**Congruency of an eating environment influences product liking: a Virtual Reality study**

Joachim J. Schouteten[1*], Arthur Van Severen[1], Daniëlle Dull[2], Hans De Steur[1], Lukas Danner[3]

[1] Department of Agricultural Economics, Ghent University, Coupure links 653, 9000, Belgium

[2] EyeQuestion Software, Nieuwe Aamsestraat 90D, 6662 NK Elst, the Netherlands

[3] CSIRO, Agriculture and Food, 671 Sneydes Rd, Werribee VIC 3030, Australia

*Corresponding author: Joachim J. Schouteten

**Abstract**

Virtual reality (VR) has been introduced as a method to improve the ecological validity of sensory tests by mimicking natural consumption contexts. VR has the advantage of being a cost-effective method that allows for sensory tests to be conducted in a more controlled and standardized environment, compared to tests performed in a natural consumption context (e.g., at home, in a restaurant, etc.). While some research with VR has been published, it remains unclear to what extent VR can truly immerse people in the intended context. One possible option to determine the effects of immersion is to investigate VR context–product congruency. As such, the aim of this study is to examine to what extent a congruent eating environment (summer/winter) impacts the perception of food congruent for that eating context. Using a between-subject design, a total of 100 participants evaluated three food products (watermelon, cracker, and chocolate truffle) either in a winter or summer VR context. Results showed that the overall liking of the congruent food (watermelon for summer and chocolate truffle for winter) was significantly higher than for the other foods when consumed in a congruent VR context. It should also be mentioned that overall, the scores were higher for the summer than the winter context, which requires further research (e.g., with other products/VR contexts). The emotions evoked by the consumed food products did not differ according to the VR context. Furthermore, high engagement scores were reported for the participants during this study, compared to the engagement reported by prior research performed in traditional sensory booths. These results illustrate the potential of VR for application in sensory research as the VR context can immerse and engage participants.

**Keywords**

Virtual Reality (VR); context; consumer; sensory evaluation; engagement; immersion

## 1. Introduction

Sensory research aims to gain insight into how consumers perceive food products through their senses. Sensory tests are typically conducted in sensory laboratories, where participants are seated in individual booths to evaluate the products under study (Jaeger et al., 2016). While the experimental environment reduces panelist bias and helps to control against confounding non-product biases, it does not mimic a realistic consumption context, which potentially results in low ecological validity (Bangcuyo et al., 2015). Prior research indicates that the evaluation context (laboratory vs. natural setting) can influence the overall acceptance of food products (Jaeger & Porcherot, 2017). The inability of traditional consumer sensory tests to reliably predict consumer decisions is believed to be a major contributor to high rates of new food product failures (Bangcuyo et al., 2015).

There is a need for improved methods to increase the ecological validity of consumer sensory science (Jaeger et al., 2016). For instance, sensory tests can be conducted in a natural consumption context (e.g., by implementing a home-use test [HUT] in which participants are asked to evaluate products in real-life situations). However, the complexity of real-life situations introduces several confounding factors (e.g., location, company, product storage, product preparation), which makes this kind of research less controlled and standardized (Jaeger et al., 2016). Recent technological developments in virtual reality (VR) allow participants to be immersed in more realistic evaluation contexts mimicking natural consumption situations in a relatively controlled and standardized—but also cost-efficient—manner (Wang, Escobar, Da Mota, & Velasco, 2021). VR can provide more contextual information and enhance data quality, as participants might be more engaged compared to in a traditional blind sensory testing (Bangcuyo et al., 2015; Hathaway & Simons, 2017).

A recent review by Wang et al. (2021) on current practices and future perspectives for the use of VR in sensory and consumer science illustrates the great potential of VR in sensory science, although the research remains in its infancy in terms of food consumption applications. Given the variety of decisions that are made in a VR study using food products (e.g., VR context, products under study, hardware, software, response methods, etc.), results are often highly study dependent, but they still offer indication that VR research could increase ecological validity. To date, the VR context has primarily been used to induce a consumption context, with the food not being represented in VR and the responses only being collected after the headset is removed. Indeed, the consumption of food products by participants within a VR context can be quite challenging with current VR methodologies (Xu, Siegrist, & Hartmann, 2021). Experimenters must often help participants by providing the food they must consume, as participants cannot see the samples when wearing head-mounted displays (HMDs). Research has thus often focused on using finger foods and straws when working with beverages (Wang et al., 2021; Xu et al., 2021). Furthermore, past VR studies have primarily measured taste and liking, with few studies examining emotions; however, some studies have shown VR contexts to affect the emotions toward elicited by consumers toward certain food products, such as chocolate (Kong et al., 2020; Torrico et al., 2021) and wine (Torrico et al., 2020). This is of particular interest, as emotional measurements might provide additional data beyond liking (Lagast, Gellynck, Schouteten, De Herdt, & De Steur, 2017); VR environments might trigger more realistic emotions by participants compared to traditional sensory testing in booths, thus providing more valid insights into their food experience.

In summary, the primary benefit of VR is that it can be a cost-effective alternative to a real-life consumption context, while providing a high degree of control and standardization (Chen, Huang, Faber, Makransky, & Perez-Cueto, 2020). Nevertheless, sensory research with VR using more realistic scenarios that better mimics a realistic consumption context, leading to higher ecological validity, is needed.

Torrico et al. (2021) have found that VR can be used to understand contextual effects with participants who evaluated two chocolate samples (one with sugar and one with artificial sweeteners) under three environments (traditional booths, VR with open-field forest, and VR with closed-space old room). Chen et al. (2020) used three artificial VR environments (sweet congruent, sweet incongruent, and neutral) to study the effect of a congruent VR environment on drinking a sweet beverage, with participants completing a questionnaire after removing the VR headset during this within-subject experiment. Their results showed that the beverage was perceived as significantly sweeter in a sweet-congruent environment compared to other VR environments but that the visual–taste congruency did not impact the liking. Song et al. (2022) found that the liking of rye bread was significantly higher in a congruent VR restaurant than in an incongruent VR cinema context. These liking assessments were also conducted after participants had removed their VR headsets.

Despite the growing interest in VR in sensory research, results are often highly study dependent. It remains unclear to what extent VR truly immerses people in a realistic virtual context when they are consuming foods appropriate for that virtual context and when measurements are occurring in the VR context. As such, the aim of the current study is to examine the impact of consuming food products that are congruent or not congruent with a virtual context, considering seasonal context (summer vs. winter). It is hypothesized that when VR can immerse participants in a context, the acceptance of food products that are congruent with a virtual evaluation context will be higher. To truly assess the impact and immersiveness of the VR context, self-report measurements (liking and emotional profiling) are conducted in a VR context. Furthermore, the engagement of the participants for the tasks is examined, as VR may lead to higher engagement and thus more reliable results.

## 2. Materials and methods

### 2.1. Participants

A total of 100 participants were recruited using direct, online recruitment procedures in Belgium. All participants were naïve product users and reportedly had no food allergies. Potential participants were excluded if they had a history of seizures, a motion sickness, or acute disease. Participants did not receive any reimbursement for their participation in the study. The study was approved by the ethical committee of Ghent University Hospital (BC-09441), and all participants signed an informed consent. This study opted to work with a between-subject design, so participants evaluated the food products in only a single VR context. Participants were randomly assigned to either the summer (n=49) or winter context (n=51). Specific information about the participants is provided in Table 1; no significant differences existed between the consumer groups regarding the various characteristics.

*Table 1*

### 2.2. Virtual context settings

For the virtual context settings, two 360° videos were selected from the internet (summer: YouTube, winter: Blend Media) to represent a summer (Fig. 1a) and winter (Fig. 1b) context. Both VR videos had a similar environment with water and trees and also included natural ambient sounds specific to the season and environment.

*Figure 1*

### 2.3. Stimuli

Stimuli were selected based on a pilot test with 60 participants (58% female, mean age: 37 years) who did not participate in the actual VR study. Participants in the pilot received pictures of each VR environment and were asked to select all foods (out of a list of 16 foods) that seemed appropriate to consume in that VR environment. The 16 food products were selected based on discussion among the authors considering the goal and practicalities of the test. For the summer VR context, a slice of watermelon was selected most often (93.3%), while a chocolate truffle (61.7%) was regarded as the most appropriate for the winter VR environment. Furthermore, plain crackers were added as a more neutral stimulus that was equally appropriate for both contexts (winter: 25%, summer: 30%). The assessment of the samples was balanced for order and carry-over effects, and participants rinsed their mouths with water before each tasting to clean their palate.

### 2.4. Experimental procedure

Participants were seated in front of a table. Due to the COVID-19 situation, testing occurred in a separate room at the participant's home. Participants were first assisted by the researcher to put on the VR headset with integrated headphones (Oculus Go, Meta). The entire experimental procedure occurred using EyeQuestion® (version 5.0.7.8, EyeQuestion Software, the Netherlands) and lasted about half an hour on average. Participants saw the questions and answer options in the VR context and answered by selecting their desired answer option using gaze tracking. Once they had the VR headset on, they first rated their level of hunger (1=totally not ⇔ 5=extremely; de-Magistris & Gracia, 2017) in a neutral VR context of a park.

The actual testing in the VR context then began. Participants had at least one minute to explore the summer or winter VR context without any distractions (Wang et al., 2021). Next, participants first rated how comfortable they felt in the virtual environment (1=totally not ⇔ 9=extremely; Chen et al., 2020). Succeeding, they evaluated each sample within the VR context, following four steps by screen prompts in the VR. First, the expected liking from the stimuli (based upon the word) was rated using a 9-point hedonic scale (1=dislike extremely ⇔ 9=like extremely). Second, participants tasted the sample and rated their experienced liking using the same 9-point scale. Third, the emotion evoked by the stimuli consumption was assessed using a single response emotion CEQ consisting of 12 emotion domains (Jaeger et al., 2020). Lastly, participants were asked how congruent the food was to the context (Chen et al., 2020).

After the assessment of all three samples, participants were allowed to take the VR headset off and completed a standardized questionnaire on a laptop. They answered questions related to sociodemographic variables (age, gender), health status (Chen et al., 2020), familiarity with VR (Chen et al., 2020), and their engagement (Hathaway & Simons, 2017). Self-reported states of health were registered on a 5-point scale (1=very bad ⇔ 5=very good), while VR familiarity was assessed using a 7-point scale (1=very unfamiliar ⇔ 7=very familiar). For engagement, statements related to usability, environmental aesthetics, novelty, involvement, and immersion were collected using a 5-point scale (−2=totally disagree ⇔ 2=totally agree) except for the statement regarding boredom which used a 5-point scale ranging from -2 = very boring to 2 = very fun. Responses related to the sensory awareness, realism, and distraction dimensions were collected using a 7-point categorical scale (0=none/not at all ⇔ 6=very; Bangcuyo et al., 2015; Hathaway & Simons, 2017). An overview of the questionnaire can be found in the supplementary material (S1).

### 2.5. Statistical analysis

Mixed-model ANOVAs (within-subject factor: product [watermelon, cracker, chocolate truffle]; between-subject factor: context [summer/winter]) were performed separately on the expected and experienced liking with Tukey's HSD post hoc test to examine whether the VR context impacted the liking scores of the food samples (Schouteten, Gellynck, & Slabbinck, 2019). Cochran's Q test was conducted to identify significant differences among the three stimuli for each of the 12 emotion domains for the winter and summer VR contexts separately (Jaeger et al., 2020). Where significant differences between stimuli were established, the sign test was used to test for pairwise differences (Jaeger et al., 2020). Fisher's exact test was performed to examine whether the emotions of each stimuli differed between the VR contexts.

The Likert data from the engagement questionnaire was analyzed in line with prior research by Hathaway and Simons (2017). First, 5-point Likert data from the engagement questionnaire was coded from −2 to 2, while categorical responses were coded from 0 to 6. Next, the coded data from relevant questions was averaged across panelists to generate dimensional scores. An overview of the items per dimension can be found in the supplementary material (S2). Finally, a total engagement score (TES) was composed for each participant by linearly summing the dimensional scores, varying from −10 (not engaged/distracted) to +28 (totally engaged). Independent t-tests were used to determine whether the dimensional and TES scores differed between the two VR contexts.

IBM SPSS Statistics (Armonk, NY) version 27 was used for the statistical analysis. All *p*-values below 0.05 were considered statistically significant.

## 3. Results

### 3.1. Product and context effects on expected and experienced liking

A significant interaction effect was found for the products and the different VR contexts, both for expected ($F(2,196)=6.10$, $p=0.003$) and experienced ($F(2,196)=8.64$, $p<0.001$) liking. For the former, the score of the watermelon was significantly higher in the summer VR context compared to the winter context. No significant difference in expected liking was found for the chocolate truffle or cracker between the VR contexts. When participants actually tasted the samples, the experienced liking was assessed with a significantly higher value for the watermelon in the summer context, while the opposite was established for the chocolate truffle in the winter context. No significant difference was established for the experienced liking of the cracker (as a control food) between the two VR contexts (Table 2).

*Table 2*

### 3.2. Product–VR context congruency

Table 3 provides the average product–VR context congruency for both the summer and winter VR contexts. The ratings for the congruency were highest for the watermelon sample and lowest for the chocolate truffe in the summer VR context. For the winter context, this was reversed, with the highest congruency observed for the chocolate truffle and the lowest for the watermelon. The average congruency for the cracker, as a more neutral sample, always lay between the other two samples. It should be noted however that crackers lay below the neutral score (5) and were deemed slightly incongruent with the winter VR context.

*Table 3*

### 3.3. Emotions

The emotions evoked by the stimuli were assessed by a single response emotion CEQ (Jaeger et al., 2020). The emotions of most stimuli lay in the lower right-hand quadrant (Figure 2a, b) covering the emotion pairs *happy–satisfied*, *secure–at ease*, and *relaxed–calm*. No significant differences were found in the emotions evoked by the stimuli in the winter VR context. For the summer VR context, a significant difference ($p=0.045$) in evoked emotions was found for the word pair *happy–satisfied*, although the sign test for pairwise differences was not significant ($p=0.10$). Fisher's exact test showed that the VR context did not lead to different evoked emotions for any stimulus (watermelon: $p=0.76$, chocolate truffle: $p=0.23$, cracker: $p=0.21$).

*Figure 2*

### 3.4. Engagement questionnaire

No significant difference ($t(98)=0.992$, $p=0.21$) was found between the average engagement scores of the participants of either VR context, with an average TES of 9.19 (SD=3.72) for the summer VR context and 8.27 (SD=3.60) for the winter VR context. When looking to the individual dimensions, as listed in Table 4, a significant difference was established only for the usability ("The testing environment assisted in my evaluations of the samples"), which was higher ($t(98)=3.56$, $p<0.001$) for the summer VR context (mean=0.63, SD=0.75) compared to the winter context (mean=0.00, SD=1.00), measured on a 5-point scale (-2=totally disagree ⇔ 2=totally agree).

*Table 4*

## 4. Discussion

While Yang et al.'s (2022) study showed that participant engagement can be improved by embedding the data collection in a VR environment, the current study examined whether the acceptance of food products changes due to congruency with the VR context to further explore whether participants were immersed in VR. The results showed that this was indeed the case, as the experienced liking (when tasting the food products) was higher for the products in the congruent VR context (watermelon in summer and chocolate truffle in winter). While this aligns with the proposed hypothesis, it also aligns with the findings of Song et al. (2022) that rye breads are more liked in a congruent VR restaurant environment then in an incongruent VR cinema environment. These results contradict an earlier study that did not find an impact of the congruency of a VR environment on a sweet beverage based upon grenadine syrup (Chen et al., 2020). However, the setup of that research differed, given that Chen et al. (2020) worked with three VR environments that were not at all natural (but relatively more abstract), used a within-subject design with only one sample (the same sample was presented in all three VR contexts), and asked participants to provide responses after removing the VR glasses. This could have contributed to the fact that an effect was only found for the perceived sweetness perception, not for overall liking.

The emotional responses for the products did not differ between the two VR contexts, which suggests that the emotional responses were quite stable. It should be noted that this study opted to work with a single response CEQ containing 12 emotion word pairs (Jaeger, Roigard, & Chheang, 2021). While this method has the benefit that participants only indicate a single emotion word pair, which reduces task time, this could have led to a lower discriminative ability (Jaeger et al., 2021). The selection of the samples (widely accepted), participants (product users), and VR environments (quite calm and relaxing) could also have played a role. This study found only a single response word pair CEQ that could discriminate between samples in the summer VR context. The higher discriminative ability of the single response emotion CEQ within Jaeger et al.'s (2021) aforementioned study could be due to their different research setup (larger number of samples and participants, only word pairs, and no tasting) or the fact that they did not work with an evaluation in a VR context. Prior research of Torrico et al. (2021) and Yang et al. (2022) with VR environments suggests that emotional responses could be more discriminating than conventional hedonic testing, which was not confirmed in this study. The type of assessment of emotions (e.g. check all-that apply (CATA) or scaling) could play a role, so future VR research might apply CATA or a scaling method if the focus lies on the assessment of emotions rather than the emotion word pair CEQ.

The TES, based upon the engagement questionnaire (Bangcuyo et al., 2015; Hathaway & Simons, 2017), did not differ between the two VR contexts. A higher engagement during the assessment might be beneficial to gain more reliable results (Bangcuyo et al., 2015). The average TES score in this study is around 9, which aligns with prior VR research by Yang et al. (2022) describing a TES of 9.43, with participants assessing beers in a VR bar context, and further strengthens that VR can—in line with other immersive technologies—lead to increased engagement compared to standard lab environments, which have reported TES scores around 0 in prior research (Bangcuyo et al., 2015; Hathaway & Simons, 2017; Zandstra, Kaneko, Dijksterhuis, Vennik, & De Wijk, 2020). However, one should consider the plural nature of engagement, as it can relate to different engagement types in research, such as the product, task, and technology (Pöyry, Parvinen, Mattila, & Holopainen, 2020). Future research should consider studying different engagement types when sensory testing is conducted with VR technology.

A strength of the current study is that the data was collected when participants were still immersed in the VR context. While the software used in this study does not require programming and can easily integrate a VR video, some improvements could still be made to increase the immersive feeling, as described in Wang et al.'s (2021) study. The questions and answers were superimposed on the VR context, which is similar to menu screens in games. However, for a more natural feeling, one option might be to incorporate speech-to-text techniques allowing participants to answer using their voice, as in an interview, or use a virtual pen to answer a virtual questionnaire. Although the data collection went smoothly and the current study received positive oral feedback from participants, the process was quite time consuming, as one person must assist each participant to provide the samples. Tracking devices combined with 3D model features could allow participants to see their own hands and foods in front of them and interact with the samples (Man, Patterson, & Simons, 2023). However, this requires support from IT specialists, which is not always available for researchers. Nevertheless, future research might look to directly compare these approaches and examine to what extent these methods lead to comparable immersive experiences by the participants. It might be interesting to explore mixed-reality technologies to not only allow a higher immersion but also collect data from multiple participants simultaneously. As the goal of VR research is to increase ecological validity, it is crucial that the overall quality of the VR context and the whole VR experience is sufficient (e.g., graphical quality, movement smoothness, etc.). As such, future research should also include such measurement. An overview of suggestions for conducting sensory testing with food products in VR with realistic methodology to increase the ecological validity of sensory tests can be found in table 5.

*Table 5*

This study has shown that VR can immerse consumers in a context in a way that influences consumer acceptance of food products, with products congruent with the consumption context being liked more in comparison to incongruent products. As such, VR could be considered an interesting, accepted option to mimic a more realistic eating environment, with the sensory evaluation of food products occurring in a standardized setting.

**References**


Bangcuyo, R. G., Smith, K. J., Zumach, J. L., Pierce, A. M., Guttman, G. A., & Simons, C. T. (2015). The use of immersive technologies to improve consumer testing: The role of ecological validity, context and engagement in evaluating coffee. *Food Quality and Preference, 41*, 84-95.

Chen, Y., Huang, A. X., Faber, I., Makransky, G., & Perez-Cueto, F. J. (2020). Assessing the influence of visual-taste congruency on perceived sweetness and product liking in immersive VR. *Foods, 9*(4), 465.

de-Magistris, T., & Gracia, A. (2017). Does hunger matter in consumer purchase decisions? An empirical investigation of processed food products. *Food Quality and Preference, 55*, 1-5.

Hathaway, D., & Simons, C. T. (2017). The impact of multiple immersion levels on data quality and panelist engagement for the evaluation of cookies under a preparation-based scenario. *Food Quality and Preference, 57*, 114-125.

Jaeger, S. R., Hort, J., Porcherot, C., Ares, G., Pecore, S., & MacFie, H. J. H. (2016). Future directions in sensory and consumer science: Four perspectives and audience voting. *Food Quality and Preference, 56*(B), 301-309.

Jaeger, S. R., & Porcherot, C. (2017). Consumption context in consumer research: methodological perspectives. *Current Opinion in Food Science, 15*, 30-37.

Jaeger, S. R., Roigard, C. M., & Chheang, S. L. (2021). The valence× arousal circumplex-inspired emotion questionnaire (CEQ): Effect of response format and question layout. *Food Quality and Preference, 90*, 104172.

Jaeger, S. R., Roigard, C. M., Jin, D., Xia, Y., Zhong, F., & Hedderley, D. I. (2020). A single-response emotion word questionnaire for measuring product-related emotional associations inspired by a circumplex model of core affect: Method characterisation with an applied focus. *Food Quality and Preference, 83*, 103805.

Kong, Y., Sharma, C., Kanala, M., Thakur, M., Li, L., Xu, D., et al. (2020). Virtual reality and immersive environments on sensory perception of chocolate products: A preliminary study. *Foods, 9*(4), 515.

Lagast, S., Gellynck, X., Schouteten, J. J., De Herdt, V., & De Steur, H. (2017). Consumers' emotions elicited by food: A systematic review of explicit and implicit methods. *Trends in Food Science & Technology, 69*(Part A), 172-189.

Man, K., Patterson, J. A., & Simons, C. (2023). The impact of personally relevant consumption contexts during product evaluations in virtual reality. Food Quality and Preference, 104912.

Pöyry, E., Parvinen, P., Mattila, O., & Holopainen, J. (2020). Engaged, but with what? Objects of engagement in technology-aided B2B customer interactions. Journal of marketing management, 36(3-4), 334-360.

Schouteten, J. J., Gellynck, X., & Slabbinck, H. (2019). Influence of organic labels on consumer's flavor perception and emotional profiling: Comparison between a central location test and home-use-test. *Food Research International, 116*, 1000-1009.

Song, X., Pérez-Cueto, F. J., & Bredie, W. L. (2022). Food Desires and Hedonic Discrimination in Virtual Reality Varying in Product-Context Appropriateness among Older Consumers. Foods, 11(20), 3228.

Torrico, D. D., Han, Y., Sharma, C., Fuentes, S., Gonzalez Viejo, C., & Dunshea, F. R. (2020). Effects of context and virtual reality environments on the wine tasting experience, acceptability, and emotional responses of consumers. *Foods, 9*(2), 191.

Torrico, D. D., Sharma, C., Dong, W., Fuentes, S., Viejo, C. G., & Dunshea, F. R. (2021). Virtual reality environments on the sensory acceptability and emotional responses of no-and full-sugar chocolate. *LWT, 137*, 110383.

Wang, Q. J., Escobar, F. B., Da Mota, P. A., & Velasco, C. (2021). Getting started with virtual reality for sensory and consumer science: Current practices and future perspectives. *Food Research International, 145*, 110410.

Xu, C., Siegrist, M., & Hartmann, C. (2021). The application of virtual reality in food consumer behavior research: A systematic review. *Trends in Food Science & Technology, 116*, 533-544.

Yang, Q., Nijman, M., Flintham, M., Tennent, P., Hidrio, C., & Ford, R. (2022). Improving simulated consumption context with Virtual Reality: a focus on participant experience. *Food Quality and Preference, 98*, 104531.

Zandstra, E. H., Kaneko, D., Dijksterhuis, G. B., Vennik, E., & De Wijk, R. A. (2020). Implementing immersive technologies in consumer testing: Liking and Just-About-Right ratings in a laboratory, immersive simulated café and real café. *Food Quality and Preference, 84*, 103934.

Table 1. Socio-demographic information (gender, age) and mean (SD) scores for hungriness (1=totally not ⇔ 5=extremely), health status (1=very bad ⇔ 5=very good), VR familiarity (1=very much unfamiliar ⇔ 9=very much familiar) and comfortability during VR (1=not at all ⇔ 9=extremely) of the participants of the summer (n=49) and winter (n=51) virtual reality context.

| | Summer (n=49) | Winter (n=51) |
|---|---|---|
| Gender (male/female) | 22/27 | 24/27 |
| Age (years) | 25.7 (9.6) | 30.0 (15.5) |
| Hungriness | 2.65 (1.11) | 2.59 (1.22) |
| Health status | 4.37 (0.60) | 4.53 (0.58) |
| VR familiarity | 3.24 (1.82) | 3.08 (2.15) |
| Comfortability during VR | 7.35 (1.18) | 7.16 (1.72) |

Table 2. Mean (SD) expected and experienced liking of the products in the summer (n=49) and winter (n=51) virtual context measured on a 9-point liking scale (1=extremely dislike ⬄ 9=like extremely).

| | **Summer (n=49)** Mean (SD) | **Winter (n=51)** Mean (SD) |
|---|---|---|
| ***Expected liking*** | | |
| Watermelon | $6.7^{Aa}$ (1.6) | $5.5^{Bb}$ (2.1) |
| Chocolate truffle | $6.1^{Aab}$ (1.9) | $6.7^{Aa}$ (1.7) |
| Cracker | $5.7^{Ab}$ (1.6) | $5.5^{Ab}$ (1.8) |
| ***Experienced liking*** | | |
| Watermelon | $7.3^{Aa}$ (1.4) | $6.4^{Bab}$ (1.9) |
| Chocolate truffle | $6.0^{Bb}$ (1.7) | $7.0^{Aa}$ (1.6) |
| Cracker | $6.5^{Ab}$ (1.6) | $6.1^{Ab}$ (1.9) |

$^{AB}$ Different uppercase letters indicate significant differences between the scores for a product between VR contexts

$^{ab}$ Different lowercase letters indicate significant differences between the products within a VR context (Summer / Winter) for either expected or experienced liking

Table 3. Mean (SD) product-VR context congruency of the products in the summer (n=49) and winter (n=51) virtual context measured on a 9-point scale (1=extremely incongruent ⇔ 9=extremely congruent).

| | **Summer (n=49)**<br>Mean (SD) | **Winter (n=51)**<br>Mean (SD) |
|---|---|---|
| Watermelon | $8.4^{Aa}$ (0.9) | $3.6^{Bb}$ (2.1) |
| Chocolate truffle | $3.3^{Ac}$ (1.6) | $5.2^{Ba}$ (2.1) |
| Cracker | $5.4^{Ab}$ (1.9) | $4.0^{Bb}$ (2.2) |

$^{AB}$ Different uppercase letters indicate significant differences between the scores for a product between VR context (Summer / Winter)

$^{ab}$ Different lowercase letters indicate significant differences between the scores of the products within a VR context (Summer / Winter)

Table 4. Mean (SD) for each dimension of the engagement questionnaire (Hathaway & Simons, 2017) experienced by the participants in the summer (n=49) and winter (n=51) virtual context.

| **Dimension** | **Summer VR context** Mean (SD) | **Winter VR context** Mean (SD) | p-value |
|---|---|---|---|
| Usability[1] | 0.6 (0.8) | 0.0 (1.0) | <0.001 |
| Environmental aesthetics[1] | 1.1 (0.6) | 0.9 (0.6) | 0.159 |
| Novelty[1] | 0.8 (0.6) | 0.8 (0.6) | 0.611 |
| Involvement[1] | 1.2 (0.4) | 1.2 (0.6) | 0.909 |
| Immersion[1] | 0.9 (0.5) | 1.1 (0.5) | 0.064 |
| Sensory awareness[2] | 3.9 (1.1) | 3.7 (1.1) | 0.396 |
| Realism[2] | 3.7 (0.9) | 3.6 (0.8) | 0.510 |
| Distraction[2] | 2.9 (0.6) | 3.0 (0.7) | 0.416 |
| Total Engagement Score[3] | 9.2 (3.7) | 8.3 (3.6) | 0.207 |

[1] coded on a 5-point scale (-2=totally disagree ⬄ 2=totally agree)

[2] coded on a 7-point scale (0=none/not ⬄ 6=very)

[3] Total Engagement score is the sum of the dimensions Usability, Environmental aesthetics, Novelty, Involvement, Immersion, Sensory awareness, Realism with the dimension Distraction subtracted. As such, the range of the Total Engagement Score is from -16 to 22.

Table 5. Considerations for the development of more realistic sensory food tests in VR

| **Suggestion** | **Relevant studies** |
|---|---|
| Include a realistic consumption context in VR | Torrico et al. (2020), Kong et al. (2020), Torrico et al. (2021), Yang et al. (2022), Man et al. (2023) |
| Ensure and assess congruency of participants, products and VR scenario | Chen et al. (2020) |
| Include measurements beyond liking (e.g. emotions, sensory profiling,…) | Torrico et al. (2020), Kong et al. (2020), Torrico et al. (2021), Yang et al. (2022), Man et al. (2023) |
| Include measurements to assess the engagement and quality of the VR environment | Song et al. (2022), Yang et al. (2022), Man et al. (2023) |
| Perform the measurements when participants are in the VR environment | Yang et al. (2022), Man et al. (2023) |
| Participants can see the samples in VR and pick them up by themselves | Yang et al. (2022), Man et al. (2023) |

Fig. 1. A) summer VR context (left), B) winter VR context (right)

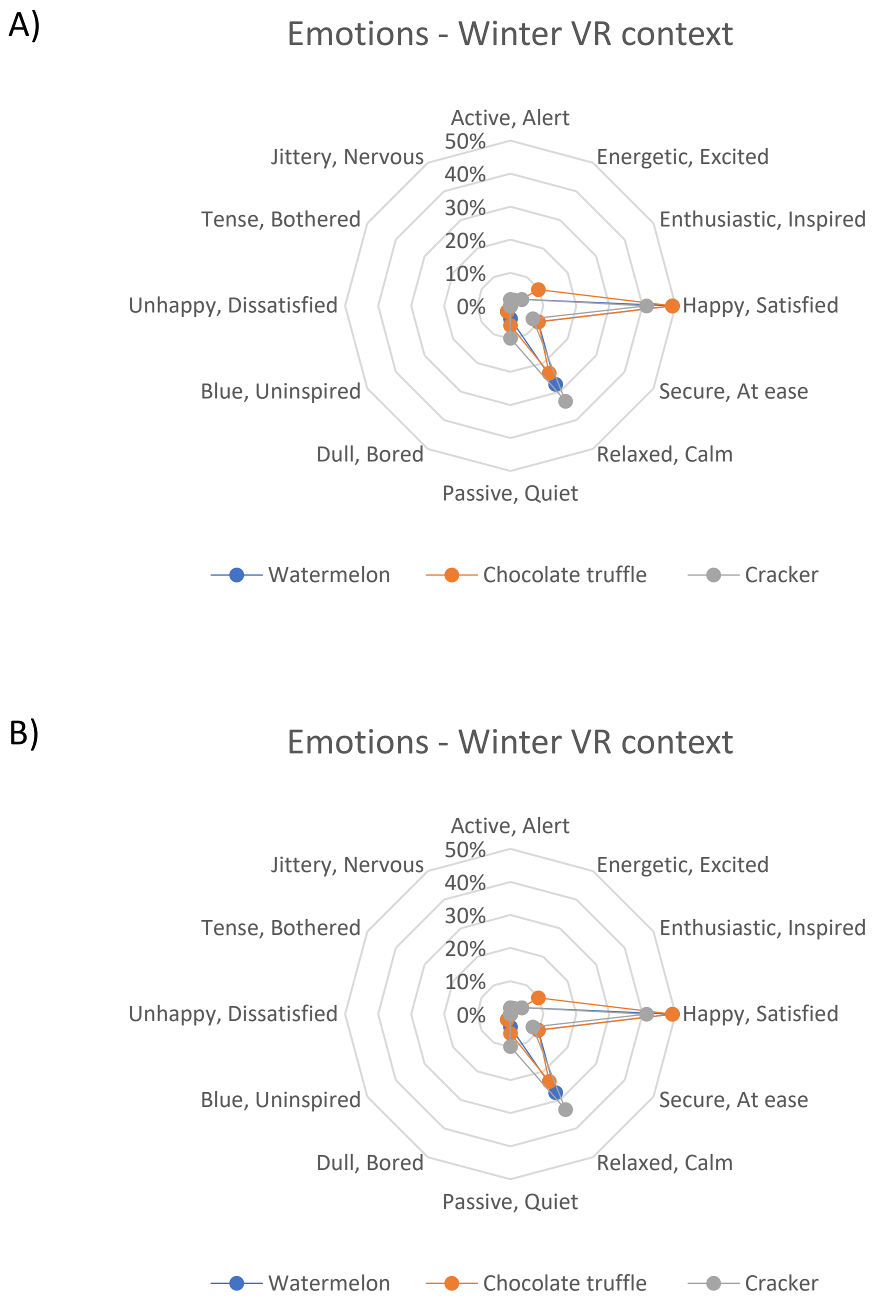


Figure 2A,B. Average frequencies of use across participants for selected word pairs for each stimuli in A) winter virtual context (n=51) and B) summer virtual context (n=49).