Food Quality and Preference 87 (2021) 104071



Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual



The comparability of consumers' behavior in virtual reality and real life: A validation study of virtual reality based on a ranking task



Chengyan Xu[a,*], Yasemin Demir-Kaymaz[b], Christina Hartmann[a], Marino Menozzi[a], Michael Siegrist[a]

[a] Department of Health Sciences and Technology, ETH Zurich, Switzerland
[b] Graduate School of Business, Marketing, Koc University, Istanbul, Turkey

ARTICLE INFO

Keywords:
Virtual reality
Healthiness evaluation
Consumer behavior
Research tool validation
Food ranking task

ABSTRACT

The comparability or similarity of consumers' behavior in virtual reality (VR) and in real life (RL) is indispensable to successfully take advantage of VR technology in consumer behavior (CB) research. This study investigated the comparability of participants' behavior in a virtual environment and in RL to validate VR for data acquisition in CB research. Participants ($N = 98$) were randomly assigned to a RL condition or a VR condition in which the settings were chosen to mimic a RL environment as closely as possible. In both conditions, participants ranked 20 types of commercial breakfast cereal by their perceived healthiness from healthy to unhealthy. We observed that the perceived healthiness of the 20 cereals in the VR and RL conditions were highly correlated ($r_s = 0.91, p < .001, N = 20$). The information-seeking behavior was not different between the two conditions, $t(96) = -1.11, p = .27$. Furthermore, the attributes employed by the participants to evaluate the healthiness of the cereals did not vary between the two conditions. The presence level of the participants in the virtual environment was also examined. The results illustrate that the participants' behavior in the virtual environment and in the real world were comparable, which provides robust evidence for the validity of VR as a beneficial and promising tool for data acquisition in CB research.

1. Introduction

The development of new technologies provides opportunities for scientific research to make improvements in experimental settings, data acquisition, and processing. Virtual reality (VR) as an emerging technology can provide not only entertainment for consumers but also numerous possibilities in scientific research. Thus far, VR has been applied successfully in data visualization (Olshannikova, Ometov, Koucheryavy, & Olsson, 2015) and confrontation therapy to treat anxiety (Lindner et al., 2019). In the food domain, VR applications have been used in the treatment of eating disorders (Ferrer-Garcia, Gutierrez-Maldonado, Treasure, & Vilalta-Abella, 2015) and for weight control (Thomas & Bond, 2014), and as a tool for food portion-size education (Celikcan et al., 2018) and for investigating food and consumer behavior (CB) (Siegrist et al., 2019).

Many VR studies in the domain of food and CB have focused on food purchasing behaviors in virtual supermarkets. Several recent studies have investigated how the way products are displayed in a virtual supermarket influences consumers' buying behaviors (Bigné, Llinares, & Torrecilla, 2016; Kim et al., 2014) and visual searching behaviors (Zhao, Huang, Spence, & Wan, 2017), how consumers react to fast-moving consumer goods innovations (Bressoud, 2013), how price changes (such as promotions) affect consumers' purchasing behaviors (Poelman, Kroeze, & Waterlander, 2017; Waterlander, Scarpa, Lentz, & Steenhuis, 2011), as well as consumers' shopping behavior regarding non-standard or imperfect fruits and vegetables (Lombart et al., 2019; Verhulst, Normand, Lombart, & Moreau, 2017).

The advantage of VR in experimental and observational research is that it enables a high level of standardization that is similar to that in laboratory experiments. Accordingly, the limitations of traditional test systems, such as low ecological validity, can be overcome while environmental cues can simultaneously be fully controlled. VR can therefore be used to test behaviors and decision-making in different environmental settings that would otherwise be difficult or impossible to realize in real life (RL). In the food domain, products can be tested early during the product development stage, and food packaging (e.g., shape, color, or label) can be changed easily, which will allow researchers to examine which factors most strongly influence consumers' perceptions and evaluations of foods. Information search behavior, for example, where different nutrition labels are used could also be more

* Corresponding author.
E-mail address: chxu@hest.ethz.ch (C. Xu).

https://doi.org/10.1016/j.foodqual.2020.104071
Received 31 October 2019; Received in revised form 13 July 2020; Accepted 24 August 2020
Available online 27 August 2020
0950-3293/ © 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

easily examined using VR compared with RL. The applications are indeed manifold; however, VR is an artificial environment and deviates from reality in terms of reaction times, for instance. Before VR can be established as a new research paradigm in CB research, it is of central importance to validate it by providing evidence that people react similarly to real and virtual stimuli and behave comparably in RL and VR environments.

Only a few studies have examined whether consumers' behavior in VR is comparable to their behavior in RL (Pizzi, Scarpi, Pichierri, & Vannucci, 2019; Siegrist et al., 2019; Ung, Menozzi, Hartmann, & Siegrist, 2018; van Herpen, van den Broek, van Trijp, & Yu, 2016; Waterlander, Jiang, Steenhuis, & Ni Mhurchu, 2015). Some of these studies used less immersive virtual environments in which two- or three-dimensional pictures of the virtual environment were displayed on one or more computer screens or a video wall, and participants sat on a chair and navigated by clicking a computer mouse or keyboard (van Herpen et al., 2016; Waterlander et al., 2015). Other researchers showed that fully immersive VR applications outperform desktop applications in terms of realism and experienced telepresence (Schnack, Wright, & Holdershaw, 2019).

For the present study, we selected a ranking task to determine whether participants showed similar behaviors in VR and RL environments. We chose a ranking task related to the perceived healthiness of cereals as an example for comparing data collection in VR and RL. This ranking task was selected for three reasons. First, it is used in psychological studies (Coxon, 1999) and food experiments (Bucher, Muller, & Siegrist, 2015). Second, it can be implemented in both VR and RL environments. Third, it enables the observation and comparison of participants' behaviors during the experimental procedure. The ranking task was therefore considered a suitable research method for assessing the applicability, validity, and acceptability of VR for data collection in CB studies. We examined whether the data (i.e., evaluations, information-seeking behaviors, and cues used to assess healthiness) collected in VR and RL were comparable. A fully immersive VR was used to enable the participants to interact with the environment and walk around.

2. Methods

2.1. Participants

The data were collected in February and March 2019 in the German-speaking region of Switzerland. A convenience sample was used. The recruitment criteria were a minimum age of 18 years, ability to speak fluent German and understand English, no pronounced motion sickness, and no training in Food Science. In total, 101 participants were recruited. Three participants had to be removed from the data analysis, because of invalid experimental results caused by technical reasons. Thus, 98 participants remained in the final sample (48 participants in the VR condition, 50 participants in the RL condition). Among them, 36 participants were male (37%), and 62 were female (63%). The mean age was 34 years (SD = 14), with a range from 19 years to 71 years. The majority of the participants (85%) had a medium to high education level (i.e., middle school certificate or higher). Most participants (87%) reported that they consumed cereals, muesli, or corn flakes regularly in their daily lives. All participants were compensated with 20 Swiss Francs for completing the experiment.

2.2. Experimental design

Participants had to rank 20 cereal products from healthy to unhealthy. In the RL condition, real cereal products were used, and participants could rank them physically on a table. The environment in the VR condition was designed to mimic the environment of the RL condition as closely as possible (Figs. 1 and 2). Participants were randomly assigned to one of the two conditions. Age, gender, and education level of participants were not statistically significantly different between the



Fig. 1. Room setting in the real-life (RL) condition. Top: The table for ranking the cereals from healthy (1) to unhealthy (20). Middle: The table with 20 cereals randomly placed in the initial positions (the same initial positions of the cereals were used for all participants). Bottom: The room arrangement for the RL condition.



Fig. 2. Room setting in the virtual reality (VR) condition. Top: The table for ranking the cereals from healthy (1) to unhealthy (20). Middle: The table with 20 3D cereals. Bottom: The room arrangement of the VR condition. The initial positions of the cereals and the room arrangement were the same as in the real life (RL) condition for all participants. Participants could walk and navigate through the VR environment in the same way that they could in RL condition.



Fig. 3. The ranking task in the virtual reality condition. Top right & left: A participant during the ranking task. Bottom right & left: The virtual environment that the participant sees and interacts with.

two conditions. In both conditions, a written consent form was given to participants to read and sign before they started the experiment. The procedures of the main ranking task were recorded by a physical camera in the RL condition and by a built-in recorder of the virtual environment design software—Unity—in the VR condition.

2.2.1. Questionnaires

Two questionnaires (in German) were used in this study. The first questionnaire, which included questions about sociodemographic variables and about the type and frequency of the participants' cereals, muesli, and corn flakes consumption, was filled out before the main ranking task.

The second questionnaire was used only in the VR condition as a subjective evaluation of the participants' presence, which was measured with the presence questionnaire (PQ; (Witmer, Jerome, & Singer, 2005). The PQ consists of 29 items which are divided into four subscales to further examine the different aspects of presence. The four sub-scales are Involvement, Sensory Fidelity, Adaptation/Immersion, and Interface Quality. Because no auditory resource was involved in the virtual environment in this study, the three items that described auditory interaction with the environment were excluded from the questionnaire. An example PQ item is "How much were you able to control events?" Participants could rate all the items by using a 7-point Likert scale ranging from "not at all" (− 3) to "totally" (3). The responses were recoded as 1–7 before scale items were combined by calculating a sum score (i.e. presence score).

2.2.2. Cereals selected for the study

All the cereals used in this study were bought from a large-scale retailer in Switzerland. There was a great variety in the cereal brands and contents. Some cereals were packed in carton boxes; others were packed in plastic bags. The Nutrition Profiling Score (NPS) also varied between the cereals used in this study. The NPSs of the cereals were calculated according to the UK Ofcom nutrient profiling model (Rayner & Scarborough, 2009) and can be considered an objective evaluation of the healthiness, or ranking, of the cereals. The Ofcom nutrient profiling score objectively evaluates the healthiness of food and drink products. For details on how to calculate the NPS, please refer to the work of Rayner et al. (2009). The cereals selected were a good representation of the cereal products found in Switzerland.

Previous research suggested that a small number of cues associated with a food product influences perceived healthiness (Bucher et al., 2015). To predict the mean rank for the perceived healthiness of the cereals, three predictors were used: sugar content (*g/100 g*), cereal contains fruits (yes, no; raisins did not count as fruit), and cereal contains chocolate (yes, no). Carbohydrates and energy content are highly correlated with the sugar content, and therefore, were not included in the analyses.

2.3. Real-life condition

The 20 ready-to-eat breakfast cereals were randomly placed on a table. The initial arrangement of the 20 cereals was identical for all the participants. Participants were asked to rank the cereals from healthy to unhealthy on a second table. Twenty rectangles representing the 20 rankings were placed on the second table, and two textual signs, "healthy" (*gesund*) and "unhealthy" (*ungesund*), on the wall at the position of Rank 1 and Rank 20, respectively (Fig. 1). Participants could decide on their own whether they wanted to check the packages or read the nutrition information, ingredients list, or any other information. After the participants completed the ranking task, the rank order was recorded by taking pictures of the participants' final placement of the cereals.

2.4. Virtual reality condition

The virtual environment was modeled using Unity game engine version 2.11f1 (Unity, San Francisco, U.S.A.). The virtual environment was designed to mimic the RL environment as closely as possible. Therefore, the objects and the structure or arrangement of the objects

were the same as in the RL condition (Fig. 2). The initial arrangement of the 20 cereals was identical for all the participants. Participants could walk in the virtual environment and interact with the virtual objects within the environment as freely and naturally as they would in RL. The environment was displayed using a Head-Mounted-Display (HMD) and was interactable by using two controllers (Fig. 3; HMD and two controllers: HTC VIVE Pro, HTC Cooperation, Taiwan). The movement of the HMD and the controllers was tracked by two base stations installed on the level of the ceiling of the room.

Generating 3D models that were as realistic as possible was important for designing a highly immersive or realistic virtual environment. We used Recap Photo (version 19.1.1.2, Autodesk Inc., California, U.S.A.), which is a cloud-based photogrammetry tool for converting aerial and object photographs into photorealistic 3D models, to generate the 3D models of 20 cereals.

The procedures in the VR condition were as follows: A training task was conducted in a virtual training environment after the participants finished filling out the first questionnaire. The virtual training environment was also modeled in Unity, in which two non-relevant cereal boxes were placed on a table. They could be interacted with in the same fashion as the cereals included in the ranking task to train participants with the skills required for the real ranking task. In the training task, participants were instructed to grab, move, and release the boxes with the controllers, and learned how to find and read detailed information on packages when they wanted to. Participants were also taught how to optimally focus on text for reading by adjusting the HMD. The duration of the training task was not strictly limited. It stopped when participants could finish the tasks mentioned above without instruction or help from instructors. The ranking task started after the training task. Participants were instructed to perform the same tasks as the participants in the RL condition. The procedures of the ranking task were recorded by the Unity recorder captured from the game view. Every time the cereal boxes were placed on the ranking table, their positions were recorded and outputted.

2.5. Statistical analyses

Non-parametric Spearman rank-order correlation analyses were conducted to investigate the relationship between the mean ranks of perceived healthiness of the cereals in the RL condition and the VR condition. Mann–Whitney *U* test (with Bonferroni correction) was used to investigate the mean rank difference of the 20 cereals between VR and RL conditions. Video data from the two conditions were manually coded. The coding included how often participants looked at nutrition information or ingredients information on a package (information-seeking behavior) and the time needed to complete the ranking task for each participant. Independent sample *t*-tests were used to test the time difference for completing the ranking task, and the difference in information-seeking behavior between the RL condition and the VR condition. Ordinal logit regression analyses were used to examine which cereal attributes influence health perceptions. Finally, the presence scores of participants in the VR condition were calculated and analyzed. The significance level was 0.05. All data analyses were conducted in SPSS 25.0 (SPSS Inc., Chicago, U.S.A.)

3. Results

3.1. Perceived healthiness

As shown in Fig. 4, the mean ranks——which represents perceived healthiness——of the cereals in the VR and the RL conditions were highly correlated ($r_s = 0.91, p < .001, N = 20$). This high association suggests that the healthiness perception was almost identical in the two conditions, and it did not matter whether data were collected in VR or RL. The mean ranks of the cereals are shown in Table 1. The mean rank differences between the two experimental conditions were small for all 20 cereals; the largest difference was only 2.10. The result of Mann–Whitney *U* test (with Bonferroni correction: nominal $p < .05$, corrected $p < .0025$) shows that the mean rank differences were not significantly different between the VR condition and the RL condition for all the 20 cereals.

The NPS was used as an objective healthiness score for the 20 cereals and was compared with participants' subjective perceived healthiness (i.e. rank). There were strong positive correlations between perceived healthiness and the NPS in both the VR condition ($r_s = 0.73$, $p < .001$, $N = 20$) and the RL condition ($r_s = 0.76$, $p < .001$, $N = 20$).

3.2. Time needed for the ranking task

The time participants took to complete the ranking task in the VR condition and in the RL condition was compared. Fig. 5 shows the time needed in the VR condition and the RL condition. The *t*-test showed that the time needed in the VR condition ($M = 17.58$ Min, $SD = 8.16$) was statistically significantly longer than the time needed in the RL condition ($M = 12.02$ Min, $SD = 4.81$), $t(96) = 4.13$, $p < .001$.

3.3. Information-seeking behavior

When participants looked at nutrition information or ingredients information, the behavior was coded as information-seeking behavior. The number of times participants looked at the nutrition or ingredients information on a cereal product was counted. The variable is labeled "total information seeking (TIS)." In addition, the number of packages at which a participant looked at least one time at the nutrition or ingredients information was counted (information seeking for single product (ISSP)). The value of this variable could range between 0 and 20. Fig. 6 shows boxplots for TIS and ISSP in both conditions. The *t*-test showed that there was no statistically significant difference between TIS in the VR condition ($M = 38.50$, $SD = 29.39$) and in the RL condition ($M = 44.56$, $SD = 24.61$), $t(96) = -1.11$, $p = .27$. The difference in ISSP in the VR condition ($M = 15.96$, $SD = 5.59$) and in the RL condition ($M = 17.42$, $SD = 5.44$) was also not statistically significant, $t(96) = -1.31$, $p = .19$. TIS and ISSP were highly correlated with each other in the VR condition ($r = 0.69$, $p < .001$, $N = 48$) and in the RL condition ($r = 0.71$, $p < .001$, $N = 50$). Furthermore, the time needed to complete the ranking task and the information-seeking behavior had a strong positive correlation in both conditions (in the VR condition: $r = 0.80$, $p < .001$, $N = 48$; in the RL condition: $r = 0.72$, $p < .001$, $N = 50$).

3.4. Attributes influence participants' perceived healthiness of cereals

Ordinal logit regression analyses were conducted to investigate whether the attributes of the cereals were statistically significant predictors of the participants' perceptions regarding the healthiness of the cereals. The regression models for the VR condition and the RL condition are shown in Table 2. In both conditions, "Sugar (*g/100 g*)" (VR: Estimate B = 0.36, SE = 0.11; RL: Estimate B = 0.57, SE = 0.15) was a strong predictor of the participants' perceived healthiness of the cereals: the higher the sugar content, the lower the perceived healthiness of the cereals. "Contains Fruit" (VR: Estimate B = −3.51, SE = 1.22; RL: Estimate B = −3.46, SE = 1.26) and "Contains Chocolate" (VR: Estimate B = 4.38, SE = 1.81; RL: Estimate B = 5.98, SE = 2.20) were also statistically significant predictors of perceived healthiness. Cereals that contain fruit and do not contain chocolate were considered healthier than the other cereals.

3.5. Presence in the VR condition

Presence is a way to measure how participants feel about the virtual environment they experienced. For each participant, a sum score of the

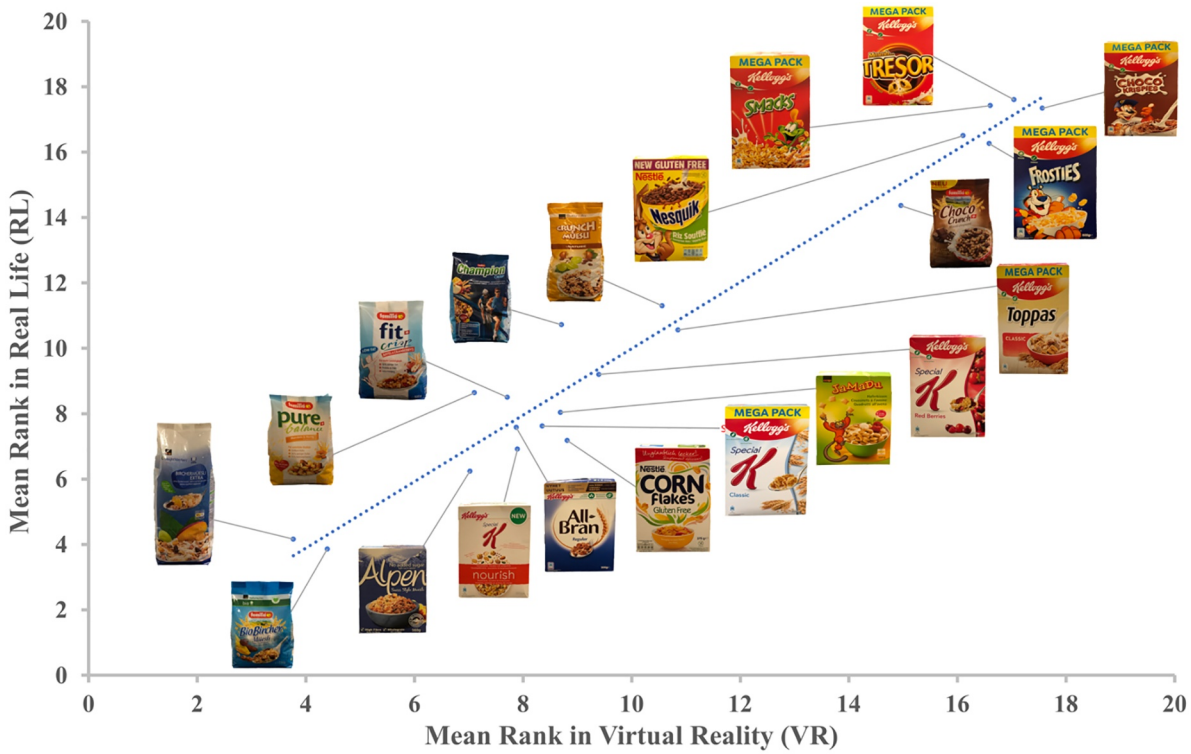


Fig. 4. The correlation between the mean ranks in the virtual reality (VR) condition and in the real-life (RL) condition. The ranks from left to right (the x-axis) and bottom to top (the y-axis) represent perceived healthiness ranging from healthy to unhealthy. The trend line in the graph shows the mean rank in the VR condition equals the mean rank in the RL condition.

Table 1
The mean rank, the mean rank difference, and the Nutrition Profiling Score (NPS) of 20 cereals in the virtual reality (VR) condition and the real-life (RL) condition.

Cereal Name	Mean Rank in VR Condition	Mean Rank in RL Condition	Mean Rank Difference	Nutrition Profiling Score (NPS)
Bio Bircher Muesli	4.40	3.86	0.54	–4
Pure Balance	7.10	8.64	–1.54	8
Red Berries	9.40	9.20	0.20	6
Champion	8.71	10.72	–2.10	10
Weightwatchers	3.77	4.16	–0.39	−3
Fit Crisp	7.71	8.50	–0.79	6
Nourish (nuts)	7.90	6.92	0.98	0
Crunch Muesli (nature)	10.56	11.30	–0.74	11
Alpen	7.02	6.24	0.78	−2
Special K Classical	8.35	7.62	0.73	0
Toppas	10.85	10.56	0.29	−3
AllBran	7.88	7.58	0.30	0
Jamadu	8.69	8.04	0.65	8
Smacks	16.60	17.42	–0.82	8
Cornflakes Gluten Free	8.81	7.18	1.63	9
Frosties	16.58	16.26	0.32	13
ChocoCrunch	14.96	14.36	0.60	13
Nesquik	16.10	16.50	–0.40	8
Choco Krispies	17.56	17.34	0.22	12
Tresor	17.04	17.60	–0.56	15

scale items (i.e. the presence score) was calculated. Overall, the participants had high presence scores (M = 141.90, SD = 13.97, possible maximum = 182 [observed minimum = 112, observed maximum = 169]) in the virtual environment. The presence scores for men (M = 144.35, SD = 12.44) and women (M = 140.55, SD = 14.76) did not differ statistically significantly, t (46) = 0.90, p = .37. Age was positively correlated with the presence score (r = 0.35, p = .016, N = 48).

The analysis of the four subscales of the PQ, which are Involvement (M = 67.56, SD = 7.39, possible maximum = 84), Sensory Fidelity (M = 15.67, SD = 3.46, possible maximum = 21), Adaptation/Immersion (M = 45.50, SD = 5.68, possible maximum = 56), and Interface Quality (M = 13.17, SD = 3.25, possible maximum = 21), indicated that the participants had high involvement and immersion in the virtual environment, and the sensory fidelity and interface quality of the virtual environment were also highly rated by the participants.

4. Discussion

The aim of this study was to investigate whether the participants would behave similarly during data collection in VR and in RL to determine whether VR is a reliable and valid method for CB research. Participants assessed the healthiness of cereals either in a VR or RL setting. The perceived healthiness in the two conditions were highly correlated, and the perceived healthiness were similarly correlated with the NPS of the cereals. Therefore, the results showed that the participants made similar decisions in virtual reality and in real life. Virtual reality as an environment or tool for the participants to perform tasks did not influence their judgment process and judgment results.

The time needed to complete the ranking task was longer in the VR condition, which was expected. VR as a new subject or environment for people is different from RL which we interact with and experience naturally every day. Participants need more time to get used to the environment and the ways of interacting in VR, although there was a training session for practice before the real task was conducted. Although more time was needed to complete the ranking task in VR than in RL, the difference was only 5 min, which is acceptable for using VR as a method to conduct experiments. This time difference was found in other studies (Siegrist et al., 2019). They found that the participants took slightly more time to look at the package of a cereal (represented by the relative fixation duration) in a VR supermarket environment

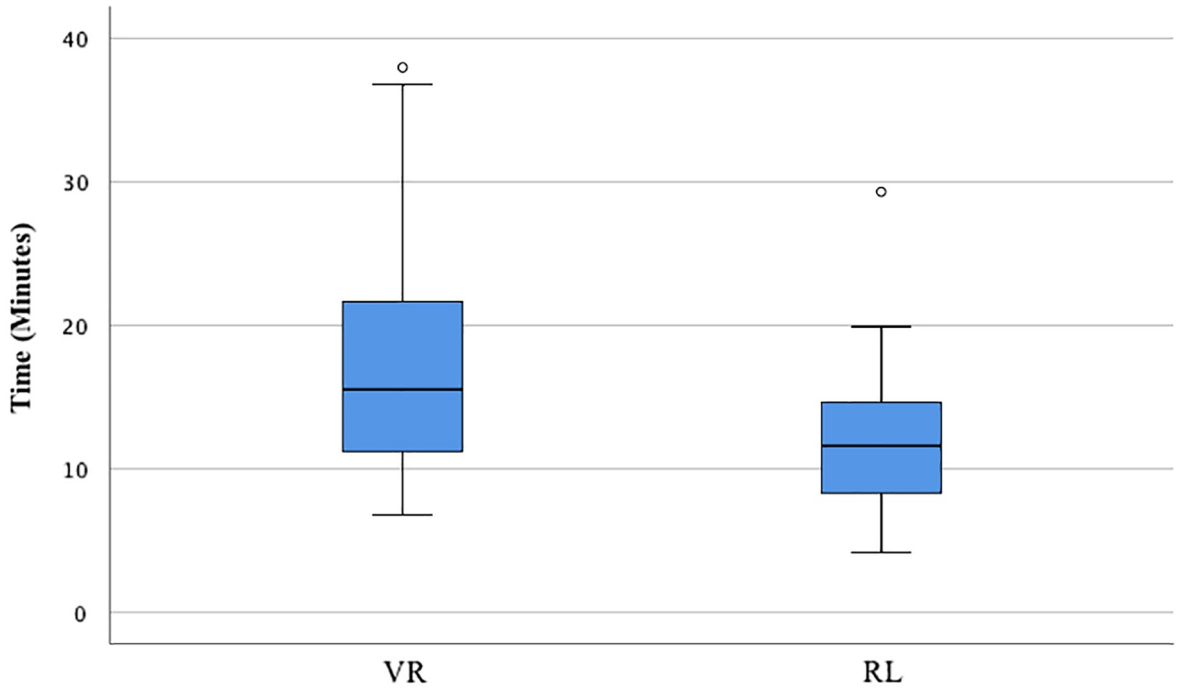


Fig. 5. Boxplot of the time needed to complete the ranking task in the virtual reality (VR) condition and in the real-life (RL) condition.

than in the real world, potentially because people were not as familiar with the virtual environment they explored.

Results of the information-seeking behavior analyses suggest that the participants showed similar information-seeking behavior in order to evaluate the healthiness of foods in both conditions. Those results reveal that the participants behaved similarly in the VR environment as they did in RL in many aspects, not only in terms of performance but also in the way they approached the task.

Consumers' approach for estimating the healthiness of a product normally relies on the information presented on the package. The results of the present study showed that the participants in the VR condition and in the RL condition used the same attributes of the cereals to evaluate their healthiness. Participants perceived cereals that have a high sugar content as unhealthier than other cereals. Furthermore, fruits shown on the package increased perceived healthiness, whereas chocolate on the package decreased perceived healthiness. Similar relationships have been found in other studies (Bucher et al., 2015). These results suggest that the participants used similar information or criteria to help them make evaluations under both the VR and RL conditions, which provides further evidence that the participants behave comparably in many respects in the VR and RL.

In the present study and by using realistic food packages, we

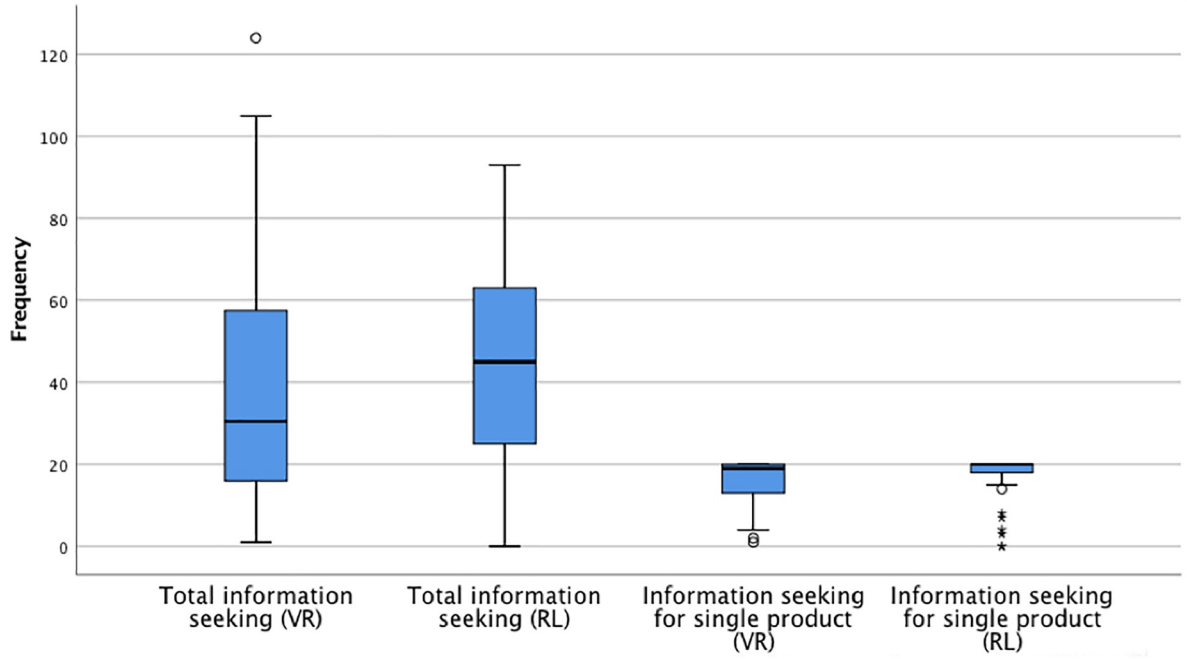


Fig. 6. Boxplot of the participants' information-seeking behavior in the virtual reality (VR) condition and in the real-life (RL) condition. Frequency = how many times the participant looked at the nutrition or ingredient information during the whole ranking task.

Table 2
Parameter estimates results of ordinal logit regression, with the mean ranks of the 20 cereals in the virtual reality (VR) condition and in the real-life (RL) condition as dependent variables.

	Virtual Reality				Real Life			
Variable	*Estimate B [95% CI]*	*Std. Error*	*Wald*	*df*	*Estimate B [95% CI]*	*Std. Error*	*Wald*	*df*
Sugar (g/100 g)	0.36** [0.15, 0.58]	0.11	11.33	1	0.57*** [0.28, 0.87]	0.15	14.37	1
Contains Fruit[a]	–3.51** [–5.91, –1.11]	1.22	8.23	1	–3.46** [–5.93, –0.99]	1.26	7.53	1
Contains Chocolate[a]	4.38* [0.84, 7.98]	1.81	5.88	1	5.98** [1.67, 10.29]	2.20	7.41	1

Note: model fit for the VR condition: −2 (Log Likelihood) = 86.88, $\chi 2 = 31.57$, df = 3, $p < .001$; model fit for the RL condition: −2 (Log Likelihood) = 79.58, $\chi 2 = 38.87$, df = 3, $p < .001$.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.
[a] Dummy coded variable: 0 = product not contain this attribute; 1 = product contains this attribute.

showed that data collection in RL and VR resulted in similar findings. Similar results have been found through different methodologies in other studies. For example, Ung et al. (2018) found a high correlation between the food choice in a virtual food buffet and in a fake food buffet. Siegrist et al. (2019) also found a similar result. They conducted two experimental studies in a virtual supermarket to replicate the findings of a RL study (Visschers, Hess, & Siegrist, 2010). In their studies, the participants' food choice in different situations were recorded and compared in a VR condition and in a RL condition. The results showed that the food selection behavior did not differ in the two conditions.

VR is potentially applicable in many research areas; however, only a few studies have investigated the validity of VR in CB research. By comparing the participantś behavioral data collected in VR and RL conditions while conducting an identical ranking task, this study provides evidence that VR is a valid data collection tool in CB research. The results indicate that the data collected in VR and RL were highly comparable in many different ways. For instance, the participants made similar decisions in both VR and RL. In addition, the information-seeking behaviors of the participants and the types of information or cues that the participants used to help them make the decisions were also comparable between the VR and RL conditions. This indicates that, although VR is a novel and unfamiliar environment compared to RL, it can be used to conduct experiments and collect data with high ecological validity. The promising results of this study also provide motivation to explore possibilities for the application of VR in future studies and thus contribute to the development and understanding of VR applied in the CB domain.

A ranking task was used in this study; however, other tasks are likely to be similarly suitable for further investigations and to expand the possibilities of VR as a research tool. Given the promising results of our study, it would certainly also be interesting to investigate whether, for example, choice tasks or conjoint analyses conducted in VR will provide similar results to those conducted in RL.

In future studies, VR could be used to investigate research questions such as how different food labels or the arrangement of food products on shelves affects participants' purchasing decisions or whether different information (e.g., positive or negative) in different formats (e.g., text or video) influences participants' food choices. VR provides the opportunity to conduct studies in a shopping setting instead of a laboratory environment without modifying or changing the real products or environments, which is usually time-consuming and expensive in RL. Compared to in RL, it is therefore easier to investigate the impact of changing the surrounding environment on participants' behaviors in VR. For example, one can study whether adding different smells (pleasant or unpleasant) or sound or light to a shopping environment would lead to different purchasing decisions or whether varying shopping environments (e.g., farmers market, discounter, supermarket) affects participantś preferences or willingness to buy certain products. It would also be interesting to investigate whether changing the surrounding environment in real-time (e.g. by showing additional information or rearranging the positions of products) based on participants' current behaviors will lead them to adapt to healthier or better choices.

VR further provides an environment in which an eye-tracking system can be easily implemented. Eye-tracker is more reliable in VR because the distance between the eyes and the objects is constant. Attention data such as the area of interest can also be recorded and analyzed more easily without requiring manual intervention by researchers. Combining VR and an eye-tracking system could be a focus in future studies as it may provide opportunities for the closer investigation of people's focus while they perform consumption or buying behaviors and may help determine what information (e.g., ingredients information and labels) influences people's decisions.

4.1. Perceived presence

Presence is the sense of "being" in the virtual environment. Studies showed that a high presence level can help participants behave more naturally in the virtual environment (Schnack et al., 2019; Waterlander et al., 2015). In this study, the participants were engaged in a fully immersive VR environment in which they could interact and behave as naturally as in a RL environment. A training task before the main ranking task helped the participants become familiar with the virtual environment, which had been proven to be important for the participants (Burke, 2018; Siegrist et al., 2019). Otherwise, participants might be more engaged in exploring the virtual environment instead of behaving spontaneously (Siegrist et al., 2019). Overall, the participants experienced a high degree of presence in the virtual environment of this study.

5. Conclusion

The present study tested the applicability and validity of VR as a data collection tool in the CB domain by comparing the participants' behavioral data collected in both a virtual environment and the real world. The findings suggest that the data collected in the RL and VR conditions were highly comparable, even though VR is a very different environment from the real world. This provides robust evidence of the validity and applicability of VR as a data collection tool. Overall, this study explored one application of VR within the CB domain, with the aim of gaining a better understanding of how people behave in an artificial virtual environment as well as the applicability and validity of VR as a data collection tool. The application of VR in CB research can be further explored by investigating other research questions using different kinds of experimental tasks. We strongly believe that VR will be a very useful tool for conducting experiments and producing results with high ecological validity.

CRediT authorship contribution statement

Chengyan Xu: Software, Investigation, Data curation, Formal

analysis, Writing - original draft, Writing - review & editing. **Yasemin Demir-Kaymaz:** Investigation, Data curation, Formal analysis. **Christina Hartmann:** Conceptualization, Supervision, Writing - review & editing, Validation. **Marino Menozzi:** Software, Supervision, Writing - review & editing, Validation. **Michael Siegrist:** Conceptualization, Supervision, Writing - review & editing, Validation.

Acknowledgements

Funding: This study was supported by the Swiss National Science Foundation (SNSF) under the project "Digital Lives" (grant number 10DL11_183171/1).

References

Bigné, E., Llinares, C., & Torrecilla, C. (2016). Elapsed time on first buying triggers brand choices within a category: A virtual reality-based study. *Journal of Business Research, 69*(4), 1423–1427. https://doi.org/10.1016/j.jbusres.2015.10.119.

Bressoud, E. (2013). Testing FMCG innovations: Experimental real store versus virtual. *Journal of Product and Brand Management, 22*(4), 286–292. https://doi.org/10.1108/jpbm-05-2012-0141.

Bucher, T., Muller, B., & Siegrist, M. (2015). What is healthy food? Objective nutrient profile scores and subjective lay evaluations in comparison. *Appetite, 95*, 408–414. https://doi.org/10.1016/j.appet.2015.08.005.

Burke, R. R. (2018). Virtual Reality for Marketing Research. In L. Moutinho, & M. Sokele (Eds.). *Innovative Research Methodologies in Management: Volume II: Futures, Biometrics and Neuroscience Research* (pp. 63–82). Cham: Palgrave Macmillan.

Celikcan, U., Bulbul, A. S., Aslan, C., Buyuktuncer, Z., Isgin, K., Ede, G., & Kanbur, N. (2018). The Virtual Cafeteria: An Immersive Environment for Interactive Food Portion-Size Education. *3rd Workshop on Multisensory Approaches to Human-Food Interaction (Mhfi*https://doi.org/10.1145/3279954.3279960.

Coxon, A. P. M. (1999). *Sorting data: Collection and analysis.* Thousand Oaks, CA: Sage.

Ferrer-Garcia, M., Gutierrez-Maldonado, J., Treasure, J., & Vilalta-Abella, F. (2015). Craving for Food in Virtual Reality Scenarios in Non-Clinical Sample: Analysis of its Relationship with Body Mass Index and Eating Disorder Symptoms. *European Eating Disorders Review, 23*(5), 371–378. https://doi.org/10.1002/erv.2375.

Kim, A. E., Nonnemaker, J. M., Loomis, B. R., Shafer, P. R., Shaikh, A., Hill, E., ... Farrelly, M. C. (2014). Influence of point-of-sale tobacco displays and graphic health warning signs on adults: Evidence from a virtual store experimental study. *American Journal of Public Health, 104*(5), 7. https://doi.org/10.2105/AJPH.2013.301723.

Lindner, P., Miloff, A., Fagernas, S., Andersen, J., Sigeman, M., Andersson, G., ... Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial (vol 61, pg 45, 2019). *Journal of Anxiety Disorders, 64*, 90. https://doi.org/10.1016/j.janxdis.2019.04.002.

Lombart, C., Milian, E., Normand, J. M., Verhulst, A., Labbe-Pinlon, B., & Moreau, G. (2019). Consumer perceptions and purchase behavior toward imperfect fruits and vegetables in an immersive virtual reality grocery store. *Journal of Retailing and Consumer Services, 48*, 28–40. https://doi.org/10.1016/j.jretconser.2019.01.010.

Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing Big Data with augmented and virtual reality: Challenges and research agenda. *Journal of Big Data, 2*(1), https://doi.org/10.1186/s40537-015-0031-2.

Pizzi, G., Scarpi, D., Pichierri, M., & Vannucci, V. (2019). Virtual reality, real reactions?: Comparing consumers' perceptions and shopping orientation across physical and virtual-reality retail stores. *Computers in Human Behavior, 96*, 1–12. https://doi.org/10.1016/j.chb.2019.02.008.

Poelman, M., Kroeze, W., & Waterlander, W. (2017). Food taxes and calories purchased in the virtual supermarket: A preliminary study. *British Food Journal, 119*(12), 11. https://doi.org/10.1108/BFJ-08-2016-0386.

Rayner, M., & Scarborough, P. (2009). The UK Ofcom Nutrient Profile Model: Defining 'healthy' and 'unhealthy' foods and drinks for TV advertising to children. Retrieved from https://www.ndph.ox.ac.uk/cpnp/files/about/uk-ofcom-nutrient-profile-model.pdf.

Schnack, A., Wright, M. J., & Holdershaw, J. L. (2019). Immersive virtual reality technology in a three-dimensional virtual simulated store: Investigating telepresence and usability. *Food Research International, 117*, 40–49. https://doi.org/10.1016/j.foodres.2018.01.028.

Siegrist, M., Ung, C. Y., Zank, M., Marinello, M., Kunz, A., Hartmann, C., & Menozzi, M. (2019). Consumers' food selection behaviors in three-dimensional (3D) virtual reality. *Food Research International, 117*, 50–59. https://doi.org/10.1016/j.foodres.2018.02.033.

Thomas, J. G., & Bond, D. S. (2014). Review of Innovations in Digital Health Technology to Promote Weight Control. *Current Diabetes Reports, 14*(5), https://doi.org/10.1007/s11892-014-0485-1.

Ung, C. Y., Menozzi, M., Hartmann, C., & Siegrist, M. (2018). Innovations in consumer research: The virtual food buffet. *Food Quality and Preference, 63*, 12–17. https://doi.org/10.1016/j.foodqual.2017.07.007.

van Herpen, E., van den Broek, E., van Trijp, H. C. M., & Yu, T. (2016). Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite, 107*, 196–207. https://doi.org/10.1016/j.appet.2016.07.033.

Verhulst, A., Normand, J. M., Lombart, C., & Moreau, G. (2017). A Study on the Use of an Immersive Virtual Reality Store to Investigate Consumer Perceptions and Purchase Behavior toward Non-standard Fruits and Vegetables. *Ieee Virtual Reality (Vr), 2017*, 55–63.

Visschers, V. H., Hess, R., & Siegrist, M. (2010). Health motivation and product design determine consumers' visual attention to nutrition information on food products. *Public Health Nutrition, 13*(7), 1099–1106. https://doi.org/10.1017/S1368980009993235.

Waterlander, W. E., Jiang, Y., Steenhuis, I. H., & Ni Mhurchu, C. (2015). Using a 3D virtual supermarket to measure food purchase behavior: A validation study. *Journal of Medical Internet Research, 17*(4), Article e107. https://doi.org/10.2196/jmir.3774.

Waterlander, W. E., Scarpa, M., Lentz, D., & Steenhuis, I. H. M. (2011). The virtual supermarket: An innovative research tool to study consumer food purchasing behaviour. *BMC Public Health, 11*. https://doi.org/10.1186/1471-2458-11-589 Artn 589.

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments, 14*(3), 14. https://doi.org/10.1162/105474605323384654.

Zhao, H., Huang, F. X., Spence, C., & Wan, X. A. (2017). Visual Search for Wines with a Triangle on the Label in a Virtual Store. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02173 ARTN 2173.