Searching for Products in Virtual Reality: Understanding the Impact of Context and Result Presentation on User Experience

Austin Ward*
austinrw@unc.edu
UNC Chapel Hill

Sandeep Avula
sandeavu@amazon.com
Amazon

Hao-Fei Cheng
hchenga@amazon.com
Amazon

Sheikh Muhammad Sarwar
smsarwar@amazon.com
Amazon

Vanessa Murdock
vmurdock@amazon.com
Amazon

Eugene Agichtein
eugeneag@amazon.com
Amazon

Figure 1: SERP conditions: (a) DETAILED MENU, (b) BASICMENU, and (c) NOMENU. Context alignment conditions: (d) HIGHDETAIL, (e) LOWDETAIL, and (f) NOROOM

ABSTRACT

Immersive technologies such as virtual reality (VR) and head-mounted displays (HMD) have seen increased adoption in recent years. In this work, we study two factors that influence users' experience when shopping in VR through voice queries: (1) context alignment of the search environment and (2) the level of detail on the Search Engine Results Page (SERP). To this end, we developed a search system for VR and conducted a within-subject exploratory study (N=18) to understand the impact of the two experimental conditions. Our results suggest that both context alignment and SERP are important factors for information-seeking in VR, which present unique opportunities and challenges. More specifically, based on our findings, we suggest that search systems for VR must be able to: (1) provide cues for information-seeking in both the VR environment and SERP, (2) distribute attention between the VR environment and the search interface, (3) reduce distractions in the VR environment and (4) provide a "sense of control" to search in the VR environment.

CCS CONCEPTS

• **Information systems → Information retrieval**; **Search interfaces**; • **Human-centered computing → Interactive systems and tools**; **Virtual reality**.

KEYWORDS

Virtual Reality, Information Retrieval, User Study

*Work done while at Amazon

SIGIR '23, July 23–27, 2023, Taipei, Taiwan

ACM ISBN 978-1-4503-9408-6/23/07.
https://doi.org/10.1145/3539618.3592057

ACM Reference Format:
Austin Ward, Sandeep Avula, Hao-Fei Cheng, Sheikh Muhammad Sarwar, Vanessa Murdock, and Eugene Agichtein. 2023. Searching for Products in Virtual Reality: Understanding the Impact of Context and Result Presentation on User Experience. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23), July 23–27, 2023, Taipei, Taiwan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3539618.3592057

1 INTRODUCTION

Immersive technologies, such as virtual reality (VR) head-mounted displays (HMD) , have seen increased adoption in recent years. As more consumers have VR devices at home, the e-Commerce industry has an opportunity to provide satisfying shopping experiences for customer needs that benefit from exploration in a virtual environment. This is especially beneficial for customers who may find it challenging to assess an item's suitability unless they've placed the item in a physical context (such as a rug for the living room, or a picture for the den). Broadly, there have been two approaches to realize shopping in VR. In the first, an immersive environment mimics a real world shopping store and customers navigate through the store in VR as they would in real life [20, 22]. In the second, the customers are situated within a plausible context for which they are planning to shop, for example, a chair for their living room or a desk for their workspace [25]. Our study focuses on the second approach, i.e., supporting users situated in a context for which they are making a purchase. In this work, we take a first step toward designing an interactive retrieval system for 3D objects in VR, by studying the influence of the context alignment in VR and the level of detail on the Search Engine Results Page (SERP) to provide a satisfying search experience in VR.

We manipulate two aspects crucial to the immersive experience: (1) context alignment and (2) Search Engine Results Page (SERP). We define context alignment as a measure of overlap between a target environment and a search environment. Target environment is the physical space rendered in 3D for which the user is purchasing the items, and a search environment is the 3D space within which they will be placed to shop for the items. We manipulate the representation of the search environment across three conditions (Figure 1): (1) **HighDetail**, (2) **LowDetail**, and (3) **NoRoom**. Prior work suggests that providing users with a high-quality environment , i.e., high fidelity image and texture quality, helps them with recall [6, 18, 19], and supports their decision-making process [10, 24]. However, rendering a high-quality environment is not always possible . For this reason, alternative approaches exist to support environment rendering, such as rendering only the object geometry using LiDAR [17]. Therefore, we are interested in studying the impact of accurately representing an environment for VR shopping. In the **HighDetail** condition, the overlap between the target environment and search environment is exact, as we provide a replica of the target environment with high-quality rendering. In the **LowDetail** condition, the search environment is devoid of colors and textures in comparison to the target environment, but retains the geometry. The **NoRoom** serves as a baseline condition, in which the environment in completely empty i.e., there is no overlap between the target environment and the search environment.

With regard to the SERP, our goal is to understand how (and if) people make use of information presented on the SERP. Prior work in Information Retrieval (IR) suggests that users employ attributes of the SERP (such as the presence of query terms in the snippet, the position of the result, and the number of relevant results) to quickly assess if the result is relevant [12, 21]. Additionally, the process of searching is dynamic in that the user gains expertise through their interactions with the search engine [5]. Our goal is to understand how the details on a SERP impact the user's experience when searching in VR. Therefore, in this study we investigate three designs (Figure 1): **DetailedMenu**, **BasicMenu**, and **NoMenu**. We modulate the level of information displayed on the SERP to study its impact on participants' user experience. In DetailedMenu, participants viewed the top three search results with the product title, product description, and thumbnail image. In BasicMenu, participants viewed the: product title and thumbnail image, with no description. Finally, in the NoMenu condition, there was no UI menu to display the search results. Instead, the top result from the search system was automatically selected for the user to place in the environment (Figure 2).

Overall, we conducted a within-subject user study with 18 participants to study the impact on participants' experience of: (1) context alignment and (2) details on the SERP. We define user experience across three dimensions: (RQ1) workload, (RQ2) perceptions of user engagement, and (RQ3) perceptions of system usefulness.

2 RELATED WORK

Shopping in VR relies on contextual memory as the VR environment may not be a perfect rendering of the the real life environment. Shin et al. [19] found that items are better recalled when the target context is aligned with the study context, a result echoed in Essoe et al. [6] studying language learning in a VR setting. In terms of e-Commerce, this suggests that a customer will be better able to recall key aspects of items they are looking for, if the visual context in the VR setting is aligned to the the target setting. However, Wälti et al. [23] examined whether the richness of the visual context in a VR setting aided in contextual memory, and found that it did not. Through our study, we extend the body of literature by investigating whether the fidelity of the search environment to the target environment is important for retrieval tasks.

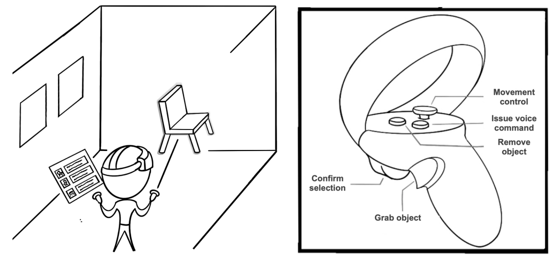


Figure 2: Using the left controller, participants moved or pinned the SERP. With the right controller, they interacted with 3D objects and issued voice queries. Controls: A - Voice query, B - Remove, Grip - Grab, Trigger - Confirm.

In the early days of search engines, Tombros and Sanderson [21] found that showing a search results summary that contained terms matching the user query improved the user's ability to assess the relevance of the document, both in terms of accuracy and speed. These findings were further strengthened with eye-tracking studies which found that users tend to acquire new terms from search results for query reformulations to improve their search queries [4]. Furthermore, Kelly and Azzopardi [11] found that the number of search results on a page influenced users exploration and their difficulty finding information. While these findings assert the importance of a SERP, translating these findings across different modalities hasn't been straightforward. For instance, in voice search, a fallback strategy for many frustrated users has been to reformulate their queries [9]. While this may suggest that voice response from a system should be rich with relevant terms, prior work has also found that users prefer short and concise responses [7], which makes it challenging to surface all the relevant keywords in a voice-friendly response.

3 SYSTEM DESIGN

We built the VR shopping experience for the Meta Quest 2 platform using the Unity game engine. Participants interacted with three major components: (1) search system, (2) SERP in VR, and (3) VR environment. Based on the study order (Section 4), participants were placed in a specific manipulation of the VR environment. Once in the environment, they could use the VR controllers and search by voice to shop for 3D objects (Figure 2). The task information and SERP were anchored to left hand controller, much like a painter's palette. The right hand controller was used for interacting with the UI menus, initiating voice search, also inserting, moving, and removing (unwanted) furniture.

To allow participants to retrieve 3D furniture objects and insert them into the scenes, we built a retrieval dataset for our study from a subset of the 3D objects from the Amazon Berkeley Objects Dataset [3]. We selected a subset of 2000 objects related to the task objectives (for example, we removed all instances of phones and shoes as these were not related to any of the search task objectives).

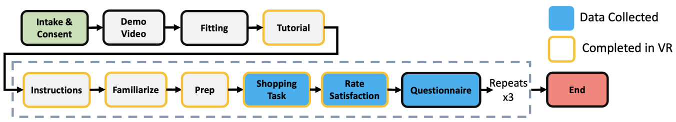


Figure 3: Study Procedure.

Participants retrieved furniture using a voice-based search system. The search system was composed of three components: (1) a ranking component based on the BM25 implementation from Pyserini framework [13]; 2) a search result diversification component that we implemented the based on the Maximal Marginal Relevance (MMR) algorithm [2]; and 3) a Query Expansion (QE) component based on the double-metaphone algorithm [16]. Finally, the context alignment and SERP conditions were introduced in Section 1 (see Figure 1 to get a view of these conditions in VR).

4 METHODS

We conducted a within-subject laboratory study with 18 participants (F=8, M=10) recruited within our organization. Participants reported their prior experience with virtual reality technologies as: None (N=2), Little (N=3), Some (N=11), and A Lot (N=2).

4.1 Study Protocol and Design

Study Procedure: In an hour-long session, participants completed three search tasks in a VR environment after watching a demonstration video and completing a step-by-step tutorial on navigation, voice search, and object manipulation. The tasks involved being temporarily placed in an empty space with instructions, then being placed in a target room for at least 60 seconds, familiarizing themselves with the room as if furniture shopping, and finally being briefed on the experimental condition before performing the search task(Figure 3).

Study Design: Each participant was exposed to three levels of context alignment and three search results presentations (i.e. within subject design). To control for learning effects of the two variables, we varied the order of exposure for each of these conditions using two Latin-squares. We used the cross-product of two Latin squares which gave us 9 orderings, and repeated this twice, thereby generating 18 orderings for all the participants[1] Each participant received a $50 gift card for their participation.

4.2 Search Tasks

To prepare participants to shop for furniture, we simulated three scenarios [1]: (1) Studio, (2) Kitchen, and (3) Office. In each of the three tasks, we asked participants to imagine that the target room they were in belonged to a friend who was still in need of a couple of pieces of furniture for that specific room. We asked participants to search for furniture that would meet the needs of the requester (friend) and match the color and materials of the other furniture in the room. In this study, we did not want to test for differences in the type of the room. We randomly ordered the rooms for each search session.

[1]After a power analysis using G*Power, we determined 18 participants were needed for 80% power in a within-subjects ANOVA with three conditions, alpha=0.05, and a large effect size (partial eta squared = 0.3).

4.3 Study Data and Analysis

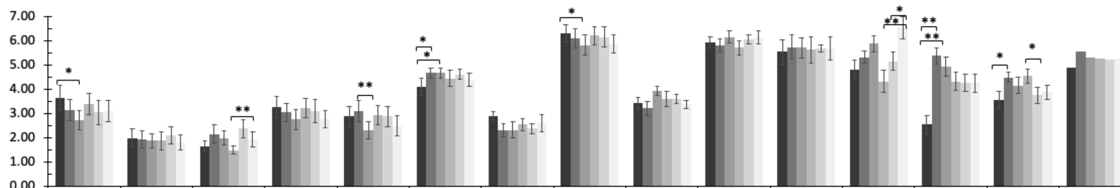
To answer our three research questions on: (RQ1) workload, (RQ2) perceptions of user engagement, and (RQ3) system usefulness, we collected post-task questionnaire responses that participants answered after each shopping task. For RQ1, we used the NASA-TLX [8]. For RQ2, we used four questions adapted from the User Engagement Scale [15]. In RQ3, we asked four questions about helpfulness of (a) the VR environment and (b) how the search results were presented in assisting them to complete the task, (c) satisfaction with the search results and (d) satisfaction with their final furniture selection. To analyze participants' responses to the post-task questions, we used mixed-effects regression models. The fixed effects were the context alignment and SERP conditions, and the random effect was the participant ID. Finally, after participants completed all the tasks and questionnaires, they were interviewed about their experience completing the three shopping tasks.

5 RESULTS

(RQ1) Workload: Across the six measures for workload, context alignment had a significant effect on temporal demand. Participants perceived to have felt rushed in the LowDetail condition ($\beta = 0.888, S.E = 0.282, p < 0.01$) in comparison to the NoRoom condition. Next, among the SERP conditions, participants expressed more frustration in the BasicMenu condition ($\beta = 0.777, S.E = 0.278, p < 0.01$) in comparison to the DetailedMenu condition. Participants also felt that they were more successful in completing their task in both the BasicMenu ($\beta = 0.555, S.E = 0.214, p < 0.05$) and DetailedMenu ($\beta = 0.555, S.E = 0.214, p < 0.05$), compared to the NoMenu condition. Participants also reported experiencing higher mental demand when exposed to the NoMenu condition ($\beta = 0.944, S.E = 0.422, p < 0.05$) in comparison to the DetailedMenu condition.

(RQ2) User Engagement: Overall, we found no significant effect for context alignment, and found the SERP to have a significant effect on focused attention. Participants reported that they felt that time they spent in the virtual shopping slipped away much faster in the NoMenu condition ($\beta = 0.500, S.E = 0.178, p < 0.01$) compared to the DetailedMenu. In other words, participants in the NoMenu condition were significantly more focused on the task and that impacted their time perception.

(RQ3) System Usefulness: In terms of context alignment, participants reported to have found the HighDetail condition to be more helpful than the LowDetail ($\beta = 1.388, S.E = 0.550, p < 0.05$) and NoRoom ($\beta = 2.222, S.E = 0.550, p < 0.001$) conditions. This is corroborated by the interviews, in which participants explained that since the HighDetail environment was a replica of the target room they *'didn't have to rely on memory" - P09* during the search. The colors, materials and the refined render of the HighDetail environment also provided cues for participants to (re)formulate their queries and add descriptive words. In a participant's word: *'You know how the room looks like right in front of me. So I know the color I should pick. Like I keep saying here like I need a chair which is like orange color or yellow color or something like that" -P15.* Although somewhat surprisingly, participants were more satisfied with the search results in the NoRoom condition than the LowDetail condition ($\beta = 0.777, S.E = 0.361, p < 0.05$).



	Mental Demand	Physical Demand	Temporal Demand	Effort	Frustration	Success	Failure	Focussed Attention	Perceived Usability	Aesthetic Appeal	Rewarding Exp.	VR Helpfulness	SERP Helpfulness	Search Satisfaction	Selection Satisfaction
NoMenu	3.67	2.00	1.67	3.27	2.88	4.12	2.88	6.33	3.44	5.94	5.55	4.83	2.55	3.55	4.89
BasicMenu	3.16	1.94	2.16	3.05	3.11	4.67	2.33	6.11	3.22	5.83	5.75	5.33	5.38	4.50	5.56
DetailedMenu	2.72	1.88	2.00	2.77	2.33	4.67	2.33	5.83	3.94	6.16	5.72	5.88	4.94	4.16	5.33
NoRoom	3.38	1.88	1.50	3.22	2.94	4.45	2.55	6.22	3.61	5.72	5.63	4.33	4.33	4.55	5.28
LowDetail	3.05	2.11	2.38	3.11	2.88	4.62	2.38	6.16	3.61	6.05	5.69	5.16	4.27	3.77	5.22
HighDetail	3.11	1.83	1.94	2.77	2.50	4.39	2.61	5.88	3.38	6.16	5.69	6.55	4.27	3.88	5.28

Figure 4: Post-task response about participants' perception across: (RQ1) Workload, (RQ2) User Engagement, and (RQ3) System Helpfulness. * and ** denote significant differences at $p < 0.05$, and $p < 0.01$ level respectively.

For the SERP, participants reported to have found the BasicMenu ($\beta = 2.833, S.E = 0.373, p < 0.001$) and DetailedMenu ($\beta = 2.388, S.E = 0.373, p < 0.001$) to be more helpful than NoMenu and were more satisfied with the search results in BasicMenu ($\beta = 0.944, S.E = 0.361, p < 0.05$) compared to NoMenu.

6 DISCUSSION

Our findings suggest that high fidelity VR room renderings were more helpful for completing the shopping task, while lower fidelity increased time pressure. Participants preferred having UI elements to interact with rather than relying solely on voice commands. Below we discuss several implications from the results.

Time Pressure and Visual Distortion: We found that participants experienced greater time pressure in the LowDetail condition compared to the NoRoom condition. Prior work in VR reports a phenomena known as "Time Compression" [14] wherein a person perceives time to pass faster in VR than in real life. We posit that participants may have felt time pressure in the LowDetail condition as this was a distortion of the high-fidelity room wherein we removed the lighting, colors, and texture. Interestingly, from the interviews we also observed that participants perception of object sizes was distorted in the LowDetail condition, i.e., they felt the same 3D objects were oversized compared to the target room. For instance, consider the following quote *"This one (LowDetail) that [the proportion of objects] didn't seem to at least with the table—the lamp was fine but the table definitely felt like it was massively large oversized and just was like that didn't fit." — P1.* The half-rendered environment may have distorted participants' perception of both the time spent and space orientation.

Environmental Distraction: Participants reported that the search results in the NoRoom condition are better than that in the LowDetail condition. From the interviews participants tried their best to fit retrieved items in the LowDetail condition, which was proven challenging with visual distortion. In the NoRoom condition participants could not evaluate the fit of the items with the target room (since the background environment was absent). Participants instead focused on the retrieved items. The reduced distraction led them to a more satisfying search experience[2].

[2]In practice, the furniture they searched and selected may not have been a good fit the target room, as they made the selection without enough information to determine suitability.

Success and Being in Control: Participants preferred to having a SERP regardless of the amount of detail, as this gave them control on the items they could inject into the environment. Interestingly, they reported greatest frustration in the BasicMenu condition, despite reporting overwhelmingly in the interviews that they did not perceive any different between both SERP conditions as their focus was primarily on the images. We speculate that this may have happened as they may have felt they system could have provided "additional details," besides the basic information about the product. Finally, in the DetailedMenu condition, participants were able to complete their tasks with less mental workload than in the NoMenu condition, and this may have contributed to feeling the search was more successful.

Trade-off between Attention and Search Experience: Participants reported having greater focused attention in the NoMenu condition in comparison to the DetailedMenu condition. This may have been due to the lack of menu options, encouraging participants to focus on the object in front of them. Additionally, if they were unsatisfied with the object, they had no additional search context, leading to greater focus. However, the gain in attention was not a positive outcome, as participants did not have enough context to improve the search results through effective query reformulations, which led to greater mental workload and negatively impacting the helpfulness of the VR system.

7 CONCLUSION AND FUTURE WORK

In this study we investigated the impact of context alignment and search results presentation on shopping in VR. Our results suggest a preference for having a SERP, without increasing workload of the participants. For the surrounding environment during shopping, a high fidelity environment helps participants with their search, but a low fidelity environment could cause negative spatial and temporal disorientation. As next steps, we plan to investigate how to design interactive search system that help users search and shop for items without having to manually formulate all attributes in the queries (e.g. retrieve items similar to objects in the viewpoint) and further investigate methods that help reduce temporal and spatial distortions when we are unable to fully recreate an 3D shopping environment.

REFERENCES

[1] Pia Borlund. 2003. The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information research* 8, 3 (2003), 8–3.

[2] Jaime Carbonell and Jade Goldstein. 1998. The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 335–336.

[3] Jasmine Collins, Shubham Goel, Kenan Deng, Achleshwar Luthra, Leon Xu, Erhan Gundogdu, Xi Zhang, Tomas F Yago Vicente, Thomas Dideriksen, Himanshu Arora, Matthieu Guillaumin, and Jitendra Malik. 2022. ABO: Dataset and Benchmarks for Real-World 3D Object Understanding. *CVPR* (2022).

[4] Carsten Eickhoff, Sebastian Dungs, and Vu Tran. 2015. An Eye-Tracking Study of Query Reformulation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'15)*. 13–22.

[5] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. 2014. Lessons from the journey: a query log analysis of within-session learning. In *Proceedings of the 7th ACM international conference on Web search and data mining*. 223–232.

[6] Joey Ka-Yee Essoe, Nicco Reggente, Ai Aileen Ohno, Younji Hera Baek, John Dell'italia, and Jesse Rissman. 2022. Enhancing Learning and Retention with Distinctive Virtual Reality Environments and Mental Context Reinstatement. (2022).

[7] Gabriel Haas, Michael Rietzler, Matt Jones, and Enrico Rukzio. 2022. Keep it Short: A Comparison of Voice Assistants' Response Behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–12.

[8] Sandra G Hart. 1986. NASA task load index (TLX). (1986).

[9] Ahmed Hassan Awadallah, Ranjitha Gurunath Kulkarni, Umut Ozertem, and Rosie Jones. 2015. Characterizing and Predicting Voice Query Reformulation. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. 543–552.

[10] Hyo Jeong Kang, Jung-hye Shin, and Kevin Ponto. 2020. How 3D virtual reality stores can shape consumer purchase decisions: the roles of informativeness and playfulness. *Journal of Interactive Marketing* 49 (2020), 70–85.

[11] Diane Kelly and Leif Azzopardi. 2015. How Many Results per Page? A Study of SERP Size, Search Behavior and User Experience. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 183–192.

[12] Diane Kelly, Xin Fu, and Chirag Shah. [n. d.]. Effects of position and number of relevant documents retrieved on users' evaluations of system performance. ([n. d.]).

[13] Jimmy Lin, Xueguang Ma, Sheng-Chieh Lin, Jheng-Hong Yang, Ronak Pradeep, and Rodrigo Nogueira. 2021. Pyserini: A Python Toolkit for Reproducible Information Retrieval Research with Sparse and Dense Representations. In *Proceedings of the ACM Conference on Research and Development in Information Retrieval (SIGIR'21)*. 2356–2362.

[14] Grayson Mullen and Nicolas Davidenko. 2021. Time compression in virtual reality. *Timing & Time Perception* 9, 4 (2021), 377–392.

[15] Heather L O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies* 112 (2018), 28–39.

[16] Lawrence Philips. 2000. The Double Metaphone Search Algorithm. *C/C++ Users J.* 18, 6 (jun 2000), 38–43.

[17] Giovanni Pintore, Claudio Mura, Fabio Ganovelli, Lizeth Fuentes-Perez, Renato Pajarola, and Enrico Gobbetti. 2020. Automatic 3D reconstruction of structured indoor environments. In *ACM SIGGRAPH 2020 Courses*. 1–218.

[18] Gabriel A Radvansky, Sabine A Krawietz, and Andrea K Tamplin. 2011. Walking through doorways causes forgetting: Further explorations. *Quarterly journal of experimental psychology* 64, 8 (2011), 1632–1645.

[19] Yeon Soon Shin, Rolando Masís-Obando, Neggin Keshavarzian, Riya Dáve, and Kenneth A Norman. 2021. Context-dependent memory effects in two immersive virtual reality environments: On Mars and underwater. *Psychonomic Bulletin & Review* 28, 2 (2021), 574–582.

[20] 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 (2017), 102:1–102:31. https://doi.org/10.1145/3130967

[21] Anastasios Tombros and Mark Sanderson. 1998. Advantages of Query Biased Summaries in Information Retrieval. In *Proceedings of ACM SIGIR*.

[22] Maria Grazia Violante, Enrico Vezzetti, and Pietro Piazzolla. 2019. How to design a virtual reality experience that impacts the consumer engagement: the case of the virtual supermarket. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 13, 1 (2019), 243–262.

[23] Michel Juhani Wälti, Daniel Graham Woolley, and Nicole Wenderoth. 2019. Reinstating verbal memories with virtual contexts: Myth or reality? *PloS one* 14, 3 (2019), e0214540.

[24] Nannan Xi, Elpida Bampouni, and Juho Hamari. 2022. How Does Extended Reality Influence Consumer Decision Making? The Empirical Evidence from A Laboratory Experiment. In *Proceedings of the 55th Hawaii International Conference on System Sciences*. HICSS Conference Office University of Hawaii at Manoa.

[25] ZDNET. 2016. IKEA launches virtual reality shopping experience. https://www.zdnet.com/article/ikea-launches-virtual-reality-shopping-experience/, Last accessed on 2022-10-06.