktiv laboration i den virtuella miljön och kontrollgruppen en motsvarande laboration i verlig miljö. Efter laborationen gjordes en semi-strukturerad intervju med varje deltagare för att få en inblick i deras förkunskaper om vätskedynamik, deras tidigare utbildning samt erfarenhet av VR-teknologi. Efter laborationen fick varje försöksperson besvara en uppsättning frågor om Arkimedes princip, tänkt att mäta både nivån av kunskap och övertygelse i förhållande till

principens innehåll (mer om detta test i Utvärdering av kunskap och övertygelse). Deltagarna var universitetsstudenter med begränsad naturvetenskaplig utbildning.

Arkimedes princip säger att ett föremål nedslikt i vätska upplever en flytkraft lika stor som tyngden på den av objektet undanträngda vätskan. Laborationen som deltagarna genomförde är en av flera möjliga demonstrationer av principen och kallas för *överflödesmetoden*:

- (a) sänk ner ett objekt i vätska och observera att volymen *undanträngd* vätska är identisk till volymen av objektet,
- (b) undersök tyngden av den undanträngda vätskan,
- (c) i enlighet med Arkimedes princip så är denna tyngd identisk med flytkraften som objektet upplever.

En sådan enkel laboration stödjer inlärning via repetition, där egenskaper av objekt och vätska kan varieras över olika exempel. För att designa en inlärningsfrämjande sekvens av sådana exempel använder vi principer från *Variationsteorin* (Kullberg m. fl., 2017). Applicerat på inlärning föreslår teorin ett särskilt mönster av variation och *invarians* för att lyfta fram *kritiska* aspekter av ett fenomen, vars uppmärksammande är nödvändiga för att studenten ska nå en fullständad förståelse. Mönstret av variation sker i två steg: 1. *separation* av kritiska aspekter genom variation (över olika exempel) i enbart en dimension åt gången, samt 2. *fusion* av tidigare separat upplevda aspekter genom variation av olika egenskapsdimensioner samtidigt (Åkerlind, 2015, s. 16-18). Laborationen kommer att instansiera detta variationsmönster: under första delen av laborationen får deltagaren utföra steg a) till c) med objekt av olika volym men med samma densitet och därefter med objekt av samma volym men med olika densitet (seperation). Därefter sammanfogas dessa aspekter; objekt varierar då både i volym och densitet samtidigt (fusion).

I utvecklingsstadiet av laborationen involverades även en gymnasielärare i naturvetenskapliga ämnen. Syftet med detta var att kunna ta hänsyn till studentperspektivet, delgett från någon som dagligen arbetar med vidareförmedlandet av olika fysiska koncept, och som därför även har stor förståelse för vilka aspekter som gymnasielever ofta uppfattar som svåra. Utifrån mycket av den feedback vi fick av läraren gjordes ett flertal förändringar av produkten.

3.2 Implementering och programmering

Applikationen utvecklades med spelmotorn Unity (Unity Technologies, 2022), som innehåller verktyg och programmeringsbibliotek som möjliggör VR. Den underliggande tekniken som möjliggör detta heter OpenXR (The Khronos Group Inc., 2022) och är en standard för VR- och AR-applikationer som tillåter olika modeller av VR-headsets att köra samma applikation oberoende av tillverkare. För att snabbt kunna få en överblick gällande vilka komponenter som krävdes för att utforma den virtuella labbmiljön skapades ett Unity-projekt med enkla interaktioner i VR, där användaren kunde plocka

upp objekt i en avskalad miljö. Därefter skapades ett Git-repo för att kunna distribuera projektet mellan projektmedlemmar och handledare. Detta möjliggjorde en distribuerad och smidigare arbetsprocess genom att alla har tillgång till allt som behövs för att lägga till komponenter.

Unity har en medföljande fysikmotor som kan simulera gravitation och kollision mellan olika objekt. Dock är simulerings av vätskor ett svårt och beräkningstungt problem som trots framgångar i metod inte är optimalt löst för realtidsanvändning (Zsolnai-Fehér, 2022). Dessutom kräver en demonstration av Arkimedes princip en tydlig och direkt koppling mellan en vätska och objekt som interagerar med vätskan. I det fallet ökar komplexiteten avsevärt då kraftberäkningar behövs åt båda håll, där vätskan och objekten kan påverka varandra i oändligt många beräkningssteg. I och med denna insikt gjordes ett beslut att förenkla många av de fysikaliska fenomen som uppstår vid en demonstration av Arkimedes princip.

Ett sådant exempel är hur vi beräknar flytkraft, $P(y)$ vilket ges av ekvation 1, där P_0 är det normala atmosfäriska trycket, ρ är vätskans densitet, g är jordens tyngdacceleration, och y det negativa avståndet till vattenytan.

$$P(y) = \begin{cases} P_0 & \text{om } y \geq 0 \\ P_0 + \rho * g * y & \text{om } y \leq 0 \end{cases} \quad (1)$$

För att sedan beräkna flytkraften vid varje punkt för ett objekt kan algoritmen beskrivas med följande pseudokod:

```
för varje yta (p_1, p_2, p_3):
    c = (p_1 + p_2 + p_3) / 3
    n = (p_2 - p_1) x (p_3 - p_1)
    normal = n / ||n||
    a = ||n|| / 2
    f = -normal * a * P(c_y - W(c_x, c_z))
    sätt kraft f vid punkt c
```

Detta är fortfarande relativt beräkningstungt men det är fysiskt baserat. Även en del dämpning tillämpas på objekten för att stilla rörelser som uppstår på grund av begränsad precision av beräkningarna.

Liukande förenklingar hittas i hur vätskorna reagerar med objekten som sänks ned, där det egentligen bara är vätskenivån som ändras baserat på objektens volym.

3.3 Timeline-system

Unitys inbyggda Timeline-system användes för att laborationen skulle genomföras på ett linjärt sätt. Olika ljudklipp med instruktioner och tillhörande text lades in i timelinan och stannades sedan med en egengjord paus-signal. För att komma vidare i laborationen behöver användaren klicka på en tydlig, lysande röd knapp. Samma interaktion upprepas för att gå vidare.

3.4 Utvärdering av kunskap och övertygelse

Deltagarna gavs efter laborationen ett formulär med frågor som testar kunskap om Arkimedes princip. Frågorna var grupperade enligt tre olika nivåer av förståelse eller djup av informationsbearbetning - vad Albus och Seufert (2022) kallar för *återgivning*, *förståelse* och *transfer* (se Appendix B) (eng. *recall*, *comprehension* och *transfer*). Frågor på nivån av återgivning är till skillnad från förståelse och transfer *inte* av flervalsform, för att kunna separera igenkänningsminne från genuin återgivning (eng. *recall*). Med dessa frågor hoppas vi kunna jämföra och utvärdera hur väl en VR-laboration kan stödja inlärning eller förståelse av Arkimedes princip.

Frågan om hur vi mäter *övertygelse* började med frågan om begreppets betydelse i den relevanta kontexten. För våra ändamål lämpar sig en tämligen folkpsykologisk definition, där 'x är övertygad om P' om och endast om (1) x har en propositionell attityd riktad mot P's sanning och (2) x har en tillräckligt stark förbindelse/anförtroende till sanningen av P. Kriterie (2) bygger på idén om *a sufficiently rich notion of belief* (Sartwell, 1992), ett element som är särskilt avgörande för att kvantitativt kunna jämföra *graden* av övertygelse mellan kontroll och testgrupp.

Ett tidigt förslag gick ut på att utsätta försökspersoner för verkliga scenarier där de tvingas göra val som "spelar roll", dvs. där utfallet av ett felaktigt val får otrevliga konsekvenser (exempelvis be personer placera en mobiltelefon på ett *troligen* flytande föremål i vatten). Sådana val kan tänkas kräva en hög grad av övertygelse och därav anförtroende till giltigheten av Arkimedes princip. Detta förslag sällades dock bort av tre anledningar: 1. idén lyckades inte bli tillräckligt precis i dess formulering för att kunna användas som mätmetod, 2. stora skillnader i riskkänslighet mellan individer kan vara en confounder, samt 3. etiska överväganden. Istället valde vi att sammanfoga kunskapsrestet med en likertska, där svar på samtliga frågor paras med en bedömning av *säkerhet* (alternativt förbindelse, anförtroende) på en skala från 1-5. Skalan vi använder liknar ett *Certainty of response Index* (CRI), något som har använts flitigt inom samhällsvetenskaperna, men även inom pedagogik i naturvetenskapliga områden (Hasan, Bagayoko & Kelley, 1999). Metoden tillåter oss (potentiellt) att kvantifiera nivån av övertygelse mellan test- och kontrollgrupp. Vidare, intressanta korrelationer mellan nivån av övertygelse och kunskap kan nu studeras.

3.5 Mätning av presence i virtuell miljö

För att mäta deltagarnas presence i VR-miljön användes *The iGroup Presence Questionnaire* (IPQ) (Schubert, Friedmann & Regenbrecht, 2001b). IPQ konstruerades av Schubert, Friedmann och Regenbrecht och är en skala som används för att mäta presence i en virtuell miljö. Den består av 14 frågor som korrelerar med tre subskalor, förutom en fråga som inte tillhör en subskala utan istället mäter en generell känsla av presence och har hög belastning på alla tre subskalor. De tre subskalorna beskrivs nedan:

- Spatial presence: En känsla av att vara fysiskt nära-

rande i den virtuella miljön.

- Involvering: Mätning av den uppmärksamhet som ägnas åt den virtuella miljön och den involvering som upplevs.
- Upplevd realism: Mätning av den subjektiva upplevelsen av realism i virtuella miljön.

Deltagarna fick fylla i frågeformuläret efter att ha utfört experimentet i den virtuella miljön. Enligt en tidigare studie ger det mer gynnsamma resultat att utföra detta frågeformulär inom den virtuella miljön istället för att behöva ta av sig VR-headsetet då det förbättrar deltagarnas upplevda presence (Schwind, Knierim, Haas & Henze, 2019). I vårt fall implementerades inte frågeformuläret i den virtuella miljön eftersom design och implementering av själva experimentet prioriterades.

3.6 Utförande, verklig laboration

Deltagarna fick komma in i rummet efter att laborationen förberetts på bordet. På bordet stod en väg, på vilken det stod en genomskinlig plastbägare, med volymmarkering (se figur 1). Till vänster om vägen hade en konstruktion byggts upp med en bägare av metall, vars utstickande rör hängde över plastbägaren. Metallbägaren var fylld med vatten enda upp till röret. I närheten fanns även två vikter av samma volym och vikt, samt två behållare fylda med mjöl, resp. vatten.



Figur 1. Laborationsbordet i den verkliga miljön

Deltagarna började med att ange information om ålder och tidigare erfarenhet av fysikutbildning och blev sedan informerade om vad som skulle ske. Instruktionerna skulle spelas upp för dem, men de kunde också alltid läsa instruktionerna på en skärm de hade framför sig. Arkimedes princip fanns dessutom uppskriven på en tavla i rummet, så att de alltid kunde referera till denna. De blev även uppmanade att göra detta så ofta de behövde. Experimentledare var den som ställde fram relevant material och såg till att exempelvis vägen var påslagen. En annan ur projektgruppen förde under tiden anteckningar om eventuella problem som stöttes på och annat som kunde tänkas vara relevant.

Deltagaren började med att sänka ner vikterna i metallbägaren för att se hur mycket vatten som trängdes undan, se

figur 2. Detta mättes i plastbägaren. Hen sänkte sedan ner ytterligare en vikt för att jämföra volym vatten som trängts undan, samt dess vikt. Efter att försiktigt ha hällt tillbaka vattnet i metallbägaren sänkte deltagaren ner vatten-, resp. mjölbehållaren för att undersöka om och, i så fall, hur volymen undantränt vatten skilde sig. Efter detta steg fick deltagaren på nytt sänka ner den första vikten, men dena gång i t-sprit, vilken har lägre densitet än vatten. Syftet med detta var att förstå att vikten av vätskan, men inte volymen, hade förändrats sedan första steget. Slutligen testade deltagaren att hänga vikten på en dynamometer, vilken mäter tyngd och sänka ner den i behållaren med vätska. Denna tyngd jämfördes senare mellan de olika vätskorna (vatten och t-sprit).



Figur 2. Nedsänkandet av en vikt i metallbägaren

Efter att deltagaren genomfört varje steg i laborationen fick hen anteckna relevanta svar på ett papper, vilket vi hade förberett.

När deltagaren utfört laborationen fick hen svara på frågorna i kunskapstestet. Hen blev informerad om att det kunde vara kluriga frågor, men att hen skulle göra sitt bästa och fick ta den tiden som behövdes. Pappret från tidigare med datan som antecknats fanns tillgänglig under hela processen, och likaså Arkimedes princip.

Efter att kunskapstestet var utfört, genomfördes en semi-strukturerad intervju där frågor om svårighetsgrad och eventuella otydligheter ställdes. Här ombads deltagaren också att med egna ord återge Arkimedes princip.

3.7 Utförande, VR-laboration

VR-laborationen utfördes på ett liknande sätt som den verkliga, med skillnaden att laborationen utfördes i VR istället för verkligheten. Deltagarna laborerade individuellt och fick börja med att ange information om ålder och tidigare erfarenhet av fysikutbildning. Därefter fick de ta på sig VR-headsetet medan handledaren förklarade grundläggande saker, som hur man interagerar med VR-världen och hur man sätter igång själva laborationen. Figur 3 visar hur detta såg ut i ett tidigare stadie av testande.

Precis som i den verkliga laborationen fanns formeln för Arkimedes princip och definitioner för tillhörande variabler



Figur 3. Testande av VR-miljö

alltid uppskrivna på en tavla i VR-världen, se figur 5. Deltagaren påbörjade laborationen genom att klicka på en knapp i VR-miljön, vilket innebar att instruktionerna började läsas upp i VR-headsetet samtidigt som de textades i VR-miljön. Efter att instruktionerna var upplästa gavs deltagaren fritt med tid att utföra de angivna instruktionerna. För att få fortsätta laborationen kunde de klicka på samma knapp som tidigare nämnts. Efterhand som deltagaren fortsatte framåt i laborationen antecknades de relevanta resultaten automatiskt på en whiteboard i VR-miljön, se figur 4



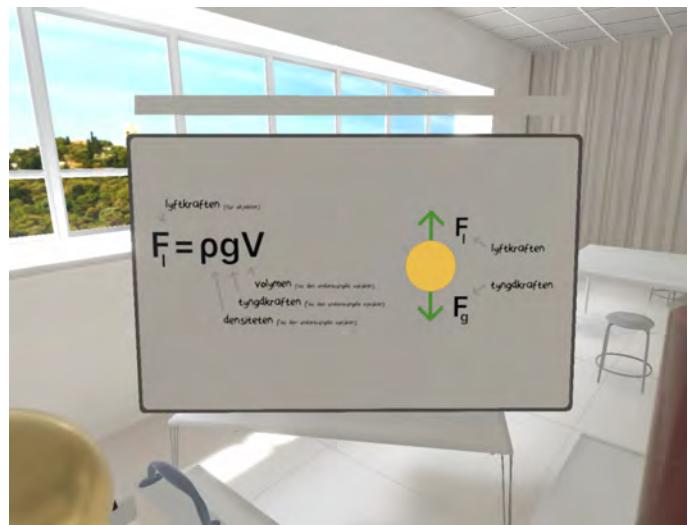
Figur 4. Virtuell laborationsmiljö

När laborationen var färdig fick deltagarna, precis som i den verkliga laborationen, svara på frågorna i samma kunskaps-test. Därtill utfördes samma semistrukturerade intervju som i den verkliga laborationen. Till skillnad från den verkliga laborationen fick deltagarna även svara på en ”presence questionnaire”, där de fick svara på frågor om hur närvaroinducerande VR-miljön var.

3.8 Deltagare

Sammanlagt samlades resultat från 13 deltagare in, 7 personer i testgruppen i VR, och 6 personer i kontrollgruppen, vilka utförde den verkliga laborationen. Deltagarna rekryterades på olika institutioner på Lunds Universitet under ett par dagar.

Genomsnittsåldern skilde sig inte nämnvärt grupperna



Figur 5. Tavlan i VR-miljön

emellan (testgrupp: $M=26.8$ $SD=5.8$, kontrollgrupp: $M=22.3$ $SD=2$). Testgruppen bestod av 2 kvinnor och 5 män, medan kontrollgruppen bestod av 3 kvinnor och 3 män. I testgruppen hade 4 av 7 läst fysik på gymnasienivå, 1 av 7 på universitetsnivå, och 2 av 7 hade enbart läst fysik på högstadienivå. I kontrollgruppen hade alla 6 deltagare läst fysik på enbart högstadienivå. Inom båda grupperna angav många att de sedan tidigare kände till principen, och hade genomfört liknande laborationer redan på mellanstadiet, men att de hade glömt bort detaljerna av principen.

I testgruppen angav 4 att de befunnit sig i en VR-miljö 1-2 gånger tidigare, 1 deltagare svarade 3-4 gånger, 1 fler än 5 gånger, och för 1 deltagare var det den första gången hen befann sig i en VR-miljö.

3.9 Dataanalys

All statistisk analys är gjord i R 4.1.3. Samtliga modeller är skapade med brms-paketet; grafer är skapade med ggplot2. Alla rapporterade modellestimeringar är posteriora medelvärden för relevanta parametrar, med 95% Bayesianska konfidensintervall. Följande fyra modeller användes i dataanalysen och är skrivna i brms/lme4 syntax:

Modell A (logistisk): Korrekt ~ Nivå + Grupp + (1|Deltagare)

Modell B (linjär): Övertygelse ~ Nivå + Grupp + (1|Deltagare)

Modell C (logistisk): Korrekt ~ Övertygelse + (Övertygelse|Deltagare)

Modell D (linjär): Övertygelse ~ Presence

Modell A är en multipel mixad logistisk regression som använder data på individuell försöksnivå, dvs. icke-aggregerad data. ’Nivå’ (kodad Återgivning<Förståelse<Transfer) och ’Grupp’ (kodad Kontroll<Test) är fixerade effekter och ’(1|Deltagare)’ varierad effekt.

Modell B använder icke-aggregerad data, delar koefficienter med föregående modell, men estimerar istället graden av övertygelse. Modellen är linjär, vilket medför att responsvariablen, som tekniskt sett är *ordinal*, antas vara normalfördelad. Två olika skäl ligger bakom detta val: (1) Studiens forskningsfråga intresserar sig för *grova* skillnader i övertygelsenivå mellan test och kontrollgrupp. En ordinal logistisk regression estimerar en likelihoodfunktionen för varje responskategori, vilket bidrar med överflödig information och tolkningssvårigheter i förhållande till forskningsfrågan. (2) I litteraturen är det vanligt att behandla $>=5$ -likertskalor som approximativt kontinuerliga (Anikin, 2022). Dessutom finns det en pågående debatt om hur man bäst ska hantera likertskalavariabler (Norman, 2010; Sullivan & Artino Jr, 2013; Zumbo & Zimmerman, 1993).

Modell C predicerar precision utifrån graden av övertygelse och dataan är tillika på försöksnivå. Övertygelsevariabeln är standardiserad sådant att:

$$x' = \frac{x - \mu(x)}{\sigma(x)} \quad (2)$$

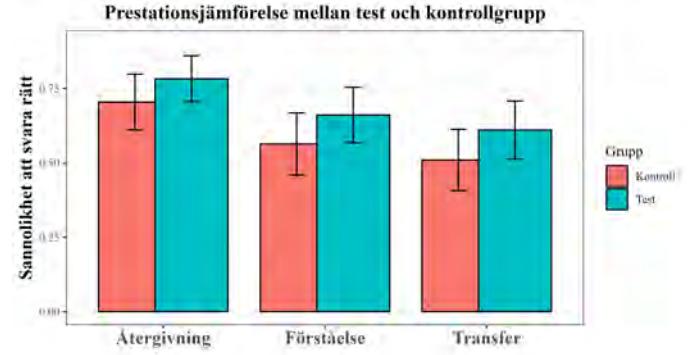
vilket transformeras övertygelsevektorn till att ha ett medelvärde på 0 och en standardavvikelse på 1. Detta underlättar både vad gäller avläsandet av resulterande parametervärden och modellens möjlighet till att konvergera. '(Övertygelse|Deltagare)' tilldelar en unik α och β till varje deltagare ($N = 13$) och är således modellens varierande effektvariabel.

Modell D undersöker återigen graden av övertygelse, men denna gång som en funktion av graden av presence. Av samma skäl som nämndes förut så behandlas övertygelseskalan som kontinuerlig och normalfördelad. Eftersom det inte finns någon korrespondens mellan presence och övertygelse på försöksnivå så har vi valt att använda det sammantagna medelvärdet per person från IPQ-subskalan *spatial presence* och *involvement*. Valet är baserat på Schubert, Friedmann och Regenbrecht (2001a) modell av närväro i virtuella miljöer: "Two cognitive processes lead to this model: the representation of bodily actions as possible actions in the VE, and the suppression of incompatible sensory input. It is hypothesized that the conscious sense of presence reflects these two components as spatial presence and involvement" (Schubert m. fl., 2001a, s. 266). Närvarovariablen är standardiserad på liknande sätt och av samma anledning som övertygelsevariabeln är i modell C (se ekvation (2)).

4 Resultat

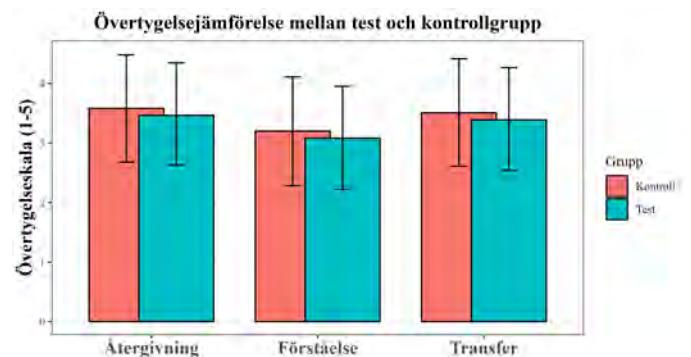
För att besvara tre huvudfrågor gällande kunskap, övertygelse och sambandet mellan de två har vi utformat tre motsvarande bayesianska mixade regressionsmodeller. Med en fjärde modell undersökte vi även ett potentiellt samband mellan graden av presence och övertygelse.

Kunskapsnivå av test och kontrollgrupp estimeras enligt respektive grups sannolikhet att svara rätt på de återgivning-, förståelse-, och transferfrågor som gavs till varje försöksperson efter utförd laboration. Resultaten enligt en logistisk regressionsmodell (se Dataanalys för detaljer) tyder



Figur 6. Error bars visar 95% konfidensintervall.

på en något högre kunskapsnivå i testgruppen, där grupp- β -koefficienten är estimerad till log-odds 0.43 95% CI[-0.49, 1.4]. En formell uträkning av genomsnittlig kontrast mellan test- och kontrollgrupp, över samtliga frågenivåer, visar att testgruppen har 9.1% (95% CI[-11.7, 32.2]) högre sannolikhet att svara rätt. Även om effekten är väldigt osäker så visar vidare uträkning att 82% av den posteriora kontrastdistributionen är större än 0, dvs. enbart 18% av modellens estimerade värden tyder på att kontrollgruppen presterade bättre. Störst skillnad mellan test- och kontrollgrupp sätts på transferfrågorna, där testgruppen har 10% (95% CI[-11.7, 32]) större sannolikhet att svara rätt (se fig. 6). Vi ser ingen märkvärd individuell variation, med en genomsnittlig varians från populationseffekten på 0.38 (log-odds), 95% CI[0.02, 1.07].

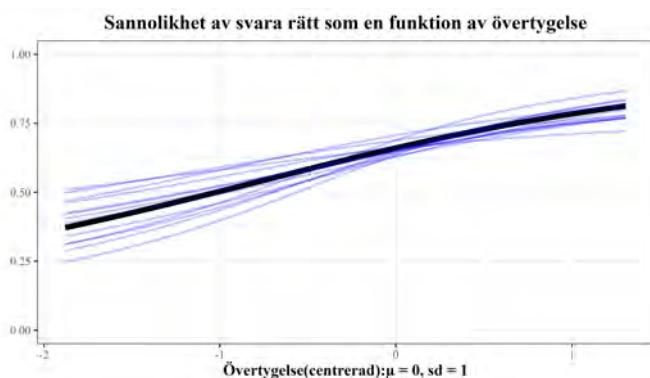


Figur 7. Error bars visar 95% konfidensintervall.

Graden av övertygelse i kontroll- och testgrupp utforskas med en linjär modell (se Dataanalys för motivering av detta): övertygelse är en beroende variabel, grupp och frågenivå är fixerade effektvariabler och 'deltagare' varierande effektvariabel. Resultatet från modellen tyder på att övertygelsenivån är likvärdig mellan test- och kontrollgrupp, med en praktiskt taget platt grupp- β på -0.12, 95% CI[-0.59, 0.35] (se fig. 7). Testgruppen har dessutom en outlier med en genomsnittlig övertygelseskattning på 1.4 (sd = 0.88). Utan denna datapunkt estimerar modellen i stället en något positiv grupp- β på 0.19, 95% CI[-0.81, 1.22]. Däremot finns det

en stor individuell variation i övertygelse-nivå; den genomsnittliga variansen från populationseffekten (dvs. från $\alpha =$ återgivning-kontrollgrupp = 3.58, 95% CI[2.68,4.48]) är en hel poäng på övertygelseskalan, 1.01 (95% CI[0.61,1.66]).

Sambandet mellan övertygelse och kunskapsnivå undersöks med en logistisk regressionsmodell där övertygelseskalan är oberoende, 'deltagare' varierande effektvariabel och sannolikhet att svara rätt är en beroende variabel (se Dataanalys för detaljer). Modellen estimerar det posteriota medelvärdet för populations- β till 0.66 (log-odds), 95% CI[0.09,1.31]. Transformerat till sannolikhet innebär detta att en standardavvikelse från medelvärdet på övertygelseskalan leder till 13% högre sannolikhet att svara rätt (se fig. 8). Även om effektstorleken är osäker kan vi vara säkra på att sambandet är robust då 99% av populationsbetans posteriota distribution är positiv.



Figur 8. Svart linje representerar populations- β . Blå linjer är varierande effekter från deltagare ($N = 13$).

Till sist utforskar vi det potentiella sambandet mellan deltagares känsla av *presence* och övertygelsenivå. En linjär modell med deltagares aggregerade presencepoäng som oberoende variabel (se Dataanalys för diskussion) estimerar ett väldigt svagt men signifikant samband: deltagare som befinner sig en standardavvikelse från medelvärdet i presence har en övertygelsenivå som är 0.33 (95% CI[-0.01, 0.67]) poäng högre.

I tabell 1 sammanfattas intercept samt tidigare nämnda populationseffekter för alla modeller (A-D).

Modell	Populations- α	Populations- β
A	.91, 95% CI [.01, 1.86]	.43, 95% CI [-0.49, 1.4]
B	3.58, 95% CI [3.1, 4]	-.12, 95% CI [-0.59, .35]
C	.68, 95% CI [.19, 1.2]	.66, 95% CI [.09, 1.31]
D	3.3, 95% CI [2.98, 3.6]	.33, 95% CI [-0.01, .67]

Tabell 1. Estemering av α och tidigare nämnda β -värden.

5 Diskussion

Med förbehåll för pilotstudiens bristande urvalsstorlek kan vi ge ett försiktigt jakande svar på samtliga av studiens tre

huvudfrågor. Resultaten tyder på att testgruppen har bättre förståelse av Arkimedes princip som ett resultat av utförd VR-laboration. Övertygelsenivån är tillika jämt fördelat mellan experimentgrupperna, vilket blir särskilt tydligt om man tar hänsyn till en av testgruppens starkt avvikande deltagare. Med betydligt större säkerhet kan vi även sluta oss till att graden av övertygelse har ett påtagligt positivt samband med kunskap och förståelse av Arkimedes princip.

Resultatet från denna studie kan tänkas ha betydande implikationer även för andra studier som har undersökt virtuella lärmiljöer. Många tidigare studier har utgått ifrån att virtuella representationer av fenomen antas möjliggöra samma nivå av övertygelse som fenomenen i verkligheten själv. Detta är ett något lättstött antagande utan hänvisning till underbyggande empiri.

Att testgruppen möjliga presterade bättre än kontrollgruppen på kunskapstestets samtliga nivåer, tyder på att förståelsen av principen på ett likvärdigt sätt kunde anamas genom de olika laborationerna. De tillkortakommanden som den virtuella miljön har (exempelvis att det de facto inte är verklighet) verkar vägas upp genom nyttjandet av inlärningsprinciper i multimediala lärmiljöer, så som text-annotering etc. Syftet med applikationen är inte enbart att skapa en miljö som efterliknar verkligheten, utan även att utnyttja de inlärningsstödjande funktioner som karakteriseras virtuella miljöer, så som visuellt representerade krafter. Miljön måste också animera studenten att hämma sin misstro genom att använda *place illusion* och *plausibility illusion* (Slater, 2009).

En annan implikation av våra resultat är vikten av en väl-designad lärmiljö. Studenter presterade bättre i den virtuella miljön, vilket vi utifrån tidigare forskningsresultat antar kan härledas till framlyftandet av kritiska aspekter nödvändiga för förståelsen av principen. Trots att några deltagare angav att de tyckte att vätskan inte såg realistisk ut, eller att några detaljer var lite oskarpa, angav många fördelarna med att kunna ändra densitet med en sådan precision som inte var möjlig i verkligheten. Detta illustreras också av flera deltagares reflektioner kring nyttan med kraftpilarna. På så sätt visar resultatet (både de kvantitativt mätta prestationerna och deltagarnas egna svar i intervjun) hur miljöns brister kan vägas upp av de många möjligheterna som en virtuell miljö medför.

I följande stycke diskuteras en annan potentiell faktor till vissa deltagares höga, resp. låga säkerhetsskattning utifrån de svar som samlades in i den semi-strukturerade intervjun som följde laborationerna. En av deltagarna som angav lägst säkerhet på svaren i post-testet (i genomsnitt 2,44 av 5), var del av testgruppen som fick utföra experimentet i den virtuella miljön. När hen blev ombedd att med egna ord förklara vad principen gick ut på, var svaret ett något vagt sådant, om vilka ämnen som flyter och vad som kan tränga bort. Nämnd är också att deltagaren inte kände till Arkimedes princip sedan tidigare. Även en annan deltagare som kände sig osäker på många av svaren (genomsnitt 2,67 av 5) hade inget konkret minne av att ha lärt sig principen.

På andra ändan av spektrat finns en deltagare som hade

angivit en säkerhet på i genomsnitt 4,11 på post-testets frågor. ”Lyftkraften på ett objekt i en vätska är lika med tyngdkraften på den undanträngda volymen vätska” blev svaret när hen ombads återge principen med egna ord. Hen hade förkunskaper kring principen. Två andra deltagare i test- resp. kontrollgrupp, som också sedan tidigare kände till principen, hade ett högt medelvärde på säkerheten med vilken de besvarade post-testfrågorna, och kunde även ge en adekvat beskrivning av principen med egna ord.

I ljuset av dessa fem deltagares reflektioner och svar på intervjun verkar det som om förkunskaper spelar stor roll för hur säker man känner sig på svaren. Om detta beror på att man lättare kan ta till sig den informationen som man försöker tillägna sig i laborationen, eller för att man uteslutande förlitar sig på tidigare kunskap kan vi utifrån svaren på intervjun inte veta.

5.1 Begränsningar och framtida arbete

Den största begränsningen med vår studie har haft att göra med begreppet *övertygelse* och hur vi kan nära oss ett sätt att mäta detta psykologiska tillstånd. En bränande fråga är om CRI eller en liknande likertskala överhuvudtaget mäter det vi är intresserade av, nämligen övertygelsen av Arkimedes princip i sig och dess giltighet. Det finns en risk för att en likertskala snarare mäter övertygelsen eller säkerheten som personen har vad gäller sin egen förmåga att kunna återgiva och applicera tidigare vunnen information. Det Intentionala objektet i denna kontext är i sådant fall inte Arkimedes princip *qua* Arkimedes princip, utan deltagarens tillräckliga eller bristande kunskapsrelation till tidigare förmedlad information. Vår kvalitativa mätmetod i form av en semi-strukturerad intervju kompenserade något för denna brist.

Som tidigare nämnt, kan presence hos deltagaren, och därav resultatet från IPQ, förbättras om man inte tar ut deltagaren ur den virtuella miljön (Schwind m. fl., 2019). Istället låter man personen göra utvärderingen medan hen fortfarande befinner sig i den virtuella miljön. Detta gör att deltagaren inte behöver anpassa sig till den verkliga världen och därmed inte utsätts för exempelvis sociala distraktioner och/eller interaktioner med den verkliga världen som orsakar störningar eller avbrott i presence. Påverkan över när och hur vi utförde utvärderingen av presence i denna studie är speciellt intressant eftersom deltagaren fick utföra flera moment innan de gjorde själva IPQ, vilket kan ha haft inverkan på personens presence och därmed resultatet av formuläret. För tillfället är det svårt att bedöma vilken inverkan metoden bakom IPQ hade, däremot kan det vara värt att ha i åtanke till framtida arbete för att mer noggrant mäta deltagarnas presence.

Eftersom pilotstudien enbart undersökte Arkimedes princip, kan egentligen bara slutsatser om deltagares övertygelse av just detta vätskedynamiska samband dras. För att vidare undersöka övertygelsenivån i virtuella världar skulle det var fördelaktigt att även undersöka andra fysikaliska koncept.

Antalet deltagare i pilotstudien var relativt låg, och i framtida studier vore det bra att öka det totala antalet deltagare för att få mer robusta resultat. Även vad gäller deltagarre-

kryteringen finns stor förbättringspotential. Mer specifikt så finns, mellan test och kontrollgrupp, en betydande skillnad i tidigare fysikutbildning. Detta leder särskilt till en större osäkerhet av resultaten från Modell A, då mycket av variationen kan tänkas reflektera en högre grad av tidigare bekantskap med fysikaliska koncept snarare än effektfullheten av VR-laborationen.

På grund av tekniska begränsningar är exakt simulering av vätskor svår att genomföra i realtid, som tidigare diskuterats. Det finns dock lösningar som kan approximera vätskesimuleringar till att köras i realtid och en möjlig förbättring skulle vara att implementera ett sådant system. Detta skulle dock förmodligen öka komplexiteten kring beräkningen av flytkraft och därav öka kravet på hårdvaran applikationen körs på. Förutom detta kan applikationen fortsatt finslipas och utvecklas för att stödja till exempel fler klot i behållaren, andra objekt, och andra miljöer.

6 Slutsats

Med hjälp av en laboration som demonstrerar Arkimedes princip i en virtuell miljö har denna studie ämnat undersöka hur övertygande verkliga fenomen ter sig i virtuella miljöer. Resultatet visar att test- och kontrollgruppen är ungefär lika säkra på de svaren de angett i kunskapstestet de presenterades med efter laborationen. Däremot presterar de deltagarna som utfört laborationen i en virtuell miljö betydligt bättre än kontrollgruppen. Följaktligen påvisas det att personer övertygas om giltigheten av principen, att lärmiljön möjliggör inlärning och förståelse av den, samt att det finns ett samband mellan nivå av övertygelse och förståelse.

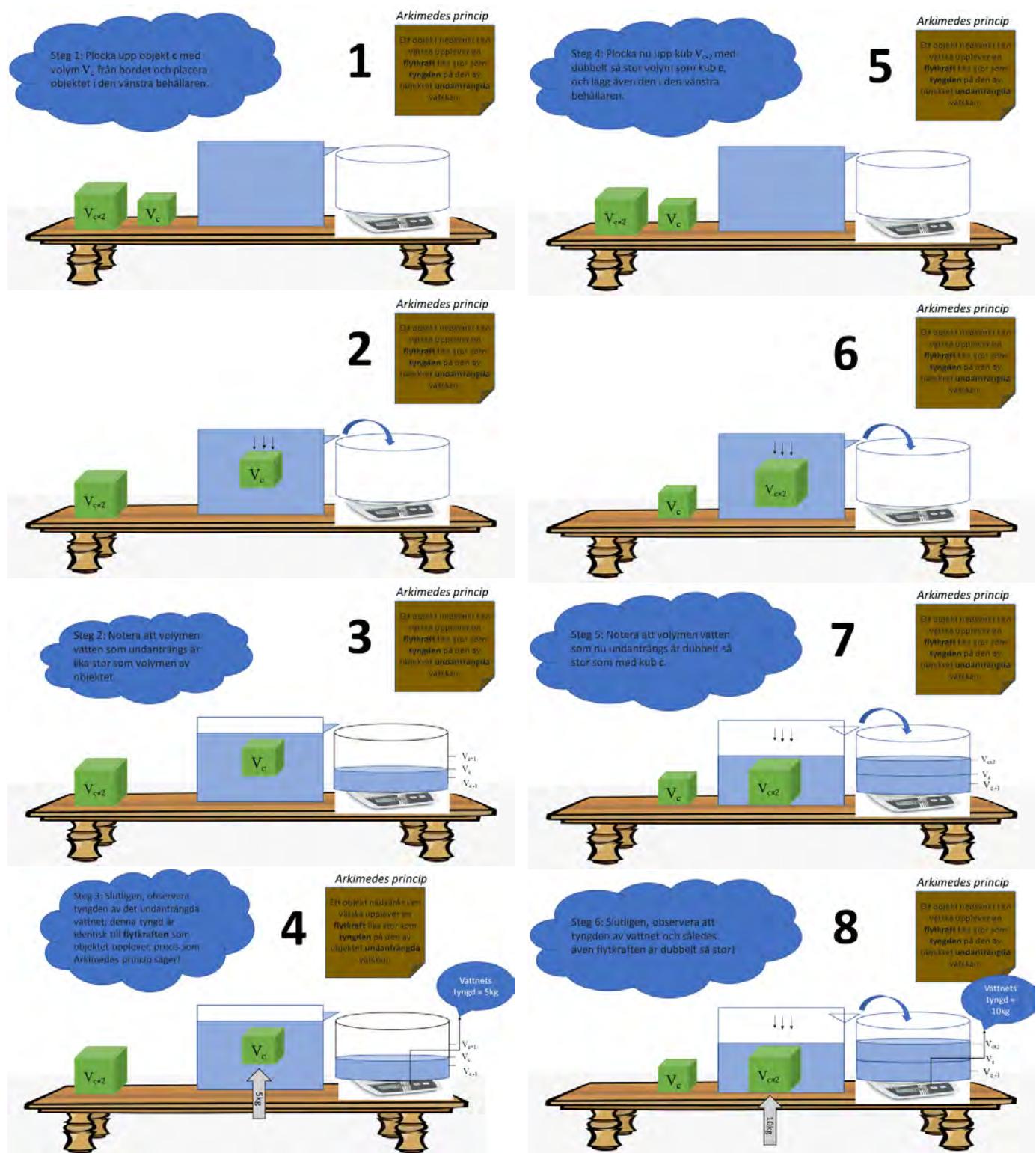
Referenser

- Åkerlind, G. (2015). From phenomenography to variation theory: A review of the development of the variation theory of learning and implications for pedagogical design in higher education. *HERDSA Review of Higher Education*, 2, 5–26.
- Albus, P. & Seufert, T. (2022, maj). Signaling in 360° desktop virtual reality influences learning outcome and cognitive load. *Front. Educ.*, 7.
- Anikin, A. (2022, 26/12). Personligt samtal.
- Baud-Bovy, G. & Balzarotti, N. (2017, november). Using force-feedback devices in educational settings: a short review. I *Proceedings of the 1st ACM SIGCHI international workshop on multimodal interaction for education*. New York, NY, USA: ACM.
- Bloom, B. (1956). *Taxonomy of educational objectives. book i: Cognitive domain*. New York: David McKay.
- Bussey, T. J., Orgill, M. & Crippen, K. J. (2013). Variation theory: A theory of learning and a useful theoretical framework for chemical education research. *Chem. Educ. Res. Pr.*, 14(1), 9–22.
- Dengel, A. (2020). Effects of immersion and presence on learning outcomes in immersive educational virtual environments. *Front. Educ.*, 9, 571126.

vironments for.

- Hasan, S., Bagayoko, D. & Kelley, E. L. (1999, september). Misconceptions and the certainty of response index (CRI). *Phys. Educ.*, 34(5), 294–299.
- Jicol, C. (2022). *A new framework for understanding user presence in virtual reality : from technical to human factors*. University of Bath.
- Johnson-Glenberg, M. C. & Megowan-Romanowicz, C. (2017, maj). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cogn. Res. Princ. Implic.*, 2(1), 24.
- Kullberg, A., Runesson Kempe, U. & Marton, F. (2017, augusti). What is made possible to learn when using the variation theory of learning in teaching mathematics? *ZDM*, 49(4), 559–569.
- Magana, A. & Balachandran, S. (2017). Unpacking students' conceptualizations through haptic feedback. *Journal of Computer Assisted Learning*, 33(5), 513–531. doi: 10.1111/jcal.12198
- Mayer, R. (2005). *The cambridge handbook of multimedia learning*. Cambridge university press.
- Mikropoulos, T. A. (2006, november). Presence: a unique characteristic in educational virtual environments. *Virtual Real.*, 10(3-4), 197–206.
- Minogue, J. & Borland, D. (2016, april). Investigating students' ideas about buoyancy and the influence of haptic feedback. *J. Sci. Educ. Technol.*, 25(2), 187–202.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in health sciences education*, 15(5), 625–632.
- O'Regan, J. K. & Noë, A. (2001, oktober). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.*, 24(5), 939–73; discussion 973–1031.
- Sartwell, C. (1992, april). Why knowledge is merely true belief. *J. Philos.*, 89(4), 167.
- Schubert, T., Friedmann, F. & Regenbrecht, H. (2001a). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3), 266–281.
- Schubert, T., Friedmann, F. & Regenbrecht, H. (2001b). *Igroup presence questionnaire (ipq) overview*. Hämtad från <http://www.igroup.org/pq/ipq/index.php>
- Schwind, V., Knierim, P., Haas, N. & Henze, N. (2019). Using presence questionnaires in virtual reality. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. doi: 10.1145/3290605.3300590
- Slater, M. (2009, december). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 364(1535), 3549–3557.
- Slater, M. (2018, augusti). Immersion and the illusion of presence in virtual reality. *Br. J. Psychol.*, 109(3), 431–433.
- Sullivan, G. M. & Artino Jr, A. R. (2013). Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education*, 5(4), 541–542.
- The Khronos Group Inc. (2022). *Openxr 1.0*. Hämtad 2022-10-31, från <https://www.khronos.org/openxr/>
- Unity Technologies. (2022). *Unity*. Hämtad 2022-10-31, från <https://unity.com/>
- Xu, X., Kang, J. & Yan, L. (2022, februari). Understanding embodied immersion in technology-enabled embodied learning environments. *J. Comput. Assist. Learn.*, 38(1), 103–119.
- Xu, X. & Ke, F. (2014, december). From psychomotor to ‘motorpsycho’: learning through gestures with body sensory technologies. *Educ. Technol. Res. Dev.*, 62(6), 711–741.
- Zsolnai-Fehér, K. (2022, oktober). The flow from simulation to reality. *Nat. Phys..*
- Zumbo, B. D. & Zimmerman, D. W. (1993). Is the selection of statistical methods governed by level of measurement? *Canadian Psychology/Psychologie canadienne*, 34(4), 390.

A Utdrag från flödesschema



B Kunskapstest

Recall:

1. Nämnn en egenskap av ett objekt som påverkar objekts **flytförmåga**, men som inte påverkar den resulterande **lyftkraften** när objektets sänks ned i vätska.

2. Volymen vatten som undanträngs av ett objekt är identisk till en särskild egenskap av objektet, vilken?

3. Nämnn en viktig egenskap av vätskor som påverkar **lyftkraften** som resulterar av att ett objekt sänks ned i vätskan.

Förståelse:

4. En guld bit med volymen 102 cm^3 och en frigolitbit med **samma** volym sänks båda ner **helt och hållat** i vätska. Hur förhåller sig lyftkraften mellan de två objekten?

- a) Objekten upplever lika stor lyftkraft.
- b) Guldbiten upplever större lyftkraft än frigolitbiten.
- c) Frigolitbiten upplever större lyftkraft än guldbiten.

5. När du placerar en potatis i en kastrull fyld till hälften med vatten så sjunker den till botten. Vilken av följande utgör den bästa strategin för att få potatisen till att flyta?

- a) Fyll kastrullen till toppen med vatten (dvs. öka volymen vatten i kastrullen).
- b) Tillsätt stora mängder med salt (ökar densiteten av vattnet).
- c) Få vattnet i kastrullen till att börja koka (sänker densiteten av vattnet).

6. Ett objekt flyter när...

- a) lyftkraften är lika stor som objektets tyngd.
- b) objektet upplever en uppåtriktad lyftkraft.
- c) objektets massa är mindre än volymen vatten som undanträngs av objektet.

Transfer:

7. Tänk dig att du sitter i en fiskebåt på en sjö, med ett väldigt tungt ankare i fiskebåten. Föreställ dig nu att du kastar ankaret i sjön. Vad kommer att hänta med vattennivån i sjön?

- a) vattennivån förblir densamma.
- b) vattennivån stiger.
- c) vattennivån sjunker.

8. Ballong A och B är fylda med syre respektive helium. Ballong A har dubbelt så stor volym som ballong B.

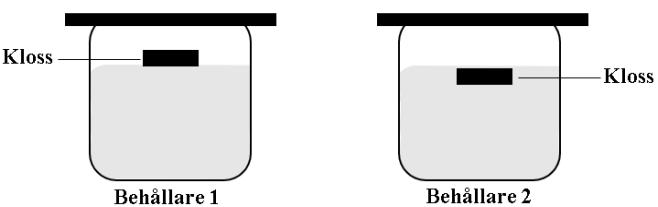
Föreställ dig att båda ballonger släpps från en höjd på 2 meter. Ballong A (med syre) faller långsamt till marken, ballong B (med helium) fortsätter lyfta från marken. Vilken är den bästa förklaringen till varför ballong B lyfter, men inte A? (OBS! Arkimedes princip gäller även för gasämnen!)

a) Ballong B har mindre volym och väger därför mindre.

b) Densiteten av helium (i ballong B) är mindre än det undanträngda luftens i omgivningen.

c) Ballong A (med större volym) tränger undan mer syre i omgivningen.

9. Kolla på bilden nedan. Den visar vad som händer när två objekt av samma volym och material sänks ner i varsin behållare med vätska. Baserat på vad du ser i bilden, vad kan du säga om klossarna och/eller vätskan i de två behållarna?



a) Klossen i behållare 2 väger mer än klossen i behållare 1.

b) Densiteten av vätskan i behållare 2 måste vara mindre än den i behållare 1.

c) Klossen i behållare 1 måste uppleva en svagare gravitationskraft.

C Manus i VR-miljön

1. Hej och välkomna till Arkimedes badrum! I dag ska du få utföra en laboration för att förhoppningsvis komma till samma insikt som Arkimedes gjorde för 2000 år sedan. Arkimedes princip säger att *"ett föremål som är nedränt i en vätska påverkas av en lyftkraft lika stor som tyngden av den volym vätska som objektet undantränger"*. För att ta dig igenom laborationen så kan du trycka på den röda knappen när den lyser upp. Till vänster av dig har du en whiteboard med formler och information som kan hjälpa dig genom laborationen. Till höger har du en whiteboard som kommer användas för att spara viktiga värden genom laborationen. Se dig omkring i labbet och när du känner dig redo att påbörja experimentet tryck på den röda knappen.
2. Vi ska börja med att undersöka hur volymen av en kropp påverkar vätskan den sänks ner i. Framför dig har du två behållare. Den ena behållaren innehåller vatten medan den andra behållaren är till för att mäta den mängd vatten som undanträngs när man sänker ner ett objekt i vattnet. Till vänster av dig har du ett antal olika klot, de större kloten har dubbelt så stor volym än de mindre kloten. Testa att plocka upp det lilla guldklotet och lägg den i den stora behållaren med vatten. När du lägger ett objekt i vatten kan du observera en pil uppåt som representerar lyftkraften, och en pil nedåt som representerar tyngdkraften. Observera hur mycket vatten som trängs undan med hjälp av att titta på den lilla behållaren samt vågen under den.
3. Notera att tyngden vätska har nu sparats på tavlan framför dig. Denna tyngd är, i enlighet med arkimedes princip, identisk med lyftkraften som objekten upplever!
4. Plocka upp klotet ur den stora behållaren. Du ska nu testa att sänka ned det stora guldklotet som har dubbelt så stor volym i behållaren. Observera återigen hur mycket vatten som trängs undan till den lilla behållaren med hjälp av vågen.
5. Volymen vätska som trängs undan bör vara cirka dubbelt så stor. Med hjälp av vågen kan du observera tyngden av den undanträngda vätskan, även denna kommer att ha fördubblats! Vidare, i enlighet med arkimedes princip, så har den sammantagna lyftkraften också fördubblats!
6. Efter att vi testat volymens inverkan på lyftkraften ska vi nu undersöka inverkan av densiteten. Testa att sänka ner det röda plastklotet i behållaren. Observera hur mycket vatten klotet har trängt undan och vad det undanträngda vattnet väger!
7. Ta bort plastklotet från behållaren och upprepa samma procedur med det stora guldklotet.
8. Observera att volymen och tyngden av det undanträngda vattnet är identisk i båda fallen. Detta innebär att lyftkraften som genereras också är densamma i båda fallen.
9. Plast och guld har inte samma densitet, dvs deras massa per volymenhets skiljer sig. Av denna anledning sjunker guldklotet till botten men plastklotet förblir flytande. Densitet har därför ingen direkt inverkan på lyftkraften som objekten upplever, men likväl spelar den en avgörande roll vad gäller objekts flytförmåga.
10. Vi ska nu testa vad som händer när vi varierar densiteten av vätskan som objekten sänks ned i.
11. På behållaren framför dig kan du reglera densiteten av vätskan med hjälp av reglaget bredvid behållaren. Sänk ner det stora frigolitklotet i den vänstra behållaren och notera tyngden av den undanträngda vätskan. Testa nu att minska densiteten av vätskan till den av etanol och observera hur tyngden av vätskan förändras.
12. Nu ska du undersöka hur lyftkraften varierar baserat på vätskans densitet. Ändra densiteten till vatten och observera lyftkraften på det lilla guldklotet.
13. Ändra densiteten nu till den av flytande guld och observera återigen lyftkraften på det lilla guldklotet.
14. Som du ser kan du få det lilla guldklotet att flyta när densiteten av vätskan är högre än densiteten av objekten, trots att objekten har samma volym och vikt.
15. Du är nu klar med laborationen - bra jobbat! Du kommer nu att få svara på ett par frågor.

Johansson, B., Gulz, A., Haake, M., Wallergård, M., Nirme, J., Ternblad, E. M., & Tärning, B. (Eds.) (2023). Intelligent, socially oriented technology VI: Projects by teams of master level students in cognitive science and engineering. *Lund University Cognitive Studies*, 185.