VRShop: A Mobile Interactive Virtual Reality Shopping Environment Combining the Benefits of On- and Offline Shopping

MARCO SPEICHER, German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus
SEBASTIAN CUCERCA, German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus
ANTONIO KRÜGER, German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus

In this work, we explored the main characteristics of on- and offline shops with regard to customer shopping behavior and frequency. Thus, we designed and implemented an immersive virtual reality (VR) online shopping environment. We tried to maintain the benefits of online shops, like search functionality and availability, while simultaneously focusing on shopping experience and immersion. By touching the third dimension, VR provides a more advanced form of visualization, which can increase the customer's satisfaction and thus shopping experience. We further introduced the Virtual Reality Shopping Experience (VRSE) model based on customer satisfaction, task performance and user preference. A case study of a first VR shop prototype was conducted and evaluated with respect to the VRSE model. The results showed that the usability and user experience of our system is above average overall. In summary, searching for a product in a WebVR online shop using speech input in combination with VR output proved to be the best regarding user performance (speed, error rate) and preference (usability, user experience, immersion, motion sickness).

CCS Concepts: • **Human-centered computing → User centered design**;

Additional Key Words and Phrases: Virtual reality, customer satisfaction, shopping experience, virtual environment, 3D user interfaces, user-centered design.

ACM Reference format:
Marco Speicher, Sebastian Cucerca, and Antonio Krüger. 2017. VRShop: A Mobile Interactive Virtual Reality Shopping Environment Combining the Benefits of On- and Offline Shopping. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 102 (September 2017), 31 pages.
https://doi.org/10.1145/3130967

1 INTRODUCTION

Current online shops may be functional and efficient, but do not provide enough of an immersive shopping experience [Buffa and Lafon 2000]. This paper focuses on developing an immersive virtual reality (VR) online shopping environment that includes the major advantages of offline and online shopping (see Figure 5). The main goal of the paper was to do the next step in investigating interaction in mobile VR shopping environments

This research was funded in part by the German Federal Ministry of Education and Research under grant number 01IS12050 (project VRShop). The responsibility for this publication lies with the authors.
Authors' addresses: M. Speicher and S. Cucerca and A. Krüger, German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus, Saarbrücken, Germany.

2474-9567/2017/9-ART102 $15
https://doi.org/10.1145/3130967

Fig. 1. This figures shows a screenshot of our Virtual Reality Online Shopping Environment.

by taking a closer look on the influence of input (head pointing, speech) and output (desktop, HMD) on task performance and user's preference and behavior. In order to provide a realistic setting, we evaluated two common hands-free interaction techniques using smartphone VR in a comparative user study.

Technological changes, in the form of online shops, pop-up stores and digitization in general, have not only brought economic benefits to the retail sector. There is also a change in strategy underway, in which retailers will put more and more emphasis on satisfying the customer. Therefore, a lot has been invested in research to improve performance and usability of online shops' user interfaces. Nevertheless, it is equally important for the effectiveness of such user interfaces as well as for the customer's satisfaction and shopping experience to provide the user with the desired information in an appropriate and supportive way.

Using VR systems in the shopping area has grown in its importance and has emerged to a new trend to create virtual stores. Retailers initiate these virtual environments to take customers along experiences. Although VR shopping is still at the very beginning, this could fill the gaps of common online shopping (e.g., customer satisfaction, and experience) and become one of the more popular ways to shop online. With the fast-growing market of VR hardware as well as the data and knowledge that retailers have gathered in the last decades through online shopping, the process is mutually beneficial. For the future of buying products online, retailers could have a huge benefit. Using VR shops, in which a 3D rendering of products is created for customers to view from every side, could lead to more satisfied customers as well as an ability to showcase detailed items, e.g., to compare them with their 3D rendered counterparts. As ordinary online shops offer 2D content, they use simple 2D interfaces with hyperlinks, labels, icons and menus. Here, only the products are important, and you need to find them quickly, for the sake of convenience and conversion rates. The products are mostly displayed in a list or grid view, i.e. data browsing is done by scrolling or paged-based navigation. While this approach can have high usability ratings, in particular for finding products, it abstracts away from the "3D world" of stores, and totally disregards the value of user experience and immersion, especially with an increasing number of products and categories. A VR shop instead could profit by its third dimension and 3D interfaces such as 3D graphics, natural metaphors or avatars. More vivid content representations might give a more positive consumer response.

But modern media requires more complex interaction techniques, which consequently cause higher levels of user instrumentation [Griffith et al. 2001]. By touching the third dimension, VR provides a more advanced form of visualization, which can increase the customer's satisfaction and thus shopping experience. It allows more natural user interfaces (e.g. camera control using head rotation) than the usual mouse/keyboard interaction in a desktop environment. But from a technological point of view, visualizations and interfaces of VR, shopping worlds, as well as customer interactions in VR were rarely studied as well as the development of VR online shopping platforms. Thus, there are only a few results that give insights on how such technologies provide compensation for the lack of multi-sensory input and output (see look-oriented interaction with objects as in 360° product views, spatial interaction with fresh food counters [Gehring et al. 2012] or interactive clothes booths) or ways to enable multi-sensory interactions online.

In this work, we explored the main characteristics of on- and offline shops. Equipped with this knowledge, we moved from the task of searching for a product to the design and implementation of a VR shop prototype. We tried to maintain the benefits of online shops, like search functionality and availability, while shopping experience and immersion should not be disregarded. Nevertheless, it should be mentioned that the experience and immersion of a mobile online VR shop is strongly influenced by its technical limitations (e.g., lower resolution and render quality, less computing performance, and possible network latency). So, our focus was more about evaluating common potential input mechanisms for selecting products in a mobile interactive VR shop and the experience of using desktop versus HMD. Our concept can be further applied to existing online shops, as they are connected to a database consisting of adequate data quality and market layout data. Therefore, we developed an immersive virtual shopping environment that includes the major advantages of e-commerce and WebVR using a smartphone. Smartphones are already available for most users, and therefore are the cheapest, most affordable scenario for virtual online shops. Additionally, WebVR applications can be used as mobile apps instead of high-resolution VR applications requiring a wired Oculus Rift or HTC Vive device plus expensive high-performance hardware. So our system allows the user to search for products by head pointing and speech input in a virtual shopping environment.

In order to evaluate our system, we conducted a comparative user study to investigate the task of searching for a product in a VR web store with different combinations of hands-free input (pointing vs. speech) and output (desktop vs. VR) types. The motivation behind the two hands-free input methods instead of using text entry was to reduce the inconvenience of shopping, because searching by typing the name can cause issues like mistyping or not knowing the correct spelling of the desired product name. We used the desktop condition as the baseline and because it is the standard output device for common 3D shopping environments. The VR setting should particularly enhance the immersion and user experience. Moreover, the textured and true-to-scale 3D representation of the product should also increase the customer's satisfaction [Ohta et al. 2015; Pine and Gilmore 1998]. The results showed that the usability and user experience of our system is above average overall. As expected, the search tasks were performed faster using the mouse and keyboard setting in the desktop mode. Particularly in VR, speech input turned out to be the essential factor for usability and performance compared to more commonly used head pointing and to both desktop inputs. The study also proves that the users had an outstanding sense of presence and immersion in VR, although motion sickness was rated as quite high compared to related VR applications, which could be due to the limitations of smartphone VR. In summary, searching for a product using speech input in combination with VR output proved to be the best regarding user performance and preference.

In the following, the structure of the remaining parts of this work is outlined. First, we assessed the current situation of prior work in the fields of virtual shopping environments and interaction through literature review (see Section 2). The next step was to introduce the VRSE model (see Section 3) as a theoretical approach and metric for evaluating VR shopping environments based on common metrics in the related fields of customer's satisfaction, 3D User Interfaces, and Virtual Reality. As part of the user-centered design cycle, we conducted a customer survey to understand the user and explore the main characteristics between on- and offline shops (see Section 4.1), followed by the definition of first design principles for designing interfaces for VR shopping environments based on related work and the survey results (see Section 4.2). These parts formed the basis for the development of a proof-of-concept in the form of a mobile interactive VR shop using head pointing and speech as input (see Sections 4.3,4.4 & 4.5). The next step was to conduct a case study and evaluate our VR shop prototype (see Sections 5) with respect to the VRSE model. Finally, we formulated design guidelines for developing VR shopping environments based on new findings from the customer survey and experimental results (see Section 6). Hence, the main contributions of this paper are as follows:

- **Exploration of the main characteristics of on- and offline shopping** with regard to customer shopping behavior and frequency.
- **A case study of a first VR shop prototype** was conducted and evaluated with respect to customer satisfaction, task performance and user preference. The study further investigated potential input mechanisms for selecting products in a mobile interactive VR shop and the experience of using desktop versus HMDs for output.
- **Introducing the *Virtual Reality Shopping Experience (VRSE)* model** based on customer satisfaction, task performance and user preference.

2 RELATED WORK

The investigation and evaluation of virtual reality shopping experience using head pointing and speech input in VR approaches from different domains. Specifically we identified (1) prior work in the field of shopping in virtual environments, (2) reducing motion sickness in virtual environments, as well as (3) interaction in mobile virtual reality applications.

2.1 Shopping in Virtual Environments

Currently, a major problem of online shops is the lack of clarity, realism and immersion. Buffa et al. [Buffa and Lafon 2000] describe advantages and limitations of 3D stores. Adding a third dimension could fill these gaps in the virtual shopping experience. In their point of view, customers will then benefit from always-open unstaffed warehouses, time-saving shopping and multi-modal product information. This applies for online shops, but not for physical stores. These findings have provided the basis on which we investigated the benefits of online versus offline shopping and tried to combine them in the form of an interactive VR shopping environment.

In the last decades, more and more VR shopping environments have emerged, but mostly very simplistic. Some projects are equipped with avatars [Zhao and Zhang 2012] or a personalized assistant. They further claimed that VR shopping environments provide a „more feel experience" compared to 2D e-commerce environments [Chen et al. 2008; Zhao and Zhang 2012] and 3D applications for e-commerce are feasible [Bei et al. 2005]. Sanna et al. [Sanna et al. 2002] presented a virtual shopping environment generated from preferred products the users chose before. This all indicates that personality and feasibility are crucial factors and should be addressed when designing VR shopping experiences. In our work, we focus more on efficiency and feasibility, but the personal aspect nevertheless will be considered as well as using speech input for searching product. AWE3D and ADVIRT [Chittaro and Ration 2000] are architectures for adaptive 3D e-commerce websites. Here the store layout, organization, and appearance were adapted to the personalized rules of the user. Nevertheless, these 3D shop concepts only focus on adaptability, not availability, for VR technology. There are different concepts of creating virtual shopping environments, but we will focus on dynamically generated environments, based on existing on- or offline shop data, because of scalability, availability and sustainability. But the technology to develop e-commerce applications for immersive VR already exists. Laver et al. [Laver et al. 2012] developed a grocery shopping simulator to investigate the interaction between the user and the program. Patients participating in neurological rehabilitation performed user tests resulting in mainly positive feedback. Although they focused on usability, their results indicated that VR could have a strong impact on immersion. So we adapted the concept of fully dynamic generated 3D online shops and extended it with immersive VR for output.

In Shop-WISE [De Troyer et al. 2007] the user is able to pick up 3D products and inspect them. A product database is maintained separately and thus it can be altered very easily. Like common e-commerce systems and online shops, we adapt this to our system to read out a product list dynamically to provide an up-to-date overview of the inventory. Besides that, Shop-WISE allows searching for products by text input and moves the user to the desired product after it has been selected from a results list. Here, the fact that all products have static locations makes the shopping experience more realistic. However, in their approach the user chooses a product from a text list without knowing if the product is the desired one, except by its title. We adapted their approach and extended it using 3D model representations of the search results, so the user gets better clarity, whether the desired product has been found, which can also persuade users to buy the product [Häubl and Figueroa 2002].

Ohta et al. [Ohta et al. 2015] propose a concept of a mixed reality shopping system to let the user explore a virtual store in VR using a smartwatch as input. Their goal was to support disadvantaged shoppers and help customers speed up their shopping trip. As one of the first ones, they provide a clear distinction between online and offline shopping and stated characteristics. They claimed that e-commerce reduces the inconvenience of shopping, but searching by name (speech or textual input) has issues like wrong inputs or not knowing the store's name for the desired product. Moreover, product images can also be insufficient for checking details, and many products change their appearance more frequently, which can lead to high maintenance effort. We entirely agree with this, so we used rotating 3D product representations in our system. In order to extend and verify their claims and characteristics of online and offline shopping with regard to the users' needs, we decided to conduct a preliminary customer survey (see Section 4.1). Magic Home [Wan 2000] is a concept prototype of a VR furniture store, which illustrates what the shopping experience could be like in future of hybrid reality systems. In their scenario, the customer walks through a local physical store, where he or she can try out and feel the furniture. And because this store provides a connection to a virtual representation of the customer's home, he or she can get a preview of how it will look like. We adapted this concept of a connection between physical and virtual world, but vice versa. In our approach of a mobile interactive VR shop, the customer can go shopping in a virtual 3D representation of a grocery store from home or anywhere else.

But also commercial solutions are appearing on the market. A very recent example, but not yet market-ready, is the eBay VR Shop Australia[1], with more than 8,000 products grouped in categories in a connected graph, instead of rendering a virtual shopping market. Here, the user can select products by head pointing using dwell time (6 seconds). They call this interaction *Sight Search*, but, however, it is not more that the standard ray-cast pointing technique for object selection using head orientation [Argelaguet and Andujar 2013]. This graph approach provides a good overview of coarse-level categories, but can quickly break down into doubt and confusion if a finer-level product search is needed, and is strongly dependent on the data quality and categorization complexity. ShelfZone VR[2] is a retail space simulator and reproduces shops dynamically generating a virtual shopping environment by the use of planogram output, with data categorized into products, shelves, and zones. When standing in front of a shelf, the customer can grab a product to get detailed information, which requires HTC Vive controllers in their approach. They also provide a speech interface which can receive voice commands and give feedback.

[1]https://vr.ebay.com.au/
[2]http://invrsion.com/shelfzone

2.2 Reducing motion sickness in virtual environments

When developing immersive VR applications using HMDs, motion sickness is one of the most crucial problems and a main factor influencing user preference and experience. The main symptoms of motion sickness are nausea, eye strain, headaches, and blurred vision. Sharples et al. [Sharples et al. 2008] showed that the symptoms are greater when using immersive VR. Although they are negligible in a desktop setting, it can be used as a reference for comparing different VR applications based on results in a desktop setting. However, there are different methods for reducing motion sickness in virtual environments. Research of offline shopping experience in HCI is either context-aware [Kim et al. 2015] or social [Garcia-Perate et al. 2013; Morris et al. 2014]. Based on findings of Black et al. [Black et al. 2010] concerning context- aware shopping experience, we integrated a virtual shopping cart into our design. Whittinghill et al. [Whittinghill et al. 2015] displayed a virtual nose to reduce motion sickness. The users didn't perceive the virtual nose, but a study showed that the presence of the virtual nose reduces motion sickness. Recent research in motion sickness indicates that high velocity could amplify the symptoms [Mourant and Thattacherry 2000; Tanaka and Takagi 2004], e.g. driving on a highway causes stronger symptoms than in the city. A non-trivial factor for motion sickness is the field of view (FOV), because a larger FOV gives the user a greater feeling of presence, but also of motion sickness [Lin et al. 2002; Seay et al. 2002]. Furthermore, they found a negative correlation between motion sickness and enjoyment. Fernandes and Feiner [Fernandes and Feiner 2016] presented a trade-off using a dynamically changing FOV. They conducted a study which showed that their approach could help to keep the sense of presence while reducing the symptoms of motion sickness. In our approach, both the default velocity and the FOV are set to a certain value to minimize motion sickness while keeping the immersion as high as possible for every participant in our study.

2.3 Interaction in Mobile Virtual Reality Applications

Today's smart-phones have multiple built-in sensors providing 3D data. The accelerometer measures the acceleration relative to the free fall, unlike the gyroscope, which measures the absolute device orientation [Liu 2013]. By combining the sensors' data, the 3D orientation of the device can be computed, but involves a certain risk of jitter. Google Cardboard[3] is a cheap head-mounted display (HMD) solution, compared to Oculus Rift[4] or HTV Vive[5] hardware, because in their setup only a state-of-the-art smart-phone is needed. Here, the computed device orientation serves as head rotation, so the user can look around in the virtual environment by moving the device, or head if it is attached to a HMD case. In our approach, we used such a mobile WebVR setup inspired by Google Cardboard, because of its high availability and low costs, which are crucial factors in common retail scenarios.

Concerning interaction with such hardware in a virtual environment, Soojeong et al. [Yoo and Parker 2015] presented some controller-less interaction possibilities especially for mobile VR applications. They claimed that gaze is one of the simplest, and has two types: instant and dwelling gaze. Instant gaze triggers object selection directly by looking at it without any confirmation. In contrast, by using dwelling gaze, the selection is triggered if the user looks at an object for a few seconds. While instant gaze or pointing cause the "Midas touch problem" [Istance et al. 2008], dwelling gaze prevents it, so we chose dwelling gaze for our approach. According to Kim et al. [Kim et al. 2015], gaze- or pointing-based interaction has several benefits: naturalism, remote controllability, and easy accessibility. They present their concept of gaze-based interaction that augments the physical world in a shopping scenario.

Another powerful sensor in smartphones is the built-in microphone, which can be used to interact with the virtual environment by speech input. Speech input is a very comfortable and natural input modality, and there has been much research in the last decades in the field of speech recognition, with major companies like Microsoft or

[3]https://vr.google.com/cardboard/
[4]https://www3.oculus.com/en-us/rift/
[5]https://www.vive.com/us/product/

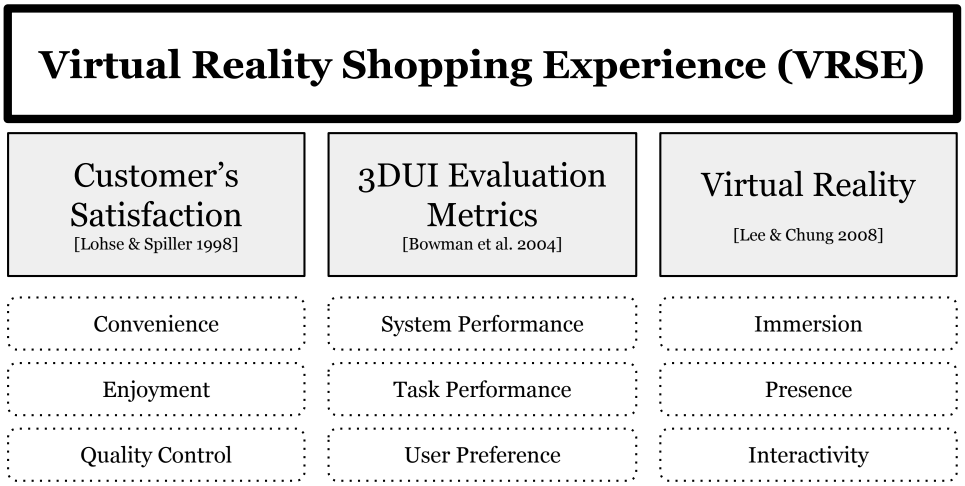

Fig. 2. The Virtual Reality Shopping Experience (VRSE) model.

Google bringing services to market. With the increasing amount of available data and computational resources, it becomes more and more accurate [Schalkwyk et al. 2010; Schuster 2010]. Finally, we decided to use the "Web Speech API"[6] in our approach, which is a speech recognition framework at word level for browsers using Google's speech recognition web service.

However, VR-based e-commerce applications require validation of such findings and answers to new questions. In particular, springing up VR shops are mostly based on web or mobile guidelines. Therefore, we introduce the Virtual Reality Shopping Experience (VRSE) model. This will enable designers to build more usable and effective VR shopping systems and help to move towards a stronger theoretical basis and more principled design guidelines for VR shops.

3 VIRTUAL REALITY SHOPPING EXPERIENCE (VRSE)

One purpose of this research is to develop principles for the design and evaluation of effective and usable VR shopping environments. Although there is lots of research in multimedia and e-commerce [Nemetz 2000], we believe that there are still open questions and especially designers of VR shopping environments are still faced with little guidance as there is no link between VR and e-commerce. So, we introduce the Virtual Reality Shopping Experience (VRSE) model in order to fill the gap between VR and e-commerce. Our model combines the metrics of customer's satisfaction, evaluation of 3D user interfaces, and the characteristics of virtual reality (see Figure 2).

3.1 Customer's Satisfaction

Most users see the interface of online shops as appropriate [Kendall and Kendall 2002], mainly because of their search functionality and opening hours. This can lead to the assumption that ordinary online shops with user-friendly and attractive user interfaces would provide a higher *customer satisfaction* including three main characteristics: convenience, enjoyment and quality control [Jarvenpaa and Todd 1996; Lohse and Spiller 1998]. Satisfied customers could tend to make not only more purchases at once, but also repeated purchases. Therefore,

[6]https://cloud.google.com/speech/

we want to address the main characteristics of customer satisfaction regarding online and offline shopping. With regard to online shops, convenience includes store navigation features like search functions, site maps or product indices which are essential for large stores [Lohse and Spiller 1998]. Enjoyment is important, because people find playful interaction intrinsically interesting, e.g. when they are involved in activities like purchasing something for pleasure and enjoyment [Agarwal and Karahanna 2000]. Quality control by customers of online shops is still a challenging task. It requires the ability to test a product before purchase or get an impression of its size and shape, which is not handled well by current online shops. While isolated online stores offer 360° images or videos of products, they cannot actually fill the gap in experience or immersion.

3.2 3DUI Evaluation Metrics

As VR applications necessarily consist of *3D user interfaces (3DUI)*, Bowman et al. [Bowman et al. 2004] differentiate between three evaluation metrics for 3DUI: system performance, task performance, and user preference. Whereas system performance metrics include benchmarking (like average frame rate, average latency or interaction time), task performance metrics include quantitative measurements such as task completion time, error rate and accuracy (or precision). These metrics indicate to what extent users are able to cope with the task and interaction method. User preference metrics usually consist of subjective feedback, such as *user experience*, *usability* and *motion sickness*. User experience should be measured by the User Experience Questionnaire (UEQ) [Laugwitz et al. 2008], an end-user questionnaire consisting of 26 very short questions to measure user experience quickly in a simple and immediate way. For the usability assessment, the System Usability Scale (SUS) [Brooke et al. 1996] should be used, which is likely the most popular questionnaire for measuring attitudes toward system usability[7]. It is a reliable and valid measure of perceived usability. Furthermore, it performs as well as or better than commercial questionnaires and homegrown internal questionnaires. For measuring motion sickness, we recommend to let participants fill out the Motion Sickness Assessment Questionnaire (MSAQ) [Gianaros et al. 2001], which is a valid instrument for the assessment of motion sickness.

3.3 Characteristics of Virtual Reality

VR, however, could be the next component to a better customer satisfaction and shopping experience. Lee and Chung [Lee and Chung 2008] formulated three important characteristics of *virtual reality (VR)*: immersion, interactivity, and presence. Immersion and presence could be combined into one characteristic, i.e. to what extent the customer's senses are isolated from the real and stimulated by the virtual world and the subjective experience of being in one environment, but physically situated in another. Interactivity indicates to what extent users can participate in manipulating virtual content in real time. We recommend to use the common immersion questionnaire for VR applications from Slater et al. [Slater et al. 1994] to measure the immersion and presence of a virtual shopping environment. Here, it is important to customize this questionnaire by changing the location description from 'office' to 'store or shop'. The âĂŹSUS (Slater Usoh Steed) CountâĂŹ shows the mean of the SUS count of âĂŸ6âĂŹ or âĂŸ7âĂŹ scores amongst the 6 questions. The âĂŸSUS MeanâĂŹ uses the mean score across the 6 questions instead.

In summary, we want to combine the VR characteristics and 3DUI evaluation metrics into a model for evaluating VR shopping experiences with regard to customer satisfaction (see Figure 2). This will enable designers to build more usable and effective VR shopping systems and help to move towards a stronger theoretical basis and more principled design guidelines.

[7]Not to be confused by the Slater-Usoh-Steed (SUS) Count [Slater et al. 1994], which measures presence and immersion.

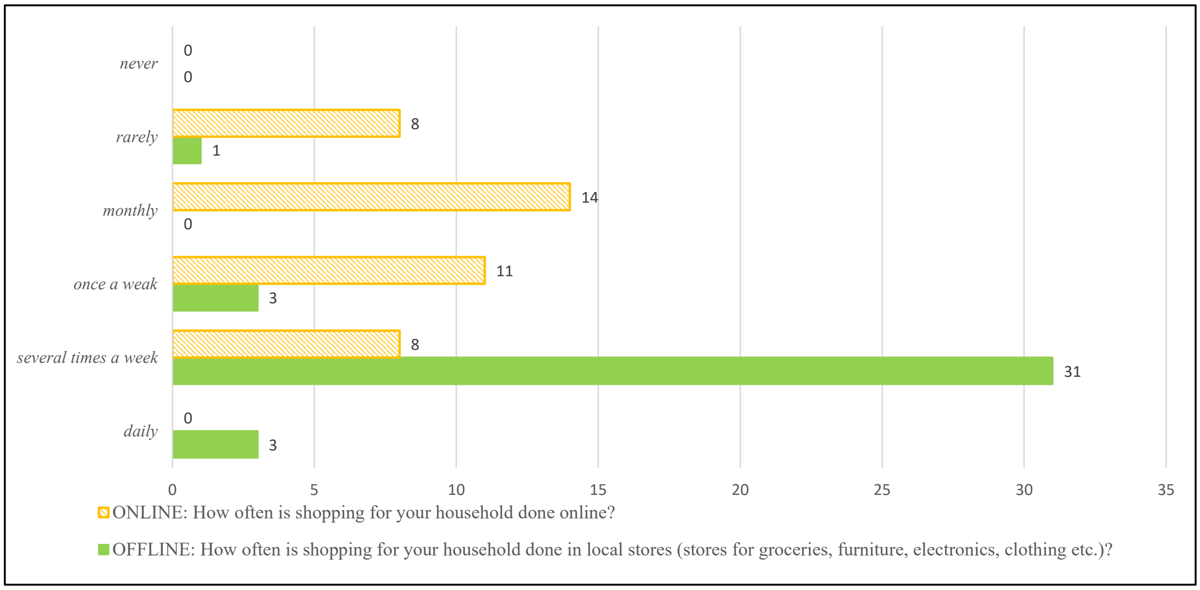

Fig. 3. Shopping frequency online vs. offline (N=41).

Fig. 4. Shopping behavior online vs. offline with regard to groceries, textiles, clothes, furniture, and electronics. (N=41).

4 MOVING FROM TASK TO DESIGN

4.1 Customer Survey

In the run-up to our work, we conducted an exploratory customer survey with a focus on positive and negative aspects of on- and offline shopping, and also with a focus on shopping behavior and frequency. Based on the results of this survey, we designed and developed a WebVR online shop addressing the issues and including the benefits of on- and offline shopping.

A total of 41 participants (19 female) took part on this survey including 13 short questions. The participants' ages ranged from 18 to 58 years ($M = 31.12$, $SD = 12.35$). Based on this age distribution and on common standards, we defined four different age groups: 18-20 (7.3%), 21-24 (31.7%), 25-39 (39.0%), and 40-59 (22.0%). The housing situation of the participants was nearly balanced, with most 31.7% living with a partner, followed by 24.4% living alone, and each 22.0% living with parents/family or in a shared apartment. They were recruited in the entrance area of a collaborated German retail hypermarket and volunteered uncompensated. After we asked a potential participant and he or she agreed to participate, we let him or her fill out the survey using Google Form on an tablet with internet connection. The survey was designed as short as possible and didn't last longer than 5 minutes. The results of the survey were analyzed by us using IBM SPSS Statistics 24.

We asked them on a scale from 0 to 5 (0: never, 5: daily), how often they go shopping on- and offline. The results indicate that customers go shopping more often in offline ($M = 3.91$, $SD = 0.62$) than in online stores ($M = 2.45$, $SD = 1.02$). 98% go shopping offline at least once per week, 83% several times a week, but only 7% daily (see Figure 3). We further asked them about their shopping behavior on a scale from 0 to 5 (0: never, 1: only offline, 5: only online) with regard to five common product categories: food, textiles, clothes, furniture, and electronics (see Figure 4). Food ($M = 1.29$, $SD = 0.55$) and furniture ($M = 1.62$, $SD = 0.86$) are bought more locally than online to a great extent, as well as textiles ($M = 2.45$, $SD = 1.13$) and clothes ($M = 2.55$, $SD = 0.97$) mostly locally but occasionally online. Asked on a scale from 1 to 5 (1: doesn't apply at all, 5: applies fully), whether they find it easy to find a desired product in offline and online shops, over 78% preferred online shops ($M = 4.05$, $SD = 0.85$), whereas the results were rather moderate for finding products in offline shops, in general ($M = 3.45$, $SD = 0.89$). A Multivariate ANOVA with all frequencies of types of goods (groceries, clothes, textiles, electronics, and furniture) as dependent variables, and family status (living alone, in a relationship, with family/parents, or in a shared apartment) and age ranges (18-20, 21-24, 25-39, and 40-59) as factors was conducted. A significant effect has been found in terms of grocery shopping frequency regarding the family status ($p < 0.03$, $F(3, 41) = 3.80$, $\eta^2 = 0.29$). Moreover, there was an interaction between age ranges and family status in terms of grocery shopping frequency ($p < 0.04$, $F(6, 41) = 2.62$, $\eta^2 = 0.36$). We conducted another multivariate ANOVA with online versus offline shopping frequency as dependent with regard to age ranges and family status, which showed no significant differences ($p > 0.05$, $F(3, 41) < 1.60$, $\eta^2 < 0.14$). Finally, an interaction has been found between age ranges and family status in terms of ease of finding desired products ($p < 0.05$, $F(6, 41) = 2.47$, $\eta^2 = 0.34$).

With regard to the positive aspects of online shops, 68% of the participants mentioned the search functionality (see Figure 5). In each case, however, 12% mentioned that the search functionality is often badly implemented, and the display of products in lists or tiles is unstructured and confusing. On the other hand, as benefits of offline shops, most mentioned employees (51%), signs (34%), and the use of departments and shelves as categories (24%).

In summary, today's online shops are generally acceptable in their search quality and functionality, but lag behind because of its often confusing and unstructured interfaces, which finally results in lower user experience and customer's satisfaction in general according to participants' comments. This is also reflected by the product's dimensions and representation in ordinary online shops. The majority of participants of the customer survey reported that the product representations in the form of colored 2D pictures in scrollable lists are hard to understand, i.e. the user doesn't get a clear sense of its size, shape, or weight. In addition, the user is forced to scroll the list or navigate through several pages, so the workload (effort, frustration, etc.) and user experience (attractiveness, stimulation, etc.) may thereby be impaired. Nevertheless, more and more 360° product pictures and videos are to be found. This is a next step, but still not an optimal solution. The lack of clarity and the aspect of experience and satisfaction still are issues.

VR could fill this gap by merging the advantages of on- and offline shops together and getting off their major limitations. In any case, our concept provides 3D model representations of the products, placed in virtual shelves and departments in a virtual store. The focus here is on user experience and customer satisfaction. But the user's

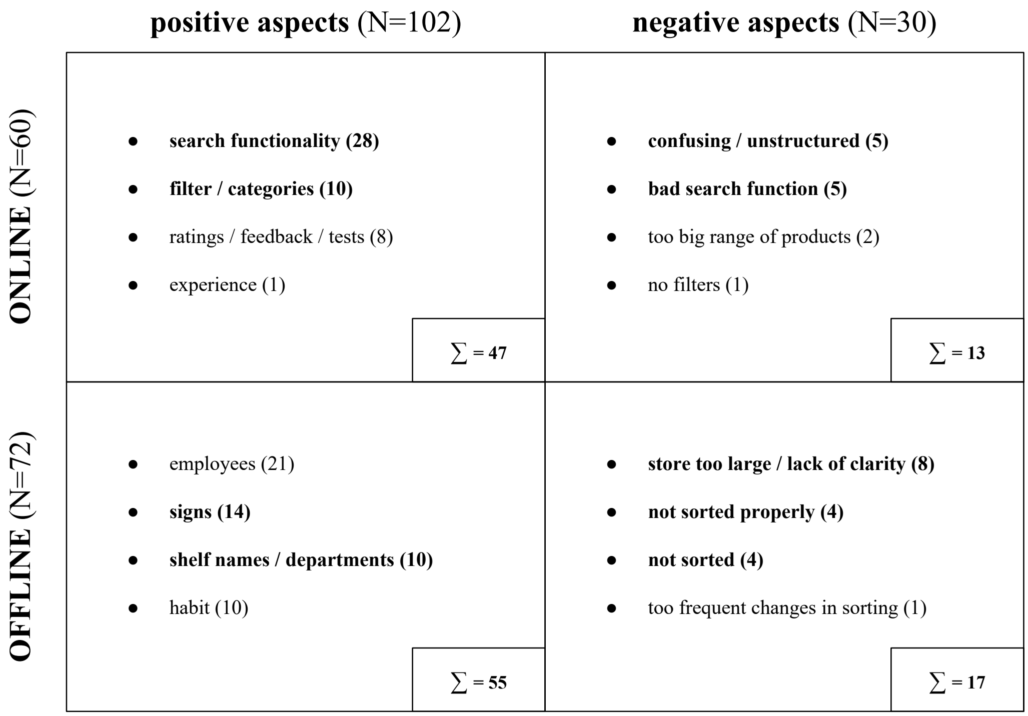

Fig. 5. This figure illustrates the main characteristics of off- and online shopping mentioned by the participants of our customer survey (N=132, with N as the number of mentioned items).

performance (error rate, speed) is still an important factor, because even an otherwise satisfied customer needs to find his/her desired products.

But nevertheless, further studies should be conducted to investigate the effects of VR in different shopping areas or instances. The main reason why people prefer to buy furniture in physical stores is that they can feel the texture of a furniture and get a better idea of how it could look like in their apartment. And thus here is the great potential for VR furniture stores, because the customer could then look at the furniture in his home environment and easily configure individual components, like color, texture, or size. IKEA has made a first step for VR furniture shopping in the form of a free VR application in the Steam Store [8], where the user can explore and interact with an IKEA Kitchen in VR. But also clothes can be tried in VR in mixed reality dressing rooms [Yang et al. 2009]. VR stores for electronics could enable the customer to test and try out products and configure their individual components in VR before purchasing them. One example could be to arrange and test a new sound system in a virtual representation of the user's home environment. However, according to our customer survey, groceries are bought above-average offline. Furthermore, groceries have, compared to other types of goods (e.g., furniture or electronics), the most different product types, categories and ranges. In addition, online grocery shopping is becoming more and more relevant with an increasing amount of offers to deliver groceries to your doorstep, but

[8]http://store.steampowered.com/app/447270/

the acceptance and feedback of online grocery shops is rather negative. Therefore, we decided to investigate grocery shopping in a mobile online VR shop including a better experience than with traditional online grocery shopping and better than the performance in physical offline stores.

In summary, the factors generated from our exploratory customer survey influenced our decision to develop our system based on WebVR simulating a online grocery shop and equipped with search functionality. Nonetheless, the VRSE model is generalizable to all the mentioned retail areas. And our experimental study includes one example scenario how to use the VRSE model for evaluating VR shopping experience of a virtual shopping environment, in our case with focus on online grocery shopping.

4.2 Designing Interfaces for Virtual Reality Shopping Experiences

When designing interfaces for VR shopping experiences, the main task of the application should be clear, so that the designer can move from task to design. In this paper, we focus on the task of searching for products. However, we recommend to conduct a preliminary customer survey preferably right in the offline store, which should gather direct feedback from the target group. The main benefits should be respected while avoiding the mistakes that surround online as well as offline shopping environments according to customer survey findings (see Figure 5). Therefore, we provide a list of potential guidelines based on the survey results that could be used to inform the design of VR shop interfaces:

i. **CLARITY.** The scene should be as simple as possible, but still keep the feeling of being in a store looking for products. Therefore, signs, shelves and departments as visual categorizations and cues can help the customers to be better oriented and feel more immersive. And an increasing number of products in a store requires customer-friendly filters (organic food, allergens, sizes, colors, etc.) and a proper product categorization to help the users finding their desired products faster.
But: *Too many categories and a lack of clarity of the filters may also influence the performance and experience. The customers wish well-structured stores, not confusing, not too large, and not a too big range of products and departments. And do not overfill the shelves with products. In common large offline stores, when customers are standing in front of a shelf searching for a product, it is like „they cannot see the wood for the trees". Keep in mind that in virtuality, there is no need to overfill the shelves with duplicates of products.*

ii. **EFFICIENCY.** Provide the user with a user-friendly and efficient search functionality. Here, search filters can help to select products by their attributes and features or by their description texts.
But: *If you are using text-based methods, there will be a strong need of auto-completion, spell and synonym checker. Otherwise the usability and performance might be influenced negatively.*

iii. **ORIENTATION.** Use signs, shelf names or numbers in combination with a proper and useful categorization to give the users a better orientation in the virtual market. In addition, depending on the size of the market, the products and their shelves should be sorted properly in visually distinguishable departments.
But: *Do not overload the user with disturbing colors, unintuitive names or „too much information". A clearly arranged assortment makes it easier and faster for the user to find and select the desired product.*

iv. **PERSONALITY.** An essential factor is the personal aspect of a virtual shop. According to the feedback of the customer survey (see Figure 5), the customers of offline stores prefer the presence and interactivity of employees. More precisely, the ability to talk to an assistance and ask questions like „where can I find the milk?" is preferred instead of typing text on a computer or terminal. Use speech input and/or output to fill this gap. The speech interface serves as the personal part of the environment. Here, the customer can use his voice to search for products instead of typing text, so she is interacting with the system in a more personal way.
But: *Today's state-of-the-art speech-based interfaces are still in an infancy stage. So the developers should*

always keep in their minds that the speech recognition ratio can have a crucial influence on the shopping experience.

v. **QUALITY.** Provide the user with sufficient information about the products (e.g. feedback, ratings and tests). Instead of providing textual detailed information about the product's dimensions, shape or weight, it can be visualized using interactive 3D object representations and gravity in the virtual environment.
But: *Incorrect size, shape or weight will influence massively the user's impression of the product and environment itself.*

4.3 Design

The virtual environment of our VR shop was designed to be as simple as possible, but still close to a real market (see Figure 6). While in reality shelves are overfilled with products and the environment is decorated all over with advertisements, our approach contains only a clearly arranged assortment for quick and easy selection. This should solve one of the most common problems the customers complained of in the customer survey, namely that they struggle more with the lack of clarity the larger the market is. The environmental colors were chosen to be as neutral as possible to prevent positive or negative impact. Nevertheless, the attention of the user should always be attracted to the products, therefore they are colorful and natural, unlike the colorless surrounded elements. The contrast is emphasized by a subtle glossy reflection of the products.

Existing market layout and floor planning data provides information about dimensions, position and size of different types of store elements (walls, shelves, departments), so they could be easily parsed and rendered. Shelves contain at least one product, products are evenly distributed, and the space between each rack is computed by the market is height and the number of tiers (levels in the shelf). A product is represented by a virtual 3D model and is placed on its corresponding shelf level according to the floor plan data. After a product has been

Fig. 6. This figure illustrates the virtual environment of our VR shop including the shopping cart, a move plate (orange), the speech interface in the back, and shelves equipped with products on the right. The colors of the environment were chosen to be as neutral as possible, and only the products, move plates and the cursor were colored, in order to improve clarity and minimize distractions.

selected, a detail view is displayed, including a larger rotating 3D model of the product, side information like price and description, and buttons for adding it to the shopping cart or closing the detail view. The cart is always located in front of the user and holds all previously selected products, which can also be removed. A control panel at the cart's handle includes the total price, a button to open the speech interface and one to reset the scene. Besides those store elements, waypoints are displayed as move plates and spread over the whole market. In order to prevent unintentional movements, the user only sees nearby surrounding move plates. Thus for every possible user position, it is ensured that there at least two move plates in range.

4.4 Controls and Interfaces

As we focus on a mobile online WebVR application without controller and hand interaction, our system provides two different input modes: speech input and head pointing. According to the customer survey results, the majority of the participants really valued the search functionality of online shops. But on the contrary, this can even have the opposite effect if it is error-prone. Therefore, and because we want to avoid text entry in VR, our speech input method should enable VR customers to find their desired products in an efficient way. The speech interface is located on a wall in the scene and provides the functionality to search for products and list the results (see Figure 8). The user starts by gazing at the button with a speech bubble icon and speaking the search term into the microphone. After the search results are displayed, the user can select the desired product by gazing for five seconds in order to be moved to the desired shelf. When the user is looking at a result, the front view of the respective product is displayed in the middle of the search interface for better decision-support.

Another way to search for a product, and for the more exploratory characters among the users, we chose a navigation through head pointing and waypoints. Here, the user is moving around using the way-points and looking for it in the shelves (see Figure 7). Our system doesn't include eye-tracking hardware, as it is the today's state-of-the-art for common VR installations and applications. But different concepts exist (e.g., PupilLabs' HTC

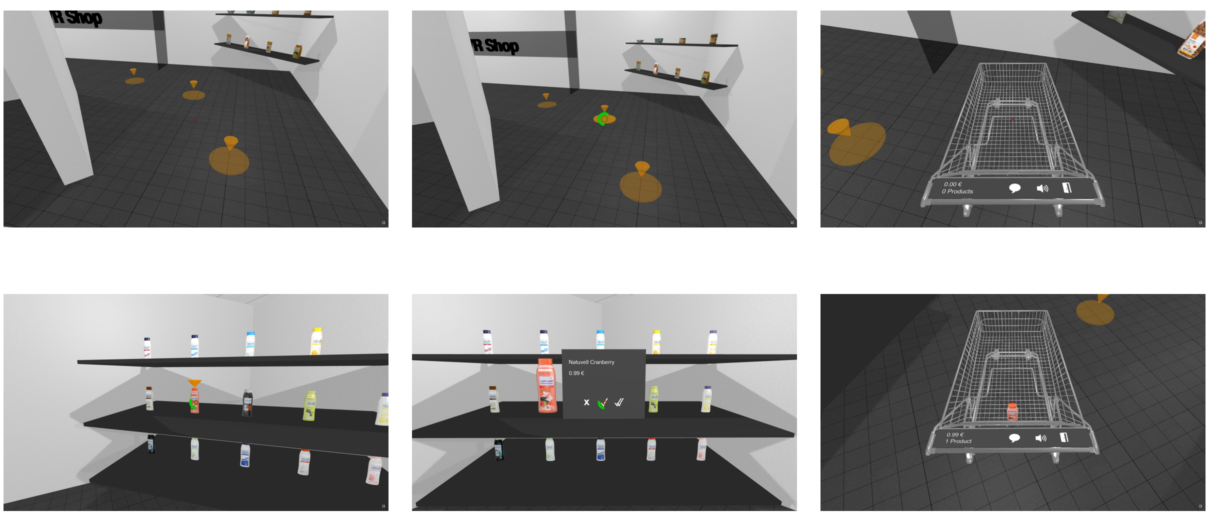

Fig. 7. **Head Pointing.** Here, the user looks at a nearby waypoint in order to be moved. Then, a product is placed into the shopping cart after it has been selected. Finally, the total sum and amount of items is displayed on the shopping cart's handle.

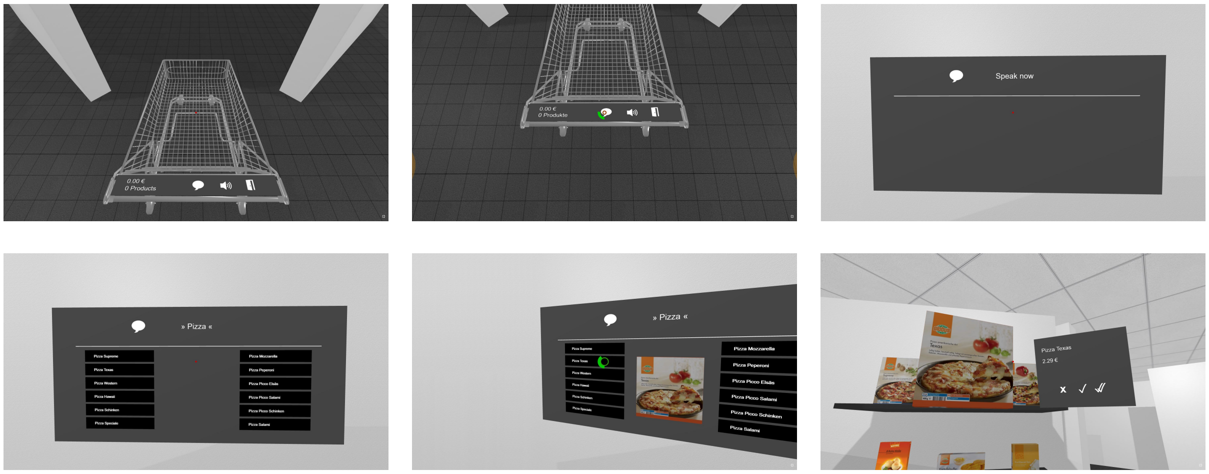

Fig. 8. **Speech input.** Here, the user starts by gazing at the speech button, and speaks the search term into the microphone. After the search results are displayed, the user can select the desired product in order to be moved to the desired shelf.

Vive or Oculus DK2 integration [9], or other commercial products like SMI [10]), so it is just a matter of time until the first mobile VR hardware equipped with eye-tracking will be affordable and usable. However, as we wanted to focus on affordable, commodity and state-of-the-art mobile VR scenarios, we decided to use head pointing interaction instead of eye gaze. Whereas the procedure of this input mode is straightforward, i.e. the user moves around in the virtual environment by using (looking at) the move plates and looking around to explore the market, it has the limitation that in the worst case scenario the user has to move through the whole virtual environment until the desired product is found. The advantage is that efficiency should increase proportional to the spatial knowledge of the market layout, i.e. once the environment is well-known, it is easy to find products by moving straight towards them. Nevertheless, it can still take some time to find the product in the desired shelf.

Concerning the output devices, the system automatically detects which output devices are currently available and displays the rendered scene to them. One way to use it is in a desktop setting with an attached monitor (see Figure 9). Of course, the render quality is better than on a mobile device, but applications in the desktop setting could be lacking in immersion and user experience. Another way is to use a mobile VR, which is cheap and affordable, in contrast to high-end HMDs (see Figure 9). Whereas in a desktop setting, head pointing and speech have to be tracked by an extra device (like Microsoft Kinect), both come for free using smart-phone sensors and a built-in microphone. Another benefit is handiness, because the smart-phone doesn't need a tethered connection, so it can be used nearly everywhere, even in public. However, the most crucial limitation for smart-phone VR is the render quality and computing performance, which is not comparable with the power of a PC.

4.5 Mobile Interactive Virtual Reality

The main benefits of WebVR are its mobility and affordability. Our choice of a WebVR application was made to gain a further advantage, namely that existing web shops could be rendered in a 3D scene instead of on a classical 2D web page. Our approach includes a link to an existing product database and its search functionality, as well as to the market layout system of the corresponding retail store. The novelty of our approach consists in the visualization and interaction of a virtual store and not in product search or market navigation. We investigated

[9] https://pupil-labs.com/blog/2016-02/eye-tracking-for-vr-and-ar/
[10] https://www.smivision.com/eye-tracking/product/mobile-eye-tracking-hmd-samsung-gearvr/

how to enhance the shopping experience, as a combination of customer satisfaction and user experience, for online shops. In an experiment to evaluate our system, the virtual environment (market layout, 3D models and locations of the products) of the VR online shop is dynamically generated based on a fictive store's product database and floor plan.

5 A CASE STUDY: VR SHOP

We conducted an experiment in order to gather insights into user performance, preference, and unmet needs. The study provides us with:

- **Metrics.** Objective and behavioral performance data that provides a usability baseline to measure future improvements of virtual reality online environments.
- **Customer insights.** Actionable insights on how to optimize usability, satisfaction, and experience for the customers.
- **Actionable improvements.** Concrete recommendations for improvements based on research findings.

The purpose of this experiment was to evaluate the end-to-end experience of users as they interact with our system using, in each case, two different input (speech vs. pointing only) and output (VR vs. desktop) methods.

5.1 Pilot study

Before turning to the first main study, we first describe a preliminary experiment we conducted to establish the trial time limit, appropriate product locations, different movement types and feedback modes. This pilot study was conducted with four unpaid university students, without giving the participants feedback regarding their trial time. The task of the pilot study was to process items on a shopping list after a short exploration phase in the environment. One participant was concerned about the speech interface's location, which initially was dynamically positioned at the user's position in his field of view. He or she complained about a lack of realism, so the speech interface was finally positioned at a fixed location in the virtual market. Another participant asked for functionality to add a product twice to the shopping cart. The participants had two options to move through the scene, by moving along a path or teleporting. In the path movement setting, the user is moved over a minimal path along way-points to the target position. The teleportation is based on concepts of dream research and uses the metaphor of closing the eyes. After the user has been teleported to the target position, the „eyes" were opened again. In our pilot study, the participants preferred the path movement, so it was used in the course of the main

Fig. 9. This figures show the **output devices** we used for our prototype. On the left: 22 inch display with Microsoft Kinect 2 for speech input and a standard mouse for camera control. On the right: a common smart-phone VR case and LG Nexus 5X.

study. With regard to the product highlighting, all pilot participants preferred the cone highlighting method, where a cone was placed above the product the user is currently looking at, fading out all other products. So in the end, the cone highlighting was chosen as the standard feedback mode for product highlighting. As we used real products from a common German retail store, product names can be very complex to detect from speech or even pronounced by the participants. So, with regard to the planned speech input method, we used the pilot study to remove product names to be searched by the participants in each task in the main experiment, which couldn't be recognized easily or impossible to be recognized using the Web Speech API from Google.

5.2 Hypothesis

- H_1 Desktop is faster than VR.
- H_2 VR is preferred by the user in terms of user experience and usability.
- H_3 Speech input is more efficient (speed, error rate) than head pointing only and is preferred by the user (usability, user experience).
- H_4 VR with speech input outperforms the others in all aspects.

5.3 Participants

A total of 16 unpaid participants (13 male, 3 female) volunteered in this experiment, aged between 17 and 55 years ($M = 23.88$, $SD = 8.52$). 69% of the participants had a university background; the rest were staff members and high-school graduates. The overall experience with VR and desktop applications, as well as the acceptance of VR shopping, were rated on a Likert-scale from 1 to 5. The level of experience with VR applications was very low overall ($M = 1.25$, $SD = 1.34$), whereas it was very high with desktop applications ($M = 4.25$, $SD = 0.93$). When asked whether they would go shopping in VR in the near future, the answers varied widely ($M = 2.88$, $SD = 1.31$). However, most of the participants tend to buy more often per month in real shops ($M = 7.56$, $SD = 5.48$) than online ($M = 2.93$, $SD = 2.23$).

Fig. 10. Left: **Desktop setting.** Here, the participant controls the camera by moving the mouse with her/his dominant hand. The speech was recorded by a Microsoft Kinect 2, placed under the display. The keyboard shown in this picture had no effect and was not used in the experiment. Right: **VR setting.** Here, the user controls the camera by moving her/his head. Speech was recorded by the smart-phone's microphone.

5.4 Apparatus

For the two desktop tasks a standard desktop computer was used with an i7 CPU, 16 GB RAM and a Nvidia GeForce GTX 980Ti graphics card. The display screen was 22 inches with a resolution of 1920x1080 pixels. In the desktop/pointing task, a standard mouse/keyboard setup was used for input devices (see Figure 10). The system had to be connected to the web at all times, because the application itself was a web page, which permanently sent logs to the a web server and was initially loaded from it. For displaying the WebVR content, a Google Chrome web browser and Microsoft Windows 10 was used. The desktop/speech task was performed with the same setup, but without the keyboard and mouse and instead using the built-in microphone of a Microsoft Kinect 2 for the speech input (see Figure 10). For both VR tasks LG Nexus 5X smartphone with a display size of 5.2 inches was used, which was put into Elegiant 3D VR glasses (see Figure 10). For using the system on the mobile device, only the Google Chrome App has to be installed, in combination with at least Android 6.0.

5.5 Design

The experiment was a within-subjects design, having two independent variables with two levels, respectively:

- Input Method (head pointing, speech input)
- Output Method (desktop, VR)

The input method conditions were counterbalanced using a Latin square. Aside from training, this amounted to 16 participants × 2 input methods × 2 output methods × 4 search terms = 256 trials. The four products were placed all over the virtual shop, and in order to ensure equal conditions for every participant, all trials started at the same position. Furthermore, the participants received only minimal instructions about the functionality of the different interaction types, so that no explicit conceptual model was assigned to them. The dependent variables were as follows:

- Performance (task completion time, error rate)
- Preference (usability, user experience, motion sickness, immersion)

It is also important to mention, that every trial had to be performed within a time limit of 120 seconds (see Pilot Study 5.1); otherwise the trial was counted as failed. The average duration of the whole experiment per participant was about 60 minutes, including an introduction and all questionnaires.

5.6 Task

The first experiment consisted of four different tasks representing all combinations of two output (desktop vs. VR) and two input (pointing vs. speech) modes. Every participant had to perform each of these four tasks with the goal to search for four specific products within the virtual shopping environment one after another, select them, and put into the shopping cart provided for this purpose. In head pointing mode, the participants performed a search task in order to find a certain product. As is standard practice, this product search task had two stages. The participant could only interact with waypoints or products at near range. However, after they stood in front of a shelf where they expected to find the desired product, the second stage of the search task began. Here, they searched for the product in the shelf's racks. When the participant had put the correct product in the shopping cart, the trial was completed successfully. In speech mode, the navigation part of the search task was omitted and replaced by a speech input task. Moreover, the participants then would not be able to move using the waypoints, so they were not displayed anymore. They started in front of a speech interface panel. After a product search was initiated by speech input, the search results were displayed. Since a search result has been selected, the participant was moved to the corresponding shelf along the shortest path.

In summary, each of the four tasks consisted of four trials, one per product. A trial was completed if the correct product was in the shopping cart within the time limit of 120 seconds; otherwise the trial was marked as failed.

5.7 Procedure

At the beginning of the experiment, the participant was welcomed by the experimenter. After that, the participant was to perform all four tasks in Latin-square order , which lasts about 20-30 seconds in average for one of four trials per task, i.e. about 5-10 minutes for all four in- and output task combinations in average. Before each task, the task goal and the handling of the interaction technique were introduced to the participant within a short warm-up phase with a maximum of 5 minutes. After the warm-up phase in the virtual environment and after each trial, the examiner reset all states in the virtual environment, so that the participant could always start from the same position. When all four trials of a task were performed, the participant was asked to fill out questionnaires collecting the subjective feedback: a System Usability Scale (SUS), User Experience Questionnaire (UEQ), Motion Sickness Assessment Questionnaire (MSAQ) and (only for VR) Presence Questionnaire. Those questionnaires took about 5 minutes after each of the four tasks. Furthermore, there was a 2-minute break after each set of post-task questionnaires and before the next task. Finally, after all 16 trials were performed and all post-task questionnaires were filled out, the participant was asked to fill out a final questionnaire collecting demographic data. Overall, the experiment took about 60 minutes per participant in total.

5.8 Results

In this results section and in the following discussion we use abbreviations for the four tasks we tested: D for desktop, V for Virtual Reality, 1 for head pointing, and 2 for speech input. Put together, they are: D1 (Desktop/Pointing), D2 (Desktop/Speech), V1 (VR/Pointing), V2 (VR/Speech). The results of the experiment were analyzed by us using IBM SPSS Statistics 24.

5.8.1 Task Completion Time & Error Rate. The task completion time is the elapsed time for the user to complete a trial within a search task. In the head pointing mode, it is defined by the elapsed time between the first pointing interaction with a way-point and the moment the searched for product was added to the cart. In the speech mode, the time measurement started with the first glance on the speech balloon symbol on the search interface panel, and ended when the searched for product was added to the cart. The error rate is defined by the ratio of the number of times that participants failed to find the desired products within a maximum period of 60 seconds. It is also worth to mention that if the speech recognition has failed, i.e. the spoken command wasn't detected correctly, the trial would have not been automatically counted as a fail. But, however, only in a very small minority of cases, which was not worth to mention due to our optimized choice of product names after the pilot study (see Section 5.1), the time limit exceeded because of recognition fails instead of just „not finding the target product". The overall task completion time of all trials ($N = 64$) was on average $24.05s$ ($SD = 11.68$) with an error rate of 21% ($SD = 0.41$) in average. There were significant differences between the four tasks concerning task completion time ($p < 0.01$, $F(1, 202) = 4.461$) and error rate ($p < 0.01$, $F(1, 252) = 26.235$). V2 was the fastest task with a mean time of $22.80s$ ($SD = 6.44$), and with an error rate of 2% ($SD = 0.13$). While D2 lasted on average $23.00s$ ($SD = 12.27$) and had an error rate of only 8% ($SD = 0.27$), D1 took $23.45s$ ($SD = 8.29$) with 22% ($SD = 0.42$) of the trials failed. Finally, V1 came off the worst both in time ($M = 31.07s$, $SD = 11.10$) and a 50% failure rate ($SD = 0.50$). When comparing the input mode, speech input was faster on average ($M = 22.90s$, $SD = 9.67$) than head pointing ($M = 25.83s$, $SD = 14.10$), and with a notably lower error rate of 5% ($SD = 0.21$) against 38% ($SD = 0.49$). But concerning the output modes, the tasks D1 and D2 were faster, with a mean time of $22.85s$ ($SD = 13.45$), than the ones in VR ($M = 25.46s$, $SD = 9.05$). The desktop tasks also had less failures ($M = 15\%$, $SD = 0.36$) than V1 and v2 ($M = 27\%$, $SD = 0.45$). Furthermore, univariate ANOVA analysis were conducted with task, input mode and output mode as factors, and the elapsed time and error rate as dependent variables. A significance was found regarding the task ($p < 0.01$, $F(3, 202) = 4.461$), as well as a significant effect of completion time with regard to the input mode ($p < 0.05$, $F(1, 202) = 5.662$) and output ($p < 0.05$, $F(1, 202) = 5.990$). In addition, an interaction was found between the input and output modes ($p < 0.01$, $F(1, 202) = 6.610$). Furthermore,

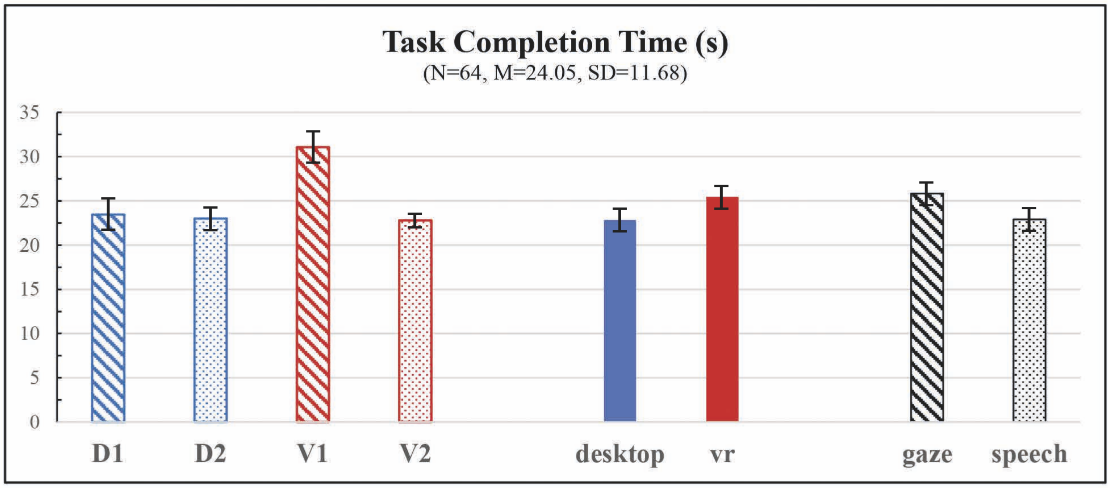

Fig. 11. Results of the task completion time measurements. The values are given in seconds (s). D1 (Desktop/Pointing), D2 (Desktop/Speech), V1 (VR/Pointing), V2 (VR/Speech).

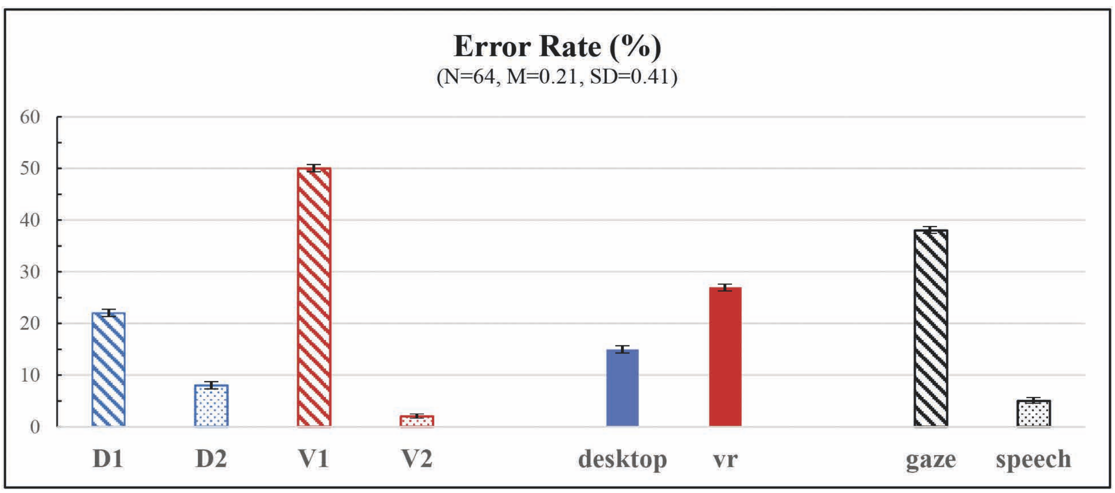

Fig. 12. Results of the error rate measurements. The values are given in percent (%). D1 (Desktop/Pointing), D2 (Desktop/Speech), V1 (VR/Pointing), V2 (VR/Speech).

there were also significant differences of error rate in general (not the recognition rate) regarding input ($p < 0.01$, $F(1, 252) = 53.480$), and output ($p < 0.01$, $F(1, 254) = 7.761$), and finally an interaction was found between input and output ($p < 0.01$, $F(1, 252) = 17.463$).

5.8.2 Usability. The System Usability Scale (SUS) [Brooke et al. 1996] is likely the most popular questionnaire for measuring attitudes toward system usability. It is a reliable and valid measure of perceived usability. Furthermore, it performs as well as or better than commercial questionnaires and homegrown internal questionnaires.

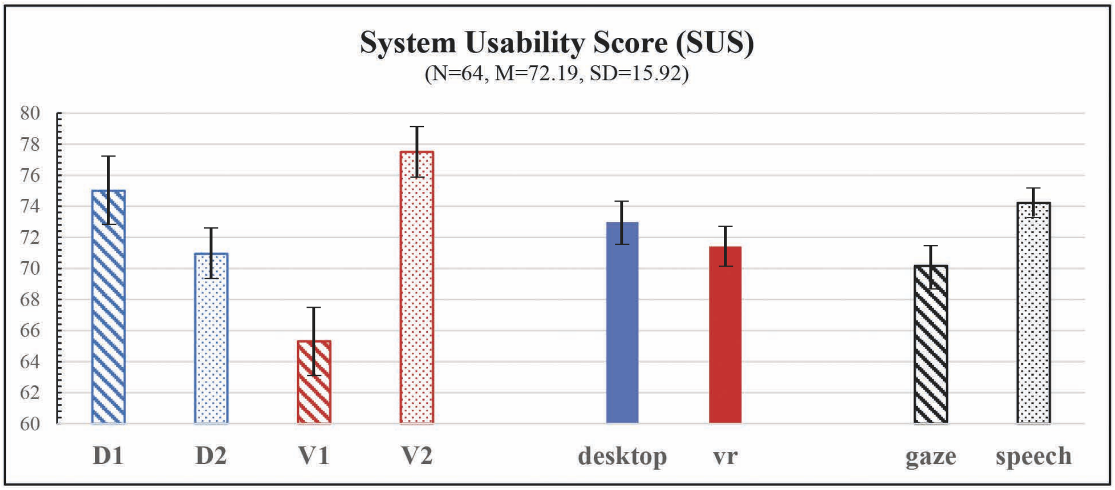

Fig. 13. Results of the System Usability Score (SUS)-Questionnaire. D1 (Desktop/Pointing), D2 (Desktop/Speech), V1 (VR/Pointing), V2 (VR/Speech).

The overall SUS score including all four tasks was on average 72.19 ($SD = 15.92$). V2 had the best SUS score in average ($M = 77.50$, $SD = 13.21$), followed D1 ($M = 75.00$, $SD = 17.18$), D2 ($M = 70.94$, $SD = 12.84$) and V1 ($M = 65.31$, $SD = 17.39$). Regarding the input modes, speech (D2 and V2) was rated higher ($M = 74.22$, $SD = 13.21$) than head pointing (D1 and V1; $M = 70.16$, $SD = 17.89$). Here, a univariate ANOVA pointed out a significance regarding each task ($p < 0.01$, $F(3, 252) = 7.82$), and a significant effect with regard to input mode ($p < 0.05$, $F(1, 252) = 4.56$). The desktop tasks (D1 and D2) had with regard to output mode an average score of 73 ($M = 72.97$, $SD = 15.25$) and the VR tasks had an average score of 71 (V1 and V2; $M = 71.41$, $SD = 16.41$), with no significant difference between them ($p = 0.43$, $F(1, 252) = 0.623$). However, an interaction between input and output was found ($p < 0.01$, $F(1, 252) = 18.22$).

5.8.3 User Experience. We chose the User Experience Questionnaire (UEQ) [Laugwitz et al. 2008] as an end-user questionnaire to measure user experience quickly in a simple and immediate way. Overall, the user experience was rated at −0.116 ($SD = 0.14$) on average on a scale between −3 to 3. A closer look into the results per single task shows that V2 and D2 were rated with a value of −0.11 on average (V2: $M = -0.11$, $SD = 0.09$; D2: $M = -0.11$, $SD = 0.13$). Nevertheless, D1 had an average of −0.13 ($M = -0.13$, $SD = 0.20$) and V1 had an average of −0.12 ($M = -0.12$, $SD = 0.11$). However, there were no significant differences between the tasks ($p = 0.84$, $F(1, 252) = 0.28$). With regard to the input methods, speech was rated with −0.11 on average ($M = -0.11$, $SD = 0.11$) and head pointing with −0.12 ($M = -0.12$, $SD = 0.16$), but also without significant differences . Finally, with respect to the output methods, VR was rated with an average of −0.11 ($M = -0.11$, $SD = 0.10$) and desktop with an average of −0.12 ($M = -0.12$, $SD = 0.17$). Unfortunately, there was no significant differences between the output modes ($p = 0.52$, $F(1, 252) = 0.41$).

However, the data was subjected to a factor analysis, which resulted in the construction of a 26-item questionnaire including the six factors Attractiveness (A), Perspicuity (P), Efficiency (E), Dependability (D), Stimulation (S), and Novelty (N), see Figure 14. A multivariate ANOVA with all six factors as dependent variables showed no significance in terms of A, N, D and E with regard to the single tasks, but P ($p < 0.01$, $F(3, 252) = 4.72$) and S ($p < 0.01$, $F(3, 252) = 5.87$). Regarding the two input methods, we found significant differences based on P

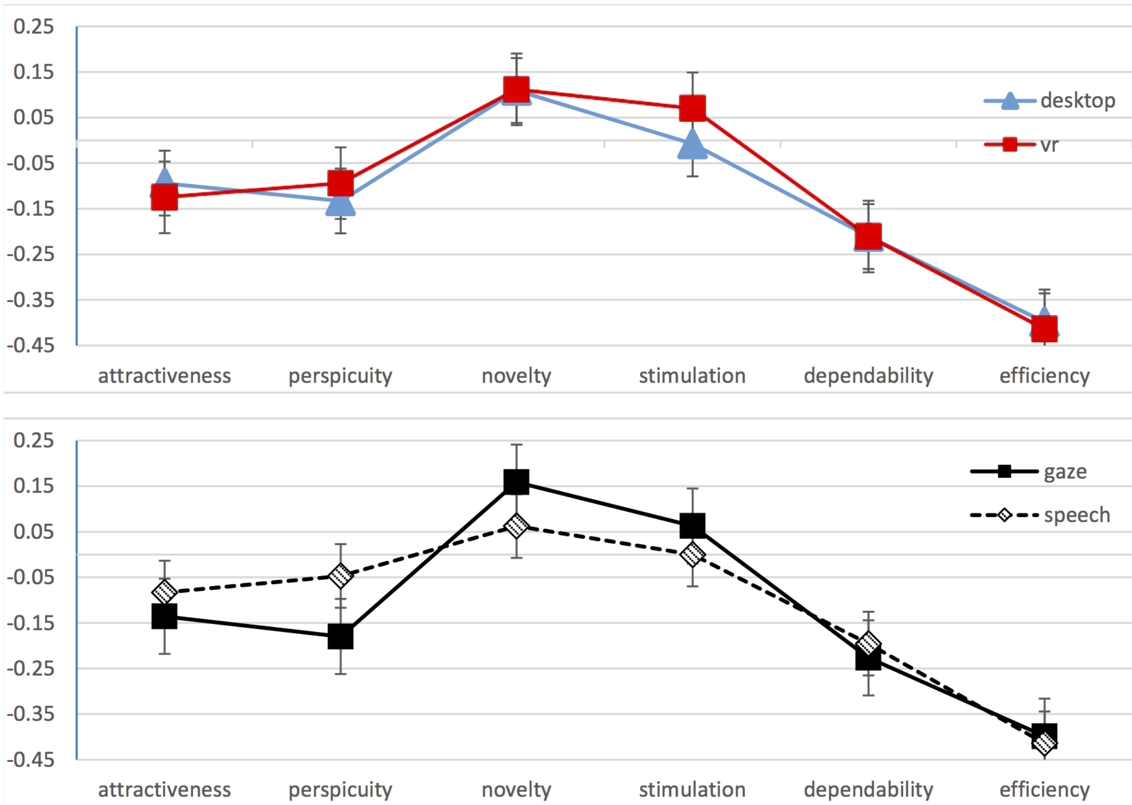

Fig. 14. Results of the User Experience Questionnaire (UEQ).

($p < 0.01$, $F(1, 252) = 9.40$) and a significant effect of N ($p < 0.05$, $F(1, 252) = 4.69$). Nevertheless, an interaction was found between input and output regarding P ($p < 0.05$, $F(1, 252) = 3.93$) and S ($p < 0.01$, $F(1, 252) = 13.15$).

5.8.4 Motion Sickness. For measuring motion sickness, we asked the participants to fill out the Motion Sickness Assessment Questionnaire (MSAQ) [Gianaros et al. 2001], which is a valid instrument for the assessment of motion sickness, and gave a total score of 5% on average ($M = 0.05$, $SD = 0.08$). The results further show that V1 has the highest motion sickness score ($M = 0.09$, $SD = 0.10$), followed by V2 ($M = 0.08$, $SD = 0.08$), D2 ($M = 0.03$, $SD = 0.04$) and D1 with the lowest score ($M = 0.02$, $SD = 0.04$). Pointing had an average score of 6% ($M = 0.06$, $SD = 0.08$) compared to speech with an average of 5% ($M = 0.05$, $SD = 0.07$), but without significant difference between them ($p < 0.01$, $F(1, 254) = 61.23$). Regarding the output, the desktop received a higher score of 2% ($M = 0.02$, $SD = 0.04$) than VR with a lower average of 8% ($M = 0.08$, $SD = 0.09$), with significant differences ($p < 0.01$, $F(1, 252) = 45.16$).

A Multivariate ANOVA with all four MSAQ factors (G: gastrointestinal, C: central, P: peripheral, S: sopite-related) as dependent variables and single task as factor was conducted. It showed, that there are significant differences regarding the single tasks (G: $p < 0.01$, $F(3, 252) = 22.79$; C: $p < 0.01$, $F(3, 252) = 21.15$; P: $p < 0.03$, $F(3, 252) = 3.29$; S: $p < 0.01$, $F(3, 252) = 18.93$). Another Multivariate ANOVA with all four MSAQ factors as dependent variables and input and output as factors was conducted. There were significant differences regarding the output (G: $p < 0.01$, $F(1, 254) = 67.10$; C: $p < 0.01$, $F(1, 254) = 61.23$; P: $p < 0.01$, $F(1, 254) = 6.81$; S: $p < 0.01$,

$F(1, 254) = 54.93$). But apart from that, there were no significances found regarding the input (G: $p < 0.33$, $F(1, 254) = 0.97$; C: $p < 0.27$, $F(1, 254) = 1.25$; P: $p < 0.10$, $F(1, 254) = 2.85$; S: $p < 0.18$, $F(1, 254) = 1.86$). There was also no interaction found between the input and output types (G: $p < 0.59$, $F(1, 254) = 0.30$; C: $p < 0.33$, $F(1, 254) = 0.98$; P: $p < 0.65$, $F(1, 254) = 0.21$; S: $p < 0.93$, $F(1, 254) = 0.01$).

5.8.5 Immersion. In order to measure the immersion and presence of the virtual environment, we used the common immersion questionnaire for VR applications from Slater et al. [Slater et al. 1994]. So, there are no results provided for D1 and D2, because this questionnaire As the SUS (Slater Usoh Steed) questionnaire is designed primarily for VR applications, it was only filled out by the participants after the VR tasks. Moreover, immersion measures are not the forefront of desktop applications and this part of the evaluation was more a plausibility test for our proof of concept including the VRSE model. So these results will serve as a basis for future studies exploring interaction in VR shopping environments. However, the 'SUS Count' shows the mean of the SUS count of '6' or '7' scores amongst the 6 questions. Here, the SUS count for V2 (speech) is slightly higher (1.13) than for V1 (1.00). The 'SUS Mean' uses the mean score across the 6 questions instead. Both VR tasks have nearly equal SUS means, whereas V2 (speech) has an average of 2.76 ($SD = 0.91$), and V1 (head pointing) 2.76 ($SD = 0.86$), which was slightly variable. Here, there was a significant effect concerning the input methods (($p < 0.05$, $F(1, 252) = 4.271$). The overall immersion score of the VR shopping environment was 2.78 ($SD = 0.88$).

5.9 Discussion

In this work the role of speech compared to head pointing input in a desktop setting as well as in VR was investigated. In the following the results of the first experiment will be discussed with a detailed consideration of the objective (task completion time, error rate) and subjective (usability, user experience, motion sickness and immersion) feedback.

5.9.1 Task Completion Time & Error Rate. The analysis of the results of the task completion time demonstrate that searching for a product in the virtual environment by speech input performs faster than using only the user's head pointing, which partly confirms H_3. So it seems to be easier for the users to simply pronounce the product name and let the system search for its location. In speech mode, the system provides a list of search results in realtime and the user only needs to choose one of these, instead of scanning the whole market for the desired product. Although there was an exploration phase before each task, searching for a product only via head pointing faces an additional handicap of spatial knowledge of the environment. However, maybe in smaller shops the input methods could equalize. But even in larger stores the customers are familiar with, they still have to move to the desired product through the whole store, which will always take longer than teleporting.

Regarding the task completion time of the two outputs, the desktop was more efficient than VR with regard to speed and error rate (H_1). On the one hand, this can be explained by the different means of pointing control (mouse vs. head pointing) and the graphical resolution due to the technical setups. Then again, it is to be expected that the desktop output tasks are faster on average than VR, because despite its more natural use of head pointing, the participants were used to controlling the camera by mouse in 3D scenes. But in practice, this is contrary to the subjective feedback of the participants (usability, user experience). The bad performance of head pointing in VR could be also explained by the technical limitations of using WebGL on a smart-phone browser and the still error-prone sensor input of mobile devices. Maybe using a more sophisticated VR HMD and hardware would provide different results. However, this experiment focuses on commodity, affordable and state-of-the-art VR hardware.

5.9.2 Subjective Feedback. Based on research, a System Usability Scale (SUS) score above 68 would be considered above average, and anything below 68 is below average [Gianaros et al. 2001; Sauro 2011]. Due to this rating the overall SUS Score of 72.2 of our VR shop application in the experiment is slightly above the average of 68, we can state that the participants understood the content, and it helped them accomplish their task. The measured SUS score also illustrates that the system adequately communicates what users are required to do with the application, namely find products in the store.

Since differences in User Experience Questionnaire (UEQ) scores for each different tasks, as well as for the input and output modes, hardly exist, all scores are still moderate on average. The hardware is still a bottleneck, so special VR smartphones with a browser specifically developed for WebVR purposes could solve most of the user issues. Nevertheless, the assumptions made in the discussion of the task completion time are reflected by the result of the SUS and UEQ. In summary, the results of the subjective feedback questionnaires shows that VR is preferred by the user in terms of user experience but not of usability. In addition, because there was also no significance found between the output mode regarding user experience and usability, H_2 could only be proved partly. Thus, V2 reached (along with the highest efficiency) the highest usability and user experience score, which can explained by its combination of in- and output methods, i.e. speech and VR. This totally confirms H_4. On closer inspection, VR outperforms the alternative especially in terms of perspicuity and stimulation, two of the six factors of the UEQ [Laugwitz et al. 2008]. This underlines once again that our approach of a hands-free virtual reality shopping environment application can be considered to have higher usefulness than a common 3D desktop variant. So the faster and more immersed the customers' product search is, the higher usability and user experience the system will have. It is worth noting, however, that the results are highly dependent on the speech interface and its recognition, as well as the data quality of the product names to be searched. Whereas head pointing is the commonly used input method for mobile VR applications, our results show that speech input is preferred by the user over head pointing regarding usability and user experience (H_3).

The results of the Motion Sickness Assessment Questionnaire (MSAQ) indicated that VR causes more motion sickness than the desktop, which is emphasized by the found significances. Although this can been seen as obvious, because 3D desktop applications should not cause motion sickness anyway, we have considered it to be very important to use the desktop as a baseline especially for motion sickness. However, with regard to the input method, head pointing causes more motion sickness than speech, which can be explained by the higher physical demand due to the higher amount and frequency of head movements. Another indicator for higher motion sickness can be the task completion time, whereas the participants stayed longer in the environment in the head pointing tasks, which can consequently lead to higher motion sickness scores. However, the overall high scores for motion sickness are largely due to WebVR and its technical limitations, like resolution, rendering performance of the smart-phone and wearing comfort of the HMD.

Finally, the results of the immersion questionnaire showed that the intensity of immersion is not affected by the input method, because both analyzed tasks had the same value. Overall, the values are located around the middle of the scale, so we could assume that the users had a good feeling of presence when using the VR output. The explanation for that is twofold. On the one hand, the render quality and computational power of the mobile device is worse than on a desktop. So low-res textures, restricted lights, and shadows can reduce immersion, which could be solved by using more powerful devices. On the other hand, the inaccuracy of the mobile device sensors can cause a temporary latency of the head movement, which could be avoided by using a device with better-calibrated sensors.

5.9.3 Observations. Besides all the results, there were some interesting and noteworthy observations during the first experiment. Overall, the participants highlighted positively our varied and innovative idea of online shopping. Most participants liked the fact, that they needn't to leave the house to have a realistic shopping experience. They further mentioned that our approach should be taken into consideration for online shopping for disabled and elderly people. Another positive aspect was the "clear and tidy design" of the shopping environment. However, almost every participant saw the speech input functionality as the most important benefit. More than 60% of the participants reported that they would use the application at home if they could. But there were also varying opinions and negative aspects reported, such as the fixed position of the speech interface. Some participants found it realistic, similar to terminals and consoles in real supermarkets. On the other hand, it requires the user to move there in order to use it, so it would be better to place it dynamically at the user's position. But based on the majority of feedback in the pilot study, we chose the fixed position. The most negative feedback was given about the product search by head pointing interaction only. In general, the head pointing search wasn't perceived to be pleasant, with almost all participants complaining about its efficiency. Despite a detailed exploration phase before each task, the participants had orientation problems in the environment and proposed a mini-map or at least signposts for better orientation. Here, it should be mentioned that the virtual shop was rather small, with an area which equals ca. 200 square meters, so the orientation drawback would increase dramatically in larger stores. Nevertheless, there were also negative aspects with regard to the VR modes. Although the idea of placing the control panel onto the handle of the shopping cart was highlighted very positively, the buttons were positioned too low, because some users had problems with looking down. In general, the main problems in VR appeared when the user had to fix their view on an object (e.g. for selection). In particular, the low render quality of WebVR and thus the blurry textures amplified this problem.

6 DESIGN GUIDELINES FOR VIRTUAL REALITY SHOPPING ENVIRONMENTS

In the field of Human-Computer Interaction (HCI) the user-centered design of interactive systems is extremely important, i.e. the needs and desires of the users constitute the starting point for all development activities. After an in-depth research, customer surveys and pilot tests should be conducted to collect and use data (quantitative and qualitative) from the target group and stakeholders regarding the planned task to determine what content the users want and where. Therefore, the methodology of developing our VR shopping environment was „moving from task to design". We wrap up this section with a set of general design guidelines for VR shopping environments, highlighting some of the important points discussed earlier (see Section 4.2) and lessons we learned while moving from the task of searching products to the design and development of a VR shopping environment. When designing for experiences in VR a new set of design considerations comes into play than when designing for 2D screens. To help coming VR designers and developers of VR shops to create experiences that doesn't frustrate or make users feel nauseous, we created the following design principles to guide the work. The following guidelines include actionable insights on how to optimize performance, usability, satisfaction, and experience for the customers of VR shops:

6.1 Choose speech as the appropriate input and VR as output modality for mobile VR shops with regard to optimal performance and preference.

Choose the appropriate input and output device by analyzing the application's goals and user's needs. As in our example of an online shopping store, consider a virtual environment that allows customers to explore a virtual store and search for products. In case of collaborative shopping, a HMD or small screen would probably not be appropriate, since a group of people are not able to see the display at the same time. Then a large-screen display would be the better choice. In contrast, using a HMD allows the customer to browse through the virtual store without being annoyed or distracted by other customers. According to participants of our customer survey, they tend to keep their shopping trip as short as possible, especially in overcrowded stores. However, as we showed in our case study, wearing a HMD influences the efficiency of finding a product in a negative way. But in terms of user experience and usability, it would be more appropriate to choose a HMD as output device. Here, the error rate is the crucial factor for the input method decision. The decision between speech and head pointing depends mainly on the speech recognition rate concerning the product name database, too. So, if the error ratio grows up above a certain threshold (about 10%, see Figure 12), it would be better to choose desktop screen for output.

Nevertheless, concerning only mobile and online scenarios, and if the data quality allows for this, speech input wearing a HMD should be chosen as it outperforms the others in all aspects (performance and preference).

6.2 Respect the main benefits while avoiding the mistakes that surrounded online as well as offline shopping environments.

It is important to pay attention to the main characteristics of on- and offline shopping. According to customer survey guidelines, the main strength of current online shops is their search functionality in contrast to physical stores, where the customer needs to find and ask employees to find products. But the use of search functionality involves serious disadvantages, if it is implemented badly or the data quality is insufficient. Moreover, bigger ranges of products require proper and clear filters and categories. In our prototype, we used shelves and sectors/departments as a visual means of structuring products. But with a look at larger physical or online stores, too many or unstructured categories and sectors can be very confusing for the customers and therefore influence their satisfaction because of a lack of clarity. Here, if sorted properly, colored signs and understandable shelf names can help. Another aspect which has been tried and tested in physical stores concerning product search, interactivity and efficiency are in-store employees and consultants or even interactive terminals. In VR this could be realized by using a virtual assistant with whom the customer can interact via speech input. Here, in case of product search, textual input in VR shops should be avoided, because text entry in VR has not been adequately researched. Especially in common mobile VR scenarios, the user hasn't even a controller (e.g. keyboard or remote) in order to entry text, apart from head pointing or speech input. In our scenario, we used the simplest form of a virtual assistant, more precisely a product search interface using speech input, which allowed the users to search for products by saying the desired product name instead of typing.

6.3 Design simple and relevant interfaces.

Virtual reality is a carefully crafted illusion. The better the experience looks, the better the illusion and immersion. So, take every opportunity to optimize the visuals of your virtual environment. But as mentioned before, the customer wishes a structured store, not confusing, not too big range of products, and the store should not be too large. So it is up on the designers to build a virtual environment as simple as possible, but still keep the feeling of being in a store looking for products. Use virtual signs and other visual means (like shelves, sectors or rooms) for a better clarity, but no overfilling with products. Here, a clearly arranged assortment and intuitive interaction technique eases a fast and easy selection of products. Furthermore, choose colors to be as neutral as possible and adjust contrast of colors to prevent positive or negative impact to let the users keep track of the products. Again, do not overfill the shelves with products, because in common large offline stores, when customers are standing in front of a shelf searching for a product, it is like they cannot see the wood for the trees. In virtuality, there is no need to overfill the shelves with duplicates of products. Moreover, the environment should be designed displaying content upholding the immersion and experience, for example, put the 3D products into a virtual shopping cart instead of using a list representation. In addition, due to user feedback in our pilot study, we decided to place the speech interface panel static on a wall in the scene and a control panel at the cartâĂŹs handle displaying the total price. Displaying a menu in front of the user's point of view has led to negative feedback with regard to immersion and user's experience.

6.4 Avoid high resolution graphics and unnecessary head movements to reduce motion sickness in virtual shopping environments.

Motion sickness is an aversive behavioral state that affects several psycho-physiological response systems. There are individual differences in the extent to which motion sickness symptoms (e.g., nausea or dizziness) are experienced and different contexts that cause motion sickness (e.g., visual simulators or vehicles) may cause a symptom. However, as motion sickness can have a strong effect on the shopping experience, all efforts should be made to prevent any motion sickness symptoms. As they are negligible in a desktop setting, there are different methods for reducing motion sickness when using a HMD. In our prototype, we rendered a virtual nose [Whittinghill et al. 2015], set the field of view to a comfortable value, and keep the velocity on a pleasant and constant level [Mourant and Thattacherry 2000; Tanaka and Takagi 2004] while moving in the virtual store. Therefore, we recommend to use speech input instead of head pointing only for the product search task whenever possible in a VR shop to further reduce unnecessary movements. Latency due to technical limitations, rendering and network performance of mobile online VR shops is another factor. If latency is high, the task should require as few head movements as possible. So, the developers and designers should use less lights, if at all possible, low quality shadows, and lower resolution textures and 3D models of the products, shelves and other shop elements. Another indicator for motion sickness is the task completion time in VR shopping scenario. According to our experimental results, users stay longer in the environment in the head pointing tasks, which consequently has led to more head movements and consequently to higher sickness ratings. Furthermore, because we decided to include movement via teleportation, we tried avoid „hard" and abrupt cuts at all costs, as this may be particularly sensitive for motion sickness and immersion. For this, we were inspired by transitions in movies, e.g., dissolves or „irises" (seen as the eyelids closing down over or opening up on a shot). All these parameters should be customizable for each user.

7 CONCLUSION

Stronger demand for immersive virtual reality (VR) applications has led to a growing market for head-mounted displays (HMDs). With this popularity, the computation and graphics power of mobile devices also increases, which could catalyze their stagnating tendency. By using smart-phone VR, it is possible to use immersive VR applications nearly everywhere, since the availability of mobile web also rises. Besides the gaming sector, use cases can be found in the constantly growing e-commerce sector, which is used by almost every smart-phone user nowadays, whether in online web shops or in-app purchases.

Current online shops may be functional and efficient, but do not provide enough of an immersive shopping experience. So this paper focuses on developing an immersive virtual shopping environment that includes and combines the major advantages of online and physical stores and tries to tear down limitations of e-commerce and VR. In order to provide a realistic setting, we evaluated hands-free interaction techniques using smart-phone VR in a comparative user study. Because the smart-phone itself is already available for most users, it is the cheapest and most affordable scenario for VR online shops. Our system allows the user to search for products by head pointing only or in combination with speech input in a virtual environment. The head pointing technique for product selection was inspired by the eBay VR shop, whereas their approaches lacked in efficiency and experience we included the speech interface into our proof of concept. Finally, in order to evaluate our system, we conducted a case study including a search task in the virtual environment using a desktop monitor and smart-phone as output devices, and head pointing and speech as input. Besides task completion time and error rate measurements, the participants were asked to fill out questionnaires about their preferences. It could be confirmed that the overall usability of the system is slightly above average, but not significantly; the VR setup with speech input was especially preferred. Regarding efficiency, speech input was more efficient than using head pointing, but it should be mentioned that proper data quality is essential to provide optimal recognition. In this respect, it is worth mentioning that our prototype of a mobile online VR shop was measured to be inferior to non-mobile and non-online solutions. So, however, whereas the task performance results regarding VR with speech input were quite well, there is still much research and development needed to improve the user's preference, more precisely usability, user experience and motion sickness, of mobile online shopping environments. Nevertheless, our concept includes availability and sustainability instead of only adaptability [Chittaro and Ration 2000]. Moreover, besides usability as a crucial factor for shopping simulators [Laver et al. 2012], we provided immersive VR for output. Here, using 3D model representations of the shelves and products turned out to be very important for the user experience, and preference at all, according to user feedback and observations in the experiment. The study also demonstrated an immersive shopping experience in VR; whereas VR causes higher motion sickness than the desktop, with technological progress there will be more ways to reduce these symptoms in the future. Overall, the results were very positive and the outcome of our study was accentuated by the throughout positive resonance and feedback of the participants; however, because we looked at a virtual online grocery store with a manageable number of products, shelves and departments, these findings may not translate to larger stores or different types of goods (e.g., furniture or clothes shopping). So, further studies should be conducted in order to explore differences of VR shop sizes, dimensions, or types of goods. However, it could be a while until such VR online stores be real and market-ready due to a very slow progress of the digitization of the retail sector, in particular with regard to groceries. More precisely, the data quality lacks of digitized market layout data or even illegible product names, descriptions and pictures prepared for the user's need. Although current product management systems offer detail product information like its dimensions, weight and location in the market, but neither the availability of 3D models of the products nor 360Âř pictures/video or even high resolution images is guaranteed. However, when breaking new ground, it is likely that there are still a lot of gaps in the knowledge base that need to be filled. Therefore, the next steps investigating alternative store visualizations and interaction concepts, based on these results, would be to build a stronger overall evidence base.

Shopping in virtual reality environments is not a new concept. Deciding whether the users need mobile VR shops is important and then following a user-centered design process offers the best chance of success. In addition, it is important to use standardized and verified metrics for evaluating and comparing future VR shops with regard to shopping experience, including task performance and user's preference. Furthermore, there are significant differences in the way that input methods as well as output modalities operate in VR compared to other and there is a need to make sure that those differences are handled with care to deliver the best result. Here, our experimental results indicate that using speech input wearing a HMD might be the best option for mobile interactive VR shops. Overall, it is worth to consider the benefits from online, such as search functionality and availability, as well as physical stores, which provide more personality and orientation.

8 FUTURE WORK

In summary, our VR shop creates compelling online virtual sensory richness through which customers can experience the value of the product information more richly and engage in a more active shopping activity, compared to ordinary online shopping applications. Similar to the work of Ohta et al. [Ohta et al. 2015], our proposed hands-free VR shopping system can assist disadvantaged shoppers, so they can do shopping as if it were done in offline stores. Using head pointing the user can move freely around the store while viewing the 3D representation of the store displayed in mobile VR. Products can be searched, selected and freely viewed. This would provide disadvantaged customers a better shopping experience, because they are limited in their ability to go out when and where they want [Ohta et al. 2015]. But with regard to the user interface and design, this scenario would require conducting a deeper customer and user analysis, specially tailored for disadvantaged shoppers.

In an ordinary shopping mall, customers have to use a rather plain user interface, leading to lower customer satisfaction. This might cause customers to become passive observers, merely observing the information. Whereas VR customers are engaged in the inspection and control of the 3D visualized target products, due to the virtual sensory richness driven by the 3D environment and user interface. But information overload in VR is more likely and should be avoided. Therefore, different layouts and representations of VR shops should be explored and compared, like graph-based approaches, or even more abstract concepts like searching in a virtual apartment.

So, another aspect what should be studied in more detail in the future is the visualization of a virtual store. Our virtual shopping environment was based on existing layout data, but due to performance issues of WebVR and smart-phone we chose a smaller retail space of about 180 square meters within one single floor. Besides different store layouts, future studies should explore differences between the store size in all three dimensions, i.e., different number of floors and sizes of the market area. Because people tend to think two-dimensionally, and even the front-back axis is more accessible than the left-right axis [Ishikawa and Montello 2006; Werner et al. 1997]. So, it might be easier for customers to orientate in virtual stores or malls with less floors and less turnings to left or right. As our work focus on evaluating common input and output modalities there is still much work to be done regarding VR user interfaces especially for shopping applications. We suspect that common UI/UX guidelines as they are existing for web shops or mobile shopping have to be adapted for VR, e.g., recommendations for colors, fonts, shapes, and sizes of UI elements. So, VR UI designers should be careful with too bright and overloading colors, because the userâĂŹs eyes are only a couple centimeters away from the screen. In addition, too thin fonts or fonts with serifs can be very hard to read, because there are simply not enough pixels for a clear view of the text. And because the VR technology itself is not yet mature, long stays in VR malls carry the implicit danger of simulator sickness. So there should be more focus on reducing motion sickness when traveling through virtual stores, for example by avoiding movement and using high-res VR hardware. However, the VR shopping mall in our experiment is a trade-off, because it is inexpensive, multi-platform and easy to use, due to its hardware and low complexity. So future studies should consider these important factors for VR e-commerce applications.

While our VRShop, the VRSE model and VR technologies in general have the potential to improve the experience of consumers in electronic shopping environments, more research is needed in order to obtain a better understanding of the factors that determine the degree of usefulness of such technologies.

REFERENCES

Ritu Agarwal and Elena Karahanna. 2000. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Q.* 24, 4 (Dec. 2000), 665–694. https://doi.org/10.2307/3250951

Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121–136. https://doi.org/10.1016/j.cag.2012.12.003

C. Han Bei Bei, N. S. Chaudhari, and J. C. Patra. 2005. Design of a virtual shopping mall: some observations. In *2005 International Conference on Cyberworlds (CW'05)*. 4 pp.–348. https://doi.org/10.1109/CW.2005.37

Darren Black, Nils Jakob Clemmensen, and Mikael B Skov. 2010. Pervasive Computing in the Supermarket: Designing a Context-Aware Shopping Trolley. *International Journal of Mobile Human Computer Interaction (IJMHCI)* 2, 3 (2010), 31–43.

Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2004. *3D User Interfaces: Theory and Practice.* Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.

John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability Evaluation in Industry* 189, 194 (1996), 4–7.

M. Buffa and J. C. Lafon. 2000. 3D virtual warehouse on the Web. In *Information Visualization, 2000. Proceedings. IEEE International Conference on.* 479–484. https://doi.org/10.1109/IV.2000.859800

T. Chen, Z. g. Pan, and J. m. Zheng. 2008. EasyMall - An Interactive Virtual Shopping System. In *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, Vol. 4. 669–673. https://doi.org/10.1109/FSKD.2008.124

Luca Chittaro and Roberto Ration. 2000. Adding adaptive features to virtual reality interfaces for e-commerce. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems.* Springer, 86–97.

Olga De Troyer, Frederic Kleinermann, Haithem Mansouri, Bram Pellens, Wesley Bille, and Vladimir Fomenko. 2007. Developing semantic VR-shops for e-Commerce. *Virtual Reality* 11, 2-3 (2007), 89–106.

Ajoy S Fernandes and Steven K Feiner. 2016. Combating VR sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI).* IEEE, 201–210.

Gonzalo Garcia-Perate, Nicholas Dalton, Ruth Conroy-Dalton, and Duncan Wilson. 2013. Ambient Recommendations in the Pop-up Shop. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13).* ACM, New York, NY, USA, 773–776. https://doi.org/10.1145/2493432.2494525

Sven Gehring, Markus Löchtefeld, Florian Daiber, Matthias Böhmer, and Antonio Krüger. 2012. Using intelligent natural user interfaces to support sales conversations. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces (IUI '12).* ACM, New York, NY, USA, 97–100. https://doi.org/10.1145/2166966.2166985

Peter J Gianaros, Eric R Muth, J Toby Mordkoff, Max E Levine, and Robert M Stern. 2001. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space, and Environmental Medicine* 72, 2 (2001), 115.

David A Griffith, Robert F Krampf, and Jonathan W Palmer. 2001. The role of interface in electronic commerce: Consumer involvement with print versus on-line catalogs. *International Journal of Electronic Commerce* 5, 4 (2001), 135–153.

Gerald Häubl and Pablo Figueroa. 2002. Ineractive 3D presentations and buyer behavior. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM, 744–745.

Toru Ishikawa and Daniel R. Montello. 2006. Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology* 52, 2 (2006), 93 – 129. https://doi.org/10.1016/j.cogpsych.2005.08.003

Howell Istance, Richard Bates, Aulikki Hyrskykari, and Stephen Vickers. 2008. Snap Clutch, a moded approach to solving the Midas touch problem. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications (ETRA '08)*. ACM, New York, NY, USA, 221–228. https://doi.org/10.1145/1344471.1344523

Sirrka L. Jarvenpaa and Peter A. Todd. 1996. Consumer Reactions to Electronic Shopping on the World Wide Web. *Int. J. Electron. Commerce* 1, 2 (Dec. 1996), 59–88. https://doi.org/10.1080/10864415.1996.11518283

K. Kendall and J. Kendall. 2002. *Systems Analysis and Design*. Prentice Hall PTR. https://books.google.de/books?id=hF8FcgAACAAJ

Mirae Kim, Min Kyung Lee, and Laura Dabbish. 2015. Shop-i: Gaze Based Interaction in the Physical World for In-Store Social Shopping Experience. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1253–1258. https://doi.org/10.1145/2702613.2732797

Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.

K. Laver, F. Lim, K. Reynolds, S. George, J. Ratcliffe, S. Sim, and M. Crotty. 2012. Virtual Reality Grocery Shopping Simulator: Development and Usability in Neurological Rehabilitation. *Presence* 21, 1 (Feb 2012), 183–191. https://doi.org/10.1162/PRES_a_00098

Kun Chang Lee and Namho Chung. 2008. Empirical analysis of consumer reaction to the virtual reality shopping mall. *Computers in Human Behavior* 24, 1 (2008), 88 – 104. https://doi.org/10.1016/j.chb.2007.01.018

JJ-W Lin, Henry Been-Lirn Duh, Donald E Parker, Habib Abi-Rached, and Thomas A Furness. 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Virtual Reality, 2002. Proceedings. IEEE*. IEEE, 164–171.

Ming Liu. 2013. A study of mobile sensing using smartphones. *International Journal of Distributed Sensor Networks* 2013 (2013).

Gerald L. Lohse and Peter Spiller. 1998. Quantifying the effect of user interface design features on cyberstore traffic and sales. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 211–218. https://doi.org/10.1145/274644.274675

Meredith Ringel Morris, Kori Inkpen, and Gina Venolia. 2014. Remote Shopping Advice: Enhancing In-store Shopping with Social Technologies. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 662–673. https://doi.org/10.1145/2531602.2531707

Ronald R Mourant and Thara R Thattacherry. 2000. Simulator sickness in a virtual environments driving simulator. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 44. SAGE Publications, 534–537.

Fabio Nemetz. 2000. The Missing Link: Multimedia and e-Commerce. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00)*. ACM, New York, NY, USA, 65–66. https://doi.org/10.1145/633292.633333

Masaya Ohta, Shunsuke Nagano, Seiya Takahashi, Hiroki Abe, and Katsumi Yamashita. 2015. Mixed-reality shopping system using HMD and smartwatch. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 125–128. https://doi.org/10.1145/2800835.2800888

B Joseph Pine and James H Gilmore. 1998. Welcome to the experience economy. *Harvard Business Review* 76 (1998), 97–105.

Andrea Sanna, Bartolomeo Montrucchio, Paolo Montuschi, and C Demartini. 2002. 3D-dvshop: A 3D dynamic virtual shop. In *Multimedia 2001*. Springer, 33–42.

Jeff Sauro. 2011. Measuring usability with the system usability scale (SUS). (2011). http://www.measuringu.com/sus.php

Johan Schalkwyk, Doug Beeferman, Françoise Beaufays, Bill Byrne, Ciprian Chelba, Mike Cohen, Maryam Kamvar, and Brian Strope. 2010. âĂIJYour word is my commandâĂİ: Google search by voice: A case study. In *Advances in Speech Recognition*. Springer, 61–90.

Mike Schuster. 2010. Speech recognition for mobile devices at Google. In *Pacific Rim International Conference on Artificial Intelligence*. Springer, 8–10.

A Fleming Seay, David M Krum, Larry Hodges, and William Ribarsky. 2002. Simulator sickness and presence in a high field-of-view virtual environment. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM, 784–785.

Sarah Sharples, Sue Cobb, Amanda Moody, and John R. Wilson. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29, 2 (2008), 58 – 69. https://doi.org/10.1016/j.displa.2007.09.005 Health and Safety Aspects of Visual Displays.

Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144.

Nobuhisa Tanaka and Hideyuki Takagi. 2004. Virtual reality environment design of managing both presence and virtual reality sickness. *Journal of Physiological Anthropology and Applied Human Science* 23, 6 (2004), 313–317.

Dadong Wan. 2000. Magic Home: Exploiting the Duality Between the Physical and the Virtual Worlds. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00)*. ACM, New York, NY, USA, 53–54. https://doi.org/10.1145/633292.633326

Steffen Werner, Bernd Krieg-Brückner, Hanspeter A. Mallot, Karin Schweizer, and Christian Freksa. 1997. *Spatial Cognition: The Role of Landmark, Route, and Survey Knowledge in Human and Robot Navigation1*. Springer Berlin Heidelberg, Berlin, Heidelberg, 41–50. https://doi.org/10.1007/978-3-642-60831-5_8

DM Whittinghill, Bradley Ziegler, T Case, and B Moore. 2015. Nasum Virtualis: A simple technique for reducing simulator sickness. In *Games Developers Conference (GDC)*.

C. Y. Yang, W. K. Chiou, D. H. Huang, Y. T. Peng, B. H. Chen, Y. H. Yang, and P. S. Tsai. 2009. Application of Virtual Reality Technology in Online Fashion Store Development. In *2009 International Conference on Computational Intelligence and Software Engineering*. 1–6. https://doi.org/10.1109/CISE.2009.5364266

Soojeong Yoo and Callum Parker. 2015. Controller-less interaction methods for Google Cardboard. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*. ACM, 127–127.

L. Zhao and N. Zhang. 2012. The virtual reality system in electronic commerce. In *2012 IEEE Symposium on Robotics and Applications (ISRA)*. 833–835. https://doi.org/10.1109/ISRA.2012.6219320

Received February 2017; revised May 2017; accepted June 2017