Transfer *Nearly* Completed

A VR shopping skills simulator for children with autism

Master Thesis

Emil Rosenlund Høeg & Luca Mangano

Aalborg University Copenhagen
Department of Architecture, Design & Media Technology
A. C. Meyers Vænge 15
2450 Copenhagen SV

Emil Rosenlund Høeg
<emil@rhoeg.dk>

Luca Mangano
<lu.mangano@gmail.com>

AALBORG UNIVERSITY
MASTER THESIS

Department of Architecture, Design & Media Technology
A. C. Meyers Vænge 15
2450 Copenhagen SV

Title:
Transfer Nearly Completed
A VR shopping skills simulator for children with autism

Theme:
Master Thesis

Project Period:
Spring Semester 2017

Participant(s):
Emil Rosenlund Høeg
Luca Mangano

Supervisor(s):
Stefania Serafin
Ali Adjorlu

Copies: 1

Page Numbers: 125

Date of Completion:
June 2, 2017

Abstract:

This Master Thesis presents a study conducted to investigate feasibility and effectiveness of Virtual Reality (VR) applied to Autism Spectrum Disorder (ASD) training, with particular emphasis on behaviour and adaptive skills. The study moved from the assessment of traditional adaptive skills training techniques, theories, and common practices, to subsequently focus on technology and VR applied to transfer of training. The system was built in collaboration with teacher from Valhøj Skole in Rødovre. A comparative experiment was conducted (n=9) with initiated training following a baseline decision-making tasks assessed in a real supermarket. After running 7 sessions over 2 weeks for the treatment group, participants were assessed again. Significant difference ($P < 0.05$) was found with a Difference in Difference regression with added predictors to explain causal interference, but due to the sample size the success of the treatment was substantiated on the basis of qualitative measures of performance.

Contents

Chapter 1

Introduction

> **Charlie:** He's not crazy, he's not retarded but he's here.
>
> **Dr. Bruner:** He's an autistic savant. People like him used to be called idiot savants. There's certain deficiencies, certain abilities that impairs him.
>
> **Charlie:** So he's retarded.
>
> **Dr. Bruner:** Autistic. There's certain routines, rituals that he follows.
>
> **Charlie:** Rituals, I like that.
>
> **Dr. Bruner:** The way he eats, sleeps, walks, talks, uses the bathroom. It's all he has to protect himself. Any break from this routine leaves him terrified.
>
> *From the movie "Rain Man" (1988) [1]*

At the 1989 Academy Awards, Dustin Hoffman won *Best Actor in a Leading Role* for his portrayal of the autistic Raymond in the movie Rain Man (1988). He maintained his presence with reinforcing conviction, basing his performance on close observations of two individuals with Autism Spectrum Disorder (ASD). The character Hoffman portrays has incredible savant skills, that allow him to conduct complex calculations within seconds. Skills that incidentally are abused by Raymond's brother Charlie (Tom Cruise) in Las Vegas, for remunerative gains.

Although Hoffman's character may accurately capture the obsessive mental preoccupation of some people with autism, also referred to as savant syndrome or savant autism, it only has an estimated prevalence of 1:10 (in people affected with ASD), but no definitive statistics exist. This emphasizes the importance of recognizing that autism is not a binary condition, and that the cognitive features displayed by Raymond are a rare occurrence. In reality autism should be viewed as a spectrum of disability, inclined to change over time, hence making people affected by ASD a population of broad diversity.

Autism has been studied since the early 1900s and its definition changed throughout the years, nowadays it is defined as a broad and varied spectrum, which ranges from low- to high-functioning. Individuals diagnosed with high-functioning autism can present exceptional, or above average, intelligence and skills -as Raymond in Rain Man-, while low-functioning can present medium to severe learning disabilities. Independent from the position in the spectrum, all individuals diagnosed with ASD can present characteristics such as repetitive behaviour, social and communication difficulties, and selective focus and interest[2]. Amongst the range of impairments ASD patients might have, it was decided to focus the scope of this study on behaviour and adaptive skills.

Traditionally, behaviour therapy would require specialists and families to develop training sessions and techniques around the individual's needs [3]. The patient would undergo a process in which they would be guided and trained to react and adapt to a wide variety of scenarios[4]; several training techniques have been found when reviewing ASD studies, and it has been highlighted how simulation-based learning can be the most beneficial when training is mediated by technologies.

Several studies have investigated and experimented the effects of technology in ASD treatment, and promising results have been highlighted by many [5; 6]. More specifically, Strickland's work [7; 8] has been key in defining guidelines and techniques to be adopted when developing Virtual Reality (VR) systems for ASD, and concluded that for the time being (1990s) the effectiveness of such systems would not overcome their prohibitive cost. With the recent widespread of commercial VR headsets, Strickland's equation can be redefined and VR technologies are increasingly more accessible by institutions caring for people affected by ASD.

The study presented in this report deals with the development and evaluation of a VR behaviour and adaptive skills training systems for teachers and pupils at Valhøj Skole in Rødovre. Amongst the hundreds of students, the school welcomes children and young adults diagnosed with ASD, independent from where they fall in the spectrum. Part of their training consist of home economics and adaptive training, which is conducted as field trips to the local supermarket Føtex in Rødovre Centrum; this specific training is aimed to help students develop skills towards an safe and independent shopping experience. Many of the students have been observed to have difficulties with money management, alternative products selection, and general discomfort in the supermarket.

In order to conduct the study, a problem statement and three research questions were formulated:

> *To what degree can a VR simulation facilitate shopping skills in adolescents with autism given the following conditions:*
>
> - *Can a VR training simulator instill comfort and confidence around a real shopping experience?*
> - **Can skills taught in a virtual supermarket be transferred to a real context?**

Entering this study, we started from the analysis of autism and the different behaviour training approaches that have been adopted throughout the years. We then focused on the analysis of adaptive skills training, and its augmentation with the use of technology; particular emphasis was put on the analysis of VR and Virtual Environment (VE) systems. Methodologies and evaluation techniques used throughout the study have been reported, together with knowledge gathered over several meetings with the collaborators. The iterative development process of the system is then illustrated, and the most relevant aspects are presented. Once fully created, the system was then tested by the pupils, and results are presented. Finally, these results have been analyzed and discussed, together with methodologies, and development.

Chapter 2

Analysis

2.1 Autism

Symptoms, cause and diagnosis

The term autism was first scientifically defined in 1943 by Leo Kanner, who identified it as the innate inability of a person to create normal, biologically determined, emotional contact with others. The definition has evolved throughout the years, autism is nowadays referred to as Autism Spectrum Disorder (ASD) so to include the whole range of behaviour, communication, and interaction deficits. The broader definition has been adopted to accommodate the personal nature of the disorder, in fact:

> *Even though there are strong and consistent commonalities, especially in social deficits, there is no single behavior that is always typical of autism or any of the autistic spectrum disorders and no behavior that would automatically exclude an individual child from diagnosis of autistic spectrum disorder.* - [9]

Although there is no universal test to diagnose autism, several guidelines and rating scales are used to assess whether a child is affected by it, and ASD can be reliably diagnosed as early as 18 months of age. As of 2017 the prognosis of the global prevalence of ASD is 62 cases per 10.000 people [10], but the estimate varies greatly from country to country. At all ages autism impairs individuals' ability in decision making; this has been attributed to the disorder's inherent poor audio-visual integration, possibly reflecting dysfunctional 'mirror neuron' systems which have been hypothesised to be at the core of the condition[11]. Whether a choice needs to be taken swiftly, it involves a change of routine, or it involves communication with other individuals, the process of decision-making has been described as exhausting, overwhelming, and — at times — anxiety-provoking [12]. Anxiety and exhaustion were found to bring adult individuals to the brink of depression, thus highlighting the need for family and experts to guide and support them through the decision-making process so to alleviate the cognitive overload and discomfort they are often subject to [12]. Traditionally, autism has been treated by reproducing or simulating factors that would trigger unexpected behaviours, and guiding the individual through a process of understanding, while training them to overcome eventual obstacles. Several behaviour-training intervention techniques were analyzed and, although some were more effective than others, no technique had outstanding favourable results [3].

Being able to distinguish between the self and others, and to be able to infer that the mental states of others may differ from our own, involves the development of Theory of Mind (ToM)[13, p 616]. It is hypothesized to be an innate ability in primates, humans included, to understand that others have minds different from our own[14]. ToM does not come as a complete package solution, but requires social, learning, and societal interactions and contexts over many years to fully develop. Within the context of ASD, it is often stated that one of the core behaviors deficits of autistic people is a lack of understanding for ToM[15]. Baron-Cohen *et al* [16] coined the term *mindblindness* to describe the ToM-deficiency in people with ASD, because their understanding and/or appreciation of other people's mental states and non-verbal cues is abnormal[13, p633]. According to Astington *et al* [15] the first 5 years of development are the most critical for establishing ToM, which explains why early efforts to diagnose, treat and train autistic children is vital to their future development[9; 17].

> *"The way he eats, sleeps, walks, talks, uses the bathroom. It's all he has to protect himself. Any break from this routine leaves him terrified."* Doctor Bruner - "Rain man" (1988) [1]

What Dr. Bruner explains in Rain Man is the symptom of insistence to sameness i.e. a general trait of autism expressed as a dedicated to routine, that quite often does not involve other individuals[13, p 617]. Any break from that routine can result in strong reactions, such as temper tantrums. This is widely regarded as a main hindrance in an autistic individual's success in generalization and adaption of behaviour across settings, time, and people [18; 19].

The analysis presented in the following paragraphs, offers more in-depth insights on ASD treatment, behaviour therapy, with particular emphasis on social, adaptive, and generalization skills training. Consecutively, an overview is presented about what technologies and methods have been used in the more recent years to augment, modulate, or influence traditional training, with Virtual Reality (VR) and its applications being the main focus.

2.2 Treatment interventions

As reported by K. Francis [3], interventions for treating people with ASD can vary from behavioural treatments, medicine (usually for sequelae conditions), or both. There is no pharmaceutical cure, and it is considered a life-long neurodevelopmental disorder, although there have been cases of people being "cured" (i.e. children may mature out of certain types of autism), treatments can be so beneficial that they fall outside autism rating scales.

2.2.1 Behavior therapy

Social skills training

Among therapy or intervention-methods is social skills training (SST), which seeks to train interpersonal responses and relations. It teaches the child rudimentary verbal and non-verbal communication techniques, that will allow him/her to adapt to the environment (e.g. *non-verbal:* eye-contact, raising the hand in school, and reading facial expressions and emotions courtesy, showing appreciation, *verbal:* politeness, showing appreciation, speech intonation and verbal disruptions)[20].

Adaptive skills training

Adaptive skills, on the other hand, refer to the social responsibilities and performance during daily activities, usually centered around the mental process of choice behavior. The process of consciously performing decision-making based on evaluation of multiple options is something that is inherently difficult for people with ASD [9]. Adaptive behavior not only focuses on choice, but also behavior conditioning such as toilet training, brushing teeth, purchasing skills, and essentially all skills required to function in a community and at home[9, pp. 103-104].

In training for shopping in a supermarket (i.e. purchasing skills), the children would need to apply both social and adaptive skills. Social skills are required to interact with employees and cashiers, navigate among and around other shoppers, while adaptive skills are required to remember to take a basket or cart, make use of external aids such as signs and aisle numbering for navigation, choice decision in product and quantity, as well as keeping track of money.

The focus of this study is the non-social aspect of purchasing skills in a supermarket, which can pose a challenge to people affected by ASD. Thus, more effort was put into researching autism from an adaptive behavior point-of-view.

2.3 Training and learning
A cognitive perspective

2.3.1 Generalization

Generalization refers to the transfer of learning from one situation to another, with no preconditioned explicit training in the unfamiliar scenario. In short, it can be described as the individual's ability to adapt preexisting acquired knowledge to a new task or context. It is a continuously described problem in autism interventions, that even though an intervention effectively teaches the individual new behaviors and skills in a specific case, it oftentimes falls short when the individual is presented with a deviation of the task [18].

Dr. Trevor Stokes and Dr. Donald Baer are recognized by many as the most prominent behavior analysts, who have affected behavioral research in regard to generalization, since 1977. Together they reviewed more than 120 studies on behavioral interventions, and identified 9 general strategies that are found to be most commonly effective in promoting generalization in individuals with ASD [4]. Furthermore, these practices have been proven to increase the treatment effects in adaptive behaviours and skills training. The basic strategy relies on a "train and hope" approach, which Stokes and Bear suggest to avoid; this, together with more advanced strategies for promoting generalized behavior were categorized under nine general groups based on the 120 individual studies:

1. *Train and hope* - basic technique focused on the hope of the natural development of generalization ability by the individual over time.
2. *Sequential Modification* - generally regarded as a common practice, it is centered around the measurement of generalization effects, and the eventual subsequent intervention in case of absence or deficiency of sought after effects.
3. *Introduce to natural maintaining contingencies* - practice based on the introduction of the individual to natural social interactions and learning environment, in order to achieve the desired generalization effects. The skill practiced needs to be rewarding to the participants within their environment.
4. *Train sufficient exemplars* - mainly focused on situation diversity, this practice required the individual to be exposed to several different vari-

ations of the same social interaction in order to develop generalization on a wide-range.

5. *Train loosely* - with this practice the individual should be trained to answer to stimuli in several different ways, rather than focusing on strictly following a script.
6. *Use indiscriminable contingencies* - individuals are trained to familiarize with intermittent reinforcement, i.e. while training not all the responses given by the individual should be rewarded, so to have a closer comparison to the real world, where we are not always rewarded for appropriate social interaction.
7. *Program common stimuli* - this practice motivates to determine which stimuli the training and the real-world environment have in common, and to include as many of them as possible into the training sessions.
8. *Mediate generalization* - individuals should be encouraged to state their intent to engage in social behaviours before doing it; this highlights a marginal behaviour that is not always present in young children.
9. *Train to generalize* - oftentimes practices focus on training generalization for precise social interactions, while it should be as important to train individuals to adopt the learned skill in several different contexts.

Training and hoping generalization of social or adaptive skills, is the typical approach in traditional child development, i.e. hoping that generalization will be facilitated naturally. However, for individuals with intellectual disabilities such as ASD, generalization seldom occurs without predetermined plans or strategies [21, p. 3150]. With this approach, effects of generalization may be documented, but are not necessarily actively investigated or pursued. However, for the most part such effects are welcomed, but "not explicitly programmed" [4]. To increase chances of generalization, Stokes and Baer recommended the implementation of explicit or systematic strategies[4]. Three of them have been reviewed more in depth due to two of them i.e. sequential modification or introduction to natural maintaining contingencies which are very well used, but also the strategy of programming common stimuli which have previously been applied to virtual contexts [22].

Sequential modification

Sequential modification studies utilize a systematic approach to generalization. Generalization-effects are measured, and in case effects are absent or deficient, predetermined procedures are put into play to facilitate change in non-generalized conditions [4]. In other words, if generalization does not occur, effective intervention strategies that are known to produce stimuli, such as settings, people, and situations are confluently utilized to achieve results (e.g. training in familiar environments — a "safe space" — with members of the family or teachers). It is generally regarded as common practice, but it is also regarded as less effective and time consuming to simply wait for an effect to occur. It also requires a coordinated effort between parents and school community, to implement strategies consistent across key influencing milieus [18].

Introduction to natural contingencies

Natural contingencies operate without social mediations, such as positive or negative reinforcements. Naturally occurring behaviors solicited in everyday life cause logical consequences of actions. Therefore, in order to facilitate generalization, the child may only need to be exposed to natural reinforcers present in natural social interactions and learning environments to properly achieve the effect [4]. Initially, this may sound like natural conditioning involving nothing but aggregation from consequences, but the strategy does not seek to expose the individual to unfamiliar experiences, without first supplying them with tools acquired through video-modelling or other simulated realities.

Program common stimuli

The strategy revolves around recognizing that there are certain similarities between the different settings where generalization will occur i.e. there are some *common stimuli* that could be included to facilitate the transfer from one environment, activity or social context. Fundamentally it is based on the assumption that, the more common, salient stimuli that are equally present in both training and non training setting, the more easily generalization will be facilitated [23]. In a study using the online virtual world *Second Life*, Mason *et al* used common stimuli between the VE and a natural environment to facilitate the transfer of training of a face-to-face tutoring [22].

2.3.2 Simulation-based learning

A simulation is a technique, not necessarily of technological nature, in which practice and learning can be achieved, while protecting the learner from dangerous or harmful situations. It can actively place learners in a realistically modulated situation, that can evoke a response-as-if-real to the virtual context. For example, training to be a pilot in a flight simulator can place the trainee in a complex and/or dangerous situation, yet in a safe and relatively risk-free virtual environment. Furthermore, it can be used to evaluate different factors of task performance through observation facilitated by the increased amount of controlled experimental factors, that might otherwise be difficult to measure. As we learned in the previous section, transfer of training (or learning) occurs when learning in one context is transferred to performance in another[24]. Training in a simulated graphical representation of reality, also known as a virtual environment (VE) [25; 8; 26], hence begs the question about how well acquired training can be transferred to a real environment, context or situation, and furthermore what tools can be applied to facilitate optimal learning. For a VE training to be successful, the transfer process should allow the corresponding real task to be completed effectively according to predetermined criteria[27] (i.e. some tasks require higher level of precision and lower tolerance for errors). According to several sources[28; 29; 30], multimodal interfaces can generally create better learning conditions through *inter alia* cognitive integration and limited cognitive load[31]. Furthermore, it has also been suggested to be especially effective in situations involving people with ASD, e.g. Williams *et al* concluded that, due to successful training and recognition of visual speech in bimodal conditions, children affected with ASD might benefit from multimodal integration in imitative therapy and language training[11]. In essence, a virtual simulation can immerse the learner in a situation where tasks can be performed and repeated in safe surroundings i.e. safer learning environments. Several studies have found convincing evidence for a positive transfer of learning from a VE to a real environment; Wilson *et al* successfully reinforced spatial abilities in children with physical disabilities, by training emergency drills and evacuation in a VE [32], Rose *et al* showed a clear positive transfer effect in a simple sensorimotor wire loop task [33], and most interesting Cromby *et al* demonstrated a successful transfer to a real world scenario of shopping skills taught in a virtual supermarket on pupils with severe learning difficulties[34] (see section 2.6). The Low Road-/High Road theory, developed by Salomon and Perkins [35] introduced the model comprised of two complementary terms and conditions that are widely used in transfer studies[27]. The taxonomy offers a categorization of transfer methods and corresponding cognitive processes, based on the assumption that transfer ensues when learning in one context enhances (i.e. a positive transfer) or decrease (i.e. a negative transfer) an associated performance in another context[24].

The following list accounts for some of the most applied elements.The model has since been further expanded by other researchers.

1. **Context outcome:**
   (a) *Positive transfer* - when learning in one context or situation renders an enhanced performance in another related context or situation.
   (b) *Negative transfer* - when learning in one context or situation renders a decreased performance in another related context or situation.
2. **Conditions of transfer:**
   (a) *Near transfer* - Transfer between overlapping contexts, formed in situations very similar to the initial context of learning.
   (b) *Far transfer* - Between dissimilar contexts that facilitate connections to contexts that, despite intuitively appearing different, forms a transfer of learning.
3. **Mechanisms of transfer:**
   (a) *Low road transfer* - Also known as reflexive transfer. Transfer through practice routines that are similar, in terms of stimuli, to those in the intended real situation.
   (b) *High road transfer* - Also known as mindful transfer. It is less dependent on superficial stimuli, as the person can often conceive deeper analogies through reflective abstraction.

In an article from 2008, Bossard *et al* [27] describe VEs and corresponding transfer as a binary definition, i.e. transfer based on either procedural or declarative knowledge. Procedural learning refers to skills acquired on the basis of proficiency, associated with efficient and automatic behaviors. Furthermore, they suggest that most training in industrial settings and motor skills focused on procedural and near transfer, which emphasizes that acquired skills generate as a consequence of repetition. Declarative learning refers to knowledge that require deeper comprehension to establish abilities and skills based on generalization and abstractions [27]. For VEs used in education, the focus appears to be on declarative knowledge and far transfer of learning[27]. Furthermore, Bossard *et al* highlight some suggestions for creating virtual training simulations *inter alia* that the context should facilitate autonomy (i.e. agency), contain various assistive elements, and that the learning environment should be rich and offer realistic responses.

2.3.3 Memory retention

Memory retention is the mental process of retaining information acquired through sensory modalities, for a longer period of time[36, p. 216]. It could also be called the preservation of memory within the mind, and can be considered a catalyst for the ability to recall and reproduce behavior. It is important to consider for training interventions, because an initial significant effect is seldom the only objective, but rather a longer, and hopefully constant, maintained effect. To achieve good maintaining performance over time, all of the distinct conceptual stages of newly acquired information i.e. *acquisition*, *encoding*, and *retention* of new information, must be accomplished successfully. A poor performance, or lack of maintenance might be attributable to a faulty process in of the stages, therefore all of them must be successfully processed to accomplish good performance [36, p. 216].

The German psychologist Herman Ebbinghaus developed, in 1885, a mathematical representation of the exponential rate for memory loss (the forgetting curve) given no intervention to retain it. He suggested that approximately 50-60% of the acquired information is lost within the first 24 hours[36; 37] (figure 2.1). Furthermore he developed a theory on learning, called the spacing effect which essentially suggests that learning retention is greater if exposure to information is spread out over time.

$$retention = e^{-\frac{time}{relative\ memory\ strength}} \tag{2.1}$$

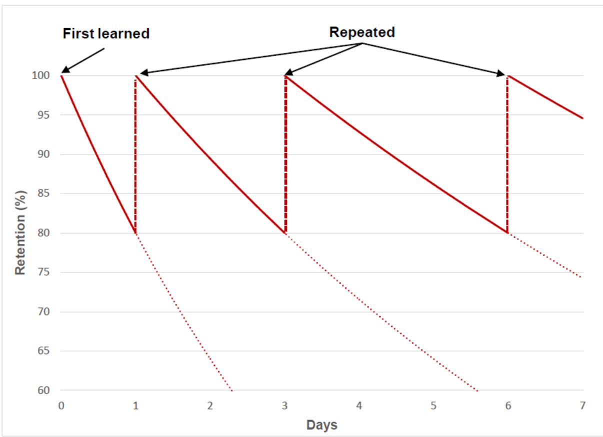


Figure 2.1: (left) Retention equation of exponential decay and (right) Ebbinghaus' forgetting curve, graphically representing the theory of memory decay when new information is first learned, and how the decay-rate decreases as the relative memory strength increases, as a result of repetition. ***Source:*** *from figure and description* [36, pp. 216-218]

The theory continually verified that spreading out exposure or training improves e.g. in free-recall and frequency-judgment tests[38], in learning and memory after traumatic brain injuries[39], and to successfully improve maintenance of training in children with ASD[40]. Some of the reviewed literature validate retention by post-training rest for a period of 24-48 hours [41]. This must, however be considered a practicality rather than an optimal way to validate retention, although it appears to be a well-grounded way to limit carryover effects. Only a very little portion of the studies go as far as to check again after extensive periods to determine to what degree the training was maintained over time. For example Nikopoulos & Keenan followed on the maintenance of generalization taught with video-modelling one month after the treatment, and again 2 months after [42],

2.4 Technology in behaviour training

Due to the vast nature of ASD, several studies can be found dealing with the attempt of behaviour and adaptive-skills training children and young adults using technology. The scope of technology employed in such studies is vary, and ranges from more traditional techniques (e.g. VHS and DVD video-help) to more cutting edge technologies such as Virtual Environments and VR [5]. In general, studies analyzed in [5] have not given conclusive answers on whether technology-based training could significantly affect participants performance in everyday activity. Overall, considering repeated measure tests, it was noted how — even though longer exposure could eventually lead to positive results — the proper administration of the treatment was a key for success; studies in which specialists in the ASD field administered the training were found to be more successful than ones in which the training was self-administered or guided by non-qualified individuals.

Video-modelling

Video-modelling techniques have been widely used in ASD social training; video clips were often pre-recorded in order to develop a storyline and have the participant train in decision making based on the visual cues. [42] used this technique to assess childrens willingness to engage with the experimenter either verbally or in an implicit manner. The study brought to the conclusion that not all the conditions would bring to generalization of learned behaviour, and often children would need cues from the environment in order to trigger the behaviour they have been exposed to. Moreover, parents found that children were more willing to engage in play with other people after the training, which brought them to a behavioural level comparable to traditionally developed siblings.

A meta-analysis [43] assessed the effectiveness of studies that adopted video-recording techniques as their main means of training; a list of guidelines was extrapolated from the evaluation of the studies, and the main concepts emerging from this guidelines are hereby reported:

- Reinforcements - several studies were found to successfully adopt error correction and reinforcement principles in order to correct and sediment concepts in the participants;
- Imitation - it is important to assess children based on their model imitation rather than only on their comprehension;
- Attention - if children are able to keep their focus on the message for more than one minute (rather than focusing on "noise" e.g. actor clothes) they are more likely to benefit from the training;
- Length - the video should not be longer than five minutes, in order to not require a too extensive focus period;
- Model - effectiveness is independent from what model the participants are exposed to.

In principle, the guidelines presented could be adapted and applied to VR applications meant to aid ASD users more effectively. Although some of the principles can be directly transposed to virtual reality (i.e. reinforcements, length, model), special care should be adopted when dealing with attention and imitation, due to the immersive nature of head mounted displays, users might have issues focusing on a single aspect of the fully virtual environment and might not be able to locate the proper model to imitate.

Computer-Assisted Technology

Lastly, the effectiveness of Computer-Assisted Technologies (CAT) has been assessed [6] and benefits have been found by the adoption of properly developed systems. Key focus of the development should be the users' abilities and eventual shortcomings; these must be taken into consideration since unexpected properties of the system could triggered unwanted reactions and eventual misuses of the technologies. Whats-more, CAT systems can — at a certain degree — provide proper training, sometimes with higher precision compared to teachers. Although, it is vital that such systems have a sound scientific approach, thus requiring an expert involvement. The analysis [6] concludes with the affirmation that although CAT could be an effective training tool, and the highest peak of effectiveness is achieved when teachers or educators are allowed to adopt such tools and mediate them to the children.

An example of CAT systems is the initiative *Unges UddannelsesCenter* launched by the Danish association for autism. They aimed to help children and young adults affected by autism in everyday life tasks; they created the smart-phone application *Scan How*[1], which allows users to scan QR codes and get a visual guide on how to accomplish a certain task. The guides offer step by step help for routine tasks such as doing the laundry, finding the correct products to buy at the supermarket, and finding one's way around the city.

Lastly, Cromby and Standen used CAT to train adaptive skills in students with severe learning difficulties, using a virtual supermarket. The importance of detail and recognizability was noted, and it was suggested that the correct balance between the important salient information should be achieved, in order to neither render the environment bland and generic, nor overload the users' perception with superfluous details. The study proved that the repeated training in a virtual supermarket can bring to a statistically significant increase in accuracy and completion time, compared to traditional training. The researchers observed that the structure and repetition of tasks helped the treatment group familiarize with the tasks and subsequently advantaging them over the control. However, they conclude that maintenance and retention of such results was unknown, expressing the need for further studies focused on the identification of critical features that would include such systems in the classroom environment [34].

2.5 Autism and VR

VR has been widely adopted as therapeutic tool in ASD training since before its introduction to the general market, Strickland in 1996 [7] already conducted studies in this field. Recently, VR has been used for treatment of a wide variety of dysfunctions derived from ASD — e.g. social phobias, communication impediment — and certain studies [44; 45] confirmed the efficacy of these methods. [44] confirmed that over an extended period of exposition, VR leads to the significant treatment of fears and phobias for a sustained period of time (up to 16 months after the intervention). The outcome highlighted the possibility of adoption of these technologies in several fields touched by ASD impairments; being phobias triggered in public settings — e.g. public transport, supermarket, classrooms — the use of VR would allow for a more flexible and controlled exposition to the different factors. Similarly, [45] dealt with the impairment ASD patients faced when compared to job interviews and, although not significant, results showed the applicability and efficacy of such a VR system.

[1]Scan How mobile application to assist people with ASD with daily routines and chores: `http://scanhow.com/`

Several studies have also dealt with the application of VR to specific settings of social interaction and training. [46] suggests how VR can be a highly helpful tool in social skill training for ASD patients. Since current training techniques would mainly involve one-to-one interaction and repetition, VR would allow for an ideal practice environment thanks to the minute possibility of variables manipulation and immediate repetition of certain settings. Voice overs, visual cues, and step-by-step help would be highly recommended as a starting point, so to train the users to accomplish certain actions and slowly freeing them from the guidance leash. Using a similar approach, studies [47; 25; 48] had participants train in a 3D VR simulation, aimed to develop interaction and communication skills. Overall results showed improvement in all the fields, but also highlighted the need for specific development for each participant, so to satisfy the particular training a participant would require [47]. On top of this, [25] proved how a more frequent exposition to the training, would increase the efficacy while being able to decrease the overall experiment period. Lastly, [48]'s preliminary results showed great potentials for the application of VR as an enabling tool to facilitate social training and interaction for children.

All the studies presented above, pertain similar results in regards to applicability and structure of VR systems in regards to ASD treatment and exercises. Each study was developed following certain guidelines, and they can all be summarized based on Strickland's insights on the advantages VR adds to traditional therapy [8]:

- **Controllable input stimuli** - the virtual environment can be simplified to the point when the auditory/visual stimuli are reduced to a bearable level for the user, thus to avoid triggering of unwanted reactions such as phobias or reluctance;
- **Modification for generalization** - similarly to the previous point, the environment can be manipulated so to allow users to generalize a concept that has been learned in a previous (similar) scene;
- **Safer learning environment** - the overall virtual environment allows to be more forgiving about mistakes and overall stimuli can be tweaked so to gradually adapt the user to a realistic environment;
- **Primarily visual/auditory world** - being ASD patients's thought patterns primarily visual, it has been shown that VR could comfortably be a mean of teaching;
- **Individualized treatment** - since individuals affected by ASD can be influenced at different degrees, VR allows for a more personal and particular experience, structured around strength and weaknesses of the user;

- **Preferred computer interactions** - being this kind of interaction deprived of the social aspect, users are more inclined to focus on the learning point other than trying to overcome possible social impairments they have;
- **Trackers** - tracking the vestibular system could be beneficial for both teaching gestural expressiveness and track the performance in the VE.

VR has been overall regarded as a positive and engaging tool for enhancing behavioural training, Casanova *et al.* [49] emphasize how this technology, especially administered through head mounted displays (HMDs), can increase the sense of immersion and presence in children and young adults with ASD. Furthermore, they they argue that virtual simulators on regular monitors cannot justifiably be called VR [49]. Presence and immersion are key factors for the success of VR simulations [50]; the first is achieved through the realistic and appropriate stimulation of senses as a consequence of one's actions within the virtual environment, while the latter refers to the VR system's ability to recreate a perceptually realistic environment. These principles apply to VR experiences independent from the particular system used to convey them, which have been several throughout the years.

Before the introduction of head mounted displays (HMDs), the term *virtual reality* was used to refer to experiences that could be explored through screens, CAVE systems, or purposely built devices. Bowman *et al.*[51; 52] evaluated human behaviour and performance in both HMDs and CAVE systems; results from both studies showed how, from a navigational perspective, HMDs significantly outperform other systems bringing them to explicitly set a base guideline that reads:

> *For VE applications involving navigation through enclosed spaces and frequent turning, choose an HMD with head tracking to provide increased efficiency and spatial orientation to users.* - [52]

To conclude, recent years have seen a widespread adoption of VR headsets in ASD treatment, not only thanks to the more affordable prices, but mainly thanks to researches showing how these technologies can improve performance in everyday tasks training [53].

2.6 An overview of related work

Shopping in a real supermarket requires a variate of different skills: memorizing the products on the list, acquire, organize and utilize the knowledge gained from e.g. the spatial environment to purposely use assistive elements to visually search for the products, while keep track of where you have already been. Naturally, prior knowledge attribute to the required skills. Those skills are usually naturally facilitated through repetition and generalization, but for people with ASD it can be difficult due to cognitive deficiencies, but also to transfer learning through training in the real environment, because such environments poses a challenge or even phobia. In a supermarket, many factors collectively attribute to a stressful experience for most people, and even more for people with ASD. Is not sufficient to prove that skills have transferred from one situation to another, either by near or far transfer, but it is also necessary to validate whether or not maintenance has occurred. The retention period is oftentimes treated as less relevant. However, it is in many aspects a phase of great importance to ensure long-term success, and to validate that maintenance has been achieved. Optimally researchers would continuously test for maintenance after relative long time intervals, but due to practicalities the effect is often validated with a 24-48 hour resting period [41].

The researched studies have been examined, and the most relevant to the purpose of the current study have been gathered in table 2.1. The table reports technology centered studies relevant to the investigation — mostly, but not exclusively VR — and sample sizes, study duration, means of experience conveyance, and reported by year in ascending order.

The sample sizes used in these studies is varying greatly (n=2-33, median = 8), assumable due to the prospected 0.0067% global prevalence [10]. It is also interesting to note that most of these studies utilize a relatively large amount of training and exposure sessions with a median-score of 22 sessions. Furthermore these sessions are usually spread out over a longer period of time (with a median experimental period of 5 weeks) often using spacing (see spacing effect, section 2.3.3) with 4 sessions per week. Also, only two of the studies used HMDs [7; 8] and two studies used CAVEs [44; 48]. 3 of the studies report to have used VR, but only used single-monitor desktop setups which are sometimes referred to as low-immersive VR [54] or according to Casanova, M. *et al*, should not even qualify to use VR-terminology [49].

Table 2.1: Overview of selected ASD experiments

Study	Author(s)	Year	Type	Sample size	VR type	Duration
Effects of Videotape Instructional Package on Purchasing Skills of Children with Autism [55]	Alcantara P. R. et al	1994	Case study	3	Not VR	3-5 sessions/week, 38 sessions
Successful transfer to the real world of skills practised in a VE [34]	Cromby, J. et al	1996	Between group design	19	Desktop	2 sessions/week, 11 weeks
Case Studies Using VR as a Learning Tool [7]	Strickland, D. et al	1996	Case study	2	HMD	4-7 sessions/week, 5 weeks
VR for the Treatment of Autism [8]	Strickland, D.	1997	Case study	2	HMD	40 sessions/six weeks
VR Social Cognition Training for Young Adults with High-Functioning Autism [25]	Kandalaft, M. R. et al	2012	Case study	8	Desktop	10 sessions/5 weeks
Reducing Phobia in Young People ASD through a VR Intervention [44]	Maskey, M. et al	2014	Case study	9	CAVE	2 sessions/20-30min.
VR Job Interview Training in Adults with ASD [45]	Smith, M. J. et al	2014	Between group design	26	Desktop	5 sessions/within a 2-week period)
Video Self-Modeling and Peer Training to Increase Social Engagement in Preschool Children with ASD [56]	Bellini, S. et al	2016	Case study	3	Not VR	7 sessions/week, 3-5weeks (individual)
Social Adaptation in Inclusive Education Settings for Children with ASD [48]	Horace, H.S. et al	2016	Case study	33	CAVE	28 sessions/14 weeks

2.7 Multimedia considerations
Developing for people with autism

As previously established (section 2.4) children and adolescents with ASD have not been found to have problematic relationship with technology. However, that does not mean that there are not certain design criteria that, unless properly considered, will have a negative impact on the experience. The research into 3D user interfaces (3DUI) and usability in ASD treatment is very limited. Some literature suggests that people with ASD display superior abilities, compared to traditionally developed people, when it comes to spatial cognition tasks. Minshew *et al* showed that children with ASD often outperformed normally developed children on visuospatial abilities [57]. The notion of enhanced visuospatial was confirmed by [58] in another study involving map learning tasks. The group with high-functioning autism outperformed the control group in accuracy abilities to recall the path through a 2D maze. 9 out of 16 of the test participants with high functioning autism learned the map faster than the fastest participant from the normally developed comparison group [58]. The notion that people with ASD generally perform better in visuospatial tasks might imply that such abilities are also present in 3DUI tasks. However, children with ASD have also been found to have a lower hand-eye coordination compared to normally developed children of similar age [59], and Falter *et a*l showed that children with ASD have a significantly lower performance, compared to typically developing children, in targeting tasks in a 2D environment [60].

2.7.1 Hand-eye coordination

Although a lower hand-eye coordination implies that interactions might be negatively affected in a 3D environment, Mei *et al* suggests that the effect may potentially be compensated by their superior spatial ability [53]. Mei *et al* compared the effect of object manipulation given different two transformation tasks (translation and rotation) with non-isomorphic mapping (keyboard) and with isomorphic mapping (a Razer Hydra controller with 6 degrees of freedom interaction). The task completion time (CT) for the ASD group were significantly higher than the control-group, which suggests that hand-eye coordination is not compensated by increase visuospatial abilities. However, Mei *et al* concluded that, additional cues and assistive technology might be necessary to alleviate the negative effects of general poor hand-eye coordination [53].

2.7.2 Locomotion

As with 3DUI, the amount of empirical research is close to non existent when it comes to VR locomotion and people with ASD. However, one new and yet uncited, study by Bozgeyikl *et al* made a comparison of 8 existing locomotion techniques (*redirected walking, walking-in-place (WiP), stepper machine, point & teleport, joystick, trackball, hand-flapping and flying*) [61] specifically for people with ASD. They found that the most enjoyable were also the ones making the user feel most in control i.e. *joystick*, *trackball* and *point & teleport*). However, for the average time it took to reach the destination with obstacles (much like one would assume in a supermarket), the most effective locomotion techniques were *redirected walking*, *stepper* and *hand flapping*[61].

The reason for the high score of conventional interaction methods such as the joystick, may be due to the fact that it is familiar and does not require a high learning curve to utilize. However, some logical delimitations should prevent the use of locomotion methods incongruent to the context. That should naturally exclude flying and hand-flapping which were also the two methods to cause most frustration[61]. The use of joysticks, trackball or stepper machine would require a stationary position, which essentially defeats the purpose of a room-scale experience. That leaves redirected walking, WiP and teleportation as suitable locomotion-solutions in this context.

2.8 Valhøj public school

Valhøj Folkeskole is a public school in Rødovre with approximately 600 students. In 2015 it joined a municipal initiative to the development of inclusive learning environments for children with ASD [62]. Previous methods of autism-treatment involved schooling in special schools or secluded classes without regular contact with "regular" pupils, but according to the National Knowledge Center for Inclusion and Exclusion (Nationalt Videncenter for Inklusion og Ekslusion), children receiving specialized schooling have a higher risk of exclusion from societal communities later in life [63]. The newly established PCA (Pedagogical Center for Autism) at Valhøj focuses on including the children in the education, actively rejecting concepts of normality and embracing individualism, as well as encourage interaction with other pupils in the regular classes [62]. It is the PCA which we will be collaborating with, to try and develop a virtual shopping tool for 7th-8th graders (approximately 13-14 years old). Currently there are 9 pupils in the two grades, 8 males, and 1 female. At the first meeting with the teachers we learned that Denmark is using ICD-10 (International Classification of Diseases), and has been using it since 1994 [64] which includes diagnoses such as Aspergers, Infan-

Figure 2.2: The front entrance of the public school Valhøj Folkeskole.

tile autism and GUA (gennemgribende udviklingsforstyrrelse, anden)[2] as overall classifiers, and then each categorized under common traits. ICD-11 classifications such as high- or low-functioning autism is therefore not in use in the danish health sector, but the teachers described all the pupils to be well-functioning. Furthermore, given that autism is a diverse and highly individual disorder, it was discussed with the teachers whether the simulation experience would have to be personalized for the pupils; generally the educators agreed that a uniform treatment would be the best approach, because 1) the pupils already have joint classes and training, and 2) the pupils may become jealous if others receive more VR-training than the others, and lastly 3) measuring the effect in a comparative study would be manageable, as long as the control-group received the same access to VR after the experiment was completed.

2.8.1 Current methods

From the meeting with teachers March 9^{th} (transcript and recording can be found in appendix G.1) it was established that current training-methods depended on introduction to real shopping environments, but seemed to appeared train and hope strategy (see section 2.3.1). Furthermore the teachers and pedagogues at Valhøj Skole utilized social stories interventions and modelling, and role-play to train the children in social and adaptive skills; all interventions are highly customized to each pupil needs.

[2]The English corresponding classification is known as Pervasive Developmental Disorder – *not otherwise specified* (PDD-NOS).

Grocery shopping

> *"We've tried shopping many, many times. And this one time, where we send 2 of the students over alone with a shopping list with 6 products on, they come back with nothing, because they couldn't find the first thing on the list."* - Claes Kiilsgaard (condensed and translated from danish)

The grocery shopping tours usually occur in relation to home-economics and cooking classes (Hjemkundskab), accompanied by teachers, and they always go to the same Føtex in Rødovre Centrum (a shopping mall in Rødovre). It does not appear to be a strategic intervention, but rather a "train and hope" approach, which seeks to expose the kids directly to real scenarios to hopefully instill training. The pupils have been going on these grocery shopping tours for 2-3 years, approximately 20 times total. However, these tours still occasionally have proven to be challenging for the pupils for a numerous reasons. Each pupil, of course, faces individual social and adaptive challenges when engaged in a shopping scenario, but some general issues applied to all:

1. They have problems with quantity (understanding measures deciliter, kilograms etc).
2. Asking for help is "dangerous".
3. Finding products
4. Populated areas / social anxiety.

Measuring and evaluation

The children are used to self-reported evaluation using likert-scales with pictorial representations instead of numbers. These pictorial elements are usually represented by smiley faces of varying mood. This method of self-evaluation for children is backed by empirical evidence [65] and have been previously used successfully for children with ASD [44]. However, the children are not confident and get nervous about being measured or observed, and even less by complete strangers. Naturally the consequence of this implies that any system and test design will have to be developed with an autonomous basis.

A focus on VR

Valhøj School and PCA have bought a VR-system for the purpose of utilizing it more in their education. Furthermore, it is expected to be employed for more content in the future as part of an ongoing collaboration between Valhøj school and Aalborg University. The purchased system consists of a high-end purposely built computer, and a tethered HTC Vive HMD.

2.9 Chapter discussion

The reviewed literature established how ASD can be difficult to diagnose due to the different repercussions it can have on an individual [9; 11]; this makes the disorder particularly challenging to treat and oftentimes therapies need to be developed around an individual's needs [3]. Behavioural therapy was the main focus of the analysis, specifically the transfer of learning involved in its process [18]. Several techniques have been used throughout the years to facilitate ASD individuals' behavioural habits retention through training [4], and out of all of them, simulation-based has been widely deemed the most appropriate to be adopted for virtual solutions [28; 29; 30]. The scope of such system has been analyzed, and samples of technology-mediated studies have been presented to describe how different obstacles have been overcome throughout the years [43]. A more specific analysis has then be conducted on VR systems developed for behaviour and adaptive skills training [45; 46]. Relevant guidelines have been extrapolated from the analysis of technology based systems [8; 43], and should be taken into consideration when developing VR systems for ASD individuals.

When reviewing the literature for specific considerations regarding multimedia applications, it became clear that people with ASD,at least high-functioning autism, does not appear to have problematic relationships with technology. This was further corroborated by Claes Kiilsgaard, at the mentioned that initial VR sessions had proven highly positive, and they had almost instantaneously picked up on functionalities such as teleportation. Although some other methods were previously considered (redirected walking and WiP) teleportation was deemed the preferred choice, since it could avoid confounding the true purpose of the experiment, with a potentially high learning curve of implementing functionalities of preestablished familiarity.

The knowledge gathered and analyzed, has then been used to establish a discussion with the stakeholders of the study (teachers and pedagogues at Valhøj Folkeskole in Rødovre), who set more precise guidelines for the development of an alternative behaviour training system. One of the issues met by the stakeholders was the difficulty in training their pupils' adaptive skills in supermarket environments; thus defining the need for a dynamic virtual alternative to the currently used train-and-hope method. The study would then address the performance of the VR system compared to their traditional methods; thus would rely on a comparative analysis of participants performance. The studies analysis 2.1 highlighted the importance of an extended intervention, based on between-groups design together with pre- and post-treatment evaluation.Transfer and memory retention should be taken into consideration as well [36; 37]. As previously shown, the median training period for transfer studies are 5 weeks (sec. 2.6), with 4 weekly session to comply with the theory of inter-session resting. Logically, we should strive to achieve a similar test-duration, but since studies with much shorter periods has achieved significant results [44] and achieved a maintenance after 12 months, it is assumable that other factors are at play as well.

> "The virtual environment needs to have sufficient detail for the learner to be able to practice skills needed in the real world (visual search, navigation) and to be recognizable as a representation of the real world. On the other hand too much detail may prevent the learner from extracting the salient features necessary for the task to be learned." [34]

Due to the limited time-span of the thesis period, a near transfer and procedural approach will be adopted, thus relying on the participants' ability to establish context through repetition. Furthermore, generalization is the transfer of learning from one situation to another (see section 2.3.1), and to avoid the risk of just implementing another "train and hope" scenario we should strive to follow Stoke and Bear's suggestions when it comes to advanced strategies [23]. For a study utilizing near transfer, we estimate that the *Program Common Stimuli*-strategy would be the most beneficial given the short time-frame. As previously mentioned the strategy revolves around establishing homogeneity between the training environment and the natural environment, by applying contingencies and common stimuli. It is unclear what level of similarity is optimal, but the consensus seems to be, that establishing as many similarities as possible will lead to a enhanced transfer of learning [23; 22; 66]. On the other hand, Cromby and Standen suggests that too many details may prevent the the learner from differentiating between non-essential and essential features meant to facilitates transfer of learning [34]. With level of detail being a relative concept, given that the article was written in 1996, we speculate that a reasonable compromise can be achieved, as long as elements meant to facilitate transfer, are salient enough to be distinguishable from non-essential elements in the surrounding environment.

Chapter 3

Methodology

This chapter will describe the iterative nature of the project, which was continuously remodelled to match the schedule of the participants. Furthermore, the chapter serves to clarify on the consulting-process with the involved teachers, primarily Claes Kiilsgaard, who were also the ones most likely to determine the validity and impact of potential research ideas on the children.

3.1 Iterative design

Research highlighted the importance of having specialists and familiar figures conduct experiments when dealing with children with ASD [4; 6], therefore the study was conducted with Claes Kiilsgaard as mediator, rather than having us interact directly with the them. Another reason was that, experimental protocols on research in alternative therapies needs to be approved by the National Science Ethics Committee[1], especially with experiments concerning children. That is, unless the experiment is mediated by an authority with preexisting authorization to conduct such experiments, such as teachers.

The following sections will describe meetings to account for the stages of the project which were decided in plenary.

3.1.1 Meetings and iterations

This section will briefly give an overview of the meetings held to establish overall stakeholder requirements and some of the iterations, reconsiderations, and changes that occurred along the way. These meetings were held to evaluate research, intended evaluation methods, and implementations.

[1]National Videnskabsetisk Komité (Forskning i alternative terapiformer): `http://www.nvk.dk/forsker/naar-du-anmelder/hvilke-projekter-skal-jeg-anmelde`

Meeting 1: Introduction

Aim of the first meeting was to initiate the collaboration with Valhøj Skole and establish fundamental issues which could be solved with VR; Claes Kiilsgaard, teacher and main contact from the school, was invited to Aalborg University Copenhagen on the 13^{th} of February, where previous work was showcased along with consumer-ready VR applications. The meeting was an unstructured conversation, to facilitate a two-way discussion about potential problems and feasible solutions. Guide-lining topics were formulated in advance in order to gather insights and initial information regarding the group sample that would undergo the study; partial details about the target group and their impairments were collected, in order to be able to start research and analyze possible solutions that could be implemented to help.

1. **We learned** that many of the children had problems with shopping. Not only the social, but also practical - e.g. adaptive skills.
2. **We decided** that the initial research idea should revolve around the system teaching generalization skills in a virtual shopping environment.

Meeting 2: Establishing framework

The second meeting was held on March 9^{th}, at Valhøj Skole, where more teachers and pedagogues participated. Researched had been conducted, and the problem of generalization had been established. The conversation was structured as an informal interview to gather further information on the participants and their disorders. Furthermore, the concepts of generalization and adaptive training were discussed, and their training approach was presented. A general procedure was also structured, and the educators' knowledge was used to pinpoint specific measurement techniques and behaviours that could be adopted during the tests.

1. **We learned** that 9 children could participate in the experiment
2. **We decided** that generalization should form the basis of the experiment, but with a focus on generalizing product types and being able to teach the children to choose alternatives if the listed product was missing.
3. **We decided** that the goal was 2-3 weekly sessions for a duration of 5 weeks in April.
4. **We decided** on a baseline-post-treatment experiment with a control and treatment group. The teachers would divide the children into two representative groups.

Meeting 3-4: First iteration(s) presented

Two meetings were held with Claes Kiilsgaard during the development of the system. On the 30^{th} of March, the first meeting consisted of a presentation of the preliminary system prototype, with an early version of the virtual Føtex. At this meeting, fundamental changes to the test framework were made on the basis that generalization of choosing alternatives to missing products would be hard to measure. Given the implementation of an entire represented supermarket, which looked approximately similar to the real counterpart, it was decided that the shopping task should depend on searching for products in the "entire supermarket" (only relevant areas containing food). This also caused the initial idea of 5 weeks of testing to be reduced to 2 weeks due to the increase demand for redesigning the test-framework. Once the new system was developed, a second meeting was held 10^{th} of April, to test the project and data logging on the test-computer. At this meeting, design ideas for an automatic "hint-system" were brainstormed with emphasis on learning to use the signs and navigational aids within the supermarket, more metaphysical cues such as direction on floor or light beams to help the child navigate the aisles.

1. **We learned** that generalization of product types would be hard to measure.
2. **We decided** to redesign the test framework, which required more work to refine the system. Therefore:
3. **We decided** the initial experimental period of 5 weeks were to be narrowed down to 2, as it was deemed that content was more important than duration (supported by prior research [44]).
4. **We decided** that an automatic hint system should be implemented as an extra layer of help in case the test participant got stuck. Furthermore, we hypothesised that hints could create an extra layer of measuring i.e. the prospect of a decrease in help need, due to learning effects.

Final meeting: Post-evaluation

A final meeting to evaluate the process was to be held in May. It was aimed to evaluate the process and any eventual observations that were vital to the interpretation of the acquired data. This meeting will be further elaborated on in the discussion (section 6).

3.2 Final experimental design

As described above, the experimental design was defined as early as the 2^{nd} meeting (3.1.1), and the final experimental design applied will be discussed in this section. Besides having a third party conduct the experiment, further ethical precautions have been taken for the execution of the experiment. For example applying a layer of anonymity by using letter labels instead of names, and conduct no video-recordings of the children without parental consent. A total of nine participants were scheduled to take part in the study, eight male and one female (age ranging from 12 to 15 years old). The experiment was between-groups, and the students were divided in two groups — control and treatment group — by the teachers. as presented in table 3.1.

Table 3.1: Test participants

Treatment	**Control**
A	E
B	F
C	G
D	H
-	I

The experiment has been conducted for 2 weeks (between April 18th and April 28^{th}). In order to assess all participants' base and final performances, both groups were required traditionally train in Føtex, as they would usually do, on April 18^{th} and April 28^{th}. Between base and final assessments, the treatment group was required to train with the VR system once a day, while the control group would receive no treatment.

3.3 Data gathering

3.3.1 Baseline and post-treatment assessment

In order to assess any difference in the participants' performance, a baseline was established in a real shopping scenario. They were each assigned a shopping list (appendix B) consisting of four products, optimally, in different areas of the supermarket.

The shopping lists were structured into 9 different shipping lists, from an overall list of products randomly distributed between the test participants i.e. one shopping list might include one product from the diary section, two from the general shelves, and one from the fruits and vegetable area (green section). To prevent researcher-bias (i.e. in this case, the observers might be inclined to note down positive traits and observations from the treatment group if they are aware of the identity of the participants) each participant has been assigned a specific personal colored paper, assigned by the teacher, which we, the observers and authors, did not know until after the post-treatment session. To facilitate correct categorization, the participants were asked to keep the lists visible at all time.

In order to have quick and reliable observations, key events were agreed upon by the observers previous to the baseline. A total of 11 events were deemed relevant for the purpose:

1. The children talk with others (each other, teachers, other shoppers, clerks)
2. Arrived at one (correct) shelf
3. Got distracted
   A. Other people
   B. Sounds
   C. Signs
   D. Colors
   E. Other [open-ended]
4. Used sign to orientate
5. Used another kind of help [open-ended]
6. Started looking at a shelf
7. Product handled
8. Product put back
9. Product put in the basket [open-ended]
10. Waited - external factor [open-ended]
11. Note [open-ended]

In order for the observer to blend in with the customers at the supermarket and not overtly show they were taking notes of other customers behaviour, a purposely built web-application was developed for use on smart-phones (section 4.5). This app would potentially allow immediate event-logging with the press of a button, or the input of a short text, while staying more incognito. Including the two authors, three additional observers were recruited to help keep track of each of the designated sections. The five observers covered different areas of the supermarket as presented in figure 3.1 so to have optimal observation ability.

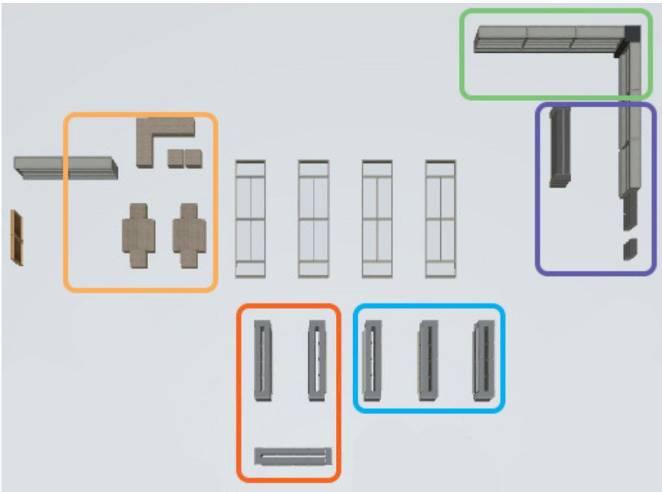


Figure 3.1: Areas covered by the five observers during baseline and final evaluation.

Additionally, after both baseline and post-treatment assessment the participants were presented with a post-test questionnaire. The purpose was to evaluate the participants' self-reported confidence level in extension of the experience, and a prospected hypothetical confidence level if they were to shop alone the next day. This was inspired in part by the pictorial confidence rating scale used by [44] to subjectively evaluate how comfortable the test participants would feel about hypothetically go buying a newspaper alone, using a 6-point likert scale. Furthermore, participants were used to self-reporting using pictorial smiley-scales. In figure 3.2 is an example of a the question regarding confidence, which at a lack of a better translation became a question regarding "safety and security".

Besides measuring confidence and prospected confidence, using a 5 point likert-scale, we asked them to report how easy or hard the shopping experience was, which assistive elements they used to locate the products (multiple answers and open-ended answer) and finally to report which product was the hardest to find and why.

Figure 3.2: An example of the smiley-face likert-scale. "How safe/confident did you feel shopping today?"

Lastly, in order to evaluate accuracy and task performance of the participants, start and end times were recorded for each participant, and the supervisors were asked to take pictures of the final baskets with products and shopping lists in them (appendix G.1). Together with the pictures, eventual feedback from the supervisors was collected in the form of written notes.

3.3.2 VR-training

The treatment group was required to train every day between the initial and final evaluation, and each participant had a time-span of maximum 20 minutes to complete the VR-simulation each time. The training would require participants to gather the four products presented on the *shopping list* user-interface (UI). The lists would change every day and they were the following:

Table 3.2: Session shopping tasks

Ses.	Date	Product 1	Product 2	Product 3	Product 4
1	19/4	2 kg hvedemel	1 fl. sojasauce	1 pk. smør	1 pk. cornflakes
2	20/4	1 ps. sukker	1 ps. pastaskruer	1/2 l. fløde	1 pk. ml.lagret ost
3	21/4	Earl Grey te	1 pk. digestive	1 dåse majs	1 l. kærnemælk
4	24/4	1 pk. knækbrød	1 pk. flormelis	1 ps. jasmin ris	1 l. minimælk
5	25/4	12 æg	1 ds. kokosmælk	1 ps. rosiner	1 ps. havregryn
6	25/4	1 pk. spaghetti	1 ds. tun i vand	1 pk. cornflakes	1 l. sødmælk
7	27/4	1 pk. nudler	1 ds. majs	1 ps. revet mozerella	1 ds rød karry

During this part of the experiment data was collected both covertly, through data logging in VR, and overtly through questionnaires for both the participants and the supervisors.

Questionnaires for the participants were aimed to assess their emotional state, their comfort during the experience, and their comfort if they were to do groceries by themselves the following day (appendix A.2).

1. **Hvor sjovt synes du det var at handle i det virtuelle supermarked i dag?**
   How fun was it to shop in the virtual supermarket today?
2. **Hvor let synes du det var at handle i dag?**
   How easy did you think it was to shop today?
3. **Hvor tryg var du ved at handle i dag?**
   How confident did you feel about shopping today?
4. **Hvor tryg ville du være ved at handle alene i den rigtige Føtex i morgen?**
   How confident would you feel about shopping alone tomorrow in the real Føtex?
5. **Har du nogle kommentarer?**
   Any other comments?

These questions were reformulated and submitted to the supervisors (appendix A.3), in order to cross-check the reliability and precision of the answers.

1. **Var eleven tilfreds med sin egen præstation? (meget utilfreds-meget tilfreds)**
   Were the student satisfied in his own performance (ver unsatisfied - very satisfied)
2. **Hvor selvsikker vurderer du at eleven er, i hans tilgang til systemet og oplevelsen? (meget usikker-meget selvsikker)**
   How confident do you estimate that the student is, in his approach to the system and the experiemce (very inconfident - very confident)
3. **Præstation ved specifikke elementer**
   Elements about specific elements
   - **Hvordan svært/let var det for eleven at bruge teleportation?**
     How easy/hard was it to use teleportation?
   - **Hvor svært/let var det for eleven at bruge cues (spor)?**
     How easy/hard was it to use cues?
   - **Hvor svært/let var det for eleven at finde rundt i VR-Føtex?**
     How easy/hard was it to navigate around the VR-Føtex?
4. **Var der nogle fejl eller mangler ved systemet der gjorde det svært for eleven at gennemføre det virtuelle indkøb?**

Were there any errors and missing things about the system that made it hard for the student to conduct virtual shopping?

5. **Nogle kommentarer fra eleven undervejs?**
   Any comments from the student during the session?
6. **Andet**
   Other ...

On top of this, the teacher conducting the experiment (Claes) was asked to fill out an observation-diary of his own estimation of the student's confidence and behavior towards the system and their performance. Additional, indirect objective measures were applied to gather data logging information about the participant's behaviour and performance:

- Product touched
- Hint asked
- Product put in the basket
- Product taken out of the basket
- Product grabbed
- Product released
- Total time spent in the experience

3.4 Chapter conclusion

The chapter presented the steps taken together with the stakeholders in order to define the scope of the project, assess the needs of the system, and to agree on data collection methods. The study was designed as a two-week long matched-pairs experiment: after a baseline evaluation in supermarket shopping, treatment group would undergo the VR training while control group would have no treatment, afterwards both groups' performance in the supermarket would be evaluated again. For baseline and final evaluations, the participants' behaviour was observed by the authors and three collaborators; additional data was recorded by the pedagogues through notes and pictures, and the participants themselves through questionnaires. During the experiment weeks, both participants and teachers were asked to answer questionnaires after each VR experience, and additionally data was overtly logged through the system. The following chapter presents the iterative development of the VR system, and how its functionalities were created.

Chapter 4

Design and implementation

The virtual simulation was developed through an iterative process described in the previous chapter, made in close, and continuous collaboration with the teachers at Valhøj School. This chapter presents the design and implementation of the final system, but also contains information about some rejected elements, because they lead to new ideas and implementations, and as such are vital to describing how the technical aspect of the project continuously evolved into the final test-ready system.

4.1 The system

The VR simulation was developed in Unity 5.5, and was meant to be run on a custom built system consisting of an Intel i7 7700k processor, GTX 1080 graphic card, and 16GB of ram. As suggested by Bowman et al. [52], it was decided to deliver the VE through a VR headset, and HTC Vive was chosen due to its unrivalled room scale tracking. The VR experience was aimed to facilitate ASD children's adaptive training in supermarket environments; the sample group considered in this study was already familiar with a specific supermarket (i.e. Føtex in Rødovre Centrum), hence it was replicated in the virtual world. Due to the high importance and need of guidance suggested by the teachers, three different sets of hints were developed to be accessed when users needed. Voice-overs, visual highlights on signs, and a compass pointing towards the product of interest were given to the users; on a second iteration of the system, the abuse of the guidance system required a redesign, thus only the compass was kept and the triggering possibility was given to the supervisor rather than the users themselves.

In all iterations, participants would undergo the test procedure as presented in the use-case diagram 4.1 below. Before starting the proper training simulation, if the user was trying the experience for the first time, they would undergo a tutorial;

this was meant to help them familiarize with the specific controls implemented. Basic interaction was developed following standard industry guidelines; teleportation triggered by the touch pad, and object interaction enabled on trigger press were deemed the most appropriate interaction techniques, since the teachers had participants play several games that used these interaction means, before the study was conducted. Additional means of interaction needed to be developed in order to interact with other parts of the experience such as shopping list toggling and cues triggering. These were assigned to different buttons on the right hand controller. Once users familiarized with the environment and interaction techniques, they would be transferred to the main experience. Here they were presented with a shopping list made up of four items; users were required to *buy* (i.e. place the product in the basket) the products following their order on the shopping list. The teachers shared a list of 32 products that participants have been trained to search for and buy during their training sessions. Virtual shopping lists were generated by selecting four products from the mentioned list, and once the fourth item was *bought*, the experience would finish. Along their experience, data was covertly collected whenever they touched, grabbed, released, or *bought* a product.

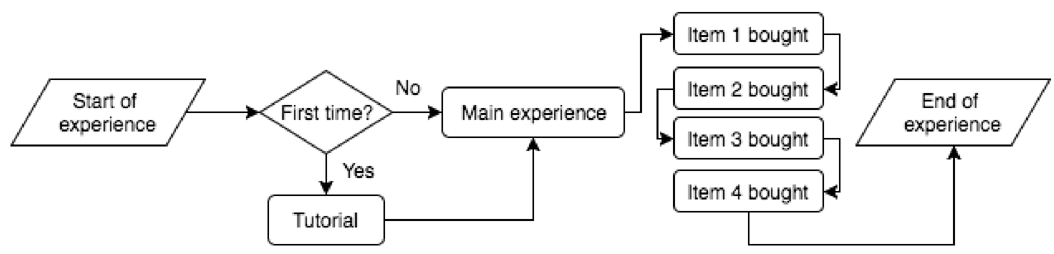


Figure 4.1: Flow chart of the VR simulation usage

4.2 First development

The development of the system started from the general assessment of the environment requirements; this was followed by the 3D modelling of a restricted amount of products and the development of procedurally populated shelves.

Most of the implementation, if not stated otherwise on following sections, was maintained in the final version of the system.

4.2.1 Creating the virtual environment

Recognizability of the environment is one of the requirements for establishing optimal conditions for training and transfer of knowledge. Two field trips were carried out to Føtex in Rødovre Centrum; the initial trip (March 20th, 2017) had the purpose of establishing looks and feels as well as the store layout. One of the initial ideas to do so, was to acquire access to floor emergency plans so to have an architectonically correct representation of the layout. The request was denied by the staff, thus store layout had to be reconstructed less accurately from images and sketches drawn in situ (fig. 4.2). Furthermore, a total of 96 images were taken of the interior layout, shelf design, signs, commercials and products to achieve optimal accuracy.

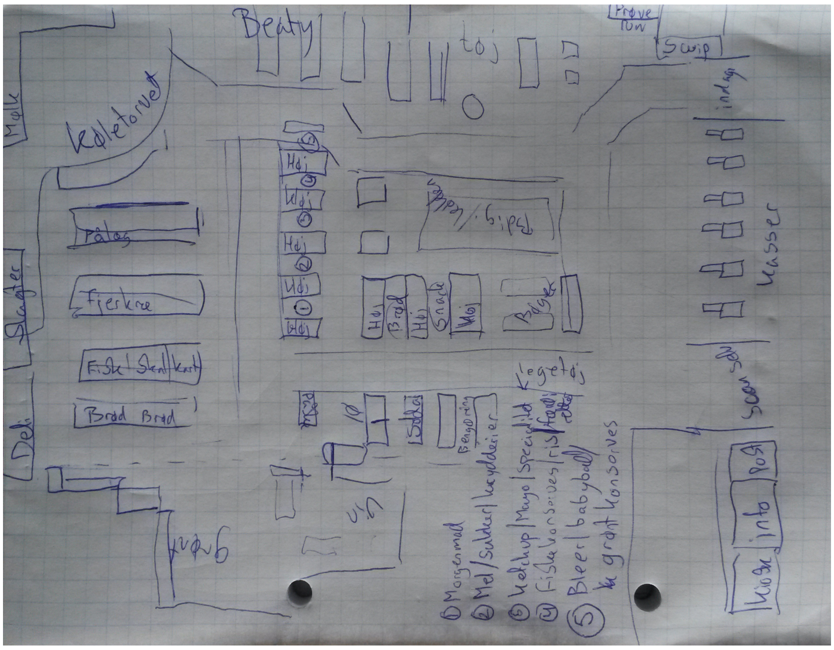


Figure 4.2: The sketch drawn 20th of March, during a field-trip to establish store layout.

The areas of the store adjacent to the entrance consist of clothing, make-up and other non-food departments. Being these areas not pertinent to the experiment, they were not included in the development of the simulation. Instead, the areas of interest included:

1. **Køletorvet (Dairy department)** - *dairy products, fresh pasta, cold cuts*
2. **Long chest freezers (parallel)** - *frozen goods, fresh/frozen fish, frozen/fresh meat, cold cuts, organic food, ice cream etc.*
3. **Tall shelves (parallel)** - *Dry foods, sugar, canned food, pasta, rice, tea, coffee*
4. **Butcher & Deli** - *Fresh meat and delicacies*

5. **Vegetable/fruit department** - *vegetables, fruits, spices and herbs*

The first trip was sufficient to virtually reconstruct the store, but a second trip was required (April 07th, 2017) to find out the exact location of products. In case any of the products the participants had to find would accidentally be misplaced, it could cause negative transfer effects and hence jeopardize the experiment. All locations of products were noted paper and transferred to a spreadsheet (fig. 4.3).

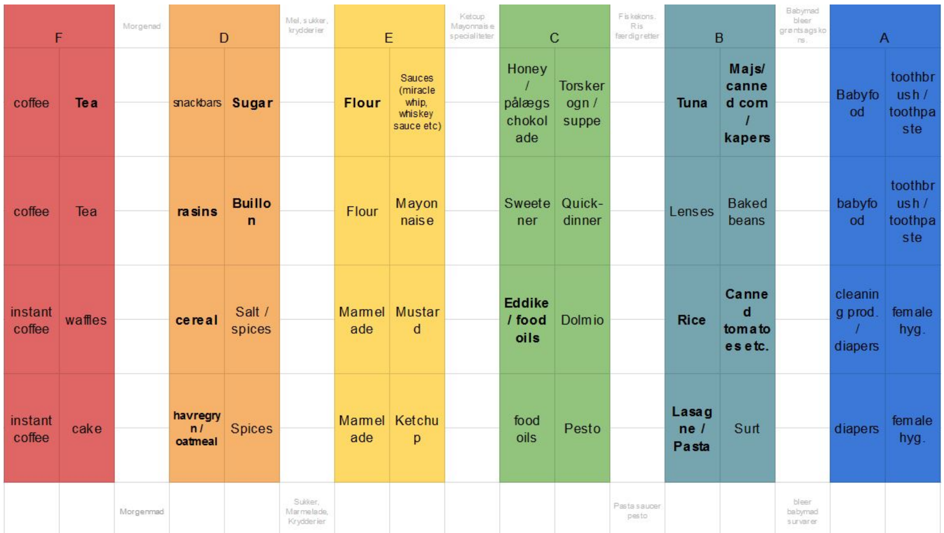
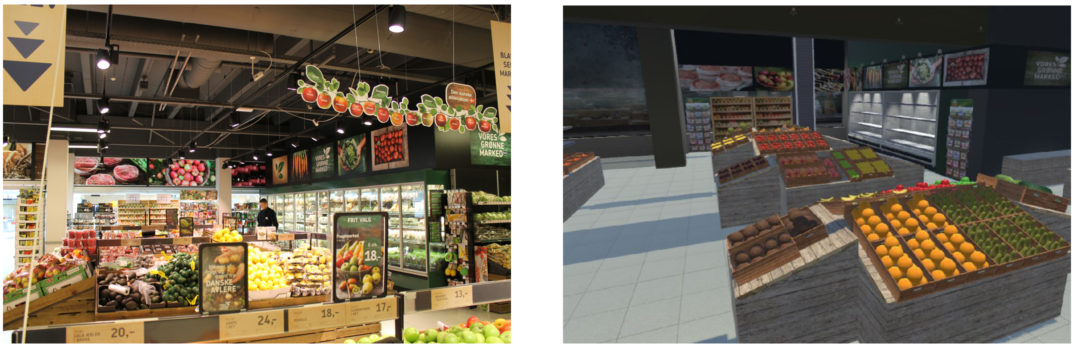


Figure 4.3: Layout of the 6 tall shelves containing dry products. In bold are reported locations of products of interest for the study. In grey is transcribed the content of the signs hung over each row.

All sections were reconstructed with custom made models specific to Føtex i.e. shelf-systems, Køletorv-canopy, organic signs, "bazar"-signs, flags, shelf signs. And a comparison between pictures taken in situ and their virtual version can be seen in figures 4.4, 4.5, and 4.6.

(a) Physical (b) Virtual

Figure 4.4: Comparison between the real freezers and "Køletorvet"/dairy department and the virtually reconstructed equivilant.



(a) Physical (b) Virtual

Figure 4.5: Comparison between the vegetable/fruit department virtually reconstructed (b) from a photo (a) taken 20-03-17.

(a) Physical (b) Virtual

Figure 4.6: (a) Shelf storing biscuits, crackers, crispbread *inter alia* and the virtual shelf (b) storing raisins, cereal and oatmeal.

4.2.2 3D Models

Not only did the interior design have to be recognizable by the participants, the products also needed to be things which can be found in Føtex. Besides the products selected for the shopping list itself, the system should include alternative and un-related products. This was aimed to also increase realistic perpetuated expectations of the users, since in a real supermarket the customer will seldom be faced with a singular choice in product.

A total of 32 individual products were listed in the list provided by the teachers; besides these, a total of 52 alternative products have been added to the database. For the sake of time, the most generic products are re-textured versions of existing assets purchased at an earlier point[1]. Unique-looking products (i.e. Kærgaard butter, Arla sliced cheese, digestive biscuits) were modelled, uv-mapped, and texturized from scratch.

Texturing

Products were texturized using three different methods depending on object shape, ranging from quick to comprehensive. Texture resolutions were either 512x512 or 1024x1024 pixels per inch (ppi).

1. **Quick** - *Mostly squared objects, could be photographed from each side and assembled into a uniform texture using Photoshop*
2. **Moderate** - *semi-squared, non-primitive shapes. Products had to be emptied or compressed, and either scanned or photographed*
3. **Comprehensive** - *Cylinder objects (mostly canned food) had to undergo an etiquette removal process.*

The quick method was utilized for the majority of the products, including noodles, crispbread (knækbrød) and powdered sugar (flormelis). The moderate solution was used for approximately 70% of the products. For the oatmeal (havregryn), an empty wrapping was used, compressed and scanned. Pasta and spaghetti packages were photographed from all angles using a Canon EOS 1100D digital camera, and underwent comprehensive post-processing in Adobe Photoshop to assemble the image into a smooth texture. Existing normal- and specular-maps were applied or generated individually using InsaneBump[2]. The most comprehensive method, etiquette removal, was adopted for canned food where labels are glued to the can. The glue was dissolved in a water bath, and afterwards carefully

[1]Supermarket Gluttony Pack, purchased from the asset store: https://www.assetstore.unity3d.com/en/#!/content/12042

[2]Insane Bumb 2.0: https://goo.gl/YZEIXH

removed and dried under press. Once dried, the labels were scanned at a resolution of 600dpi. The digitized version could thereafter be used as texture once post-processsed (see fig. 4.7). After all of the products had been modelled and texturized, they were imported in the main scene (see fig. 4.8) in Unity3D and assigned necessary scripts (interaction scripts and product-info script).

(a) Labels removal (b) Labels drying (c) 3D model

Figure 4.7: A cylinder shape is hard to photograph smoothly from all angles, so etiquettes were removed from the can by first (a) wetting the can to dissolve the glue (b) drying it to be able to scan it. Once scanned the images can be applied to a UV-map and (c) added as a texture to the final 3D product.

Figure 4.8: The 3D products created and inserted in Unity3D as game-ready prefabs.

4.2.3 Product info

In order to streamline actions that could be procedurally augmented (e.g. shelf population), each product *GameObject* was assigned a script which contained its essential information: id, name, price, category, stackable, original position, and original rotation. On top of this information, each product was assigned a method that would make it re-spawn in its original location if it collided with the floor *GameObject* or with the *FallDetector GameObject*. The full script can be found in appendix C.2.

4.2.4 Product list generator

In order to have an overview at all times of the products present in the simulation and their attributes, a script was created in order to write to a text file all the information needed. The script's logic is contained in the start function (Snippet 4.1) and after opening the *StreamWriter*, it collects all the *GameObjects* with tag *"product"* and after sorting them by id, it generates a string made up of *id*, *price*, and *name* that will then be written to the text file. The full output of the script can be found in appendix C.1.

```
StreamWriter stream_writer = new StreamWriter (
  Application.dataPath + "/SetupFiles/products.txt", false
);
stream_writer.Write("!!!!!! WARNING !!!!!! \r\n");
stream_writer.Write("This is an auto-generated file, modifying this will not affect the simulation in any way.\r\n");
stream_writer.Write("Product_id, price, name\r\n");
GameObject[] prods = GameObject.FindGameObjectsWithTag("product");
List<GameObject> products = new List<GameObject> ();
foreach (GameObject prod in prods) {
  List<GameObject> results = products.FindAll(x => x.GetComponent<objectInfo>().id == prod.GetComponent<objectInfo>().id);
  if (results.Count == 0) {
    products.Add (prod);
}}
products.Sort (delegate(GameObject a, GameObject b) {
  return (a.GetComponent<objectInfo>().id).
    CompareTo(b.GetComponent<objectInfo>().id);
});
foreach (GameObject product in products) {
  objectInfo info = product.GetComponent<objectInfo> ();
    stream_writer.Write(info.id+" "+info.price+" "+info.name+"\r\n");
  }
stream_writer.Close ();
```

Snippet 4.1: outputAllProducts.cs - *start()*

4.2.5 Shopping list

In order to present the users with the list of products that they are required to buy, a virtual shopping list was implemented. The list object was assigned to the right controller and positioned so that it would hover on top of it, as can be seen in figure 4.9

Figure 4.9: The virtual shopping list hovering on top of the right controller.

In order to show users their progress throughout the simulation, a checkbox system was implemented for the items in the shopping list. This system is managed by the method *SetBought()* (snippet 4.2) in the *shoppingListManager.cs* script. As can be seen, the script checks if the id of the product considered (i.e. one that has just been placed in the basket) is the same as the one that needs to be bought, or if it falls under the same category of the one that has to be bought. The script was divided in this manner for future eventual development, in which different scores could be given to the participants if they bought the exact product, or a similar one. Either way, the script changes the visibility of *GameObjects* indicators by adding a *green tick* to the currently bought item, and adds an *indicator* to the next product to be purchased.

```
public void SetBought(int id){
  if (shopping_list_id [current_to_buy] == id) {
    bought [current_to_buy] = true;
    guideManager.GetComponent<GuidingManager> ().UpdateHints ();
    GameObject.Find ("icon" + current_to_buy).transform.Find
        ("yes").gameObject.SetActive (true);
    GameObject.Find ("icon" + current_to_buy).transform.Find
        ("toBuy").gameObject.SetActive (false);
    if (current_to_buy < 3) {
      current_to_buy += 1;
```

```
      GameObject.Find ("icon" + current_to_buy).transform.Find ("toBuy").gameObject.SetActive (true);
    } else if (current_to_buy == 3) {
      current_to_buy += 1;
    }
  } else {
    foreach (List<int> category in categories) {
      if (category.Contains (id) && category.Contains(shopping_list_id[current_to_buy])) {
        bought [current_to_buy] = true;
        guideManager.GetComponent<GuidingManager> ().UpdateHints ();
        GameObject.Find ("icon" + current_to_buy).transform.Find ("yes").gameObject.SetActive (true);
        GameObject.Find ("icon" + current_to_buy).transform.Find ("toBuy").gameObject.SetActive (false);
        if (current_to_buy < 3) {
          current_to_buy += 1;
          GameObject.Find ("icon" + current_to_buy).transform.Find ("toBuy").gameObject.SetActive (true);
        } else if (current_to_buy == 3) {
          current_to_buy += 1;
        }
      }
    }
  }
  if (current_to_buy == 4) {
    scripts.GetComponent<endOfExperience> ().ShowEndButton ();
  }
}
```

Snippet 4.2: shoppingListManager.cs - *loadList()*

Shopping lists can be changed though text files for better clarity, although there are two separate files that need to be changed. Due to the structure of the system prototype, the shopping list instantiates the product names from one text file, while the functional list of product ids is loaded from another file. Since this system might have been challenging or confusing for the supervisors to change daily, an automated loading script was created. The script was particular for the days that the VR experiment was conducted and can be seen in snippet 4.3. The function *loadList()* is aimed to load the product names from text files named "s_*n*", where "n" is a crescent number based on the amount of experiment days. The method creates a buffer variable for the text file name, which is then defined based on the number of the day that the script is being run. Each day has been assigned a different text file, thus on April 20th the text file name would be "s_2". The variable is then used to access the file and load the correct shopping list for the day. The list of product ids is loaded in a similar fashion, but the filename is changed from "s_n" to "i_n".

```
private void loadList(){
  string name = "s_0";
  int day = DateTime.Now.Day;
  if (day == 19) {
    name = "s_1";
  } else if (day == 20) {
    name = "s_2";
  } else if (day == 21) {
    name = "s_3";
  } else if (day == 24) {
    name = "s_4";
  } else if (day == 25) {
    name = "s_5";
  } else if (day == 26) {
        name = "s_6";
  } else if (day == 27) {
        name = "s_7";
  } else {
    name = "s_8";
  }

  FileInfo theSourceFile = new FileInfo (Application.dataPath + "/SetupFiles/ShoppingLists/"+name+".txt");
  StreamReader reader = theSourceFile.OpenText();
  [...]
}
```

Snippet 4.3: shoppingListManager.cs - *loadList()*

4.2.6 Shelf population

Features for the shelves design and product placement, were dictated by the possibility of the system to be used by the educators independently after development completion; thus, the need for procedurally populated shelves based on pre-defined user inputs. This feature was envisioned keeping in mind the limited development time-frame and the technology educators could be comfortable with. Users were allowed to change the products present on each shelf by simply inserting products IDs in a text file. Although different implementations were considered (i.e. developing a whole unity scene in which users could drag and drop objects), this was regarded as the best compromise between ease of use and development time.

Products placement is handled by the script named *populateShelf1.cs* (full version can be found in appendix C.3), which is assigned to each shelf that needs to be procedurally populated. The shelves are identified by a unique id number, which will be in return used as text file name for the products id input, thus having one text file for each shelf. The script setup is being used to initialize essential variables for proper execution, and the most important part is presented in the snippet below. The function *loadList()* (Snippet 4.3) is being used to read the product ids from the specific shelf's text file and convert them into a list of lists of integers for further use; as can be seen, the function reads the whole text line and splits based on the ";" and "," characters. The semicolon is being used to separate each level of the shelf, while the comma is being used to separate the different products on the same level; it can be noticed that the amount of products that can be placed on a level can be infinite, but this problem is addressed in another part of the script, where the amount of products on one level is limited by the inherent size of the product models and the pre-set latitudinal distance between different models. Thus, the text files input should adhere these three requirements: levels divided by a semicolon, products on the same level divided by a comma, and all input must be on the same line. Finally, an example input string can then be *1;5,4,3;3;2,4* and the resulting shelf schematic is shown in figure 4.10 below.

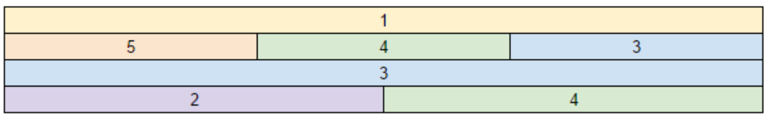


Figure 4.10: Shelf schematic. Shelf ids and colors represent the area covered by the products

```
private List<List<int>> loadList(){
  FileInfo theSourceFile = new FileInfo (
    Application.dataPath + "/SetupFiles/Shelves/"+fileName
  );
  StreamReader reader = theSourceFile.OpenText();

  string text = reader.ReadLine();

  string[] rows = text.Split(';');
  List<string> rows_list = new List<string> (rows);

  List<List<string>> products_list = new List<List<string>> ();

  foreach (string row in rows_list) {
    string[] products = row.Split (',');
    List<string> temp = new List<string> ();
    foreach (string product in products) {
      temp.Add (product);
    }
    products_list.Add (temp);
  }

  List<List<int>> final_list = new List<List<int>> ();
  foreach (List<string> products_row in products_list) {
    List<int> temp_list = new List<int> ();
      foreach (string product in products_row) {
```

```
        int tmp;
        if (int.TryParse (product, out tmp)) {
          temp_list.Add (tmp);
        }
      }
      final_list.Add (temp_list);
    }
  return final_list;
}
```

Snippet 4.4: populateShelf1.cs - *loadList()*

The population is executed by the *fill()* method. The function iterates through the amount of level colliders each shelf has, and for each one of them start and final products positions are found depending on the orientation. Once initial variables are set up, the script enters the core population logic which is run for each product that has to be placed (Snippet 4.5). Since shelves can be rotated along the y *axis* by 90 degrees interval, there are two pairs of orientations (i.e. 0°/180°and 90°/270°) in which x and z coordinates do not need to be inverted to maintain a cohesive population. In order to place the products, the script runs a while loop which adds a product in incremental positions in both x and z directions, until the position of the product (including user defined offsets) is larger than the width or depth of the shelf collider. In case one shelf level contains more than one product, the surface covered by products is calculated proportionally to the amount of products to be placed, thus having two products means they would cover half level each.

```
private void Fill(List<List<int>> products){
  //[...]
  while (product_position.z + product.GetComponent<Collider>().bounds.size.z + productsDistanceZ <
        start_position.z + final_positions [row_list.IndexOf(id)]) {
    while (product_position.x + product.GetComponent<Collider>().bounds.size.x <
          final_position.x) {
      int amount;
      if (product.GetComponent<objectInfo> ().stackable) {
        amount = Mathf.FloorToInt(maxStackHeight /
            product.GetComponent<Collider> ().bounds.size.y);
      } else {
        amount = 1;
      }
      for(int k = 0; k < amount; k++ ){
        Instantiate (
          product,
          new Vector3 (
            product_position.x + xOffset +
                product.GetComponent<Collider> ().bounds.size.x /
                2,
```

```
            product_position.y + yOffset +
                (product.GetComponent<Collider>().bounds.size.y*k)
                + (0.05f),
            product_position.z + product.GetComponent<Collider>
                ().bounds.size.z / 2
          ),
          Quaternion.Euler (offsetRotation.eulerAngles.x,
                            offsetRotation.eulerAngles.y +
                                offsetAngle,
                            offsetRotation.eulerAngles.z),
          product_parent
        );
      }
      product_position.x += product.GetComponent<Collider>
          ().bounds.size.x + productsDistanceX;
    }
    product_position.x = start_position.x;
    product_position.z += product.GetComponent<Collider>
        ().bounds.size.z + productsDistanceZ;
  }
}
```

Snippet 4.5: populateShelf1.cs - *fill()*

4.2.7 Guidance

The importance of guidance and help in virtual systems has been underlined by both research (2.5) and teachers (3.1.1); thus the following six guidance hints were brainstormed in the 3^{rd} and 4^{th} meeting with Claes Kiilsgaard, and subsequently conceptualized:

1. **Map** - meant to help users orientate around the supermarket by having a miniature view of the whole area;
2. **Voice help** - so to give users voices directions towards the product they need to buy;
3. **Blinking objects** - applied to all objects in the same category of the product of interest, this would help users generalize the product they are looking for;
4. **Arrow on the floor** - similarly to voice help, this was meant to guide the users through the supermarket directly to the product of interest;
5. **Light-beam** - a cylinder of light placed in front of the product of interest, would help the users find the shelf from any point of the supermarket;
6. **Føtex helper** - a humanoid AI assistant that would answer questions or simply give guidance indications.

The hints were then analyzed and four of them were selected for development mainly based on their complexity. The Føtex helper was discarded due to the high complexity of the mechanics its creation would involve, and similarly the navigation arrow was discarded because of the path-finding algorithm implementation was regarded too time consuming considered the time-frame of the project.

Light-beam

The light-beam (or "God Light") is often seen in games, utilized as a salient cue, to guide the player towards items or areas of interests. It is beam of light descending from the sky, which makes it ideal for the supermarket, as it will be visible from anywhere in the environment. It can either be used to mark the location of a specific product, or to draw attention to areas, shelves, or signs.

The implemented light-beam was made from a cylinder primitive shape with top and bottom removed, the bottom circumference being slightly larger. A greyscale texture with cloud rendering was generated to create light variations. Furthermore, the texture had a transparent to black gradient in the top, and a transparent to white gradient in the bottom, to make the light fade out towards the top, and increase in intensity towards the base. This can be integrated seamlessly in Unity by using a Particle/Additive shader, which decreases opacity as pixel value goes towards 0 (black). Furthermore the shader also disables culling, which effectively draws front and back faces of the cylinder. This means that the light-beam will also be visible from 'within' (see fig. 4.11).

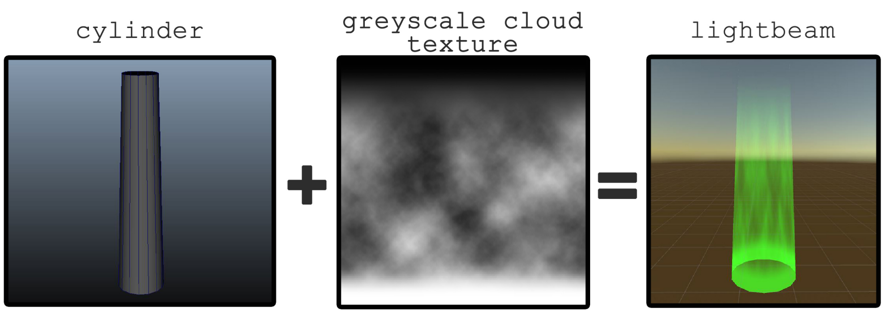


Figure 4.11: Light-beam made from a cylinder primitive with added texture (with a green tint).

In Unity the cylinder was duplicated 3 times on top of each other, and a green tint was added to the texture. A small script was added to each of them to create a rotating animations, adding to the interplay of the light (see snippet 4.6). Finally, a particle system with vertical cone emission was added to the base.

```
public class BeamLightBehavior : MonoBehaviour
{
  public Vector2 AnimationRate = new Vector2(1.0f, 0.0f);
  Vector2 uvOffset = Vector2.zero;
  public Renderer rend;

  void Start(){
    rend = GetComponent<Renderer>();
    rend.enabled = true;
  }

  void Update(){
    uvOffset += (AnimationRate * Time.deltaTime);
    if (rend.enabled = true){
      rend.materials[0].SetTextureOffset("_MainTex", uvOffset);
    }
  }
}
```

Snippet 4.6: BeamLightBehavior.cs

Map

The map has been developed once the supermarket structure was final; the models of the overall environment were duplicated and scaled. The appearance of the map was exactly the same as the environment, so to allow users to reliably recognize areas and shelves in the miniature. Figure 4.12 presents a top view of the map as could be seen by users.

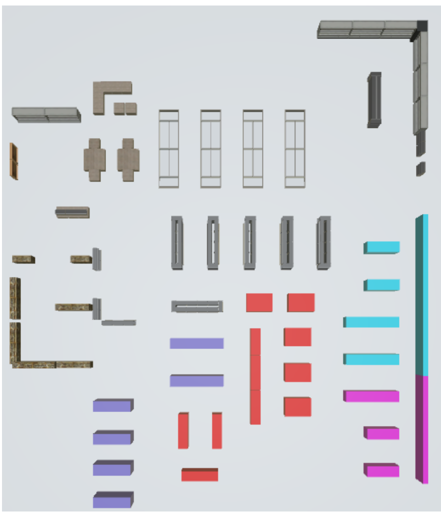


Figure 4.12: Top view of the map guiding tool

After preliminary testing the map appeared to have a fundamental lack of functionality. In order to solve this issue, the script used for the blinking products was adapted to enable shelf blink on the mini map, to properly identify the product of interest's proxy-location and offer more guidance to the users. Further tests brought to the realization of the need for a user location indicator, so to have a fully functional guidance map system, which exceeded the allocated time. Therefore it was decided to remove the map, and focus on improving the other hints.

Blinking products

In order to enable users to certainly find products they are looking for, the blinking system was implemented so to give direct and noticeable visual feedback. On user request, the products of interest would be made blink until decided otherwise, or when on the products is bought. The blinking mechanic consisted of simply changing material color depending on a pre-set time frame; as can be seen in snippet 4.7, every frame the script checks if the blinking cue is enabled, if so it executes the highlighting logic. The script initially calculates if at the current time the product of interest should be highlighted or not, consequently it iterates though all of the *GameObjects* instantiate from the product of interest and, if the time is correct, the material color is changed to the highlighted material, otherwise is set to white (which sets the material to the original aspect).

```
if (blinking) {
  var current_time = Time.time;
  if (current_time - initial_time >= 0.65) {
    highlight = !highlight;
    initial_time = Time.time;
  }
  foreach (GameObject target in targets) {
    if (highlight) {
      target.GetComponent<Renderer> ().material.SetColor
          ("_Color", highlighted);
    } else {
      target.GetComponent<Renderer> ().material.SetColor
          ("_Color", Color.white);
    }
  }
} else {
  foreach (GameObject target in targets) {
    target.GetComponent<Renderer> ().material.SetColor
        ("_Color", Color.white);
  }
}
```

Snippet 4.7: BlinkingObject.cs - *Update()*

Compass

The compass was an impulsive idea meant to substitute the more extensive pathfinding system; the latter would require rather complex algorithms, while the compass would only need to rely on the relative position of the object currently being searched for. It was purposely built as a classical compass needle rotating around the y-axis. Furthermore, underneath the needle, a range-finder was added to give the user an idea about not only the direction, but also the distance to the product (see fig. 4.13).

Figure 4.13: The compass situated on top of the controller on button-activation

In snippet 4.8 the *Update()* function displays the simple functionalities of the compass. The *CompassCanvas* is the UI of the of compass i.e. the white encirclement and the range-finder. Since the compass is attached to the right controller, the range-finder needs to be readable at all time, from all possible angles. Therefore the *CompassCanvas* is constantly pointing towards the camera using the *LookAt()* function. The *target* variable is a public accessible *GameObject*, which is always the current product being searched for on the shopping list. The *target* direction-vector is calculated by subtracting *target.position* from the current position of the compass. To make a smooth rotation, the *RotateTowards()* function is used, which essentially performs a linear interpolation at a predefined rate of change with a user-defined constant, *arrowRotationSpeed * Time.deltaTime*.

```
if(target != null && this.name == "Needle"){
  float step = arrowRotationSpeed * Time.deltaTime;
  targetDir = target.position - transform.position;
  targetDir = new
     Vector3(targetDir.x,targetDir.y*0.0f,targetDir.z);
  newDir = Vector3.RotateTowards(transform.forward,targetDir,
     step, 0.0f);
  transform.rotation =
     Quaternion.LookRotation(newDir,Vector3.up);
```

```
}
if(this.name == "CompassCanvas"){
  point = Camera.main.transform.position;
  if(target != null){
    distance = (float)System.Math.Round((double)Vector3.Distance
      (target.transform.position,this.transform.position),0);
    distText.text = distance.ToString()+" m";
  } else {
    distText.text = "? m";
  }
  this.transform.LookAt(point,Vector3.up);
}
```

Snippet 4.8: CompassBehavior.cs - *Update()*

Hints level

Hints toggle was assigned to the menu button on the right controller; pressing said button would toggle between the hints cyclically. Due to the different levels of assistance of each hint, they were divided on five levels in crescent order of aid. The levels were structured as follows:

- Level 0 - no hint
- Level 1 - Light beam
- Level 2 - Map
- Level 3 - Blinking products
- Level 4 - Compass

4.2.8 Teleportation

Being the simulation set in a 1:1 representation of Føtex, a locomotion system had to be implemented. During the preliminary meeting, different locomotion ideas were considered, of which redirected walking, walk in place (WiP), and teleportation were the preferred options. According to previous research (see section 2.7.2), redirected walking requires a lot of space to be perceived as real, and also a great deal of effort. WiP-methods could be feasible through sensors in the shoes or inertial measurement (e.g. measure the distinct motion of the HMD during walking), but accuracy is often compromised. Teleportation was preferred for more than one reason, not only would it be quick to implement with preexisting toolkits (e.g. VRTK[3]), but it would also make it quick to navigate around the

[3] Virtual Reality Toolkit (VRTK): https://vrtoolkit.readme.io/

virtual supermarket. The greatest concern was how the kids would react to the more abstract way of navigating a virtual environment. A conclusive decision was reached when the teachers assured us that the students have already been playing VR games and practiced teleportation mechanics. According to them, they picked up on those functionalities almost instantly even down to the lower grades.

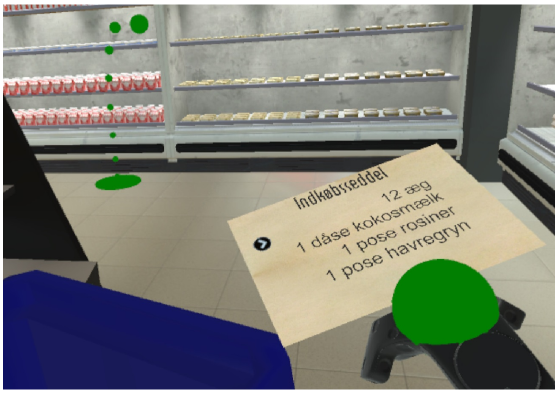


Figure 4.14: The VRTK teleportation utilizes a bezier curve as a pointer to teleport the user to a new location.

Teleportation mechanics were implemented using the VRTK toolbox. The teleport was customized to utilize a bezier curve (instead of a linear ray-casting) for better precision, with a maximum travel distance of 10 meter. The bezier curve pointer is visualized with green orbs from starting point (tip of controller) to the destination point (marked with a flat green cylinder). The curve pointer is activated on the trackpad-button and visible as long as the trackpad is pressed (see fig. 4.15), and teleportation occurs when the thumb is lifted. Teleportation relies on impact with a *GameObject*'s collider, meaning that it is not restricted to a singular level. This implied that users could teleport not only on the floor level, but on shelves, refrigerators, and other environment objects. All attempts to restrict the maximum height of the target point were unsuccessful, so instead all shelves and refrigerators were supplied with an extra trigger collider around them. This solution is not entirely reliable since, as soon as the user is inside the collider, it will be possible to teleport on top of other objects. This was, however, found to be an acceptable solution.

4.3 Final version

During this iteration of the development, several implementations dealt with the implementation of interactions based on button clicks, thus the correlation between buttons and actions has been visualized and presented in figure 4.15 for clarity.

Figure 4.15: The custom controller mapping used in the application.

4.3.1 Tutorial

Due to the high learning curve required by the system observed during the evaluation, it was agreed to structure a tutorial so to teach users the basic interaction techniques and products handling. The following list presents the actions required by the users in a chronological order:

1. Close shopping list - starting with a visible shopping list hovering on the right controller, users are instructed to press one of the two glowing grip buttons in order to toggle the list.

2. Open shopping list - in order to sediment the concept of toggling the shopping list, users are required to press the highlighted grip buttons once again.

3. Press trigger button - key point of the simulation is interaction with products and in order to properly instruct the users, the learning process was separated in two parts. Thus, the first part was aimed to have users practice

the proper trigger button press, which due to its nature, is only registered when the button clicks at the end of the travel length.

4. Place three products in the basket - once the users got familiar with the trigger press, they are instructed to touch a product and press the trigger; this would enable them to grab a product and move it around. Doing so, they were required to move three blank products from the shelf to the basket. Product models were texturized with blank materials (Figure 4.16), so to not bias users in any way when practicing.

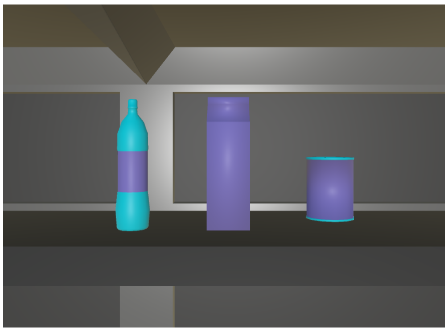


Figure 4.16: The three blank products used in the tutorial to teach users object manipulation

5. Press the hint button - another key feature of the experience is the possibility of having cues help users along the way, thus they were taught to activate them by simply pressing once the menu button on the right controller.

6. The subsequentsteps presented the user with the different hints (voice assisted help, blinking signs, and compass.

4.3.2 Hints

Hint evaluation suggested that the previously implemented guides might be overwhelming for participants, and would have a too high level of abstraction from the physical environment. It was then agreed to develop tools that would invite participants to look for hints in the environment and try to rely on them. On this baseline, the hints level were reduced from 5 to 4: light cone, map, and blinking products were removed. In their place, two similar hints were added. Level 1 was substituted with a simple voice-over that would invite users to look around them, so to identify any element that could help them find the area they were looking for.

This hint was aimed to have users independently search for signs hanging from the ceiling, or placed on top of shelves, without any external help. If the cue would not help them, level 2 was implemented so to have the relevant signs blink in order to attract users' attention and then locate the proper area. Level 3 was considered the lifeline and was aimed to direct users in a reliable and precise way to the product of interest, thus the compass was maintained.

4.3.3 Logging

The need for user behaviour data brought to the development of a script that would log specific actions performed by the user. The script was simply composed of one method for each event, which could be called from external scripts when the action occurred. Each event was assigned an integer id as follows:

- 0: "product_touched"
- 1: "hint_asked"
- 2: "product_put_in_basket"
- 3: "product_out_of_basket"
- 4: "product_grabbed"
- 5: "product_released
- 9: "total_time_spent"

Each method writes a string to a line in the log text file following a base structure:

event id, *product id*, *hint level*, *time since beginning*

Whenever one of the values might not be available, due to different events requiring different features, the value would be substituted with a *-1*.

4.3.4 Shopping basket

In the real shopping experience the users would enter Føtex with a shopping basket, so to facilitate common stimuli, a real basket was borrowed from Føtex, to be used in the simulation in the test participant's off-hand. The same basket was implemented in the virtual supermarket to further increase immersion through valid reinforced sensorimotor actions.

Physical implementation

The first iteration was created as a proof-of-concept, but its durability and reliability allowed to not have to redesign or reprogram it. A small holster was made from duct-tape to store the Vive controller, to make it easily removable for recharging and other applications purposes. For security purposes, a rubber strap was added to hold the controller in place, in case of exposure to extreme use (see fig. 4.17(a)). The handle of the basket was locked in a perpendicular position by two angle brackets, which served two purposes: **1)** it prevented the handle from falling down under the added weight of the controller, and furthermore made the handle easy to grab, and **2)** it prevented the virtual model from rotating 90 degrees either left or right, due to being parented to the controller-prefab.

(a) Physical

(b) Virtual

Figure 4.17: (a) physical basket, handle is locked in place by two angle brackets, and the pocket for the controller made from duct-tape with a rubber-strap holds it in place. (b) the virtual basket modelled using Maya.

Virtual implementation

The virtual version was modelled and UV-mapped in Autodesk Maya. Within Unity it was made a child-object of the VRTK *LeftController* prefab, and positioned approximately at the same location as the physical basket. The model was constructed in two parts: handle and body. To make the basket functional as a storage-device, a mesh collider was added to the body to prevent collision with the handles which might cause unwanted irritation during use. Furthermore a box collider was added to the inside of the basket, as a trigger, to detect products entering it. During preliminary usability testing, it was found that the basket quite often accidentally collided with shelves and refrigerators, which caused products to fall inside. This was solved by adding the basket to an independent *interaction layer*. Both things are hence rendered simultaneously, but layer-collisions are disabled. A script was created to transfer picked-up products to the *interaction layer* (see snippet 4.9). Furthermore, the controller model was hidden using *leftControllerModel.SetActive(false)* to allow only the basket to be visible in the participant's hand. In the *Update()* function, the *currentObjectHandled* by the *RightController* is transferred to the *interaction layer* i.e. layer 8. In the *OnTriggerEnter()* function, the currently held product is made into a child-object of the basket. Also, the *col.attachedRigidbody.mass* stores the mass of the product to be used for the haptic feedback. It can be notice that the retrogression to default layer is not coded in this script. That is because the *OnTriggerExit()* did not properly detect when a product was removed from the basket with the controller. Instead this event was coded in the *objectInfo.cs* script (appendix C.2), which is attached to all products.

```
private GameObject grabbedObject;
private GameObject InteractGrabController;
private GameObject currentObjectHandled;
private VRTK_InteractGrab VRTK_InteractGrabScript;
private VRTK_ControllerActions controllerActions;

void Start () {
  leftControllerModel.SetActive(false);
  controllerActions = VRTK_LeftController.GetComponent
  VRTK_ControllerActions>();
  InteractGrabController = GameObject.Find("RightController");
  VRTK_InteractGrabScript =
      InteractGrabController.GetComponent<VRTK_InteractGrab>();
}
void Update () {
  currentObjectHandled =
      VRTK_InteractGrabScript.GetGrabbedObject();
  if(currentObjectHandled != null){
    currentObjectHandled.layer = 8;
  }
}
public void OnTriggerEnter(Collider col){
  if(col.gameObject.tag == "product" || col.gameObject.tag ==
      "tutorial_product"){
    col.transform.parent = this.transform;
    float weight = col.attachedRigidbody.mass;
    RumbleController(weight);
  }
}
```

Snippet 4.9: ShoppingBasketBehavior.cs - *Start(), Update() and OnTriggerEnter() functions*

Haptics

Mass (kg) is frequently mistaken for weight, while in reality weight is rather the force exerted by gravity on an object of a certain mass (default in Unity is that of the Earth 9.81 Newton). Therefore, mass being an easily accessible constant unique to *GameObjects*, with an attached *Rigidbody*, it can be used to approximate the weight of an object by using mass as a constant for computing a weighted value. To simulate object weight, the relative mass of products acquired from their packaging, was used for haptic feedback in the off-hand. The haptic feedback in the Vive controllers is achieved with linear actuators, and though documentation is sparse, the function TriggerHapticPulse(*strength*, *duration*, *pulseInterval*) in Unity can be used to activate haptics given the parameters *strength* (value from 0-1) and *duration* (between 0 and 3999ms), and *pulseInterval* which is the delay-interval between each haptic pulse (in microseconds), but it is very poorly documented[4], and does not appear to vibrate longer than the duration of a frame. A long coherent vibration can be achieved with a coroutine (i.e. similar to multithreading, composed of independent, executing computations). The *RumbleController()*-function takes object-mass (*ObjWeight*) i.e. a value between 0 and 1kg to determine duration of the haptic pulse, in the hope that it will subjectively translate into a perceived sensation of object weight (see snippet 4.10).

```
void rumbleController (float objWeight)
{
  StartCoroutine(LongVibration(objWeight));
}
IEnumerator LongVibration(float objWeight) {
  float strength = Mathf.Clamp(objWeight, 0f, 1f);
  float adjust = strength/10.0f;
  for(float i = 0; i < adjust ; i += Time.deltaTime) {
    controllerActions.TriggerHapticPulse(Mathf.Lerp(strength,
       3999.0f,0.01f));
      yield return null;
  }
}
```

Snippet 4.10: ShoppingBasketBehavior.cs - *RumbleController-function*

[4] https://vrtoolkit.readme.io/docs/vrtk_controlleractions

4.3.5 VR optimization

The human ability to orient ourselves with, spatial skills, in both real and virtual environments is assumed to be a cognitive skill, yet it can also be described as the confluence of vision, vestibular cues and proprioception [67]. In the theory of multisensory integration the Medial Superior Temporal area (MST) within the cerabral cortex is very sensitive to sensory discrepenacies e.g. between vestibular and visual cues [67],[68]. This can cause a contradiction of sensory-input also known as visual-vestibular sensory conflict, which is believed to be the main cause of cyber-sickness (or VR-sickness) [69]. Within a VR-application, the user is especially sensitive to cyber-sickness because more senses are at play, and the core aspect being self-generated visual motion, any discrepancy between the senses may lead to discomfort. That also means, as opposed to conventional media, high latency on the refresh-rate (flicker) and high latency on frame-rate (lag) will be more likely to cause an unpleasant experience, loss of immersion or decreased task performance [70]. Therefore, several sources suggest that the lower framerate-limit of VR-applications should be at 60 frames-per-second (fps), and ideally around 90fps.

There are a number of steps that can be taken to improve the overall framerate of the system, besides delimit the level of object-quality (limited vertices and texture resolution), things like static objects, light baking or occlusin culling can also help delimit the amount of computations per frame, and hence take a load off the GPU (Graphic Processing Unit).

Static objects

Essentially, things that will not move, or could potentially be moved by the user, should not be dynamically updated with each frame. Static objects can be determined in the editor, and and can prevent wasting resources on physics calculations. Furthermore, static objects are also the only objects affected by light-baking i.e. lighting and shadow computations can actively be precomputed.

Light baking

As previously mentioned, light baking is the precomputation of lighting and shadows of .e.g environmental lighting, global illumination or static light sources within the scene. Dynamic light sources, susceptible to change (light intensity or position displacement) are computed every frame. That is of course required for light sources that need to change e.g. a flashlight, but for static un-moving lights, it is a waste of resources. Light-baking essentially creates a precomputed light-map stores just like a texture, which can be preloaded and hence optimize the performance [71, pp 104-105].

Occlusion culling

In VR the system is loaded even more, because the GPU has to render everything within the view frustrum twice (i.e. for the two seperate monitors in the HMD needed for stereoscopic rendering). Occlusion culling is a method of preventing the rendering of objects that are partially or fully occluded by other objects within the scene, or not currently visible within the camera's view frustrum (i.e. frustrum culling).

Reflection probes

By default, reflective materials within the scene, make use of the virtual sun (i.e. directional light) and the skybox to compute and emulate the behavior of light on reflective objects [71, pp 312-31]. However, this would create incoherent lighting effects, since the VE consists solely of an indoor environment. One solution would be to apply a skybox which resembles the environment, but that cannot be accomplished without e.g. photographing the environment with a 360°-camera. Alternatively, light probes can be applied, which samples the color of static objects within a predetermined area and resolution, and builds a local cubemap for that specific area This creates a more realistic impression of reflection without further loading the GPU.

A total of 2 reflection probes have been deployed in the VE, one for the tall row of shelves, and another for dairy department, both with a resolution of 128-bit.

4.3.6 Performance issues

Once shelf population and 3D models have been fully implemented in the system, a preliminary performance test showed a drastic decrease in framerate when several shelves entered the field of view of the headset (an 88.8% drop in frame-rate i.e. from approximately 100FPS to 40FPS on a GTX1080 graphics card), even with performance-optimization. Being the framerate significantly lower than the acceptable 60FPS standard, it was decided to remove the vegetable and fruit department products from the test environment. This meant less computations to be made in regards to light and physics calculations, and minimized the latency-problem creating a more steady and smooth system-performance.

4.4 In-test changes

Once the first week of testing was completed, an initial assessment was conducted; results showed that participants were feeling not challenged by the experience, consequently the supervisors were afraid that the loss of challenge would convert into less interest, thus nullifying the results of the second week of testing. An exchange of emails with the supervisors clarified that the system would need to be more challenging on both the sensory and functionality levels. It was then agreed to add background noise to the supermarket, and to remove the hints for the exception of the compass.

4.4.1 Ambient sound

Being the experience solely visual in order to avoid sensory overload, the supervisors considered appropriate to challenge the participants by adding an extra layer of distraction. Another field trip was conducted to a another Føtex to record a realistic soundscape. The soundscape was captured with an old minidisc device, but unfortunately the recording turned out so choppy due to dynamics and compression, that it was found unusable. Instead an ambient supermarket soundscape was found online [5]. It is a soundtrack of background noise recorded in a supermarket was shared by the teachers and was included in the simulation as a looping audio source.

4.4.2 Removed hints

Hints were considered to be too accessible and participants were often observed to skip the first two options by triple-pressing the assigned button. On top of this, participant appeared to be disoriented by the first two levels of hints, thus the supervisors suggested their removal. The only hint maintained was the compass, which could not be activated by users anymore; it was instead bound to a keyboard press, which was accessed and handled by the participants' supervisor. It was agreed that the hint would be activated at supervisor's discretion, when the participant would need it.

[5] Supermarket soundscape used for the simulation: https://www.youtube.com/watch?v=WoxnL5dakyA&t=43s

4.5 Observations app

The web-app presented here was created as an addon to the observation sessions for the experiment. The app was developed using Angular 1.6 framework and the Angular Material library, with an auxiliary Apache2 server running Slim Framework for database requests. As can be seen in figure 4.18, the app has a one-page structure and based on button presses displays different dialogs. Figure 4.18 (a) presents the main screen on which events can be logged by pressing the light blue buttons, or settings can be changed pressing the circular buttons. The left circular button allows to change the names of the participants that were observed in order to reduce the association effort while trying to couple a participant with the correct button. The left round button allows to choose which observer is using the app. Concerning the event buttons, once one of them was pressed, the app prompted the user with a dialog in order to select the participant that performed the action, and once the "user button" is pressed, the event is logged in the database. Some of the buttons, as presented in section 3.3.1, need more input arguments other than the participant id, thus requiring an intermediate dialog in which users could input different kind of data, ranging from notes to simple button presses.

Figure 4.18: Three screenshots of the observation web-app

Chapter 5

Evaluation

5.1 Methods

Since the goal was to establish whether learning in the virtual context had been transferred to performance in the real context, a variety of measures was applied to facilitate validation of the data through cross verification. Although some of the measures seek to compare performance through quantitative methods, we are aware that results would not directly translate to tools the participants will benefit from; task completion and completion time do not necessarily translate to a safe shopping-experience through instilled self-confidence. However, we do believe it can be an appropriate strategy for establishing whether or not changes have occurred, and whether changes in performance, self-confidence rating, and impression indicate the treatment had an effect. The treatment sessions were conducted at Valhøj School by Claes Kiilsgaard, in an approximately 7x7m large room in the basement. The room is normally used as clubroom and has Playstation, pool-table and flat-screens. The test participants were taken one at a time from the home economics/cooking-classes and each participant was predicted to complete the training and self-reporting in 20 minutes.

The baseline observation session was carried out Tuesday, April 18th, with a total of 5 observers using the purposely built web-application to quickly and discretely note down key events and observations. The schedule for the baseline was as follows:

- **09:30** - Rendezvous at Føtex, Rødovre Centrum, divide positions
- **09:50** - The participants begin walking from Valhøj Skole
- **10:10** - Estimated arrival at Føtex
- **11:30** - Shopping trip scheduled to end

The final evaluation was carried out Friday 28^{th}, and the time of arrival of the participants had been changed by the teachers to 12:45, hence shifting the schedule. A distinct increase in customers population level was noted between the two sessions, and several factors such as last bank day of the month, outlet in Føtex, later time of the day, were deemed responsible for such discrepancy.

5.1.1 Performance measures

Three distinct methods have been used to evaluate the participants' performance in both the baseline and post-treatment assessment: task completion time (TCT), effectiveness, and time-based efficiency. The purpose of these methods was to compare the Treatment Group (TG) and Control Group (CG) accordingly. TCT is the time used to complete a task [72, p. 224], and it is also known as time to completion or task performance time. Besides being a frequent measure in usability testing [73] it has also been used to compare performance between a baseline and post-treatment in ASD interventions [34; 42; 55; 58].

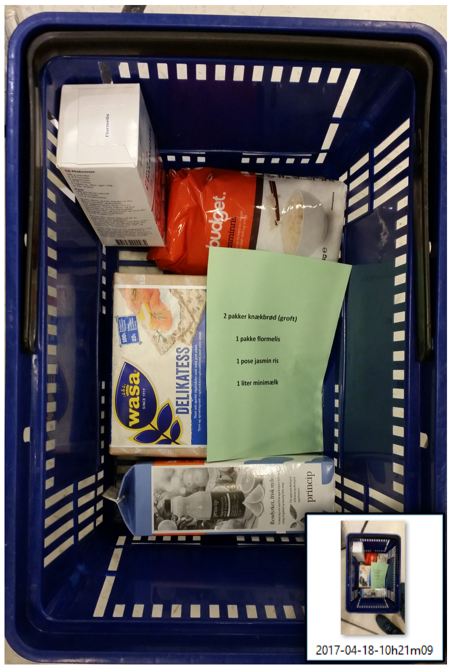


Figure 5.1: Test participant H's shopping basket after the baseline shopping trip with timestamp generated namexif (lower right corner) showing picture taken 10:21:09.

Task Completion Time

For this experiment, TCT has been measured as individual performance for all 9 participants, as the onset of the shopping (the exact time when they entered Føtex established by observing and timestamping the time of arrival) until they met up with the accompanying teachers. The offset of the shopping was flagged by a photo taken with the teacher's phone, in which the EXIF (Exchangeable Image File Format) encoding also contains information about exact time the image was taken in the format YY/MM/DD/HH/MM/SS. The seconds are not viewable in properties and details, but a timestamp can be decoded through metadata-extraction. The free software Namexif[1] was used to extract the information, to get a more precise

[1] Namexif homepage: http://www.digicamsoft.com/softnamexif.html

measure of the TCT (fig. 5.1). The image itself was subsequently used to measure shopping effectiveness and efficiency. Furthermore, due to the difference between the baseline and post-treatment conditions i.e. primarily due to a significant increased the number of other customers, we will apply a difference-in-difference (DiD) estimate. DiD is frequently used in econometrics and social sciences to estimate treatment effects. DiD is especially useful when both CG and TG are mutually influenced by external factors 5.2) [74]. In our case, both groups have been exposed to the same increased amount of other shoppers, who subsequently caused a more challenging experience.

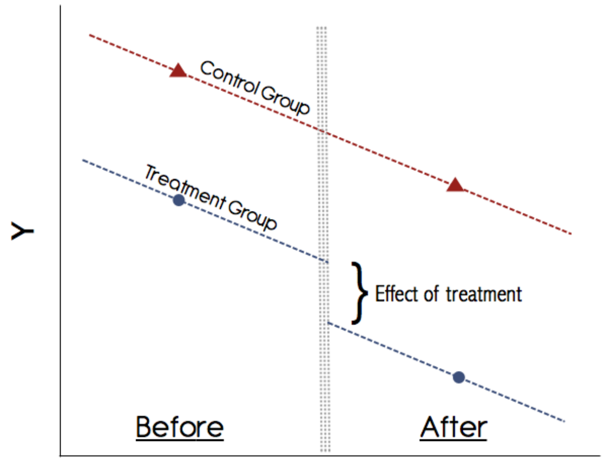


Figure 5.2: The difference-in-difference (DiD) measures the effect of the treatment group against the control group, while accounting for external factors.

Essentially, DiD calculates the difference in the differences as (change in TG - change in CG), where the change is computes as value after treatment minus value before treatment. The DiD estimate can also be obtained using a linear regression and *Ordinary Least Squares* (OLS) to estimate the coefficients (eq. 5.1):

$$y = \beta_0 + \beta_1 \boldsymbol{T} + \beta_2 \boldsymbol{S} + \beta_3(\boldsymbol{T} \cdot \boldsymbol{S}) + \epsilon \qquad (5.1)$$

Where T is a dummy-variable (0 or 1) for baseline / post-treatment and S is a dummy-variable (0 or 1) for CG or TG. The β_3-estimate equals the DiD-estimate, and by applying a regression, the standard deviation can be calculated and hence facilitate the computation of the level of significance. Another advantage of the regression approach is that it allows to add additional independent variables that could have had an effect on the DiD-estimate [74].

Effectiveness

In ISO 9241-11, effectiveness is defined as *"the accuracy and completeness with which users achieve specified goals"* [75], and one way of measuring it is by the amount of successful tasks performed by the individual, divided by the total number of tasks performed [73, p. 13]. It is frequently used in Human-Computer Interaction (HCI) and usability testing to evaluate a user's interaction with the system, where it's often referred to as Completion Rate (CR) [73]. Usually CR is a binary measure of task success (coded as 1) or task failure (coded as 0), but it's also possible to define a partial completion criteria (e.g. 0.5) [73, p. 13]. Furthermore, it's also used in cognitive sciences to evaluate mental processes and reactions to stimuli such as reaction time (RT), accuracy which are similarly computed [36, pp. 38-42]. For example, in eq. 5.2 the CR is presented as a simple equation which results in a percentage of successfully completed tasks:

$$\text{Completion Rate} = \frac{\text{Successful tasks}}{\text{Total amount of tasks}} \times 100\% \tag{5.2}$$

In ASD research effectiveness (referred to as accuracy) has been applied as a performance measures *inter alia* in video modelling to compute the error rate in number of correct tasks in a 32-step Hierachical Task Analysis (HTA)[55], Schaller and Rauh used accuracy-rates to evaluate ToM-items such as facial emotion categorization[76], and finally Mei *et al* evaluated success rate by number of blocks divided by total number of kicks in a football scoring game [77].

For evaluating effectiveness, we have used the CR equation (5.2) with the following criteria:

- **1** Task success: *correct product bought.*
- **0.5** Partial completion: *Wrong quantity or wrong type.*
- **0** Task failure: *incorrect/no product bought.*

Efficiency

> For example, many treatment approaches and demonstration projects have disseminated information, yet most have not yet provided appropriate, scientifically rigorous documentation of effectiveness and efficiency. [9, p. 21]

In ISO 9241-11, efficiency is defined as *"the resources expended in relation to the accuracy and completeness with which users achieve goals"* [75]. We have not yet come across a study that applies this, even though it could be an interesting perspective e.g. how many resources (effort, energy or time) does the person expend on a given task. This becomes increasingly more interesting given the above quote from *Educating Children With Autism* which calls for a more rigorous documentation for these types of measures in developmental psychology. It is unclear whether or not this is exactly what they mean, but given the clear link between usability methods of evaluating effectiveness and ASD performance measures, we believe it validates the application of such measures in this aspect as well.

For evaluating efficiency, we will use the only possible resource we have at our disposal i.e. time. A relative efficiency measure for time (eq. 5.3) was defined as follows:

$$\text{Relative Efficiency} = \frac{\sum(\text{Completion Rate} * \text{Task Completion})}{\sum(\text{Task Completion})} \times 100\% \tag{5.3}$$

5.2 Results

Of the 9 children (8 male, 1 female) 8 participated in both baseline and post-treatment assessment (table 5.1). Child **F** in the CG participated in the baseline, but was replaced by another child in the post-treatment assessment. The data was forfeited in the comparison of performance to avoid confounding the results. Furthermore, due to human error shopping list 8 was incidentally used twice in the baseline session.

Table 5.1: Test participants

Treatment	Control
A	E
B	F*
C	G
D	H
-	I

* *Participant F will not be used in groupwise-comparison.*

5.2.1 Baseline and post-treatment assessment

In table 5.2 the assessment results are presented showing the shopping list (List No.) in the baseline i.e. Before (B) and post-test i.e. After (A), TCT, and computed effectiveness and relative efficiency.

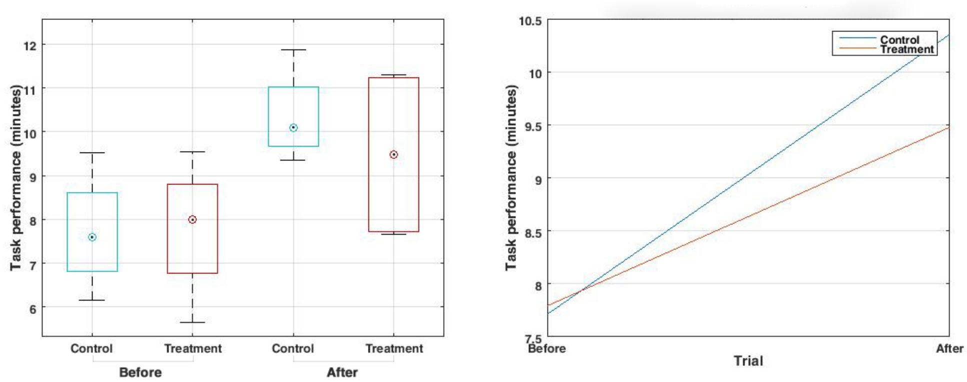
Table 5.2: Assessment Result Table

Part.	List No. B/A	TCT (minutes) Before	TCT (minutes) After	Effectiveness (percentage) B/A	Efficiency (percentage) B/A
A	7 / 1	9.9	11.18	100 / 100	
B	8[1]/ 2	8.08	7.78	100 / 100	
C	8[1]/ 4	5.65	7.65	100 / 100	
D	3 / 7	7.9[2]	11.3	100 / 100	
		avg.=7.90 std.=1.61	avg.=09.48 std.=2.04	100 / 100	100 / 100
E	9 / 5	7.71	9.35	100 / 75	
G	1 / 3	7.48	10.00	100 / 100	
H	5 / 6	6.15	11.86	87.5 / 100	
I	4 / 9	9.51	10.20	100 / 87.5	
		avg.=7.71 std.=1.38	avg.=10.35 std.=1.07	96.87 / 90.62	97.51 / 89.18

[1] *Note two similar shopping lists given due to human error.*
[2] *Image was not taken upon shopping-completion, but the teacher reported D arrived as number 5, so TCT has been estimated as the avg. of B & E.*

Results are reported in figure 5.3(a). Both CG and TG had an increase in time elapsed between baseline and final evaluation. The average TCT for the CG increased with 2.6 minutes (34%) and with 1.38 (18%) for the TG (see fig. 5.3). On average TG were 11 seconds slower than CG in in the baseline assessment, but 52 seconds faster in the post-treatment assessment.

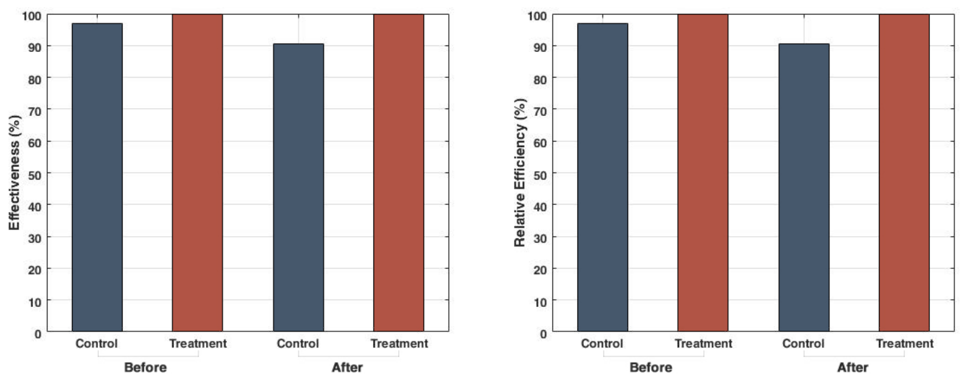
For the effectiveness, the CG scored an average of 97% at the baseline because participant **H** scored 3.5/4 (partial completion reason: bought only one pack of crispbread/knækbrød instead of 2). At the post-treatment assessment the CG



(a) Boxplot comparison (b) Plotted task performance

Figure 5.3: (a) boxplot showing Task-completion time (TCT) between baseline and final shopping experiences, and (b) TCT line plot

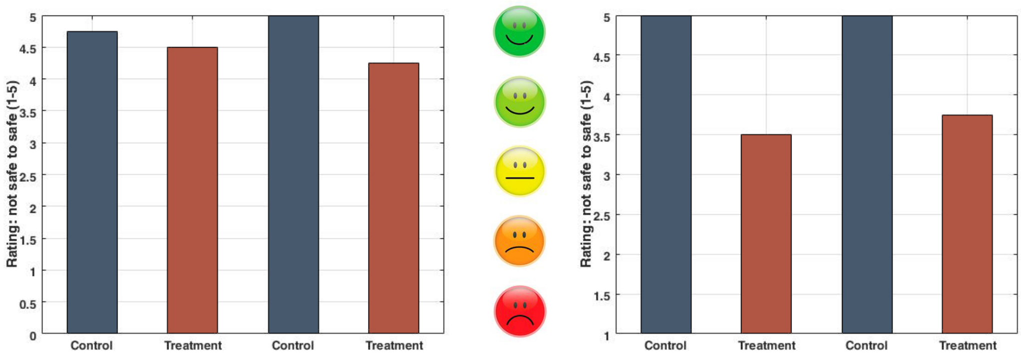
effectiveness score decreased to 91% because participant **E** scored 3/4 (task failure reason: did not buy crispbread/knækbrød) and participant **I** scored 3.5/4 (partial completion reason: bought mild cheese instead of mature). For the TG there was no difference in between the assessments, as all participants scores 100% in both effectiveness and efficiency (see fig. 5.4).



(a) Effectiveness (b) Efficiency

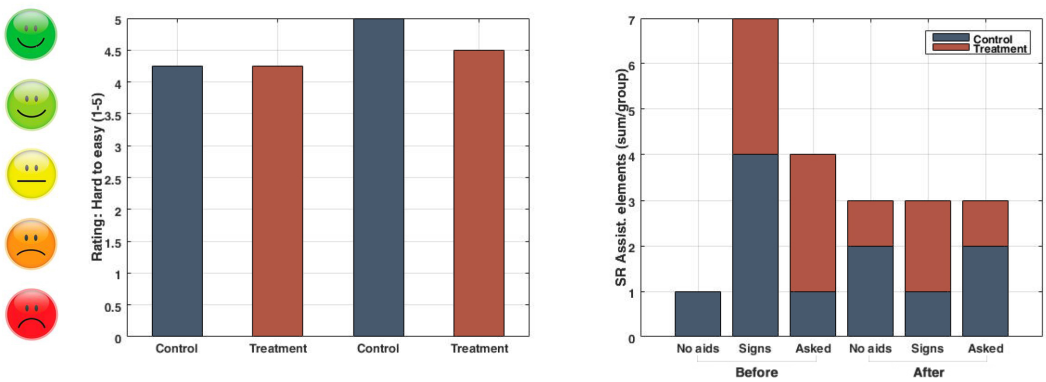
Figure 5.4: (a) average shopping effectiveness (%) of tasks completed correctly between CG *(blue)* and TC *(red)* and (b) average shopping efficiency (%).

For the self-reported (SR) confidence and prospected confidence (5-point likert-scale from *Not Safe - 1* to *Safe - 5*), the CG baseline rating was 4.75 and 5 in the post-treatment i.e. a 0.25 increase (12%) for confidence. The CG showed no changes in the prospected confidence between the two trials. The TG rated 4.5 in the baseline and 4.25 in the post-treatment i.e. a 0.25 decrease (5.5%) in confidence-rating, and a 0.25 (7%) increase in prospected confidence (see fig. 5.5).



(a) Confidence-level (b) (c) Prospected Confidence-level

Figure 5.5: (a) Self-evaluation of the avg. confidence per group in the shopping situation today, and (c) of the prospected confidence in shopping alone the next day.



(a) (b) Ease of shopping (c) Assistive elements used

Figure 5.6: (b) SR of how easy it was for the participant to shop today, and (c) the summed SR assistive elements used (before and after)

On the question of how easy the participants perceived the shopping (Hard to easy) the CG saw a increase in average perception of easiness from 4.25 to 5 between the assessments (a 5% increase), while the TG saw a increase in SR ease of shopping (see fig. 5.6(b)) from 4.25 to 4.5 (a 6% increase). For the questions regarding assistive elements (fig. 5.6(c)), 7 of the 9 participants reported that they used signs for help (CG = 4, TG = 3) in the baseline, where the number decreased to 3 (CG = 1, TG = 2) in the post-treatment assessment.

4 participants reported that they asked for help (CG = 1, TG = 3) in the baseline and when asked again after the post-treatment assessment only 3 participants had asked for help (CG = 2, TG = 1). Only a single participant in the baseline reported that he/she did not use any assistive elements No Aids (hints), but in the post-assessment the reporting had tripled (CG = 2, TG = 1). Overall the reported assistive elements used decreased between the baseline and the post-treatment from 11 to 6 (a decrease of 45%), but the data does not outline the direct reason for this.

Difference in Difference

The DiD regression-estimates were tested for significance with significance levels of 5, 10 and 15% ($\alpha = 0.05$, $\alpha = 0.1$, & $\alpha = 0.15$) due to the small sample size. The results are presented in table 5.3, where the β_3-coefficient is the standardized coefficient resulting from the DiD-regression analysis.

Table 5.3: Difference in difference (DiD) results

Predictor(s)	Coeff. (β_3)	Lower CI	Upper CI	*p*-value
1) Raw DiD-estimate	-1	4.53	2.44	0.53
2) Effectiveness	-1.58	-5.13	1.97	0.36
3) Effectiveness Required help	-3.92	-7.56	-0.29	0.054**
4) Effectiveness Required help Est. Traverse	-3.11	-6.27	0.05	0.103*
Effectiveness Required help Green section	-3.92	-7.56	-0.29	0.036***

Inference: **** p<0.05; ** p<0.10; * p<0.15*

The predictors are independent variables added to attempt to explain causal inference i.e. elements that might explain the variation in the dependent variable (y). 5 different scenarios have been defined, 1) Raw DiD-estimate of the treatment-effect with no explanatory (independent) variables as predictors, 2) Adding effectiveness as predictor, 3) Furthermore adding *Required help* as predictor (i.e. SR assistive element *"asked for help"*), 4) With the addition of traverse (measured as the optimal route to get all products for all shopping lists. See fig. B.1 in appendix D) and finally 5) switching traverse with whether or not the shopping list contained *Green section* as a predictor. This estimate was based on the assumption that the treatment group did not get any training in this specific section, so any spatial transfer effect would most likely not have been facilitated. Furthermore, it is a section with significantly more products in small area, in a more unstructured and disorderly setup with less space to navigate. Therefore, shopping lists containing elements from that section are expected to experience lower performance.

The raw DiD-regression without independent variables resulted in a DiD-estimate $\beta_3 = -1$. The result is not significant at a 5, 10 and 15% ($P > 0.05, P > 0.10, P > 0.15$) level of significance. Nor was there any significant difference when adding *Effectiveness* as predictor (β_3 = -1.58).

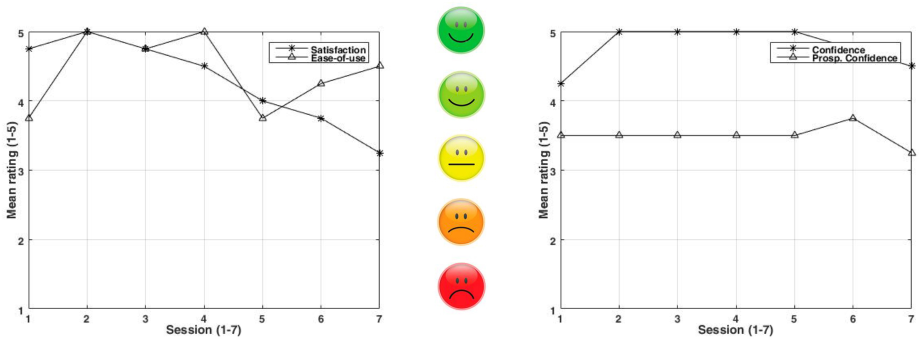
However, the DiD-estimate TG had significant ($P < 0.10$) change in performance with the *Required help*-predictor, and significantly ($P < 0.15$) better performance with added *estimated traverse*-predictor.

Most significant difference ($P < 0.05$) in performance was achieved with the *Effectiveness*-, *Required help*- and *Green section*-predictor.

5.2.2 VR-session evaluation

Self-assessment

The self-assessment questionnaire conducted after the participants finished the VR experience are presented in figure 5.7. The usability-inspired satisfaction and ease-of-use over time (note that they were based on single questions, and not the System Usability Scale (SUS)) are presented in fig 5.7(a), as an average for all participants. Ease-of-use in the 1^{st} session is slightly below 4, increases to 5 and remain approximately steady through session 2-4, where it decreases to starting-level in the 5^{th}, and steadily increases again in session 6-7. Overall satisfaction appears to decrease steadily from around 5 in the two last sessions, and ending a little over 3 in the last session. Confidence ratings presented in figure 5.7(c) are stable from the 2^{nd} to the 5^{th} session, but decreases slightly in the last two sessions, a similar trend can be noticed in prospected confidence, which remains unchanged from session 1-5, but makes a slight fluctuation in the two last sessions.

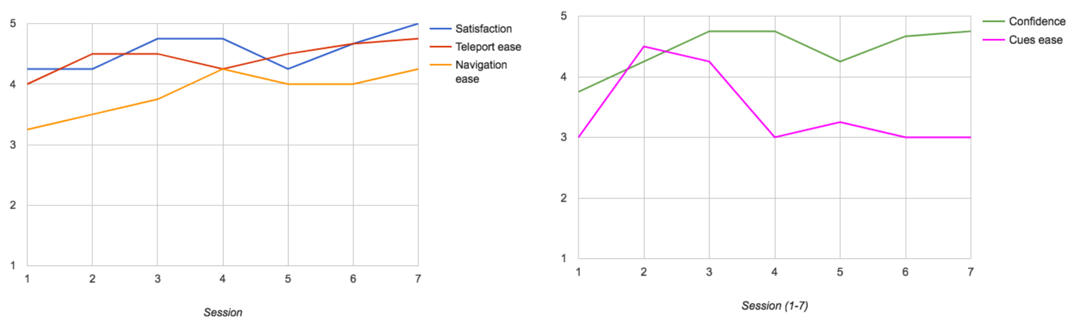


(a) Ease-of-use/satisfaction (b) (c) Confidence-rating

Figure 5.7: (a) satisfaction (asterisk) and ease-of-use (triangle) of the system. (b) is the pictorial scale used, and (c) confidence in the VR experience (asterisk) and prospected confidence in a real shopping scenario (triangle).

Teacher evaluation VR

The teacher was asked to submit a questionnaire after each participant terminated their VR experience. Figure 5.8 shows the evolution of the results means over time. The graph 5.8(a) reports satisfaction (blue), confidence (red), and teleportation technique ease (orange), while graph 5.8(b) presents cues usage ease (green), and navigation ease (pink). Each category has been represented based on means of the rating each participant was evaluated by the teacher. All of the categories have seen an upwards trend, with the exception of cues ease. The latter dropped on the 4th session, i.e. the beginning of the second week of training. On the same day, it can be observed that teleportation ease slightly dropped, but returned and surpassed to previous values the following sessions. Confidence has seen an increase during the sessions (3.75 in session 1 and 4.75 in session 7), with the exception of session number 5, in which it dropped by half a point. Satisfaction had a similar behaviour, but started from 4.25 and ended at 5.0. Navigation ease steadily rose since session 1 (3.25) and reached its maximum peak in session 4 (4.25), afterwards it dropped half a point, but it was recovered on the last session.



(a) Satisfaction, teleport ease, navigation ease (b) Confidence, cues ease

Figure 5.8: (a) Satisfaction (blue), teleport ease (red), navigation ease (orange) (b) confidence (green), cues ease (pink)

Teacher notes VR

The teacher's notes regarding the participants' experience are fully reported in appendix E.2. Hereafter are presented the filtered and summarized notes:

- Hard to go from "look at signs to compass, but the compass was extremely helpful and easy for the students to use. I haven't seen if the signs are blinking where they need to go, or all of them.
- Generally it's weird that you can teleport in top of shelves.
- I experienced that **A** had a hard time with the handle to the shopping list. He pressed teleport instead of taking groceries. He ended up on top of a shelf.
- **B** found the compass by double clicking on the cue-button. he found the"shortcut" himself
- The pupil is starting to get bored in the "game". There's missing development through challenges. I try next week to ask him not to use the compass.
- I noticed that none of the first three students threw objects around. Maybe it was more realistic with sound.
- Asked for compass, but I said to look at the shelves. He found tuna himself.
- Had a problem finding "Rød Karry" because the product label was in english and said "Red Curry"

5.2.3 Data logging

Data was logged each day and subsequently extracted from the testing computer. The initial idea was to track the number of hints used to evaluate overall performance, but since hints were partially removed[2] between session 3 and 4, measures of use are only existing for the first three days.

TCT was logged for each participant in the treatment session 1-7 (Note that participant A was not present the last day, and could therefore not participate in the session). The fastest time was B(4) at 2.5 minutes, and the slowest TCT (8.6 minutes) by the same participant the following day. From session 1-4 there is a declining trend in TCT, abrupted by session 5, where-after the TCT declines again (see fig. 5.9).

[2]They were allowed to activate on request, but could no longer do it personally

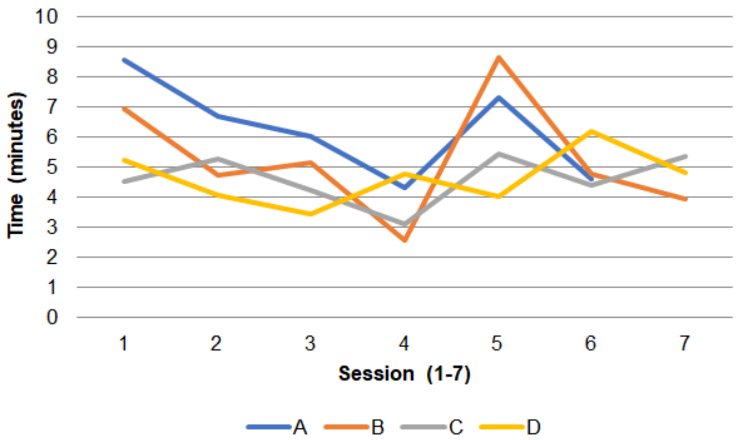
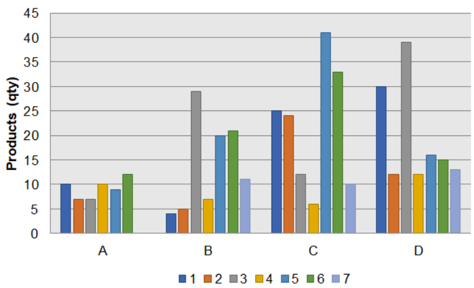
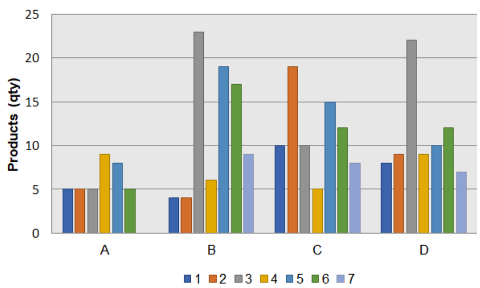


Figure 5.9: The time *(min)* spend in each session by each participant.

Fig. 5.10(a) shows the number of products touched (i.e. controller touches the products collider) by each participant, and fig. 5.10(b) shows the amount of products grabbed. On average, across all sessions, participant **A** touched 9 products and grabbed 6, **B** touched 14 and grabbed 12, **C** touched 22 and grabbed 11, and **D** touched 19 and grabbed 11. Least products touched and grabbed were **B** in session 1 (i.e. not more than 4 products both touched and grabbed) and most objects touched were **C** in session 5 (41 products).



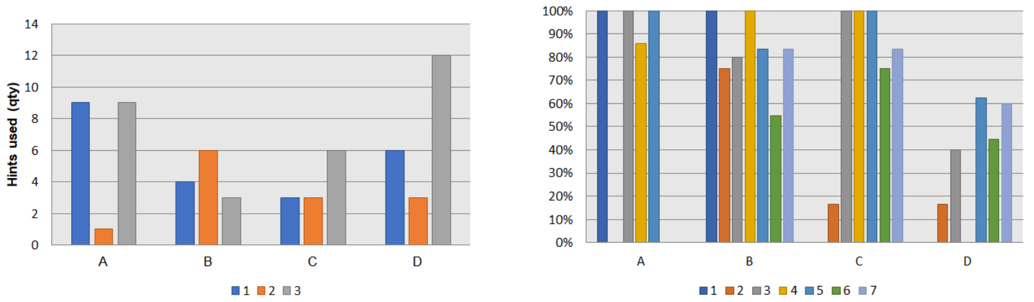
(a) Products touched



(b) Products grabbed

Figure 5.10: (a) amount of products put in shopping basket for participant A-D and (b) the amount of correct products put in shopping basket.)

Hints were just available for the first three days. **A, B & C** activated on average 4 hints per session, and **D** activated 7. For all the participants, the average use of hints the first day was 5.5 hint activation, 3 the second and 7 the third (see fig. 5.11(a)). Not all of the participants used the hints the same amount of times. On average **A** used 6 per session, both **B** and **C** used 4 per session, and **D** used 7 per session.



(a) Hints used/sessions *(qty)*

(b) Effectiveness (%)

Figure 5.11: (a) amount of products put in shopping basket for participant A-D and (b) the effectiveness-score at the end of the program. Note that this method does not include products who are no longer in the basket at test-termination.)

Fig. 5.11(b) shows the effectiveness rating for all participants across sessions. The effectiveness-score depends on the objects being in the virtual basket at check-out. The best overall effectiveness, as an average of all sessions, was achieved by **B** (81%) followed by **A** (75%), and **C** (68%) and with the lowest rating achieved by **D** (40%).

5.2.4 App Observations

The observations recorded for baseline and final shopping experience are reported in the following paragraphs. They have been structured and filtered, the full list of notes can be found in appendix D. Observations recorded though the app are presented in chronological order, while additional handwritten observations were taken after the experience ended in order to eventually complement or add on top of previous notes.

Table 5.4 reports all the times a participant was observed to be distracted by something in the environment. It can be noted how for the final session no observations were made.

Table 5.4: Distraction Events

Baseline	Final
B - got distracted	-
C - got distracted	-

Figure 5.12 presents quantified amount of help the participants were observed to use during each evaluation session. Treatment group was never observed to ask teachers for help, but asked several times (3) other customers for help. They have also been found to ask once Fotex clerks for help. On the other hand, during the baseline evaluation, control group has been observed to use all three entities' help (1-clerks, 2-teachers, 2-customers). During the final evaluation, control group was observed to not ask teachers or clerks anymore, and only twice other customers were inquired.

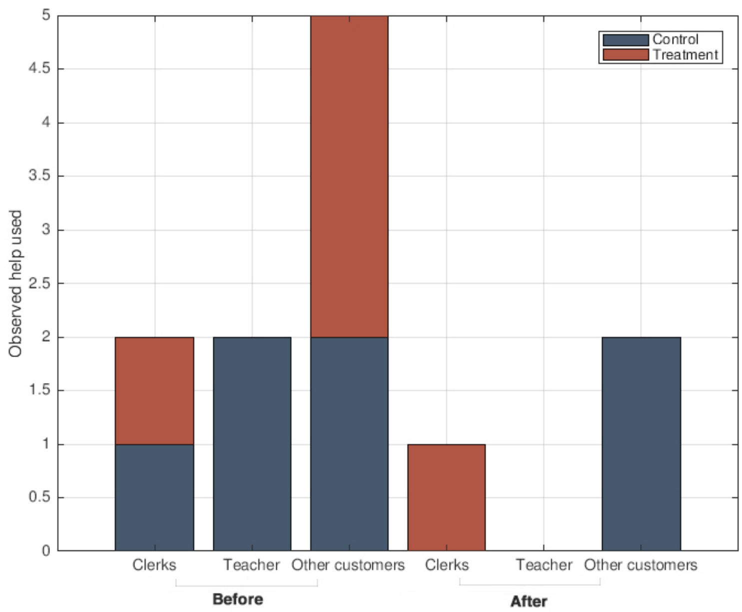


Figure 5.12: Help asked by CG and TG to Føtex clerks, teachers, other customers or other test participants

Table 5.5 shows events of what participants used environmental signs to orientate and find the product they were looking for.

Table 5.5: Sign usage Events

Baseline	Final
C - used a sign	-
H - used a sign	-

Notes

Notes that were input through the apposite event button in the app, have been gathered and presented in table 5.6. Although some of them are corrections or delucidations for other entries, some of the observations report participants' behaviour, e.g. "**F** - shouts", or participants' usage of environmental help.

Table 5.6: Notes

Baseline	**Final**
C - Look at product	**F** - Nothing found
C - Was corn	**F** - Nothing. Found two shelves
A - Tuna	**A** - Nothing found
A - Can't find	**B** - Crackers
F - Shouts	**B** - No found
I - Lost	**B** - Looking for something he can't find
F - Green looked for butter and used help from skolelærer	**E** - Prev dark blue
D - No look at signs	**E** - light green pock object from
I - Fotex guy helping	**F** - Walking aimlessly
I - Lots of tuna can't find	**I** - My user nine has no color
I - Tired to tomato sauce	**D** - Help with tuna don't know other color
D - Checks date on product	
D - Found piskefløde meget hurtigt og gik vider	
A - All at the same time	
D - Read the expiration date on piskefløde	

Chapter 6

Discussion

This chapter presents the analysis of results gathered and reported in the previous chapter. The acquired data, the adopted methods, and performance of the system are critically evaluated and discussed. Finally, we perspectivate the findings to the empirical foundation which formed the basis of our research.

The purpose of the experiment was to validate the treatment-method of using VR as a tool to facilitate transfer of learning to a real shopping scenario. Looking at the results, we cannot find a significant treatment effect on the TCT. Although the treatment group did outperform the control group, it was not a significant difference ($P > 0.15$) when considering the raw Difference in Difference estimate. However, the results can be further analyzed, and many factors could hypothetically have contributed to the performance. Several variations in conditions, ranging from the broad diversity of the spectrum, to the possible unavailability or limited accessibility of products, could be accounted for variations in performance. Additionally, the given search tasks might have varied in the degree of difficulty — although sharing rudimentary qualities — requiring participants to walk at different speeds and distances. All these factors lead to the realization that task conditions, let alone the groups themselves, are not mutually comparable unless the same task-based evaluation is followed for all participants, or repeatedly measured until all combinations of participants and shopping lists are carried out.

To account for these factors, we assumed that variations in Task Completion Time can be explained by other measured independent variables such as effectiveness, approximated travel distance (traverse), and increased difficulty (i.e. the occurrence of search task elements in areas estimated to increased completion time). Specifically, we weighed these independent variables as predictors in the Difference in Difference estimate as four explanatory assumptions that affected performance:

1. Individual task effectiveness
2. The addition of self-reported required help to achieve task completion
3. The addition of an approximated crude distance-to-travel estimate (traverse-estimate)
4. Binary traverse-estimate of difficulty based on products found in the *Green Section* of the supermarket.

Significant results were found in 3 of the 4 predictions: in scenario 3 ($p < 0.15$), scenario 2 ($p < 0.1$), and scenario 4 ($p < 0.05$). If we accept the assumption of the best resulting scenario, that these factors contributed to Task Completion Time, we can accept the alternative hypothesis that a significant difference was found between the Treatment Group and the Control Group at 5% level of significance ($p = 0.04$). However, although the factors discussed could plausibly explain fluctuations in the results, the evidence of causality is far from valid enough be taken as an absolute truth. Overall, the small sample size results in a wide confidence interval, that prevents us from making strong assumptions based on the Difference in Difference alone. Therefore all the other factors measured, including self-reported subjective opinions and that of the teachers must be considered as well.

6.1 Results

Besides the variations in the task itself that could explain differences in performance, elements of the treatment could have possibly affected the outcome of the post-treatment assessment. Despite strong evidence that the decrease in task performance was due to the difference between the assessment conditions itself, since it affected both treatment and control group, there could be other contributing factors as well. A reason for worse performance in treatment group in the final assessment, could be that VR system limitation forced the participant to follow the shopping list in order, rather than allowing them to find products in a free manner. The lack of elasticity in the system might have had a negative transfer effect (2.3.2) by forcing them to strictly follow the shopping lists in descending order, rather than following a subjectively estimated optimal route. This feature hindered the participants' training to the degree that the supervisor noted that one of the participants could not understand that *"you could not take butter before flour"* (E.3). Furthermore considering effectiveness 5.4(a) and efficiency 5.4(b), the results report the treatment group having constant perfect effectiveness, meaning that both baseline and final evaluation they collected 100% of the products on their shopping lists. On the other hand, control group had a decline in both effectiveness and efficiency and they dropped from around 97% to approximately

91%. It could be hypothesized that having trained with a more structured shopping process and on a more frequent basis, could have helped participants retain the ability of finding products' locations more accurately and confidently. It could also be argued that products on the control lists might have not been accessible or available to the participants; further analysis of the missing products and observations showed that products of interest were in fact available (e.g. "knækbrød") but not found by the participants in the control group.

Another assumption was that an effect would be measurable by the self-reported sense of confidence potentially instilled by the practice gained in the VR-treatment. However, results report a slight advantage for the control group. Confidence (fig. 5.5(a)) had a negative development for treatment group in self-report on their sense of confidence in the shopping task they just conducted. This could be imputed to environmental discrepancies between the supermarket and the VR counterpart. Participants trained in a empty and relatively un-interactive environment, and could have adapted to having an uninterrupted experience; once immersed in an environment with a high number of customers around them, they might have felt distracted and intimidated.

This theory can be confirmed when comparing these results to the self-evaluation; it was observed that once auditory information was added — thus straining the cognitive load on the participants — to the experience, confidence levels slightly dropped, as well as satisfaction. On top of this, the supervisor noted how sound stressed one of the participants (participant **D**), but also affirmed that "sound stresses him a little, not as much like real people do". This issue has been addressed in the preliminary meetings, the teachers suggested the possibility of creating virtual mannequins that would give the impression of being amongst other customers in the supermarket. The idea was finally not implemented, due to the high technical difficulty the implementation of several non-player characters would require, and the potential of inducing negative transfer if implementations were too uncanny.

The drop in confidence can be associated with the drop in self-reported assistive elements used; treatment group reported an overall large decrease in help asked or used. This could be imputed to the lackluster implementation of level 1 and level 2 hints, i.e. voice cues and signs highlights. The implementation of these hints was deemed confusing and disorienting by the supervisor himself, who commented several times that it was "hard to use" or "I haven't seen the signs blink" E.1. The two levels were altogether skipped several times until the point that a redesign of the system was needed to keep participants' engagement. While negative results were gathered for the confidence level, other confidence measurements reported contrasting results. Prospected confidence grew slightly from baseline to final evaluation for the treatment group, and confidence rates have been noted to grow throughout the VR training. These measurements can strengthen the hypothesis that environmental variables could have affected participants' performance during the specific final evaluation. Lastly, the discrepancy in confidence ratings could be due to the poor formulation (or translation) of the true meaning of the word. In the Danish formulation of the questionnaire, the word *tryg* was used which translates to *safe*. It is assumable that the students understood this as a measure of comfort rather then self-assurance which could potentially point towards a rudimentary difference between the groups (we will discuss this in 6.2).

Considering the VR training specifically, it can be observed that participants constantly reported high levels of self-confidence. The decline during the last two sessions can be attributed to the fact that, after hints removal, the supervisor tried to push the participants towards the boundaries of their comfort zone by temporarily negating access to the compass — "He asked for compass, but I told him to look at the shelves. He found tuna himself" E.2 — Such approach can justify the final increase in self-reported and teacher-reported satisfaction; it can be assumed that the achievement of finding a product without the use of hints could boost a participant's self-confidence. Similarly to confidence, satisfaction saw a clear decline when interactive hints were removed. This can be index of an unsettling experience for the participants.

In the analysis we learned that routine is an integral part of ASD (2.1), thus its disruption caused by the changes in the system might have triggered unwanted responses from the participants. On top of this, the simple increased effort that participants were required to do while searching for products, can have negatively influenced their experience. The teacher's notes report a clean differentiation between the two weeks, while at first participants grew bored of the experience and started playing with the virtual products, the second week they were observed to adhere more rigorously to the rules. The teacher suggested that this change might have been triggered by the introduction of sound, thus creating a more realistic and stressful environment.

Difference in the two experiment environments can also be noted in reports of the general ease of use of the system. When hints were removed between session 3 and 4, the teacher started pushing the participants' performance. However, we did not see a decrease in performance or reported confidence until session 5 where ease-of-use drastically declined, rose again the following sessions. This implies that it was due to conditions in that specific shopping list that caused lower performance, possibly because participants had trouble finding the product "Red Curry". The discrepancy between the name on the shopping list "Rød Karry" and the name on the product label "Red Curry" was observed to disorientate the participants, hence causing them to handle several products before finding the correct one.

Furthermore, all participants' amount of time spent in VR saw a strong increase in correspondence of the same session. Both data-sets trends for the second week show a possible return to the same values that were registered at the end of the first week; this could indicate a possible similarity in the learning curves of the two different versions of the system. While the improved version of the system was considered more effective and appropriate by the teacher, it can be hypothesized that the same decline seen during the first week, can be observed once the participants grow accustomed to the new version.

6.2 Method

The study was designed as between-groups, and was conducted as such. The final meeting with the participants' supervisor highlighted the fact that participants were in fact not divided in randomized groups, but rather purposefully divided. Teachers assigned to the treatment group pupils that were deemed to be gaining the most from the VR training. This highlights a possible lack in validity of between-groups measurements, primarily because the assumption of homogeneity is violated. Having nine participants, an equal amount of shopping lists were created for the baseline and evaluation sessions. The lists were assigned randomly to the participants and were then circulated around the second time they were used. All lists were structured so to allow participants to not hinder each other in the search, however the lists might not have been equally complex causing disparity in completion times. As a way of solving this problem, the assumed distance travelled was logged for each shopping list, and has been used as predictors in the regression, as presented in the previous paragraphs. Optimally, all combinations of participants and shopping lists would have to be tested so to have a proper performance comparison. Alternatively, all participants would have to undergo the test one at a time following the same shopping list; this was not deemed suitable for this study, due to time constraints imposed by the project framework. Another flaw that has been detected in the structure of the test concerns carry over effects. A minimum of 24-48 hours post-test rest period is suggested [41], but due to both the school schedule and the system development period, only a restricted amount of test days could be agreed upon and they were only in consecutive order. This strongly delimits the assumptions that can be made regarding eventual difference that can be found in the data gathered during the final evaluation. Data collected throughout the training might be unreliable because of the insecurity as whether the effect presented was due to recent training sessions or proper retention. On top of this, the study was conducted for a total of seven days over two weeks, as of the conclusion of the study it cannot be possible to define whether any effect measured will last; optimally a more extensive training period should be structured, and further evaluations conducted. Lastly, some of the data gathering methods lacked specificity and effectiveness. The observation web-app that has been developed for baseline and final evaluations, was aimed to track participants behaviour in the supermarket, and these results were not properly achieved for a series of reasons. Firstly the application had not been thoroughly tested on all devices that would have had to run it, thus some bugs and usability issues had been found while conducting the observations. Secondly, with hindsight, the purpose of the app was too general; the expected data logged was significantly different than the effective data collected. The app was ideally aimed to be able to log all participants' interactions in the supermarket, so to be able to restructure their behaviour and assess their performance. In reality, the data collected is limited to observations, and parts of participants' behaviour that resulted mostly unusable.

6.3 Development

The study has been cumbersome from a developmental point of view; the scale of the environment and the functionalities that had to be implemented have been time consuming and at times challenging. In spite of having developed the prototype in close contact with one of the teachers, the system resulted to be somewhat unsatisfactory the more it was used by the participants. The ease of access to the hints trigger can be deemed responsible for this; they quickly learned how to skip hints so to instantly trigger the compass. In order to avoid this, the hint system was disabled completely between the two test weeks, if more time had been available, a time-based hint system could have been implemented, so to have participants wait a pre-determined amount of seconds before they could access the following hint. Additionally, a significant mistake was made when changing the hint system: the line of code that triggered the log for the "hint asked" event was not included in the new version, thus hindering part of the data analysis.

6.4 Participant D

As reported by Claes Kiilsgaard, participant **D** has always had low confidence when visiting public environments. The participant was observed to need help when going to the gym, or to the supermarket; **D** never wanted to go alone. This was confirmed during the baseline evaluation, when **D** was observed to explore the supermarket together with a pedagogue; said pedagogue confirmed that the participant explicitly asked them to be accompanied. Additionally, **D** was observed to show discomfort throughout the baseline evaluation. This pattern in the participant behaviour was broken during the final session. Observations and teachers notes reported that the participant decided to approach the supermarket exploration alone. This result can be considered extremely positive, also considering that **D**'s effectiveness was 100%. The participant has been observed to be able to overcome obstacles without the help of teachers; he has only been noted to ask another participant for help in finding one product. Participant **D**'s positive development can also be observed during the VR training period; satisfaction and confidence levels raised to very positive values by the last training session.

6.5 Addressing the research questions

> *Can a VR training simulator instill comfort and confidence around a real shopping experience?*

To evaluate whether or not confidence was achieved we applied a direct subjective measure both after the baseline assessment and the post-treatment. The method was inspired by [65], who utilized similar pictorial questionnaires to evaluate the participants own impression of confidence. Furthermore, we continuously evaluated confidence during VR sessions using the same measure of confidence and prospected confidence. Lastly, due to the fact that people are rarely the best evaluators of their own actions, we asked Claes Kiilsgaard to evaluate, subjectively, how he estimated their development during the training-process.

As previously mentioned, judging by the participants own opinion, we saw a small decline in self-reported confidence for the treatment group, and a small increase for the increase. Although this could be attributed to the change situation in the assessments, it is a result that recurring in the self-reported confidence measure from the VR-sessions (fig. 5.7(c)) where we saw a decease in rating during the last two sessions. This could be due to high susceptibility to changes in difficulty, which both had a high influence on the VR Task Completion Time (5.3, effectiveness 5.4 their self-reported ease-of-use 5.7(a). Furthermore, when looking at Claes' observation-diary 5.8 he furthermore reported a decline in confidence.

Although cannot conclusively say that the VR instilled confidence, we cannot say that it did not either. It was observed that the Treatment Group's prospected confidence increased slightly in the post-treatment session, but since we only have a sample size of 8 participants for comparison, any variation in one participants score will greatly affect the overall score and therefor should be regarded as highly unpredictable, and the margin of error will be equally high. However, it was interesting to observe how susceptible to variations in the level of difficulty actually affected the participants sense of confidence and overall performance.

Can skills taught in a virtual supermarket be transferred to a real context?

Aim of the repeated usage of the system was to allow low road transfer [35] of adaptive shopping skills from VR to real shopping experience. This was achieved through the repetitions of product-search processes in the safety of the simulated environment [27], and stripping the environment from additional obstructive and indimidating aspects, allowed participants to focus on the improvement of their skills [31].

The Difference in Difference analysis (table 5.3), as discussed previously in this chapter, showed significant difference in the performance of treatment over control groups. *Effectiveness* and *self-reported help requrired* have played a fundamental role in this analysis; treatment group had constant outstanding effectiveness performance in both baseline and final evaluations, although the significant difference in environmental variables. In spite of this, participants self-reported lower comfort and confidence results at the end of the final evaluation compared to the baseline. A peculiar contradiction to this is participant **D**'s behaviour, which was outstandingly different between baseline and final evaluation (i.e. approached the shopping alone). Based on this analysis, we confidently confirm the possibility of near transfer from the virtual to the real experiences, although it might be covert to the participants.

Lastly, with this study has not been possible to asses participants' retention of the newly transferred skills, and it could be interesting to assess whether results improve with a longer rest period between training sessions, and follow-up on the maintenance of such skills.

To what degree can a VR simulation facilitate shopping skills in adolescents with autism?

Finally, the applied generalization-method of programming common stimuli seemed straightforward enough, and in compliance with existing theories of near transfer of learning and low road transfer. However, the notion that repetition is usually appreciated by people with ASD ultimately did not apply in these circumstances. The self-reported satisfaction declined 35% during the experiment, and would likely have continued to decline. Participants simply were not stimulated enough, and the lack of motivation could ultimately cause negative transfer of training. During the final evaluation session, Claes remarked about this, and suggested that some elements to keep the children motivated should be implemented if they were to continue training in VR; individual competitive scores could motivate them to perform better over time, trying to improve their previous scores.

These suggestions of gamified elements could possibly be achieved with another generalization method i.e. introduction to natural maintaining contingencies(2.3.1), where transfer of training is facilitated by rewarding the participants within their environments.

Chapter 7

Conclusion

The study presented in this report set out to investigate the possibility of application of virtual reality systems to behaviour and adaptive skill training in ASD children and young adults. The research firstly focused on the analysis and understanding of autism and its implications[3; 18]. Subsequently, traditional *analog* training techniques [4] were analyzed and juxtaposed to technology-based ASD treatment. The feasibility and effectiveness of virtual reality applied to such training has been assessed by several studies since the 1990s, and oftentimes positive effects have been reported [7; 45; 53].
The system was developed through an iterative process in close collaboration with teachers and pedagogues at Valøj Skole in Rødovre. The main goal was to augment their home economics classes and adaptive skill training sessions. In order to assess its efficacy, an in-between group experiment was conducted; the treatment group was required to undergo VR training for a total of seven session spread over two weeks. Measurements pre- and post-treatment, did not provide significant results, while between groups Difference in Difference evaluation reported significant difference.

This study finds that the change in Task Completion Time for the treatment group was significantly lower compared to the control group, hence the treatment group was almost 4 minute faster than the control group in the post-treatment assessment, when taking a number of other variables into account.There was a decline in the effectiveness in the control group between the baseline and post-treatment assessments, yet the treatment group maintained their score. This indicates that the treatment also had an effect on performance. The self-reported confidence of the treatment group declined in the post-treatment assessment, hence indicating, that the treatment did not increase the participants' confidence. Despite this, one of the participants, who usually is uncomfortable with shopping alone and refused to shop by himself in the baseline, accomplished the heavily populated post-treatment shopping assessment without assistance. This might not be a measurable significant result, but it indicates, that there might be effects on confidence, that the questionnaire did not capture.

All in all, the study found positive or indications of positive effects in all three measures (time, effectiveness, and confidence). The results are short term and do not provide evidence of potential long term effects. Furthermore the small sample size and the fact that the treatment and control groups were not randomly selected, calls for further and more specific investigation.

Bibliography

[1] B. Levinson (Director), “Rain man,” 1988.

[2] “Autism spectrum disorder.” https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd/index.shtml. Accessed: May 30th, 2017.

[3] K. Francis, “Autism interventions: a critical update,” *Developmental Medicine & Child Neurology*, vol. 47, no. 07, pp. 493–499, 2005.

[4] T. F. Stokes and D. M. Baer, “An implicit technology of generalization,” *Journal of applied behavior analysis*, vol. 10, no. 2, pp. 349–367, 1977.

[5] O. Grynszpan, P. L. Weiss, F. Perez-Diaz, and E. Gal, “Innovative technology-based interventions for autism spectrum disorders: a meta-analysis,” *Autism*, vol. 18, no. 4, pp. 346–361, 2014.

[6] B. O. Ploog, A. Scharf, D. Nelson, and P. J. Brooks, “Use of computer-assisted technologies (cat) to enhance social, communicative, and language development in children with autism spectrum disorders,” *Journal of autism and developmental disorders*, vol. 43, no. 2, pp. 301–322, 2013.

[7] D. Strickland, L. M. Marcus, G. B. Mesibov, and K. Hogan, “Brief report: Two case studies using virtual reality as a learning tool for autistic children,” *Journal of Autism and Developmental Disorders*, vol. 26, no. 6, pp. 651–659, 1996.

[8] D. Strickland, “Virtual reality for the treatment of autism,” *Studies in health technology and informatics*, pp. 81–86, 1997.

[9] N. R. Council *et al.*, *Educating children with autism.* National Academies Press, 2001.

[10] J. Berg, “Data in public health,” 2017.

[11] J. H. Williams, D. W. Massaro, N. J. Peel, A. Bosseler, and T. Suddendorf, "Visual–auditory integration during speech imitation in autism," *Research in developmental disabilities*, vol. 25, no. 6, pp. 559–575, 2004.

[12] L. Luke, I. C. Clare, H. Ring, M. Redley, and P. Watson, "Decision-making difficulties experienced by adults with autism spectrum conditions," *Autism*, vol. 16, no. 6, pp. 612–621, 2012.

[13] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, *Cognitive Neuroscience: The Biology of the Mind.* W.W. Norton & Company, Inc., 3 ed., 2009.

[14] S. Baron-Cohen, "Precursors to a theory of mind: Understanding attention in others," 1991.

[15] J. W. Astington and M. J. Edward, "The development of theory of mind in early childhood," *Social Cognition in Infancy*, vol. 5, p. 16, 2010.

[16] S. Baron-Cohen and M. K. Belmonte, "Autism: a window onto the development of the social and the analytic brain," *Annu. Rev. Neurosci.*, vol. 28, pp. 109–126, 2005.

[17] P. Howlin, "The effectiveness of interventions for children with autism," in *Neurodevelopmental disorders*, pp. 101–119, Springer, 2005.

[18] J. L. Matson, *Applied behavior analysis for children with autism spectrum disorders.* Springer, 2009.

[19] R. H. Horner, G. E. Dunlap, and R. L. Koegel, "Generalization and maintenance: Life-style changes in applied settings," in *Lifestyle Changes for Persons with Autism and Severe Handicaps*, Paul H. Brookes Publishing, 1988.

[20] J. L. Matson, M. L. Matson, and T. T. Rivet, "Social-skills treatments for children with autism spectrum disorders: an overview.," Sep 2007.

[21] G. Timler, *Train-and-Hope Strategy*, pp. 3150–3150. New York, NY: Springer New York, 2013.

[22] L. L. Mason, T. K. Jeon, P. Blair, and N. Glomb, "Virtual tutor training: Learning to teach in," *Design, Utilization, and Analysis of Simulations and Game-Based Educational Worlds*, p. 51, 2013.

[23] D. M. Baer, *How to plan for generalization.* Pro Ed, 1999.

[24] D. N. Perkins, G. Salomon, *et al.*, "Transfer of learning," *International encyclopedia of education*, vol. 2, pp. 6452–6457, 1992.

[25] M. R. Kandalaft, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman, "Virtual reality social cognition training for young adults with high-functioning autism," *Journal of autism and developmental disorders*, vol. 43, no. 1, pp. 34–44, 2013.

[26] P. Kortum, *HCI beyond the GUI: Design for haptic, speech, olfactory, and other nontraditional interfaces*. Morgan Kaufmann, 2008.

[27] C. Bossard, G. Kermarrec, C. Buche, and J. Tisseau, "Transfer of learning in virtual environments: a new challenge?," *Virtual Reality*, vol. 12, no. 3, pp. 151–161, 2008.

[28] M. Gellevij, H. Van Der Meij, T. De Jong, and J. Pieters, "Multimodal versus unimodal instruction in a complex learning context," *The Journal of Experimental Education*, vol. 70, no. 3, pp. 215–239, 2002.

[29] S. Oviatt, "User-centered modeling and evaluation of multimodal interfaces," *Proceedings of the IEEE*, vol. 91, no. 9, pp. 1457–1468, 2003.

[30] S. Oviatt, R. Coulston, and R. Lunsford, "When do we interact multimodally?: cognitive load and multimodal communication patterns," in *Proceedings of the 6th international conference on Multimodal interfaces*, pp. 129–136, ACM, 2004.

[31] R. Moreno and R. E. Mayer, "Cognitive principles of multimedia learning: The role of modality and contiguity.," *Journal of educational psychology*, vol. 91, no. 2, p. 358, 1999.

[32] P. N. Wilson, N. Foreman, and M. Tlauka, "Transfer of spatial information from a virtual to a real environment in physically disabled children," *Disability and rehabilitation*, vol. 18, no. 12, pp. 633–637, 1996.

[33] F. Rose, E. Attree, B. Brooks, D. Parslow, P. Penn, and N. Ambihaipahan, "Transfer of training from virtual to real environments," in *2nd European Conference on Disability, Virtual Reality and Associated Technologies*, pp. 69–75, 1998.

[34] J. Cromby, P. Standen, J. Newman, and H. Tasker, "Successful transfer to the real world of skills practised in a virtual environment by students with severe learning difficulties," in *Proceedings*, 1996.

[35] D. N. Perkins and G. Salomon, "Teaching for transfer," *Educational leadership*, vol. 46, no. 1, pp. 22–32, 1988.

[36] M. H. Ashcraft, *Cognition.* Upper Saddle River, New Jersey, 07458: Person Education Inc., 4 ed., 2006.

[37] H. Ebbinghaus, *Über das gedächtnis: untersuchungen zur experimentellen psychologie.* Duncker & Humblot, 1885.

[38] J. J. Shaughnessy, "Long-term retention and the spacing effect in free-recall and frequency judgments," *The American Journal of Psychology*, pp. 587–598, 1977.

[39] F. Hillary, M. Schultheis, B. Challis, S. Millis, G. Carnevale, T. Galshi, and J. DeLuca, "Spacing of repetitions improves learning and memory after moderate and severe tbi," *Journal of Clinical and Experimental Neuropsychology*, vol. 25, no. 1, pp. 49–58, 2003.

[40] T. Smith, "Improving memory to promote maintenance of treatment gains in children with autism," *The Psychological Record*, vol. 44, no. 4, p. 459, 1994.

[41] G. Wallet, H. Sauzéon, J. Rodrigues, F. Larrue, and B. N'Kaoua, "Virtual/real transfer of spatial learning: Impact of activity according to the retention delay," *Stud Health Technol Inform*, vol. 154, pp. 145–149, 2010.

[42] C. K. Nikopoulos and M. Keenan, "Using video modeling to teach complex social sequences to children with autism," *Journal of Autism and Developmental Disorders*, vol. 37, no. 4, pp. 678–693, 2007.

[43] S. Shukla-Mehta, T. Miller, and K. J. Callahan, "Evaluating the effectiveness of video instruction on social and communication skills training for children with autism spectrum disorders: A review of the literature," *Focus on Autism and Other Developmental Disabilities*, vol. 25, no. 1, pp. 23–36, 2010.

[44] M. Maskey, J. Lowry, J. Rodgers, H. McConachie, and J. R. Parr, "Reducing specific phobia/fear in young people with autism spectrum disorders (asds) through a virtual reality environment intervention," *PloS one*, vol. 9, no. 7, p. e100374, 2014.

[45] M. J. Smith, E. J. Ginger, K. Wright, M. A. Wright, J. L. Taylor, L. B. Humm, D. E. Olsen, M. D. Bell, and M. F. Fleming, "Virtual reality job interview training in adults with autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 44, no. 10, pp. 2450–2463, 2014.

[46] S. Parsons and P. Mitchell, "The potential of virtual reality in social skills training for people with autistic spectrum disorders," *Journal of Intellectual Disability Research*, vol. 46, no. 5, pp. 430–443, 2002.

[47] F. Ke and T. Im, "Virtual-reality-based social interaction training for children with high-functioning autism," *The Journal of Educational Research*, vol. 106, no. 6, pp. 441–461, 2013.

[48] H. H. Ip, S. W. Wong, D. F. Chan, J. Byrne, C. Li, V. S. Yuan, K. S. Lau, and J. Y. Wong, "Virtual reality enabled training for social adaptation in inclusive education settings for school-aged children with autism spectrum disorder (asd)," in *International Conference on Blending Learning*, pp. 94–102, Springer, 2016.

[49] M. F. Casanova, A. El-Baz, and J. S. Suri, *Autism Imaging and Devices*. CRC Press, 2016.

[50] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[51] D. Bowman, A. Datey, U. Farooq, Y. Ryu, and O. Vasnaik, "Empirical comparisons of virtual environment displays," 2001.

[52] D. A. Bowman, A. Datey, Y. S. Ryu, U. Farooq, and O. Vasnaik, "Empirical comparison of human behavior and performance with different display devices for virtual environments," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 46, pp. 2134–2138, SAGE Publications Sage CA: Los Angeles, CA, 2002.

[53] C. Mei, L. Mason, and J. Quarles, "Usability issues with 3d user interfaces for adolescents with high functioning autism," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pp. 99–106, ACM, 2014.

[54] E. Klinger, P. L. Weiss, and P.-A. Joseph, "Virtual reality for learning and rehabilitation," in *Rethinking physical and rehabilitation medicine*, pp. 203–221, Springer, 2010.

[55] P. R. Alcantara, "Effects of videotape instructional package on purchasing skills of children with autism," *Exceptional Children*, vol. 61, no. 1, pp. 40–55, 1994.

[56] S. Bellini, L. Gardner, R. Hudock, Y. Kashima-Ellingson, and C. P. Schools, "The use of video self-modeling and peer training to increase social engagement in preschool children on the autism spectrum.," in *School Psychology Forum*, vol. 10, 2016.

[57] N. J. Minshew, G. Goldstein, and D. J. Siegel, "Neuropsychologic functioning in autism: Profile of a complex information processing disorder," *Journal of the International Neuropsychological Society*, vol. 3, no. 04, pp. 303–316, 1997.

[58] M.-J. Caron, L. Mottron, C. Rainville, and S. Chouinard, "Do high functioning persons with autism present superior spatial abilities?," *Neuropsychologia*, vol. 42, no. 4, pp. 467–481, 2004.

[59] A. Crippa, S. Forti, P. Perego, and M. Molteni, "Eye-hand coordination in children with high functioning autism and asperger's disorder using a gap-overlap paradigm," *Journal of autism and developmental disorders*, vol. 43, no. 4, pp. 841–850, 2013.

[60] C. M. Falter, K. C. Plaisted, and G. Davis, "Visuo-spatial processing in autism—testing the predictions of extreme male brain theory," *Journal of autism and developmental disorders*, vol. 38, no. 3, pp. 507–515, 2008.

[61] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey, "Locomotion in virtual reality for individuals with autism spectrum disorder," in *Proceedings of the 2016 Symposium on Spatial User Interaction*, pp. 33–42, ACM, 2016.

[62] "Gruppeordninger på valhøj skole." `http://valhoj-skole.skoleporten.dk/sp/28384/file/Inpage/ed4a9e3f-5f6f-4cce-b0a1-638b4f54304a`. Accessed: February 16th, 2017.

[63] "Inklusion i rødovre kommunes skoler." `https://www.rk.dk/index.php?eID=tx_rkdoc_attachment&uid=229139&filename=/1205117-1471362-1.PDF`. Accessed: March 9th, 2017.

[64] M. Schmidt, S. Schmidt, J. L. Sandegaard, V. Ehrenstein, L. Pedersen, and H. T. Sørensen, "The danish national patient registry: a review of content, data quality, and research potential," *Clin Epidemiol*, vol. 7, no. 449, p. e490, 2015.

[65] L. Hall, C. Hume, and S. Tazzyman, "Five degrees of happiness: Effective smiley face likert scales for evaluating with children," in *Proceedings of the The 15th International Conference on Interaction Design and Children*, pp. 311–321, ACM, 2016.

[66] R. E. Ferdig, *Design, Utilization, and Analysis of Simulations and Game-based Educational Worlds*. IGI Global, 2013.

[67] J. J. LaViola Jr, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.

[68] Y. Yang, S. Liu, S. A. Chowdhury, G. C. DeAngelis, and D. E. Angelaki, "Binocular disparity tuning and visual–vestibular congruency of multisensory neurons in macaque parietal cortex," *The Journal of Neuroscience*, vol. 31, no. 49, pp. 17905–17916, 2011.

[69] H. Akiduki, S. Nishiike, H. Watanabe, K. Matsuoka, T. Kubo, and N. Takeda, "Visual-vestibular conflict induced by virtual reality in humans," *Neuroscience letters*, vol. 340, no. 3, pp. 197–200, 2003.

[70] C. Hänel, B. Weyers, B. Hentschel, and T. W. Kuhlen, "Interactive volume rendering for immersive virtual environments," in *3DVis (3DVis), 2014 IEEE VIS International Workshop on*, pp. 73–74, IEEE, 2014.

[71] J. Birn, *Digital lighting and rendering*. Pearson Education, third ed., 2013.

[72] J. Lazar, J. H. Feng, and H. Hochheiser, *Research methods in human-computer interaction*. John Wiley & Sons, 2010.

[73] J. Sauro and J. R. Lewis, *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2016.

[74] O. C. Ashenfelter and D. Card, "Using the longitudinal structure of earnings to estimate the effect of training programs," 1984.

[75] T. Jokela, N. Iivari, J. Matero, and M. Karukka, "The standard of user-centered design and the standard definition of usability: analyzing iso 13407 against iso 9241-11," in *Proceedings of the Latin American conference on Human-computer interaction*, pp. 53–60, ACM, 2003.

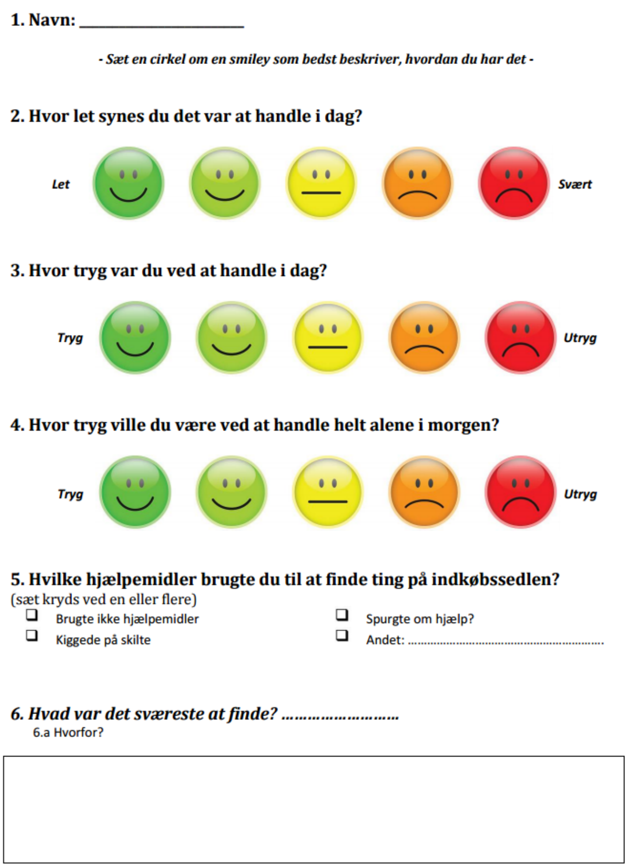
[76] U. M. Schaller and R. Rauh, "What difference does it make? implicit, explicit and complex social cognition in autism spectrum disorders," *Journal of Autism and Developmental Disorders*, pp. 1–19, 2017.

[77] C. Mei, L. Mason, and J. Quarles, "How 3d virtual humans built by adolescents with asd affect their 3d interactions," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, pp. 155–162, ACM, 2015.

Appendix A

Questionnaires

A.1 Baseline and final sessions



1. Navn: ___________________

- Sæt en cirkel om en smiley som bedst beskriver, hvordan du har det -

2. Hvor let synes du det var at handle i dag?

Let … *Svært*

3. Hvor tryg var du ved at handle i dag?

Tryg … *Utryg*

4. Hvor tryg ville du være ved at handle helt alene i morgen?

Tryg … *Utryg*

5. Hvilke hjælpemidler brugte du til at finde ting på indkøbssedlen?
(sæt kryds ved en eller flere)

- ❑ Brugte ikke hjælpemidler
- ❑ Kiggede på skilte
- ❑ Spurgte om hjælp?
- ❑ Andet: ……………………………………………………

6. Hvad var det sværeste at finde? ………………………

6.a Hvorfor?

A.2 Post-VR participants

Navn: *

Your answer

Hvor sjovt synes du det var at handle i det virtuelle supermarked i dag? *

	1	2	3	4	5	
Sjovt	○	○	○	○	○	Kedeligt

Hvor let synes du det var at handle i dag? *

	1	2	3	4	5	
Let	○	○	○	○	○	Svært

Hvor tryg var du ved at handle i dag? *

	1	2	3	4	5	
Tryg	○	○	○	○	○	Utryg

Hvor tryg ville du være ved at handle alene i den rigtige Føtex i morgen? *

	1	2	3	4	5	
Tryg	○	○	○	○	○	Utryg

Har du nogle kommentarer?

Your answer

A.3 Post-VR supervisor

Dagens VR-session *

Choose

Elevens navn: *

Choose

Var eleven tilfreds med sin egen præstation?

	meget utilfreds	utilfreds	Hverken / eller (ved ikke)	tilfreds	meget tilfreds
Eleven var	○	○	○	○	○

Hvor selvsikker vurderer du at eleven er, i hans tilgang til systemet og oplevelsen? *

	meget usikker	usikker	hverken / eller (ved ikke)	selvsikker	meget selvsikker
Eleven var	○	○	○	○	○

Præstation ved specifikke elementer *

	Meget svært	Svært	Hverken/eller	Let	Meget let
Hvordan svært/let var det for eleven at bruge teleportation?	○	○	○	○	○
Hvor svært/let var det for eleven at bruge cues (spor)?	○	○	○	○	○
Hvor svært/let var det for eleven at finde rundt i VR-Føtex?	○	○	○	○	○

Var der nogle fejl eller mangler ved systemet der gjorde det svært for eleven at gennemføre det virtuelle indkøb?

Your answer

Nogle kommentarer fra eleven undervejs?

Your answer

Andet: *

Your answer

Appendix B

Shopping lists

Lists	Item 1	Item 2	Item 3	Item 4
1	1 liter minimælk	1 fl. solsikkeolie	2 kg kartofler	1 dåse hakkede tomater
2	2 kg hvedemel	1 fl soyasauce	1 pakke kærgården (smår)	1 pk cornflakes
3	1 pose sukker	1 pose pasta penne	6 tomater	1/2 liter fløde
4	Tebreve Earl grey	1 pakke digestive kiks	1 dåse majs	1 agurk
5	2 pakker knækbrød (groft)	1 dåse bagepulver	1 pose jasmin ris	1 liter minimælk
7	12 æg	1 dåse kokosmælk	1 pose rosiner	1 pose havregryn
8	1 pakke spaghetti	1 dåse tun	4 æbler	1 liter sødmælk
9	1 pakke bouillon terninger (grønsager)	1 pakke mellemlaget ost i skiver	1 pose løg	1 pose mandler ca. 200g

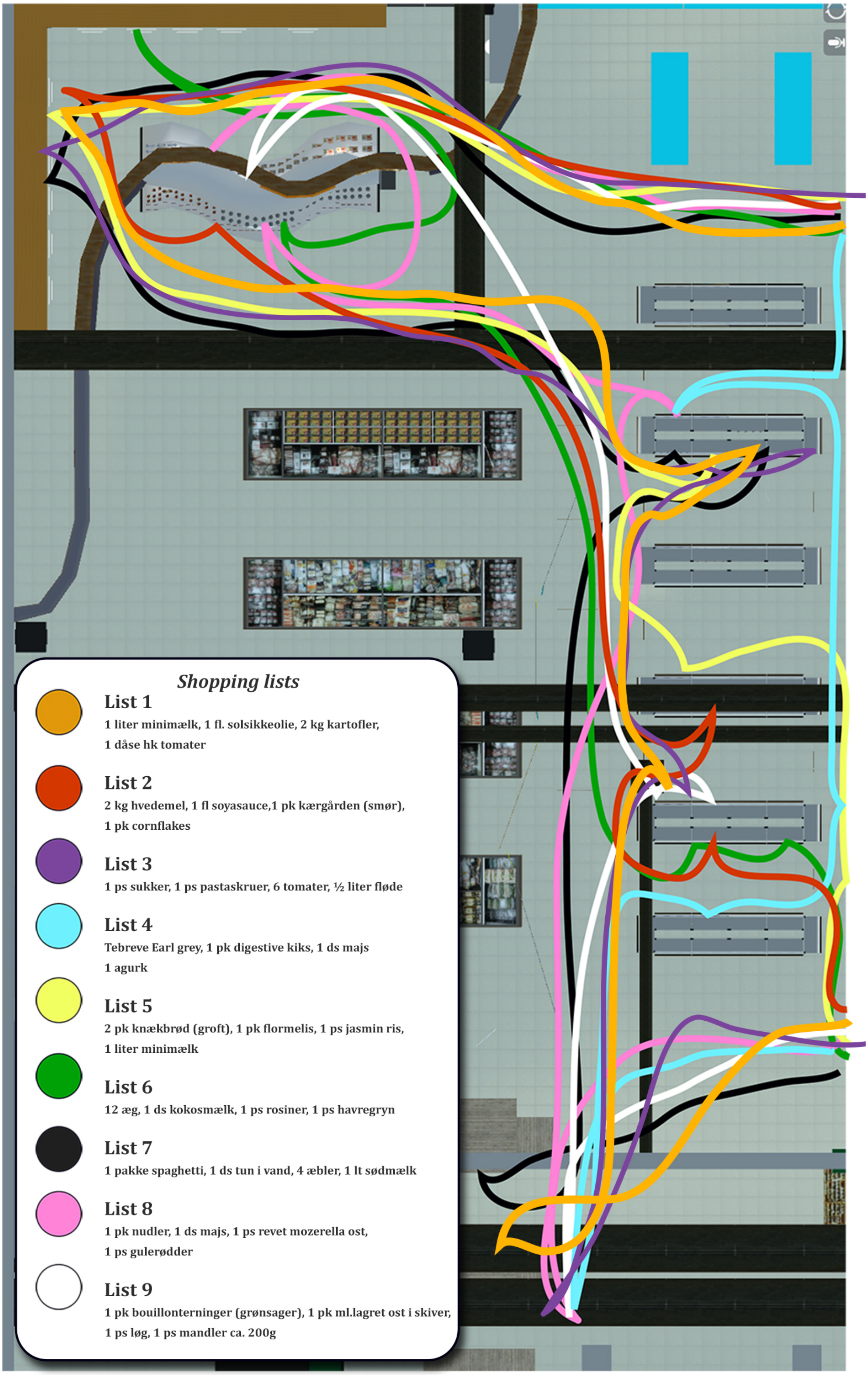


Figure B.1: Approximated traverse routes measured as length of vector path in Adobe Illustrator.

Appendix C

Scripts

```
public class outputAllProducts : MonoBehaviour {
  void Start(){
    StreamWriter stream_writer = new StreamWriter (Application.dataPath + "/SetupFiles/products.txt", false);
    stream_writer.Write("!!!!!! WARNING !!!!!! \r\n");
    stream_writer.Write("This is an auto-generated file, modifying this will not affect the simulation in any way.\r\n");
    stream_writer.Write("Product_id, price, name\r\n");
    GameObject[] prods = GameObject.FindGameObjectsWithTag ("product");
    List<GameObject> products = new List<GameObject> ();
    foreach (GameObject prod in prods) {
      List<GameObject> results = products.FindAll(x => x.GetComponent<objectInfo>().id == prod.GetComponent<objectInfo>().id);
        if (results.Count == 0) {
      products.Add (prod);
        }
    }
    products.Sort (delegate(GameObject a, GameObject b) {
      return (a.GetComponent<objectInfo>().id).
          CompareTo(b.GetComponent<objectInfo>().id);
    });

    foreach (GameObject product in products) {
          objectInfo info = product.GetComponent<objectInfo> ();
          stream_writer.Write(info.id+" "+info.price+" "+info.name+"\r\n");
    }

    stream_writer.Close ();
  }
}
```

Snippet C.1: outputAllProducts.cs

```
public class objectInfo : MonoBehaviour {
  public int id;
  public string name;
  public float price;
  public int category;
  public bool stackable;
  private Vector3 original_position;
  private Quaternion original_rotation;

  void Start(){
    original_position = transform.position;
    original_rotation = transform.rotation;
  }

  void Update(){}

  void OnCollisionEnter(Collision col){
    if (col.gameObject.name == "Floor" || col.gameObject.name == "FallDetector") {
      if(col.gameObject.name == "FallDetector"){
        this.transform.SetParent(null);
        this.gameObject.layer = 0;
      }
      transform.position = original_position;
      transform.rotation = original_rotation;
      gameObject.GetComponent<Rigidbody>().velocity = Vector3.zero;
      gameObject.GetComponent<Rigidbody>().angularVelocity = Vector3.zero;
    }
  }

  void OnTriggerExit(Collider col){
    if(col.gameObject.name == "Shopping_basket"){
      this.transform.SetParent(null);
      this.gameObject.layer = 0;
    }
  }
}
```

Snippet C.2: objectInfo.cs

```
public class populateShelf1 : MonoBehaviour {
  public enum State
  {Fridge1x4 = 0,}
  public State ShelfType;
  private int type;
  [Range(0,100)]
  public int id;
  [Range(0,1)]
  public float padding;
  [Range(0,1)]
  public float productsDistanceX;
  [Range(0,1)]
  public float productsDistanceZ;
  [Range(0,2)]
  public float yOffset;
  [Range(-2,2)]
  public float xOffset;
  [Range(0,360)]
  public float offsetAngle;
  public float maxStackHeight;
  private string fileName;
  private List<Collider> colliders = new List<Collider> ();
  private List<GameObject> all_products = new List<GameObject> ();
  private Transform product_parent;
  private float collider_width;
  private float collider_depth;
  private bool inverted;
  Quaternion offsetRotation;

  void Start () {
    type = (int)ShelfType;
    fileName = type.ToString()+"_"+id.ToString()+".txt";
    product_parent = GameObject.Find ("product_duplicates").transform;
    GameObject[] all_temp_products = GameObject.FindGameObjectsWithTag ("product");
    foreach (GameObject temp_product in all_temp_products) {
      all_products.Add (temp_product);
    }
    GetColliders ();
    List<List<int>> products = loadList ();
    Fill (products);
  }

  private void GetColliders(){
    int children = gameObject.transform.childCount;
    for(int i =0; i <= 4; i++){
      Transform child = gameObject.transform.Find ("Shelf_"+i.ToString()).transform.Find("Shelf_body");
      colliders.Add (child.GetComponent<Collider> ());
    }
  }
```

```

  private void Fill(List<List<int>> products){
    for (int i = 0; i < colliders.Count; i++) {
      Collider collider = colliders [i];
      getSizeCollider (collider);
      Vector3 start_position = new Vector3 (
        collider.transform.position.x - collider_width/2 +
           productsDistanceX,
        collider.transform.position.y,
        collider.transform.position.z - collider_depth/2 +
           productsDistanceZ
      );
      Vector3 final_position = new Vector3 (
        collider.transform.position.x + collider_width/2 -
           productsDistanceX,
        collider.transform.position.y,
        collider.transform.position.z + collider_depth/2 -
           productsDistanceZ
      );
      Vector3 product_position = new Vector3 (
        start_position.x,
        start_position.y,
        start_position.z
      );
      List<int> row_list = products[i];
      List<float> final_positions = new List<float> ();
      float step;
      if (!inverted) {
        step = collider_width / row_list.Count;
      } else {
        step = collider_depth / row_list.Count;
      }
      foreach (int id in row_list) {
        final_positions.Add (step * (row_list.IndexOf (id) + 1));
      }
      foreach (int id in row_list) {
        GameObject product = getGameObjectById (id);
        if (!inverted) {
          while (product_position.x +
             product.GetComponent<Collider> ().bounds.size.x +
            productsDistanceX < start_position.x +
               final_positions [row_list.IndexOf (id)]) {
            while (product_position.z +
               product.GetComponent<Collider> ().bounds.size.z
               + padding < final_position.z) {
              int amount;
              if (product.GetComponent<objectInfo>
                 ().stackable) {
                amount = Mathf.FloorToInt(maxStackHeight /
                   product.GetComponent<Collider>
                   ().bounds.size.y);
              } else {
                amount = 1;
              }
```

```
                    for(int k = 0; k < amount; k++ ){
                      Instantiate (
                        product,
                        new Vector3 (
                          product_position.x + xOffset +
                              product.GetComponent<Collider>
                              ().bounds.size.x / 2,
                          product_position.y + yOffset +
                              (product.GetComponent<Collider>().bounds.size.y*k)
                              + (0.05f),
                          product_position.z +
                              product.GetComponent<Collider>
                              ().bounds.size.z / 2
                        ),
                        Quaternion.Euler (
                          offsetRotation.eulerAngles.x,
                          offsetRotation.eulerAngles.y + offsetAngle,
                          offsetRotation.eulerAngles.z
                        ),
                        product_parent
                      );
                    }
                  product_position.z +=
                      product.GetComponent<Collider> ().bounds.size.z
                      + productsDistanceZ;
                  }
                          product_position.z = start_position.z;
                      product_position.x +=
                          product.GetComponent<Collider>
                          ().bounds.size.x + productsDistanceX;
              }
            } else {
              while (product_position.z +
                  product.GetComponent<Collider> ().bounds.size.z +
                productsDistanceZ < start_position.z +
                    final_positions [row_list.IndexOf (id)]) {
                  while (product_position.x +
                      product.GetComponent<Collider> ().bounds.size.x
                      < final_position.x) {
                    int amount;
                    if (product.GetComponent<objectInfo>
                        ().stackable) {
                      amount = Mathf.FloorToInt(maxStackHeight /
                          product.GetComponent<Collider>
                          ().bounds.size.y);
                    } else {
                      amount = 1;
                    }
                    for(int k = 0; k < amount; k++ ){
                      Instantiate (
                        product,
                        new Vector3 (
                          product_position.x + xOffset +
                              product.GetComponent<Collider>
```

```
                            ().bounds.size.x / 2,
                          product_position.y + yOffset +
                            (product.GetComponent<Collider>().bounds.size.y*k)
                            + (0.05f),
                          product_position.z +
                            product.GetComponent<Collider>
                            ().bounds.size.z / 2
                        ),
                        Quaternion.Euler (
                          offsetRotation.eulerAngles.x,
                          offsetRotation.eulerAngles.y + offsetAngle,
                          offsetRotation.eulerAngles.z
                        ),
                        product_parent
                      );
                    }
                      product_position.x +=
                        product.GetComponent<Collider>
                        ().bounds.size.x + productsDistanceX;
                  }
                  product_position.x = start_position.x;
                  product_position.z +=
                    product.GetComponent<Collider> ().bounds.size.z
                    + productsDistanceZ;
              }
            }
          }
        }
      }
      private GameObject getGameObjectById(int id){
        foreach (GameObject product in all_products) {
          if (product.GetComponent<objectInfo> ().id == id) {
            return product;
          }
        }
        //print ("Fallback used");
        return GameObject.FindGameObjectWithTag
          ("return_error_fallback");
      }
      private List<List<int>> loadList(){
        FileInfo theSourceFile = new FileInfo (Application.dataPath
          + "/SetupFiles/Shelves/"+fileName);
        StreamReader reader = theSourceFile.OpenText();
        string text = reader.ReadLine();
        string[] rows = text.Split(';');
        List<string> rows_list = new List<string> (rows);
        List<List<string>> products_list = new List<List<string>> ();
        foreach (string row in rows_list) {
          string[] products = row.Split (',');
          List<string> temp = new List<string> ();
          foreach (string product in products) {
            temp.Add (product);
          }
          products_list.Add (temp);
```

```
    }
    List<List<int>> final_list = new List<List<int>> ();
    foreach (List<string> products_row in products_list) {
      List<int> temp_list = new List<int> ();
      foreach (string product in products_row) {
        int tmp;
        if (int.TryParse (product, out tmp)) {
          temp_list.Add (tmp);
        }
      }
        final_list.Add (temp_list);
    }
    return final_list;
  }
  private void getSizeCollider(Collider collider){
    if ( (int)gameObject.transform.eulerAngles.y == 90 ) {
      inverted = true;
      offsetRotation = Quaternion.Euler (0,
          -(int)gameObject.transform.eulerAngles.y, 0);
    } else if ( (int)gameObject.transform.eulerAngles.y == 270) {
      inverted = true;
      offsetRotation = Quaternion.Euler (0,
          (int)gameObject.transform.eulerAngles.y, 0);
    }  else {
      offsetRotation = Quaternion.identity;
    }
    collider_width = collider.bounds.size.x;
    collider_depth = collider.bounds.size.z;
  }
}
```

Snippet C.3: shelfPopulate1.cs

```
public class CompassBehavior : MonoBehaviour {
  [Tooltip("The target transform of the gameobject to point towards.")]
  public Transform target;
  private Vector3 point;
  [Tooltip("arrow rotation speed/s (x * Time.deltaTime)")]
  public float arrowRotationSpeed;
  private Vector3 targetDir;
  private Vector3 newDir;
  private Text distText;
  private float distance;
  private string distanceInMeter;
  // Use this for initialization
  void Start () {
    distText = GetComponentInChildren<Text>();
  }
  // Update is called once per frame
  void Update () {
    if(target != null && this.name == "Needle"){
      //distance = Vector3.Distance(
        target.transform.position,
        this.transform.position
      );
      float step = arrowRotationSpeed * Time.deltaTime;
      targetDir = target.position - transform.position;
      targetDir = new Vector3(targetDir.x,targetDir.y*0.0f,targetDir.z);
      newDir = Vector3.RotateTowards(transform.forward,targetDir, step, 0.0f);
      transform.rotation = Quaternion.LookRotation(newDir,Vector3.up);
    }
    if(this.name == "CompassCanvas"){
      point = Camera.main.transform.position;
      if(target != null){
        distance = (float)System.Math.Round(
          (double)Vector3.Distance(
          target.transform.position,
          this.transform.position
        ),0
      );
      distText.text = distance.ToString()+" m";
    } else {
      distText.text = "? m";
    }
    this.transform.LookAt(point,Vector3.up);
    }
  }
}
```

Snippet C.4: compassBehaviour.cs

Appendix D

Observations

D.1 Participant A

Baseline	**Final**
Arrived at a shelf	Started looking at a shelf
Started looking at a shelf	Product handled
Product handled	Product put back
Started looking at a shelf	Product handled
Note - Tuna	Product put in the basket - Oil sunflower - handled olive
Note - Can't find	Started looking at a shelf
Talk with others	Note - Nothing found
Product handled	Talk with others
Arrived at a shelf	Started looking at a shelf
Product handled	Product handled
Started looking at a shelf	Product handled
Started looking at a shelf	Product put in the basket - potatoes
Note - All at the same time	-

D.2 Participant B

Baseline	Final
Product put in the basket - Noodles	Arrived at a shelf
Used another kind of help - Worker asked	Product handled
Got distracted - Sound	Talk with others
Started looking at a shelf	Talk with others
Talk with others	Started looking at a shelf
Started looking at a shelf	Product put in the basket -
Talk with others	Started looking at a shelf
Product handled	Note - Crackers
Arrived at a shelf	Product handled
Product handled	Product put in the basket - Cucumber
Product put in the basket - Nudle	Note - No found
Arrived at a shelf	Started looking at a shelf
Product handled	Note - Looking for something he can't find
Product put in the basket - Carrot	Used another kind of help - Asked fotex man for sygar
Started looking at a shelf	-

D.3 Participant C

Baseline	Final
Used signs to orientate	Talk with others
Started looking at a shelf	Product handled
Product handled	Arrived at a shelf
Note - Look at product	Product handled
Product put back	Product put in the basket - Butter
Note - Was corn	Started looking at a shelf
Got distracted - Sound	Product handled

D.4 Participant D

Baseline	Final
Started looking at a shelf	Started looking at a shelf
Product handled	Product handled
Product put in the basket - Pepper	Product put in the basket - Spagh
Talk with others	Product handled
Used another kind of help - Random people	Talk with others
Started looking at a shelf	Note - Help with tuna don't know other color
Product handled	-
Note - No look at signs	-
Arrived at a shelf	-
Arrived at a shelf	-
Arrived at a shelf	-
Note - Checks date on product	-
Note - Found piskefløde meget hurtigt og gik videre	-
Note - Read the expiribg date on piskefløde	-

D.5 Participant F

Baseline	Final
Arrived at a shelf	Started looking at a shelf
Talk with others	Note - Nothing found
Used another kind of help - Teacher	Started looking at a shelf
Product handled	Note - Nothing. Found two shelves
Note - Shouts	Started looking at a shelf
Product handled	Product put in the basket - Soya
Talk with others	Note - Walking aimlessly
Note - Green looked for butter and used help from skolelærer	-
Used another kind of help - Looked and read whats on the products	-
Used another kind of help - Teacher	-
Product handled	-
Product put in the basket - Green	-
Product handled	-
Product put in the basket - Flour	-
Product handled	-
Talk with others	-
Used another kind of help -	-

D.6 Participant E

Baseline	Final
Talk with others	Note - Prev dark blue
Product put in the basket -	Note - light green pock object from
-	Started looking at a shelf

D.7 Participant G

Baseline	Final
Arrived at a shelf	Product handled
Talk with others	-

D.8 Participant H

Baseline	Final
Started looking at a shelf	Product handled
Used signs to orientate	Talk with others
Started looking at a shelf	Talk with others
Product handled	-
Talk with others	-
Talk with others	-

D.9 Participant I

Baseline	Final
Arrived at a shelf	Used another kind of help - asked
Product handled	Talk with others
Talk with others	Started looking at a shelf
Note - Lost	Note - My user nine has no color
Used another kind of help - Random people	Product handled
Product put in the basket - Flour	Product put back
Talk with others	Used another kind of help - Asked
Used another kind of help - Worker	Product put in the basket - Pasta
Note - Fotex guy helptin	-
Started looking at a shelf	-
Product put in the basket - Apple	-
Note - Lots of tuna can't find	-
Note - Tired to tomato sauce	-
Started looking at a shelf	-
Started looking at a shelf	-
Product handled	-
Product put in the basket - Mais	-

D.10 Additional notes

D.10.1 Baseline

- Everybody was very fast (in general)
- 2 or 3 kids going together - many were talking to each other
- Red (**I**) approached the shelf three times, the first two he went away, the third found the item he searched for
- Light blue (**B**) looked at expiration date
- Orange (**A**) looked at each apple before taking them
- Girl+2 boys asked for help. The girl was asking more than anybody else
- One might have kept the list in the pocket?
- Similar route for everybody (shop structure?)
- Looking for specific brands (maybe what they know) - ignored other brands even if they were right in front of them

D.10.2 Final

- Way more crowded than the first time
- One of the children *(control)* asked for help to another *(treatment)*

Appendix E

Teacher VR notes

E.1 Errors, mistakes, system faults

Errors, mistakes, and missing things in the system that would have made it easier for the student to perform virtual shopping
Var der nogle fejl eller mangler ved systemet der gjorde det svært for eleven at gennemføre det virtuelle indkøb?

- **A** - It was not easy to go from "look at signs" to the "compass".
- **B** - Hard to go from "look at signs" to compass.
- **C** - Hard to go from "look at signs to compass, but the compass was extremely helpful and easy for the students to use. I havent seen if the signs are blinking where they need to go, or all of them.
- **D** - Hard to go from [...]. but **D** was faster than the others to find his way with the compass.
- **A** - Generally it's weird that you can teleport in top of shelves.
- **A** - I experienced that **A** had a hard time with the handle to the shopping list. he pressed teleport instead of taking groceries. He ended on top of shelf.

E.2 Teacher's VR notes

- **B** - found the compass by double clicking on the cue-button. He found the "shortcut" himself.
- **D** - told me that he did not want to go to Føtex due to the people (other shoppers. I think the VR experience should be expanded to also include a variety of other challenges, like people, noise etc. Before **D** would develop skills to go to the real Føtex.

- **A** - Even though I explained to double click to get the compass out, **A** was not good at it. He moved around in circles and did not reckognize which shelves contained what.
- **B** - The pupil is starting to get bored in the "game". there's missing development through challenges. I try next week to ask him not to use the compass.
- **D** - It's starting to get boring. I ask to find stuff net week without compass.
- **C** - Missing development/challenge in the "Game".
- **B** - looked as signs to navigation.
- **C** - I noticed that none of the first three students threw objects around. Maybe it was more realistic with sound.
- **D** - He tossed the products a little bit around today as the only participant.
- **D** - Got help finding coconut milk.
- **C** - missed the coconut milk, but got help when he asked an employee (me) I answered it was next to the soy sauce.
- **B** - also needed help finding coconut milk
- **A** - Coconut milk is not used alot in Rødovre.
- **D** - asked for the compass, but I asked he should look at the shelves instead.
- **A** - he was not here today
- **B** - Asked for compass, but I said to look at the shelves. He found tuna himself.
- **A** - I explained before the test that *rød karry* is the same as *red curry.* I don't believe the danish name is on the label.
- He used compass to find red curry.

E.3 Teacher reported comments from the students

- **A** - It was fun, and I felt like being in the Føtex for real
- **B** - I was happy to know where the butter was already
- **C** - First time he tried VR, but he was very good at it
- **D** - he did not understand why he couldn't take butter before flour
- **A** - He seemed satisfied with his performance
- **B** - It was easier today
- **C** - It was fun to throw stuff around, but he knows he's not allowed to do that in the real føtex
- **A** - I cannot get the compass out
- **B** - He just wanted to throw stuff around for a little while
- **A** - It was more realistic with sound, and he used signs to find the groceries
- **B** - Sound was more realistic, but in the real Føtex there are no bip bip sounds in this end
- **D** - He would have a hard time in the real Føtex due to the people. The sound was a bit stressing, but not as much as real people.
- **C** - It's starting to get boring

Appendix F

Virtual Reality Evaluation (PCA)

Velfærdsteknologievaluering The table on the following page presents the "Velfærdsteknologievaluering" that was submitted by Claes Kiilsgaard to the Chief Purchasing Consultant at Rødovre Kommune

Velfærdsteknologivurdering

Projekt Virtual Reality i PCA. (indkøb 2. maj 2017)

Rødovre Kommune

Område:	Kategori:	Spørgsmål: (Svar: Ja=2, Delvist=1, Nej=0)	Svar (2,1,0)	Uddybning/kommentar:
Teknologi	Funktionalitet	Kan teknologien det, som det forventes, at den skal kunne?	**2**	
		Fungerer teknologien hver gang den anvendes?	**2**	
	Brugervenlighed	Er teknologien let og intuitiv at anvende?	**1**	De to joystick kunne integreres bedre i hænderne
		Er manualer gode/instruktionen været god?	**2**	
Økonomi	Investering	Kan teknologien benyttes uden ressourcetung kompetence-udvikling?	**2**	
		Kan teknologien benyttes uden ombygning/indkøb af andre produkter?	**2**	
	Drift	Frigiver teknologien medarbejderressourcer? (f.eks. til andre opgaver)	**0**	det kræver meget personaletid idet der kun kan undervises en elev ad gangen.
		Erstatter teknologien andre tilbud eller anden teknologi?	**0**	
Borger	Værdi	Er borgerne mere selvhjulpen med brugen af teknologien?	**2**	
		Er teknologien årsag til bedre kvalitet? (f.eks. livskvalitet, værdighed, integritet, inklusion m.v.)	**2**	
	Anvende-lighed	Opfylder teknologien et behov hos borgeren?	**1**	
		Ønsker borgerne at anvende teknologien?	**1**	
Organi-sation	Ledelse	Er der ledelsesmæssig opbakning til brug af teknologien?	**2**	
		Understøtter teknologien gode arbejdsgange?	**1**	
	Personale	Er teknologien årsag til et forbedret arbejdsmiljø?	**1**	
		Er der motivation for at bruge teknologien?	**2**	

Appendix G

Virtual Appendix

G.1 Url

https://goo.gl/NRWuNs

G.2 Content

1. Report
2. Data
    2.1 Shopping Questionnaires
    2.2 Observation Data
    2.3 VR Questionnaires
    2.4 Observer Diary
    2.5 Data Logging
3. Digital Appendix
    3.1 Meetings
4. Application (Desktop)
5. Application (VR)