Food Research International 187 (2024) 114349



Contents lists available at ScienceDirect

Food Research International

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





Online consumer testing beyond central location tests: A case study for brewed coffee

Seyeong Park [a,b], Cho-Long Lee [a], Jungmin Oh [c], Manyoel Lim [a], Seo-Jin Chung [b], Han Sub Kwak [a,d,*]

[a] Food Processing Research Group, Korea Food Research Institute, Wanju-gun 55465, Republic of Korea
[b] Department of Nutritional Science and Food Management, Ewha Womans University, Seoul 03760, Republic of Korea
[c] Enterprise Solution Research Center, Korea Food Research Institute, Wanju-gun 55465, Republic of Korea
[d] KFRI School, University of Science and Technology, Wanju-gun 55465, Republic of Korea

ARTICLE INFO

Keywords:
Home-use test
Central location test
Non-face-to-face
Sensory evaluation
Remote test
Liking
Sensory profiling
Emotional response

ABSTRACT

Due to the constraints of the COVID-19 pandemic, conducting sensory evaluations requiring direct interactions became challenging. In response, researchers have been motivated to devise non-face-to-face testing methods as alternatives. This study aimed to compare two non-face-to-face home-use tests (HUT) with the traditional face-to-face central location test (CLT). Both HUTs involved online recruitment and sample delivery to participants' homes. One HUT provided a written protocol with no direct interaction (contactless HUT; C-HUT), whereas the other included an online meeting with a researcher for live guidance (online HUT; O-HUT). Four coffee samples were evaluated on the basis of liking and sensory and emotional attributes. The comparison between CLT and O-HUT showed RV coefficients of 0.92, 0.93, and 0.98 ($P < 0.05$) for liking and sensory and emotional attributes, respectively. In addition, based on the RV coefficient, the CLT results showed a significantly greater similarity to those of O-HUT compared to those of C-HUT. The O-HUT also outperformed the C-HUT in its ability to significantly discriminate between samples. Hence, real-time interactions between researchers and participants, as facilitated by O-HUT, may be more suitable in certain scenarios compared to C-HUT, which relies solely on a written protocol. Overall, these findings suggest that C-HUT and O-HUT are suitable methods for collecting sensory data and overcoming geographic and face-to-face contact limitations, providing greater flexibility, and reducing the time and cost associated with traditional sensory evaluations.

1. Introduction

Conducting sensory evaluations was limited during the COVID-19 period but is gradually becoming a distant memory. The lifestyles, working conditions, social interactions, and consumption behaviors of millions worldwide have undergone significant changes (Gómez-Corona et al., 2021). During the COVID-19 pandemic, face-to-face contact has been limited to prevent the spread of contagious diseases. This has resulted in many employees shifting to remote work (Galanti et al., 2021), students receiving more distance education (Toquero, 2021), and consumers purchasing products online and receiving them through delivery services (Ali et al., 2020). Despite the recent end of the COVID-19 pandemic, there has been a rapid growth in information technology-enabled non-face-to-face services, encompassing any form of communication or interaction without physical proximity.

In the field of sensory science, the pandemic period has affected the way researchers conduct sensory and consumer testing, requiring the switch to non-face-to-face methods. Home-use tests (HUTs) offer a good alternative to face-to-face tests, such as the central location test (CLT) (Lee & Lee, 2021; Romeo-Arroyo et al., 2022; Shi et al., 2021). In addition, since the onset of the COVID-19 outbreak, research on sensory evaluation has advanced, with the utilization of drive-in booths (DIB) in vehicles and online meeting platforms. Seo et al. (2021) conducted a comparison of hedonic, sensory, and emotional responses for four commercially available beverages in CLT and DIB settings. Overall, there were no significant differences in hedonic, sensory, and emotional responses between the two testing conditions. Kim et al. (2023) compared the acceptance and sensory profiles for plant-based jerky products across three distinct testing conditions: CLT, DIB, and HUT. Discrepancies in acceptance ratings were observed for some samples,

* Corresponding author at: Food Processing Research Group, Korea Food Research Institute, Wanju-gun 55465, Republic of Korea.
E-mail address: hskwak@kfri.re.kr (H.S. Kwak).

https://doi.org/10.1016/j.foodres.2024.114349
Received 13 December 2023; Received in revised form 5 April 2024; Accepted 17 April 2024
Available online 18 April 2024
0963-9969/© 2024 The Authors. 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/).

and the sensory profiles of the samples from the three test conditions were almost identical. These two studies showed that DIB can be used as an alternative to CLT, owing to the risk of face-to-face contamination during sensory evaluation. The Italian Sensory Science Society introduced a study on "remote sensory testing" to compare the performance of various sensory evaluation methods conducted at sensory facilities and at home (Dinnella et al., 2022). In their study, remote sensory testing was performed using discrimination tests, descriptive analysis, temporal dominance of sensation, and responses to check-all-that-apply questions. Consequentially, except for the tetrad test, which is a discrimination test, the results of the sensory tests were similar to those observed using CLT conditions. Tapia and Lee (2022) compared three different HUT conditions for cookie samples: a live test using an online meeting platform, asynchronous recorded video guidance, and a fully written protocol. The product acceptance ratings were consistent across the three methods. The panelists demonstrated a similar degree of sensory engagement under all the conditions. However, their study did not compare HUT to CLT, which is a traditional face-to-face test method. Recently, Park et al. (2023) compared CLT, HUT, and no-contact HUT. The no-contact HUT involved participant recruitment to sensory evaluation without any face-to-face contact. In particular, advancements in information technology (IT) including the widespread adoption of smartphones and the implementation of online contract systems and survey platforms for sensory data collection have been notable. Furthermore, logistic technologies, such as the development of fast and sophisticated courier systems, have facilitated sample delivery within two days in the Republic of Korea. These technological developments enable sensory evaluations to be conducted without face-to-face interaction.

HUT has some disadvantages, including challenges in sample control (e.g., managing short shelf-life, ensuring proper temperature control for refrigerated or frozen samples, and difficulties in concealing brand identities on the packaging), unregulated test conditions, and potential issues involving consumers' focus during evaluation. In addition, it is relatively difficult to determine whether consumers have correctly performed the evaluation (Boutrolle et al., 2007). These factors can lead to differences in sensory attributes and consumer acceptance of various food items depending on the testing condition (CLT or HUT) (De Wijk et al., 2019; Schouteten et al., 2021; Zhang et al., 2020). Performing HUT using an online meeting platform can serve as a viable approach to bridge the gap between HUT and CLT, addressing the challenges associated with uncontrollable factors during HUT. Online meeting platforms allow researchers to directly demonstrate the sample preparation process and, if necessary, immediately answer questions to ensure samples are prepared correctly. It can also help standardize evaluation procedures like CLT by clearly guiding participants through the evaluation process to ensure they are conducted correctly. Real-time interaction between researchers and participants can improve participant concentration by fostering greater engagement and active involvement in the evaluation process (Dinnella et al., 2022; Tapia & Lee, 2022).

Owing to recent advances in information and logistics technologies, various types of HUTs can be applied in the sensory evaluations. In this study, a non-face-to-face contact approach was applied to conduct the HUT. This study aimed to compare two non-face-to-face HUT test conditions with the conventional face-to-face CLT conditions. From recruitment to the sensory evaluation of HUTs and panel compensation, a non-face-to-face testing set-up was implemented. The participants were recruited online, and samples were delivered to their homes using a parcel delivery service for both HUTs. One HUT provided consumers with a written list of sensory evaluation protocols in a box with no direct interaction between researcher and participant. The other HUT additionally asked participants to participate in an online meeting to prepare samples and conduct sensory evaluations by following a live guideline by the researcher using an online meeting platform.

2. Materials and methods

2.1. Selection and preparation of samples

Sample selection focused on suitability for sensory evaluation. Coffee samples were appropriate considering the sample preparation difficulty. Brewing coffee is a familiar process, but participants had to consider water temperature and volume, sample preparation for brewing, brewing time, etc. Additionally, evaluating four samples was the maximum limit, considering the 1-h testing time and participants' ability to maintain focus during evaluation.

Four coffee samples were prepared based on bean type (Arabica and Robusta) and roasting degree (medium and dark) to maximize different sensory characteristics in coffee. Arabica beans are often described as sweeter and more acidic than Robusta, which tends to be harsh and bitter, making them distinctly different from each other (Di Donfrancesco et al., 2019). Furthermore, dark roasted coffee typically exhibits a higher level of bitterness compared to medium roasted coffee (Seninde et al., 2020). The samples used in this study were prepared according to the protocol used in a previous study (Park et al., 2023). Briefly, Colombia Supremo (GSC International, Co. Ltd., Seoul, Republic of Korea) and Vietnamese Robusta coffee (GSC International, Co. Ltd., Seoul, Republic of Korea) beans were roasted to medium and dark levels by a professional coffee roaster. Roasted coffee beans were individually packaged by a third-party coffee packaging company (Deuck-eun Co. Ltd., Gapyong-gun, Republic of Korea). Briefly, the coffee was ground and placed in a filter bag. This coffee drip bag was placed in PET/AL/OPP packaging with an oxygen absorber to maintain coffee quality. The samples were refrigerated at 4 °C until delivery to the participants. Arabica medium roast, Arabica dark roast, Robusta medium roast, and Robusta dark roast were labeled as AM, AD, RM, and RD, respectively.

2.2. Participants and procedure

2.2.1. Participants

All participants were native Korean, aged between 20 and 60 years old, with no health problems related to coffee consumption, and drank coffee at least once a week. Inclusion criteria were specified in the recruitment letter, and participants who met the inclusion criteria were selected by checking the online pre-questionnaire completed by volunteers. This study was approved by the Institutional Review Board of the Korea Food Research Institute (KFRI 2022–04-022).

2.2.2. Central location test

For the central location test, participants were recruited through an online bulletin board. A total of 110 participants were enrolled. Participants provided electronic consent using a legally recognized on-line contract system (Modusign, co. Ltd., Seoul, Republic of Korea). Participants were invited to participate in the CLT, and 101 (36 male, 65 female) of the 110 participants (91.8 %) attended the evaluation. The CLT was carried out at KFRI's sensory facilities. Participants received a coffee sample in a 384 mL paper cup with a coffee filter. Hot water (100 mL, 95 °C) was served in a 300 mL plastic beaker (MEP International, Goyang-si, Republic of Korea). Participants poured hot water for coffee brewing. After waiting for three minutes, the coffee and filter were removed, and the samples were evaluated. The coffee's approximate drinking temperature was maintained at 68–70 °C, as determined through preliminary experiments. The participants evaluated samples in a sequential monadic manner, and the samples were presented based on a Williams Latin square design. To minimize any potential carry-over effects, participants were given filtered water and crackers as palate cleansers.

2.2.3. Contactless home-use test

Participants for the contactless home-use test (C-HUT) were recruited by posting a flyer at Daangeun Market, and using a goods

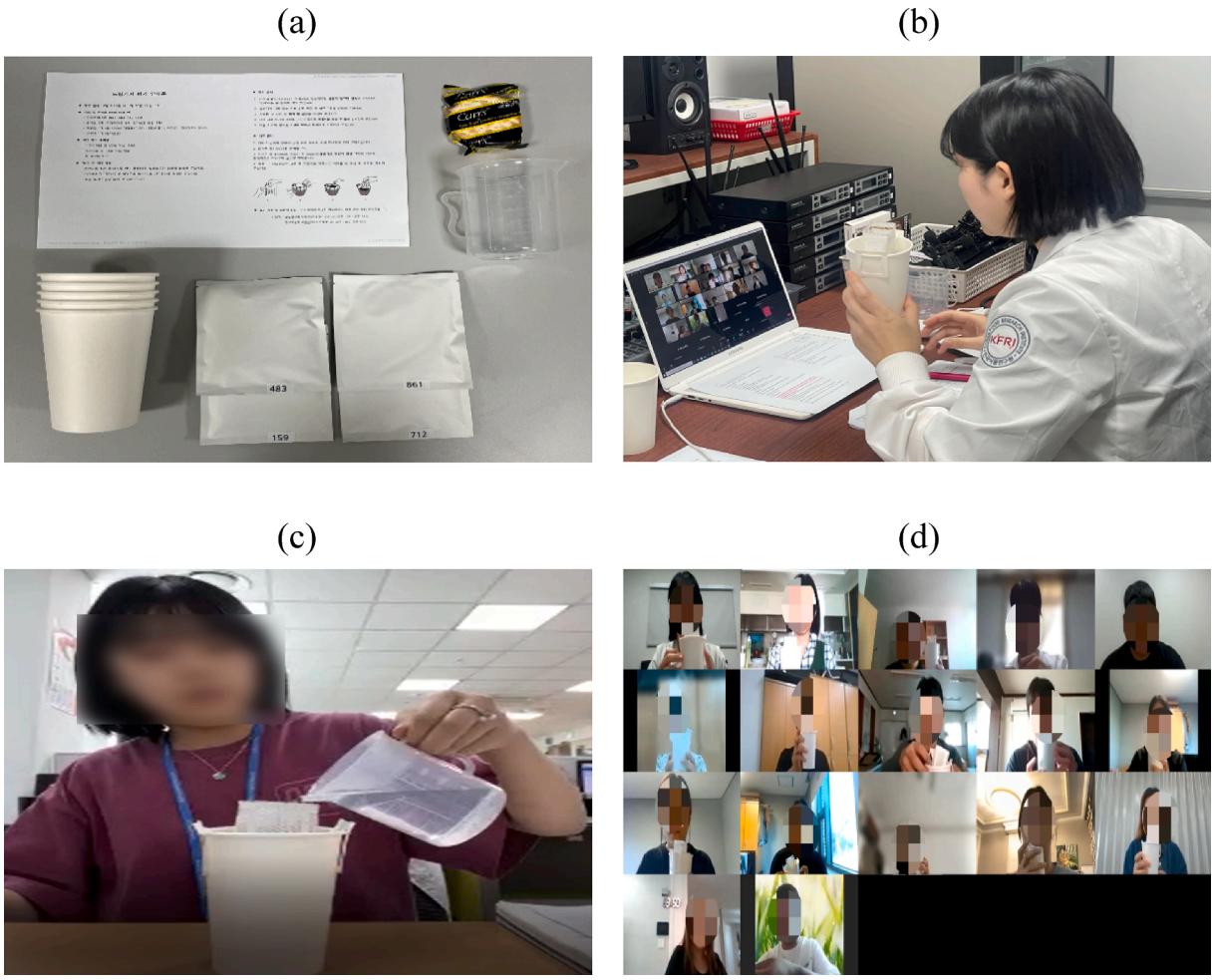


Fig. 1. (a) Evaluation materials delivered by a national parcel delivery service, (b) leading researcher explaining sensory evaluation to using the online meeting platform, (c) a panelist preparing a coffee sample observed via the online meeting platform., and (d) panelists using the online meeting platform in a session.

trading application (Daangeun Market Inc., Seoul, Republic of Korea). Recruitment was primarily conducted among the residents of Jeonju-si, Wanju-gun, and Gimje-si in the Republic of Korea, with a population of approximately 720,000. To be eligible for participation, respondents were required to express their intent to engage in an at-home coffee evaluation in the pre-survey through an online survey tool (Naver Co.). The eligible 120 participants were asked to fill out a consent form through a legally recognized on-line contract system (Modusign, Co. Ltd.). Out of 120 participants, 112 submitted their personal information and consent forms and were sent sample packages for sensory evaluation.

Sample packages were shipped using the Korean public postal system called ePOST (Naju-si, Republic of Korea). Each package contained four coffee samples, each labeled with a three-digit randomized code, along with supplementary items including four 384 mL white paper cups, a 300 mL heat-resistant plastic beaker (MEP International), Original Carr's table water crackers, and an instruction paper providing detailing the evaluation deadlines, sample preparation instructions, a mouthwash method, as well as the evaluation sequence and method for conducting sensory evaluations at home (Fig. 1a). All parcels were delivered within two days, as confirmed by tracking services. The sensory evaluation protocol was aligned with that of the CLT. The difference was that the participants were required to prepare boiling water, measure 100 mL, and brew coffee by themselves. Finally, 105 (93.8 %) out of the 112 participants completed their evaluations within four days. When evaluating the samples, the sample order was designated by the sensory data collection system using a Williams Latin square design.

2.2.4. *Online home-use test*

Similar to the recruitment method for C-HUT, 130 potential participants were recruited through an online used goods trading application (Daangeun Market, Inc.). Their intention to participate in the evaluation, sample delivery address, and access to an online meeting platform (Zoom; Zoom Video Communications Inc., San Jose, CA, USA) were collected using an online survey tool (Naver Co.). Of the 130 participants, 120 (92.3 %) signed the online consent form. The sample package was the same as the C-HUT, and instructions for participation in an online evaluation were delivered by a government public postal system (ePOST) (Fig. 1a).

The participants were asked to attend a scheduled session lasting 1 h and 30 min, occurring between 10:00 and 16:00 on weekdays. During the online meeting, participants were asked to use an appropriate device (tablet or laptop computer) to participate in an online video meeting for coffee evaluation and to use a smartphone to access an online ballot. One researcher led the live online evaluation while two other researchers monitored the study by handling unexpected issues during the evaluation. In each session, a maximum of 20 participants were invited, and their behavior was monitored during the evaluation (Fig. 1b). The online evaluation protocol closely followed the procedures outlined in the "Remote Sensory Testing" research project, conducted to assess the effectiveness and viability of remote sensory testing compared to that of traditional laboratory evaluations (Dinnella et al., 2022). The leading researcher initiated the online meeting and verified the presence of all participants' using the participant list. Individual audio and video connections were tested to ensure that participants' faces were visible on

the screen (refer to Fig. 1b and c). Subsequently, the primary investigator guided participants to access the data collection software and arrange the provided evaluation materials and water-boiling equipment on a suitable work surface, such as a table or desk. A review of the evaluation objectives and procedures was conducted, emphasizing that participants should refrain from speaking during the assessment. Instead, they were instructed to utilize the chat window on the online meeting platform for communication with the research team if needed. The participants were instructed to set the screen to show only their own face and the face of the leading researcher, and to mute the microphone (Fig. 1d). All participants participating in the session were asked to boil water, prepare the sample, brew coffee, wait for 3 min, drink the coffee, and complete the questionnaire. This procedure was repeated once per sample. The sample order was designated by the sensory data collection system using a Williams Latin square design. Each session lasted for approximately 1 h. Ultimately, 100 (83.3 %) of the initial 120 participants successfully completed the evaluation.

2.3. Questionnaire

The questionnaire comprised various inquiries concerning the acceptability of the product, sensory characteristics, and emotion responses elicited by each coffee sample (Park et al., 2023). Participants assessed the ratings of overall color, aroma, odor, taste, and mouthfeel of each coffee sample on a 9-point hedonic scale, ranging from "1 = dislike extremely", "5 = neither like nor dislike' to "9 = like extremely". A 5-point scale labeled as "1 = I would definitely not buy", "3 = I may or may not buy" to "5 = I would definitely buy" was used to measure purchase intent. Rate-all-that-apply (RATA) was used to generate sensory profiling and emotional response of the coffee samples. In the RATA method, participants were asked to check all the terms (check-all-that-apply, CATA) that they thought were relevant to the sample and rate the strength of the terms they selected using a three-point structured scale (1='weak', 2='medium', 3='strong'). Based on previous studies (Cotter et al., 2021; Seo et al., 2009; World Coffee Research, 2017), the 16 sensory attributes were selected in consultation with six sensory scientists. Attributes were aroma: citrus fruit, sour, burnt; taste and flavor: sour, bitter, sweet, burnt, citrus fruit, grain, dark chocolate, nutty; and mouthfeel: longevity, thickness, mouth drying, and astringent. The emotional terms were selected based on the overlap between the coffee drinking experience profiles (Kanjanakorn & Lee, 2017; Pinsuwan et al., 2022) and EsSense25® (Bhumiratana et al., 2014). The emotional responses were measured using seventeen emotional terms: energetic, free, peaceful, pleased, satisfied, merry, joyful, warm, understanding, guilty, worried, wild, good, pleasant, active, bored, and disgusted (Park et al., 2023). All sensory evaluation results were collected between June and September 2022 using the Compusense Cloud system (Compusense Inc., Guelph, Canada).

2.4. Data analyses

To examine the effects of samples and testing conditions on liking and purchase intention, the data set containing both metrics was analyzed using a generalized linear mixed model (GLMM) ($P < 0.05$). Participants were deemed as a random factor. A GLMM was applied: liking or purchase intent = sample + testing condition + sample × testing condition + testing condition × participant. Liking, purchase intention, and emotional ratings underwent one-way analysis of variance (ANOVA) to explore significant differences ($P < 0.05$) across the samples or testing conditions. If the difference in means across samples or testing conditions was significant ($P < 0.05$), Tukey's honest significant difference test was applied to identify mean separations. In the data matrices of the RATA method, unchecked attributes were replaced with zero as part of data preprocessing (Meyners et al., 2016). For each test condition, mean ratings of the sensory and emotional attributes for every sample were computed. Principal component analysis (PCA) was conducted to present the relationships between attributes and samples. Furthermore, to assess the configurations across the three testing conditions, multiple factor analysis (MFA) was conducted using three active variables, acceptability and sensory and emotional attributes. Within the three testing conditions, RV coefficients were calculated to generate the degree of similarity between two conditions (Robert & Escoufier, 1976). Permutation tests (5000 random) were used to test the significance ($P < 0.05$) of the RV coefficients (Josse et al., 2008). The analysis in PCA and MFA encompassed the first two dimensions. All data analyses were conducted using XLSTAT, version 2022 (Lumivero, Denver, CO, USA).

Table 1
Demographic profiles of subjects.

Category	CLT[1)]	C-HUT	O-HUT
Number of subjects			
Number of recruits	110	120	130
Number of filled out online consent forms	110 (100 %[2)])	112 (93.3 %)	120 (92.3 %)
Number of participants in the evaluation	101 (91.8 %[3)])	105 (93.8 %)	100 (83.3 %)
Age range[4)] (years)			
19–29	30 (29.7 %)	66 (62.9 %)	84 (84.0 %)
30–39	54 (53.5 %)	26 (24.8 %)	10 (10.0 %)
40–49	13 (12.9 %)	8 (7.6 %)	5 (5.0 %)
50–59	4 (4.0 %)	5 (4.8 %)	1 (1.0 %)
P-value5)	<0.0001		
Gender[4)]			
Male	36 (35.6 %)	41 (39.1 %)	36 (36.0 %)
Female	65 (64.4 %)	64 (61.0 %)	64 (64.0 %)
P-value 5)	0.85		

1) CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively.
2) Percentage of number of filled out online consent forms compared to the number of recruits.
3) Percentage of number of participants in the evaluation compared to number of filled out online consent forms.
4) Information on the participants in the evaluation
5) P-value obtained from the chi-square test.

3. Results

3.1. Demographic information

The demographic information of the participants is presented in Table 1. Consequently, the proportions of participants in the CLT, C-HUT, and O-HUT were 91.8 %, 93.8 %, and 83.3 %, respectively, based on the number of participants who finally agreed to participate in the test. In the CLT group, most participants were in their 30 s (53.5 %), whereas in the C-HUT and O-HUT groups, most participants were in their 20 s (62.9 % and 84.0 %, respectively) ($P < 0.001$). The distribution of male and female participants was relatively even across all three groups, with male participants accounting for approximately 40 % and

Table 2
Effect of the sample, testing location, and their interactions on the rating of liking and purchase intent for the coffee samples.

	Factors and their interaction			
	S[1)]	T	S × T	T × C
Overall liking	***[2)]	n.s.	n.s.	**
Color liking	n.s.	n.s.	n.s.	***
Aroma liking	n.s.	n.s.	n.s.	**
Taste and flavor liking	***	n.s.	n.s.	*
Mouthfeel liking	***	n.s.	n.s.	***
Purchase intent	***	n.s.	n.s.	**

[1)] S, T, and C denote sample, testing condition, and consumers, respectively.
[2)] Significance levels are as follows: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$; n.s. meant non-significant.

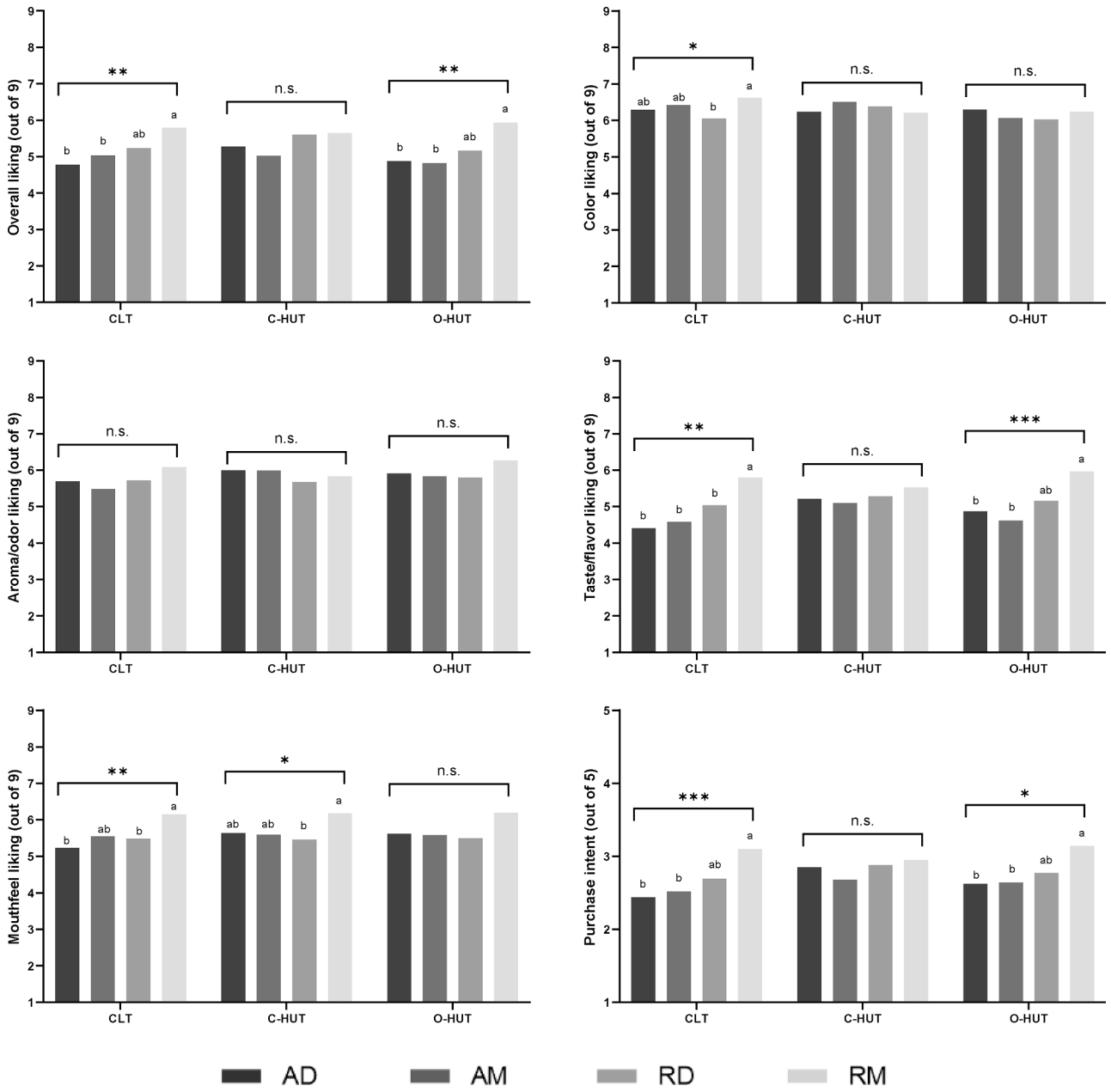


Fig. 2. Acceptability and purchase intent for the coffee samples on the three different testing conditions. Significance levels are as follows: (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$; n.s. meant non-significant. CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively. AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.

female participants approximately 60 % of the study population ($P > 0.05$).

The total labor hours were 44, 17, and 38 h for the CLT, C-HUT, and O-HUT, respectively (Table S1). For CLT, a major portion of labor time was allocated to sample serving, while for O-HUT, it was attributed to panel monitoring. The total costs, excluding labor and common materials, were USD 765.6, USD 2,364.6, and USD 2,632.1, respectively. The per capita costs for each person completing the evaluation were USD 7.6, USD 22.3, and USD 26.3, respectively.

3.2. Effect of testing conditions on liking and purchase intent

Table 2 presents the results of the ANOVA using GLMM for the samples and testing conditions regarding liking and purchase intent. There were significant differences across the samples for overall, taste/flavor, and mouthfeel liking, and purchase intent ($P < 0.05$). However, for liking and purchase intent, testing condition and sample × testing condition had no significant effect ($P > 0.05$), whereas significant differences were found for testing condition × participants ($P < 0.05$). The results of the ANOVA for the liking scores and purchase intent ratings across the three testing conditions without considering the sample types are shown in Table S2. There were no significant differences in the ratings for any of the items among the three testing conditions ($P > 0.05$). The comparison of acceptability and purchase intent across samples under each test condition is shown in Fig. 2. In terms of overall liking, significant differences were observed in CLT and O-HUT ($P < 0.01$), with the following liking rating order: RM, RD, AM, and AD. In the C-HUT group, although no significant differences were found between samples, the overall liking ratings for RM (5.7) and RD (5.6) were slightly higher than those for AD (5.3) and AM (5.0). The CLT showed significant differences in the ratings of four items: color, taste/flavor, mouthfeel liking, and purchase intent ($P < 0.05$). In the C-HUT group,

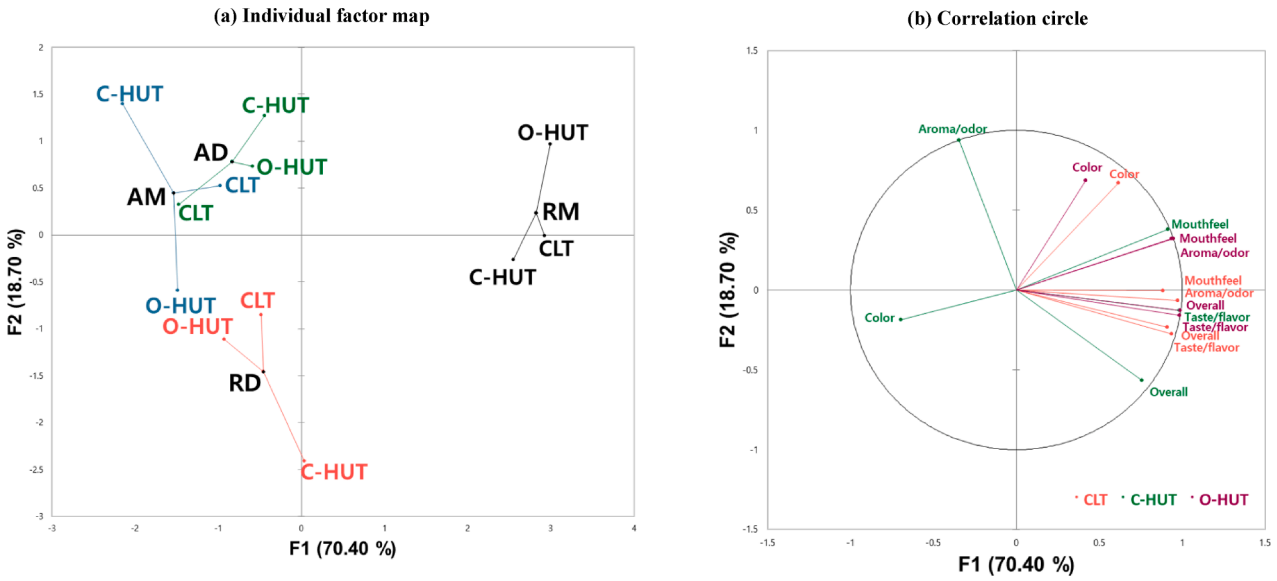


Fig. 3. MFA graph of (a) individual factor map and (b) correlation circle of acceptability for the coffee samples under the three different testing conditions. CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively. AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.

Table 3
The rating of sensory attributes for the coffee samples from RATA method.

Sensory attributes[3]	CLT[1] (N = 101)					C-HUT (N = 105)					O-HUT (N = 100)				
	AM[2]	AD	RM	RD	*P*-value	AM	AD	RM	RD	*P*-value	AM	AD	RM	RD	*P*-value
O_sour	$1.63^{a5)}$	1.37^{a}	0.46^{b}	0.42^{b}	***[4]	1.31^{a}	1.24^{a}	0.71^{b}	0.64^{b}	***	1.38^{a}	0.83^{b}	0.52^{bc}	0.41^{c}	***
O_citrus fuit	0.61^{a}	0.54^{ab}	0.32^{b}	0.34^{ab}	*	0.69^{ab}	0.74^{a}	0.43^{bc}	0.31^{c}	***	0.40^{a}	0.34^{a}	0.22^{ab}	0.08^{b}	**
O_burnt	1.23^{c}	1.62^{b}	1.20^{c}	2.02^{a}	***	1.22^{b}	1.17^{b}	1.22^{b}	1.66^{a}	**	0.71^{b}	0.97^{b}	0.73^{b}	1.50^{a}	***
T_bitter	1.43^{bc}	1.72^{ab}	1.03^{c}	1.98^{a}	***	1.20^{b}	1.45^{ab}	1.13^{b}	1.69^{a}	**	1.02^{b}	1.20^{ab}	0.81^{b}	1.57^{a}	***
T_sour	1.97^{a}	1.66^{a}	0.54^{b}	0.56^{b}	***	1.66^{a}	1.62^{a}	0.83^{b}	0.68^{b}	***	1.70^{a}	1.26^{b}	0.64^{c}	0.51^{c}	***
T_sweet	0.24	0.32	0.37	0.27	n.s.	0.69	0.52	0.68	0.60	n.s.	0.13	0.10	0.27	0.16	n.s.
F_roasted	1.05^{b}	1.22^{b}	2.06^{a}	1.76^{a}	***	1.14^{ab}	1.07^{b}	1.50^{a}	1.35^{ab}	*	0.33^{b}	0.68^{b}	1.18^{a}	1.06^{a}	***
F_dark chocolate	0.50	0.50	0.51	0.63	n.s.	0.67	0.61^{a}	0.62^{a}	0.84	n.s.	0.21	0.27	0.31	0.39	n.s.
F_burnt	1.11^{c}	1.60^{ab}	1.21^{bc}	1.93^{a}	***	1.10^{ab}	0.95^{b}	1.18^{ab}	1.49^{a}	**	0.76^{bc}	0.99^{ab}	0.58^{c}	1.20^{a}	***
F_citrus fruit	0.60^{a}	0.52^{ab}	0.32^{bc}	0.23^{c}	**	0.78^{a}	0.76^{a}	0.49^{ab}	0.40^{b}	**	0.40^{a}	0.25^{ab}	0.16^{b}	0.14^{b}	*
F_nutty	0.69^{c}	0.68^{c}	1.64^{a}	1.19^{b}	***	0.80^{bc}	0.70^{c}	1.30^{a}	1.13^{ab}	***	0.42^{b}	0.53^{b}	1.00^{a}	0.66^{ab}	***
F_grain	0.45^{b}	0.44^{b}	1.61^{a}	1.31^{a}	***	0.68^{bc}	0.51^{c}	1.25^{a}	1.01^{ab}	***	0.12^{c}	0.15^{c}	0.81^{a}	0.44^{b}	***
M_thickness	1.62^{bc}	1.93^{ab}	1.39^{c}	2.20^{a}	***	1.26^{b}	1.48^{ab}	1.20^{b}	1.77^{a}	***	0.88^{bc}	1.11^{ab}	0.64^{c}	1.38^{a}	***
M_longevity	1.27	1.37	1.04	1.44	n.s.	1.20	1.20	1.21	1.27	n.s.	1.10	0.93	0.92	1.01	n.s.
M_astringent	1.14^{a}	1.18^{a}	0.56^{b}	1.25^{a}	***	0.95	1.04	0.81	1.16	n.s.	0.74	0.71	0.55	0.80	n.s.
M_mouth drying	0.65	0.72	0.43	0.73	n.s.	0.72	0.76	0.70	0.74	n.s.	0.41	0.40	0.32	0.52	n.s.

5) Different superscripts within a row meant significant difference at $P < 0.05$ by Tukey's honest significance test.
[1] CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively.
[2] AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.
[3] O: odor, T: taste, F: flavor, M: mouthfeel. RATA method was asked using 3-point structured scale (1 = weak, 2 = medium, and 3 = strong).
[4] Significance levels are as follows: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$; n.s. meant non-significant.

only mouthfeel liking exhibited a significant difference among the samples ($P < 0.05$). The O-HUT group presented significant differences in the ratings for taste/flavor liking and purchase intent ($P < 0.05$).

To investigate the similarities and differences in the configurations of the three sensory evaluation conditions for the four coffee samples based on the acceptability results, MFA was performed (Fig. 3). The MFA results revealed that factors 1 (F1: 70.4 %) and 2 (F2: 18.7 %) accounted for 89.1 % of the total variance. In the CLT and O-HUT, the same items loaded closely together. However, in the C-HUT, some items, such as color, aroma/odor, and overall liking, loaded far apart from the same items in other conditions (Fig. 3b). The RV coefficients exhibited values of 0.80 ($P > 0.05$), 0.92 ($P < 0.05$), and 0.92 ($P < 0.05$) between CLT and C-HUT, CLT and O-HUT, and C-HUT and O-HUT, respectively (Table S3).

3.3. Effect of testing conditions on sensory profiling

The mean intensity ratings of 16 sensory attributes rated by the RATA method are shown in Table 3. Under each condition, the sensory attributes used to distinguish the samples were similar. In addition, the post-hoc analysis revealed a consistent trend in the rank order of the samples for each identical attribute under all conditions.

The PCA biplots in Fig. 4 display the results for the three test conditions based on RATA ratings. The total variance of PCA biplots was 99.3 % for CLT, 97.8 % for C-HUT, and 98.6 % for O-HUT. In the CLT analysis (Fig. 4a), Arabica coffee samples (AM and AD) were positively located on dimension 1 of the PCA biplot and characterized by citrus fruit and sour odors, citrus fruit flavors, and sourness. In contrast, Robusta coffee samples (RM and RD) were located on the negative side of the PCA biplot and characterized by graininess and nutty and roasted flavors. Dimension 2 divided samples based on the roasting degree. Dark

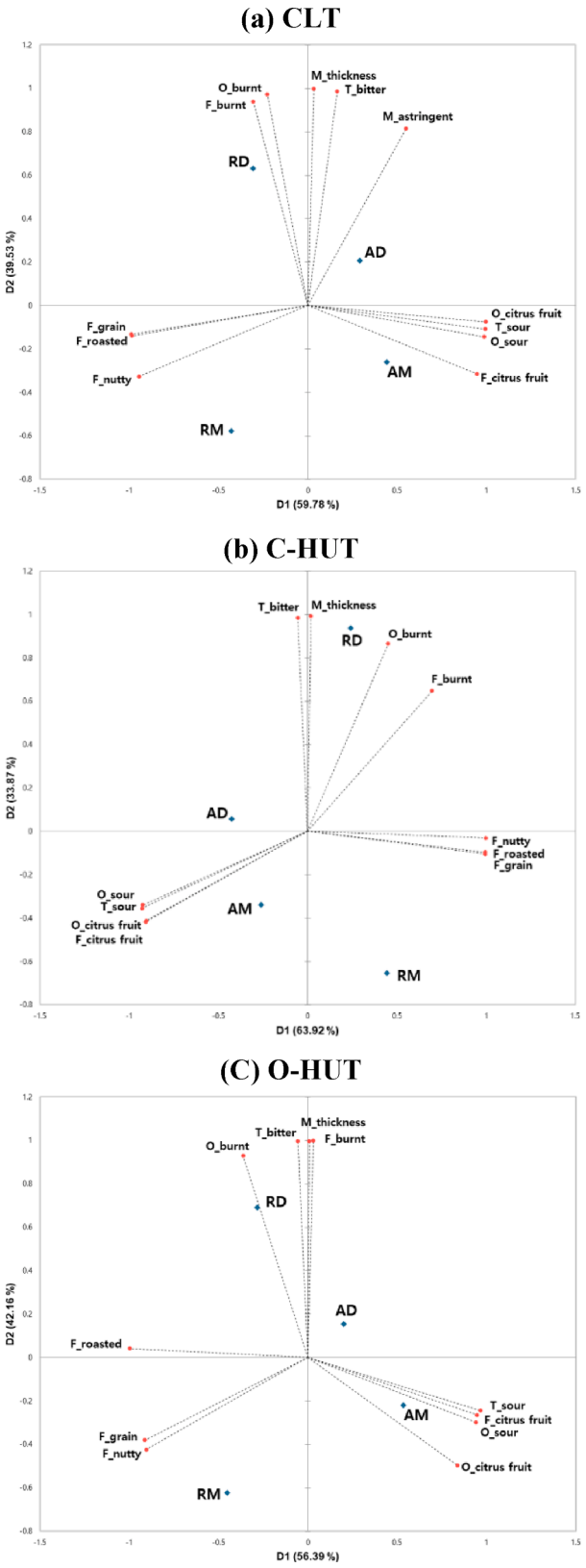


Fig. 4. PCA plot of the sensory attribute terms for the coffee samples under the three different testing conditions. O, T, F, and M stand for odor, taste, flavor, and mouthfeel. CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively. AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.

roasted samples (AD and RM) were characterized by burnt odor and flavor, bitterness, and thickness and mouthfeel sensation of astringency. The PCA biplot of C-HUT (Fig. 4b) also showed a similar pattern to that of the PCA biplot of CLT; however, the left and right sides were symmetrical. Similarly, the sensory characteristics of each sample in O-HUT (Fig. 4c) showed similarity with CLT (Fig. 4a) and C-HUT (Fig. 4b); however, astringency was not perceived by the participants in O-HUT.

Fig. 5 presents the results of the MFA for sensory attributes across the three test conditions. Factors 1 (50.0 %) and 2 (42.3 %) collectively accounted for 92.3 % of the total variance. Notably, when evaluating individual samples, it was evident that the O-HUT condition exhibited greater similarity to the CLT condition compared to the C-HUT condition (Fig. 5a). The RV coefficients between test conditions were 0.93 ($P > 0.05$) for CLT and C-HUT, 0.93 ($P < 0.05$) for CLT and O-HUT, and 0.85 ($P > 0.05$) for C-HUT and O-HUT (Table S3). Additionally, most of the same sensory attributes were closely clustered across all three test conditions (Fig. 5b).

3.4. Effect of testing conditions on emotion responses

Table 4 shows the rating of emotional terms for the coffee samples from RATA method. The number of terms that were significantly ($P < 0.05$) used to distinguish the samples in each testing condition was 9 for CLT, 3 for C-HUT, and 11 for O-HUT.

The results presented in Table 5 indicate significant differences in RATA ratings for emotional terms among three test conditions, with C-HUT scoring higher than the other conditions for all terms except "wild" ($P < 0.05$). In addition, the ratings for most emotional terms were significantly lower in the O-HUT group ($P < 0.05$).

The results of the MFA analysis applied to the similarity matrices of emotional responses within the CLT, C-HUT, and O-HUT settings, revealed a significant degree of similarity among all three test conditions, as depicted in Fig. 6. The first two factors collectively explained 85.2 % of the total variance, with Factor 1 and Factor 2 accounting for 55.7 % and 29.5 % respectively. The findings consistently indicated that O-HUT exhibited the highest similarity to both CLT and C-HUT across all samples, a trend supported by the RV coefficients. Specifically, the RV coefficients for comparisons between CLT and C-HUT, CLT and O-HUT, and C-HUT and O-HUT were calculated as 0.68 ($P > 0.05$), 0.98 ($P < 0.05$), and 0.78 ($P > 0.05$), respectively (Table S3).

4. Discussion

4.1. Comparison of participation rate and resources in the testing conditions

Not all recruits participated in the evaluation for each testing condition. In CLT, despite the participants scheduling their visiting time at least a day in advance, unexpected personal reasons prevented them from attending the evaluation at the facility. Participation rates were highest for the C-HUT conditions, where participants completed the evaluation at a location and time of their choosing. However, the participation rate under the O-HUT conditions was the lowest (83.3 %). Participants were required to log into the online meeting platform at the time they signed up for during recruitment. However, since participants did not have face-to-face contact with the researchers, they might have felt less obligation to attend the online meeting after receiving the samples.

The majority of participants were aged 20–30 years, possibly reflecting the preference for smartphone-based testing procedures. Young individuals, who are typically more familiar and comfortable with smartphone usage, might find the testing procedures such as participant recruitment, obtaining content forms, and evaluating the samples more accessible on this platform. Notably, more than half of the CLT participants were in their 30 s. KFRI is located on the outskirts of the city with poor transportation links. It would take approximately an hour

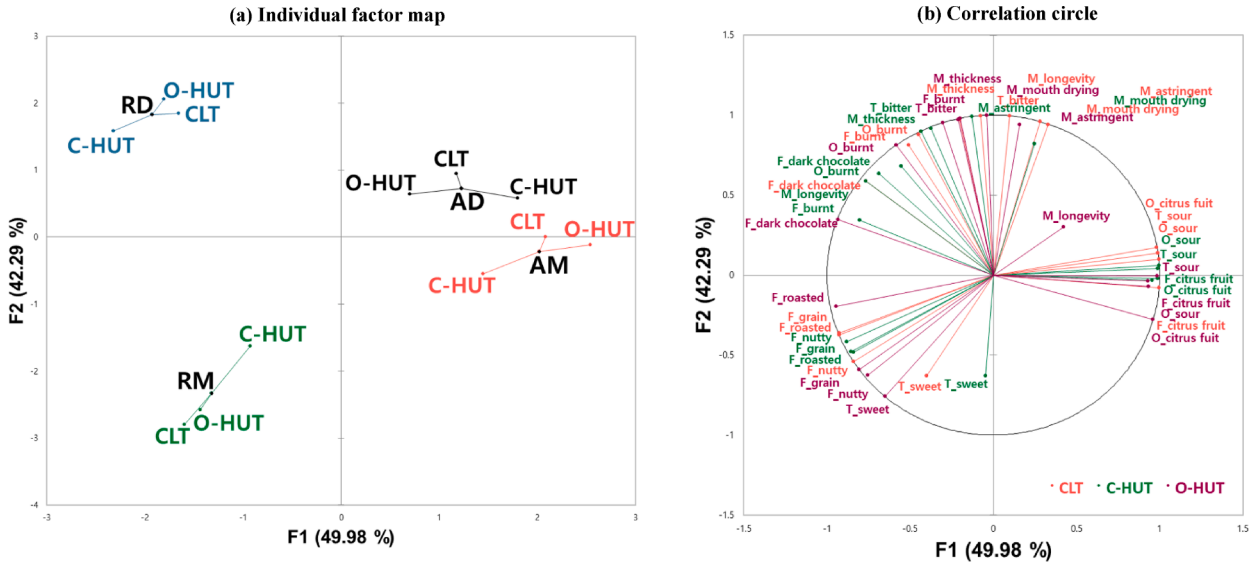


Fig. 5. MFA graph of (a) individual factor map and (b) correlation circle of sensory attributes for the coffee samples under the three different testing conditions. O, T, F, and M stand for odor, taste, flavor, and mouthfeel. CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively. AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.

Table 4
The rating of emotional terms for the coffee samples from RATA method.

Emotional terms[3]	CLT[1] (N = 101)					C-HUT (N = 105)					O-HUT (N = 100)				
	AM[2]	AD	RM	RD	*P*-value	AM	AD	RM	RD	*P*-value	AM	AD	RM	RD	*P*-value
active	0.79	0.61	0.49	0.61	n.s.[4]	0.80	0.83	0.71	0.56	n.s.	0.55^{a}	0.52^{ab}	0.28^{b}	0.27^{b}	*
merry	0.66^{a5}	0.60^{ab}	0.38^{b}	0.38^{b}	*	0.74	0.77	0.66	0.45	n.s.	0.60^{a}	0.38^{ab}	0.5^{ab}	0.21^{b}	*
wild	0.94^{bc}	1.18^{ab}	0.52^{c}	1.40^{a}	***	0.86^{ab}	0.63^{b}	0.60^{b}	1.18^{a}	***	0.64^{b}	0.78^{ab}	0.26^{c}	1.10^{a}	***
warm	0.81^{b}	0.55^{b}	1.31^{a}	0.90^{b}	***	0.86	0.88	1.15	1.10	n.s.	0.53^{b}	0.47^{b}	1.08^{a}	0.8^{ab}	***
bored	0.22	0.34	0.43	0.25	n.s.	0.43	0.43	0.48	0.54	n.s.	0.14	0.06	0.22	0.10	n.s.
peaceful	0.47^{b}	0.47^{b}	1.12^{a}	0.62^{b}	***	0.74^{b}	0.79^{b}	1.21^{a}	0.90^{ab}	*	0.56^{b}	0.47^{b}	1.17^{a}	0.63^{b}	***
worried	0.53	0.58	0.32	0.53	n.s.	0.43	0.48	0.46	0.46	n.s.	0.15	0.21	0.13	0.13	n.s.
understanding	0.22^{b}	0.25^{ab}	0.49^{a}	0.34^{ab}	*	0.46	0.55	0.52	0.53	n.s.	0.02^{b}	0.07^{b}	0.24^{a}	0.11^{ab}	**
disgusted	0.42^{a}	0.44^{a}	0.18^{b}	0.39^{a}	*	0.53	0.41	0.35	0.51	n.s.	0.29	0.18	0.13	0.32	n.s.
joyful	0.39	0.34	0.45	0.32	n.s.	0.63	0.58	0.50	0.55	n.s.	0.33^{a}	0.18^{ab}	0.29^{ab}	0.09^{b}	*
good	0.50^{b}	0.44^{b}	0.86^{a}	0.43^{b}	**	0.81	0.79	0.97	0.78	n.s.	0.44^{b}	0.49^{b}	0.91^{a}	0.46^{b}	**
energetic	0.59^{a}	0.39^{ab}	0.24^{b}	0.48^{ab}	*	0.78	0.64	0.57	0.54	n.s.	0.38^{ab}	0.43^{a}	0.15^{b}	0.22^{ab}	*
pleased	0.22	0.17	0.21	0.18	n.s.	0.57	0.51	0.43	0.44	n.s.	0.12	0.11	0.13	0.13	n.s.
pleasant	0.40	0.40	0.26	0.28	n.s.	0.68	0.58	0.48	0.55	n.s.	0.43^{a}	0.23^{ab}	0.17^{b}	0.15^{b}	*
guilty	0.39	0.43	0.17	0.28	n.s.	0.31	0.35	0.31	0.41	n.s.	0.14	0.11	0.07	0.21	n.s.
free	0.53	0.46	0.47	0.43	n.s.	0.87^{a}	0.72^{ab}	0.65^{ab}	0.46^{b}	*	0.34	0.33	0.24	0.18	n.s.
satisfied	0.50^{b}	0.51^{b}	0.89^{a}	0.63^{ab}	*	0.80	0.70	1.01	0.90	n.s.	0.39^{b}	0.34^{b}	0.93^{a}	0.63^{ab}	***

5) Different superscripts within a row meant significant difference at $P < 0.05$ by Tukey's honest significance test.
[1] CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively.
[2] AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.
[3] RATA method was asked using 3-point structured scale (1 = weak, 2 = medium, and 3 = strong).
[4] Significance levels are as follows: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$; n.s. meant non-significant.

by bus from the center of the university town, where many individuals in their 20 s reside, to reach the evaluation sites. Therefore, it is possible that more individuals in their 30 s participated due to their access to cars. Notably, 84 % of O-HUT participants were in their 20 s. The test required them to connect to online meetings using their laptop or desktop computer with a camera while simultaneously providing feedback on liking, attributes, and emotions related to each sample using their smartphone. This study's significant age distribution, with a majority in their 20 s and 30 s, suggests that IT familiarity and evaluation location are key factors influencing participation. It is important to consider IT, logistical, and demographic factors in study design to optimize participant engagement and representation in sensory evaluations.

The total labor time for the O-HUT was shorter than that of the CLT. The KFRI has sensory evaluation criteria using the CLT condition for various foods, whereas O-HUT, being a novel approach, commenced without predefined criteria. Therefore, the O-HUT procedure and findings in this study will serve as important references for future evaluation of O-HUT for other foods. Additionally, the time required to complete the evaluation procedure would be reduced compared to that for CLT. In terms of costs, similar to labor hours, the KFRI had a monetary compensation standard (approximately USD 7) for conducting CLT. The per capita costs for C-HUT and O-HUT were similar except for participant compensation. While delivery costs do not account for a large enough portion of the total cost to have a significant impact, participant compensation is an important financial consideration.

4.2. Comparison of liking among the testing conditions

The testing condition and sample × testing condition had no

Table 5
The rating of emotional terms for the three different testing conditions from RATA method.

Emotional terms[2)]	CLT[1)]	C-HUT	O-HUT	*P*-value
active	0.63[a4)]	0.73[a]	0.41[b]	***[3)]
merry	0.50[ab]	0.65[a]	0.42[b]	**
wild	1.01[a]	0.82[b]	0.69[b]	***
warm	0.89[ab]	1.00[a]	0.72[b]	**
bored	0.31[b]	0.47[a]	0.13[c]	***
peaceful	0.67[b]	0.91[a]	0.71[b]	**
worried	0.49[a]	0.45[a]	0.16[b]	***
understanding	0.32[b]	0.52[a]	0.11[c]	***
disgusted	0.35[a]	0.45[a]	0.23[b]	***
joyful	0.37[b]	0.57[a]	0.22[c]	***
good	0.55[b]	0.84[a]	0.58[b]	***
energetic	0.42[b]	0.63[a]	0.30[b]	***
pleased	0.19[b]	0.49[a]	0.12[b]	***
pleasant	0.34[b]	0.57[a]	0.25[b]	***
guilty	0.31[a]	0.35[a]	0.13[b]	***
free	0.47[b]	0.67[a]	0.27[c]	***
satisfied	0.63[b]	0.85[a]	0.57[b]	***

[4)]Different superscripts within a row meant significant difference at $P < 0.05$ by Tukey's honest significance test.
[1)] CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively.
[2)] RATA method was asked using 3-point structured scale (1 = weak, 2 = medium, and 3 = strong).
[3)] Significance levels are as follows: (***) $P < 0.001$; (**) $P < 0.01$.

significant effect on acceptance and purchase intent for the samples (Table 2). Previous studies have reported similar acceptance ratings between CLT and HUT for products such as snacks and wine (Lee & Lee, 2021; Sinesio et al., 2021). However, the number of attributes that significantly distinguished the coffee samples varied among the conditions. Specifically, the CLT condition had the highest number of significant attributes (five out of six), followed by O-HUT (three out of six) and C-HUT (one out of six) (Fig. 2). The participants tended to evaluate coffee samples more discriminately in the following order: CLT, O-HUT, and C-HUT. In our previous study comparing traditional HUT, no-contact HUT, and CLT, we observed a similar sample discrimination. HUT and no-contact HUT did not show statistically significant differences in acceptance and purchase intent across coffee samples when compared to CLT (Park et al., 2023). Although the C-HUT did not show significant differences ($P > 0.05$) in overall liking scores between samples, the RM sample had the highest overall liking scores across the three test conditions. This suggests that the participants in the C-HUT may have exhibited lower discrimination because of the lack of a controlled environment and potential distractions (Boutrolle et al., 2005).

In a previous empirical study conducted by Kim (2022), factors influencing immersion and learning outcomes in real-time Zoom classes were investigated. The findings suggested that interest and teacher-student interaction played crucial roles in determining immersion which, in turn, had a direct effect on academic performance. Applying these results, it is possible that the interaction between the leading researcher and panelists in O-HUT increased the immersion levels during the sensory evaluation process, leading to a heightened tendency to evaluate samples with greater discrimination. In addition, Klein et al. (2012) suggested that the presence of an experimenter can affect participants' performance. For example, there is the phenomenon of participants guessing what the experimenter wants or trying to be a good participant (Nichols & Maner, 2008). The similarity between CLT and C-HUT was 0.80 ($P > 0.05$), while the similarity between CLT and O-HUT and between C-HUT and O-HUT was 0.92 ($P < 0.05$) with values closer to 1 indicating greater similarity between test conditions. When measuring consumer impact, one argument is that data from HUTs have more ecological validity than data collected in controlled and sterilized environments like CLTs (Plaza et al., 2019). In the comparison above, in terms of acceptability, our RV coefficients suggest that O-HUTs are as ecologically valid as C-HUTs because participants engage and drink coffee in familiar spaces, produce better quality data than C-HUTs, and are significantly similar to CLT.

4.3. Comparison of sensory profiling among the testing conditions

The sensory attributes used to significantly differentiate the samples in each test condition were similar, and similar patterns were observed in the rank order of the samples for identical attribute terms under all conditions (Table 3 and Fig. 5). Sensory qualities are more likely to be consistent compared to liking, which is inherently subjective. The impact of the test conditions was limited when evaluating the inherent properties of the samples. This is consistent with previous findings, which showed a similar trend in the sensory attributes used to differentiate between the samples in CLT and HUT (Niimi et al., 2022; Park et al., 2023; Schouteten et al., 2019). In our study, the RV coefficients

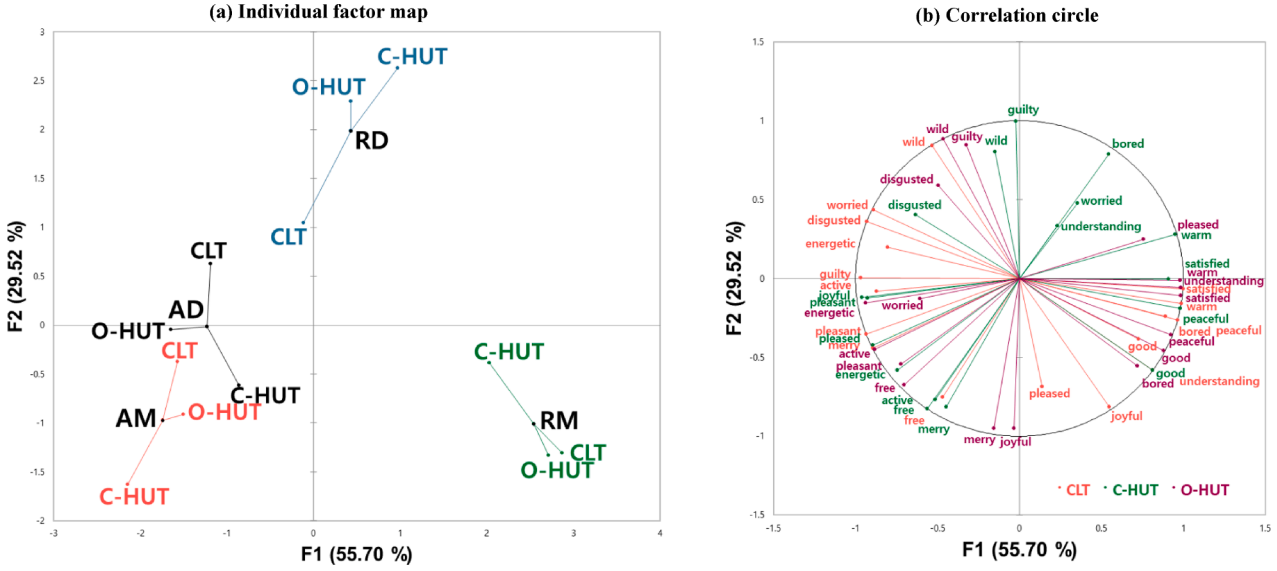


Fig. 6. MFA graph of (a) individual factor map and (b) correlation circle of emotional terms for the coffee samples under the three different testing conditions. CLT, C-HUT and O-HUT mean central location test, contactless home-use test, and online home-use test, respectively. AM, AD, RM and RD represent Arabica medium roasting, Arabica dark roasting, Robusta medium roasting, and Robusta dark roasting, respectively.

between CLT and both C-HUT and O-HUT were high, at 0.93. However, the RV coefficient between CLT and C-HUT was not significant ($P > 0.05$), and the coefficient between CLT and O-HUT was significant ($P < 0.05$). According to a previous study that compared sensory characteristics of four wines across evaluation conditions (CLT, CLT-evoked, and HUT), an RV coefficient of 0.8 or higher was considered as high similarity (Sinesio et al., 2021). This suggests that CLT yields comparable data to both HUTs. However, when trying to obtain similar data to CLT, O-HUT can offer more reliability and produce higher-quality data.

4.4. Comparison of emotion response among the testing conditions

Overall, emotion intensity was stronger in the uncontrolled condition (C-HUT) than in the controlled conditions (CLT and O-HUT) ($P < 0.05$) (Table 5). In the C-HUT, the researcher was not present and there were no time constraints other than the time control during coffee brewing. Therefore, consumers in the C-HUT may have felt less pressure to be a good participant (Nichols & Maner, 2008) and to complete the assessment on time than during the CLT and O-HUT. Participants' relaxed disposition may have influenced their emotional responses. In addition, the participants in the C-HUT used the term "free" to distinguish the samples. This may reflect the unrestricted nature of the C-HUT environment that is not bound by specific time or location constraints. In contrast, in both CLT and O-HUT, the experimental procedure was controlled by a leading researcher, with slight differences in the coffee preparation process. In the CLT, researchers provided boiled water (100 mL) for coffee, whereas in the O-HUT, the participants prepared and measured the boiling water themselves. In O-HUT, the necessity for the self-preparation of boiling water might have created a sense of time pressure during the online meeting, potentially resulting in lower emotion intensity, even if various emotions were experienced. Specifically, emotions characterized by low intensity, such as "bored," "worried," and "guilt," might have arisen because the panelists in O-HUT adhered strictly to the instructions given by the panel leader, without engaging in other thoughts. These differences in sample preparation and level of involvement may have affected the emotions experienced by the panelists under each condition. Park et al. (2023) employed the same four coffee samples as this study, comparing groups of consumers in the CLT, traditional HUT, and no-contact HUT conditions. The number of terms used to distinguish ($P < 0.05$) the samples was eight for CLT and four for the no-contact HUT, using the same protocol as the C-HUT. In our study (Table 4), the number of terms was nine for CLT and three for C-HUT, presenting similar results to the previous study (Park et al., 2023). This means that when repeating a test of the same nature, the types of terms may change. However, the number of terms that show significant differences between samples remains similar. In particular, the O-HUT used 11 terms to differentiate the samples with a similar set of emotional terms as the CLT. This suggests that the presence and control of the researcher may be a more important factor influencing the emotional response of participants to the sample than the testing location. This is supported by the RV coefficient between CLT and O-HUT (0.98, $P < 0.05$), indicating a high correlation. The variability in 'emotions' may arise because they are not always appropriate for describing food samples and because they can be influenced by the participant's state, independent of the food product.

4.5. Limitations

This study had some limitations. A reliable logistics network is essential to facilitate the delivery of the evaluation sample. However, this prerequisite may limit the accessibility of C-HUT and O-HUT, particularly in under developed countries. In O-HUT, the requirement for participants to interact with two separate devices (a laptop or tablet PC, and a smartphone) simultaneously could introduce usability challenges, potentially impacting the ease with which individuals can engage with and complete the evaluation process. This aspect of test ergonomics can further limit the inclusivity of O-HUT, such as the potential for participation by older adults, as it assumes a comfort and familiarity with multi-device interaction that not all individuals may possess. Generalizability of this study is limited by its focus on evaluating only four coffee samples. The selection, type, and preparation of the samples, frequency of evaluations, and the complexity of the sensory space could influence the outcomes. This necessitates further research to explore how these factors might affect sensory evaluation results across a wider range of products and contexts, ensuring a more comprehensive understanding of acceptability sensory profiling and emotional response.

5. 5 Conclusions

This study investigated the acceptability, purchase intent, and sensory and emotional profiles of four coffee samples across CLT, C-HUT, and O-HUT conditions. The O-HUT demonstrated a superior ability to significantly discriminate between the samples compared to that of C-HUT. Considering RV coefficients, compared to C-HUT, O-HUT exhibited a significantly closer resemblence to CLT ($P < 0.05$). The O-HUT allowed participants to evaluate the product in a familiar environment while the researcher maintained control. Considering the advantages of C-HUT has advantages, such as the high ecological validity, reducing the workload for researchers, and fewer geographical constraints, and disadvantages such as challenges in evaluation control and reduced consumer attention span. Therefore, O-HUT can be a useful alternative, combining the advantages and compensating for the drawbacks of CLT and C-HUT to improve data quality while increasing ecological validity. In addition, this study emphasized the importance of modern logistics technology (such as courier systems) and IT (such as online meeting platforms) to enable non-face-to-face interactions between researchers and participants during sensory evaluation.

Funding

This research was supported by the Main Research Program (Grant number: E0211300) of the Korea Food Research Institute (KFRI) funded by the Ministry of Science and ICT, Republic of Korea.

Ethical statement: All participants voluntarily participated and submitted informed consent forms prior to participation. This study received ethical approval for conducting research with human subjects from the Institutional Review Board (IRB) of the Korea Food Research Institute (approval number: KFRI 2022–04-002).

CRediT authorship contribution statement

Seyeong Park: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Cho-Long Lee:** Resources, Methodology, Investigation, Formal analysis. **Jungmin Oh:** Resources, Methodology, Investigation. **Manyoel Lim:** Writing – review & editing. **Seo-Jin Chung:** Writing – review & editing, Conceptualization. **Han Sub Kwak:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank Ji-sun Hwang and Bo Ram Kim for their assistance during the sensory evaluations.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2024.114349.

References

Ali, S., Khalid, N., Javed, H. M. U., & Islam, D. M. Z. (2020). Consumer adoption of online food delivery ordering (OFDO) services in Pakistan: The impact of the COVID-19 pandemic situation. *Journal of Open Innovation: Technology, Market, and Complexity, 7*(1), 10.

Bhumiratana, N., Adhikari, K., & Chambers, E., IV (2014). The development of an emotion lexicon for the coffee drinking experience. *Food Research International, 61*, 83–92.

Boutrolle, I., Arranz, D., Rogeaux, M., & Delarue, J. (2005). Comparing central location test and home use test results: Application of a new criterion. *Food Quality and Preference, 16*(8), 704–713.

Boutrolle, I., Delarue, J., Arranz, D., Rogeaux, M., & Köster, E. P. (2007). Central location test vs. home use test: Contrasting results depending on product type. *Food Quality and Preference, 18*(3), 490–499.

Cotter, A. R., Batali, M. E., Ristenpart, W. D., & Guinard, J. X. (2021). Consumer preferences for black coffee are spread over a wide range of brew strengths and extraction yields. *Journal of Food Science, 86*, 194–205.

De Wijk, R. A., Kaneko, D., Dijksterhuis, G., Van Zoggel, M., Schiona, I., Visalli, M., & Zandstra, E. (2019). Food perception and emotion measured over time in-lab and in-home. *Food Quality and Preference, 75*, 170–178.

Dinnella, C., Pierguidi, L., Spinelli, S., Borgogno, M., Toschi, T. G., Predieri, S., Lavezzi, G., Trapani, F., Tura, M., & Magli, M. (2022). Remote testing: Sensory test during Covid-19 pandemic and beyond. *Food Quality and Preference, 96*, Article 104437.

Di Donfrancesco, B., Guzman, N. G., & Chambers, E., IV (2019). Similarities and differences in sensory properties of high quality Arabica coffee in a small region of Colombia. *Food Research International, 116*, 645–651.

Galanti, T., Guidetti, G., Mazzei, E., Zappalà, S., & Toscano, F. (2021). Work from home during the COVID-19 outbreak: The impact on employees' remote work productivity, engagement, and stress. *Journal of Occupational and Environmental Medicine, 63*(7), e426.

Gómez-Corona, C., Rakotosamimanana, V. R., Sáenz-Navajas, M. P., Rodrigues, H., Franco-Luesma, E., Saldaña, E., & Valentin, D. (2021). To fear the unknown: Covid-19 confinement, fear, and food choice. *Food Quality and Preference, 92*, Article 104251.

Josse, J., Pagès, J., & Husson, F. (2008). Testing the significance of the RV coefficient. *Computational Statistics & Data Analysis, 53*(1), 82–91.

Kanjanakorn, A., & Lee, J. (2017). Examining emotions and comparing the EsSense Profile® and the Coffee Drinking Experience in coffee drinkers in the natural environment. *Food Quality and Preference, 56*, 69–79.

Kim, M. R., Heo, J., & Kwak, H. S. (2023). Exploring differences in Korean consumers' perceptions of plant-based jerky at different test locations: A comparison of the central location test, drive-thru test, and home-use test. *International Journal of Food Science & Technology, 58*, 4790–4799.

Kim, N. R. (2022). An empirical study on factors affecting immersion and learning outcomes in real-time non-face-to-face classes using Zoom. *Journal of the Korea Industrial Information Systems Research, 27*(2), 129–141.

Klein, O., Doyen, S., Leys, C., de Saldanha, M., da Gama, P. A., Miller, S., Questienne, L., & Cleeremans, A. (2012). Low hopes, high expectations: Expectancy effects and the replicability of behavioral experiments. *Perspectives on Psychological Science, 7*(6), 572–584.

Lee, N., & Lee, J. (2021). Comparison of home use tests with differing time and order controls. *Foods, 10*(6), 1275.

Meyners, M., Jaeger, S. R., & Ares, G. (2016). On the analysis of rate-all-that-apply (RATA) data. *Food Quality and Preference, 49*, 1–10.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology, 135*(2), 151–166.

Niimi, J., Collier, E. S., Oberrauter, L.-M., Sörensen, V., Norman, C., Normann, A., Bendtsen, M., & Bergman, P. (2022). Sample discrimination through profiling with rate all that apply (RATA) using consumers is similar between home use test (HUT) and central location test (CLT). *Food Quality and Preference, 95*, Article 104377.

Park, S., Heo, J., Oh, J., Chung, S.-J., & Kwak, H. S. (2023). Consumer testing away from a sensory facility: Application of home-use test and no-contact home-use test. *Food Quality and Preference, 104905*.

Pinsuwan, A., Suwonsichon, S., Chompreeda, P., & Prinyawiwatkul, W. (2022). Sensory Drivers of Consumer Acceptance, Purchase Intent and Emotions toward Brewed Black Coffee. *Foods, 11*, 180.

Plaza, A. G., Delarue, J., & Saulais, L. (2019). The pursuit of ecological validity through contextual methodologies. *Food Quality and Preference, 73*, 226–247.

Robert, P., & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: The RV-coefficient. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 25*(3), 257–265.

Romeo-Arroyo, E., Mora, M., Pazos, N., Deba-Rementeria, S., & Vázquez-Araújo, L. (2022). Effect of product properties and context on the perception of sweetness and liking: A case study with butter cookies. *Journal of Sensory Studies, 37*(3), e12740.

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.

Schouteten, J. J., Gellynck, X., & Slabbinck, H. (2021). Do fair trade labels bias consumers' perceptions of food products? A comparison between a central location test and home-use test. *Sustainability, 13*(3), 1384.

Seninde, D. R., Chambers, E., IV, & Chambers, D. (2020). Determining the impact of roasting degree, coffee to water ratio and brewing method on the sensory characteristics of cold brew Ugandan coffee. *Food Research International, 137*, Article 109667.

Seo, H.-S., Buffin, K., Singh, A., Beekman, T. L., & Arroyo, S. E. J. (2021). Stay safe in your vehicle: Drive-in booths can be an alternative to indoor booths for laboratory sensory testing. *Food Quality and Preference, 94*, Article 104332.

Seo, H. S., Lee, S. Y., & Hwang, I. (2009). Development of sensory attribute pool of brewed coffee. *Journal of Sensory Studies, 24*, 111–132.

Shi, M., Stelick, A., Licker, S., & Dando, R. (2021). On the validity of longitudinal comparisons of central location consumer testing results prior to COVID-19 versus home use testing data during the pandemic. *Journal of Food Science, 86*(10), 4668–4677.

Sinesio, F., Moneta, E., Di Marzo, S., Zoboli, G. P., & Abbà, S. (2021). Influence of wine traits and context on liking, intention to consume, wine-evoked emotions and perceived sensory sensations. *Food Quality and Preference, 93*, Article 104268.

Tapia, M. A., & Lee, S.-Y. (2022). Variations in consumer acceptance, sensory engagement and method practicality across three remote consumer-testing modalities. *Food Quality and Preference, 100*, Article 104616.

Toquero, C. M. (2021). Emergency remote education experiment amid COVID-19 pandemic. *IJERI: International Journal of Educational Research and Innovation, 15*, 162–176.

World Coffee Research. Sensory lexicon – unabridged definition and references. (2017). https://worldcoffeeresearch.org/resources/sensory-lexicon. Accessed 13 July 2022.

Zhang, M., Jo, Y., Lopetcharat, K., & Drake, M. (2020). Comparison of a central location test versus a home usage test for consumer perception of ready-to-mix protein beverages. *Journal of Dairy Science, 103*(4), 3107–3124.