Food Quality and Preference 116 (2024) 105149

Contents lists available at ScienceDirect

Food Quality and Preference

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

Assessing the visual appeal of real/AI-generated food images

Giovanbattista Califano [a,*], Charles Spence [b]

[a] Department of Agricultural Sciences, University of Naples Federico II, Naples, Italy

[b] Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, UK

ARTICLE INFO

Keywords:
NOVA
ChatGPT
AI-generated images
Food
Ultra-processed food

ABSTRACT

A study designed to investigate the ability of individuals to differentiate between AI-generated and authentic food images, as well as the impact of disclosing this information on the consumer perception of the appeal of these images is reported. Two online experiments were conducted with real and AI-generated food images stretching across the unprocessed, processed, and ultra-processed food continuum. Study 1 was designed to assess the accuracy with which people could identify AI-generated food images while Study 2 explored how the disclosure of an image's origin influenced the appeal of the depicted food. The participants in Study 1 found it very easy to recognize the AI-generated images, particularly in the case of ultra-processed foods. Notably, without disclosure, the AI-generated images were often preferred. At the same time, however, disclosing that a food image was genuine significantly boosted its appeal, whereas the revelation that it had been generated by AI mitigated this effect. These insights help to understand consumer psychology in the rapidly-evolving digital food marketing landscape, highlighting the nuanced effects of technological advancements in AI image-generation on human perception.

1. Introduction

Over the last decade or so, researchers working in the field of artificial intelligence (AI) have made remarkable progress, fundamentally altering various sectors of society, and blurring the distinction between reality and artificiality. This shift has been particularly notable with the introduction of generative AI models. Generative models are sophisticated AI systems that learn to create new content that is similar to the datasets on which they have been trained (Sætra, 2023). These models, such as OpenAI's ChatGPT, have garnered widespread public attention, while at the same time also raising widespread concern (e.g., Prem, 2023; Sallam, 2023).

In digital media, AI's influence in reshaping content creation and user interaction is profound. The field of AI-generated food imagery is evolving rapidly, with implications for everything from online grocery platforms, direct-to-consumer services, and the hospitality sector. For example, research commissioned by UK-based technology firm Slerp in 2023 examined public reactions to AI-generated versus authentic food images. Given that the visual presentation of food images on menus has been found to enhance the likelihood of dish selection (Hou et al., 2017; Spence, 2017a), the research sought to determine whether AI-generated food imagery could effectively substitute for traditional photographs in digital menus (Jackson, 2023). This question is especially pertinent for those organizations that are facing constraints in terms of their resources, time, and/or budget that impede the creation of authentic photographs. Startups such as Swipeby and Lunchbox are similarly venturing into AI-generated visual solutions for online menus (Walhout, 2023), targeting restaurants and delivery services. These developments underscore the changing dynamics of digital marketing and consumer engagement, with AI taking an increasingly central role.

The integration of AI in food imagery, while innovative, brings with it notable concerns. A primary issue is its potential impact on 'visual hunger', particularly if such images are disseminated widely and rapidly. Visual hunger refers to the way in which seeing images of food triggers appetite and food cravings while potentially also creating inappropriate internal references for appropriate portion size (Spence et al., 2016, 2022). AI-generated images, with their hyper-realistic and enhanced appeal (Miller et al., 2023), could potentially intensify this effect. This is especially relevant when considering the principles of gastrophysics, an emerging discipline that studies how various sensory elements influence the multisensory perception of food (Spence, 2017a; Velasco et al., 2021). AI has the capability, through analysis of patterns in data, to identify and perhaps even to amplify those specific features in food imagery that gastrophysics has shown to enhance the appeal of

* Corresponding author.
E-mail address: giovanbattista.califano@unina.it (G. Califano).

https://doi.org/10.1016/j.foodqual.2024.105149
Received 6 December 2023; Received in revised form 26 February 2024; Accepted 26 February 2024
Available online 28 February 2024

food, such as symmetry, shape, freshness, glossiness, dynamic-presentation, background/ambient lightning/colour, etc. (Spence et al., 2022). Additionally, the research shows that people are both fast and accurate in terms of estimating the energy–density of foods (Motoki et al., 2021), and that people's attention is rapidly drawn to energy-dense foods (Cunningham & Egeth, 2018), which are generally considered to be more appealing/palatable (Spence et al., 2022). This raises an intriguing question: can this human evolutionary trait potentially be exploited by AI? Such enhancements could potentially lead to a more pronounced influence on visual hunger, possibly influencing unhealthy eating behaviors or creating unrealistic expectations about food. While food photographs in advertisements and product packaging frequently deviate from the food's true appearance, or recommended portion size (Lazard et al., 2018; Petit et al., 2018), the widespread availability and ease of use of AI could lead to even more prevalent dissemination and 'deception'.

Another related concern is the need for clear disclosure policies regarding the AI-generated nature of these images. AI's ability to create images that are virtually indistinguishable from real photographs raises ethical issues concerning consumer deception and the importance of transparency in the context of digital marketing (Van Esch & Stewart Black, 2021). Furthermore, the inherent complexity or elaboration of the foods depicted in these images may significantly affect people's ability to distinguish between real and AI-generated content, as well as their impact on visual hunger. One method to categorize foods according to their level of manipulation is by their level of processing. The NOVA food classification system, originally proposed by Monteiro et al. (2019), offers a framework for this categorization. It divides foods into four categories: (1) unprocessed or minimally processed foods, (2) processed culinary ingredients, (3) processed foods, and (4) ultra-processed foods. Monteiro's system primarily considers the number and type of ingredients in a food product, rather than focusing on the actual method or extent of processing (Astrup & Monteiro, 2022; Hässig et al., 2023). From this perspective, the processing level could affect how effectively AI-generated imagery replicates the visual appeal and complexity of different food types. For instance, AI-generated unprocessed foods might be less acceptable than ultra-processed ones, given that the latter have already been manipulated extensively (see also van Tullen, 2023). This differential response could manifest itself in the ability of individuals to distinguish between real and AI-generated content and their perceptions of the appeal of these images.

These emerging concerns necessitate a deeper exploration of the interplay between AI-generated food imagery and consumer perception. While some researchers have begun to explore people's perception of, and opinions about, AI-generated content, such as images of human faces (e.g., Miller et al., 2023) and human speech (e.g., Herrmann, 2023), to the best of our knowledge, no study has yet investigated consumer perceptions of AI-generated food images. Therefore, the aims of the present study are twofold: First, the research explores whether individuals are able to distinguish between AI-generated and authentic images of food and whether this ability is affected by the degree of food processing. Second, it will investigate whether AI-generated images of food differ in perceived appeal compared to their real counterparts, whether the level of food processing modifies these perceptions, and finally, whether disclosure of the photo's nature (real or AI-generated) influences these assessments. Exploring these questions will help to dissect the complex interaction between AI-generated content and the response of consumers, shedding light on both the potential benefits and challenges inherent in this technological innovation in the realm of digital food marketing. At the same time, however, given the rapid advance of technology in this area, it is important to note that these results can reflect no more than a snapshot of current opinion (at the end of 2023) in what is a rapidly changing landscape (both in terms of the technological developments but also in terms of the consumer response).

2. Study 1

2.1. Stimuli

Five foods were selected as base items, with an unprocessed (or minimally processed), processed, and ultra-processed variant of each (see Table 1).

Following the selection of the 15 items, a representative photograph was obtained for each item from copyright-free databases. The criteria mandated that the photographs exclusively showcased the food product, without any human element. Additionally, we chose those foods that are typically consumed in their unprocessed state (for instance, we chose boiled rather than raw potatoes). Subsequently, for each authentic photo, an AI rendition was created using OpenAI's DALL-E 3 integrated into ChatGPT-4 between October and November 2023. The original content was uploaded with the prompt: "Replicate this photo" (see Supplementary Material). The model then generated a description for each photo, which was used by DALL-E to recreate the image from scratch, based on its training data (thus, the original image was not *modified* but entirely reconstructed). This process resulted in a pair of images for each of the 15 selected items: one real and the other AI-generated (see Table 2). The 30 photographs served as stimuli in Studies 1 and 2.

2.2. Participants and procedure

Study 1 evaluated the ability of participants to recognize AI-generated food photographs and was divided into three distinct sub-studies with independent samples (refer to Table 3 for participants' characteristics):

Study 1A ($N = 100$; $M_{age} = 41.3$ years, $SD_{age} = 14.1$): Participants were presented with the original and its AI-generated counterpart side-by-side (both position and item were randomized), and they were asked to identify which image was AI-generated.

Study 1B ($N = 100$; $M_{age} = 40.7$ years, $SD_{age} = 13.1$): Each participant viewed a single image per trial (either real or AI-generated), in a randomized order, and had to determine whether it was an actual food image, or one that had been created by AI.

Study 1C ($N = 99$; $M_{age} = 45.8$ years, $SD_{age} = 13.4$): The participants evaluated a single food image per trial (either real or AI-generated), in a randomized order, and rated their confidence in its authenticity on a 9-point differential scale (1 = "Definitely AI-generated"; 9 = "Definitely Real").

Table 1
Selected food items.

Food	Processing level	Selected item
Milk	Unprocessed	Milk
	Processed	Sweetened yogurt
	Ultra-processed	Chocolate milkshake
Potato	Unprocessed	Boiled potatoes
	Processed	Mashed potatoes
	Ultra-processed	Potato fries
Apple	Unprocessed	Apple
	Processed	Apple sauce
	Ultra-processed	Apple pie
Carrot	Unprocessed	Carrots
	Processed	Carrot juice
	Ultra-processed	Carrot cake
Peanut	Unprocessed	Peanuts
	Processed	Peanut butter
	Ultra-processed	Chocolate peanut treats

Table 2
Stimuli examples (carrot).

	Original photo	AI version
Unprocessed		
Processed		
Ultra-processed		

Note: Please refer to Supplementary Material for the full presentation of all 30 visual stimuli.

Table 3
Participants' characteristics in Studies 1A (N = 100), 1B (N = 100), and 1C (N = 99).

	1A	1B	1C	UK Census
Age				
≤ 15 years	0 %	0 %	0 %	19 %
16–24 years	11 %	13 %	4 %	11 %
25–34 years	24 %	22 %	18 %	13 %
35–49 years	37 %	40 %	31 %	19 %
50–64 years	21 %	22 %	40 %	19 %
≥ 65 years	7 %	3 %	6 %	19 %
Sex assigned at birth				
Male	37 %	42 %	31 %	49 %
Female	63 %	58 %	69 %	51 %

Note: Studies were balanced regarding sex ($\chi^2(2) = 2.45$, $p = .294$), but not age ($F(2, 396) = 4.53$, $p = .011$). However, no direct statistical comparison was conducted in Study 1 amongst the participants from sub-studies A, B, and C. Data for the latest UK Census (2021) were retrieved from: https://www.ons.gov.uk/census.

The hypothesis was that the three methods of investigation would yield similar patterns of results. However, while Study 1A was based on a joint evaluation mode, Studies 1B and 1C were based on a single evaluation mode. Since evaluators, in joint evaluations, can use the attributes of one image as a reference to evaluate the other image, evaluability should be higher in Study 1A (Hsee, 1996; Hsee & Zhang, 2010). We were interested in whether juxtaposing real with AI-generated images could accentuate discernible differences, thus potentially making participants more sensitive to the images' artificial nature in Study 1A. Additionally, we investigated whether any subtle sense of artificiality might be detected when participants were queried about their confidence in Study 1C.

Data collection took place in November 2023 via Google Forms, with UK participants recruited through Prolific Academic. Only those participants with normal or corrected-to-normal vision were invited, and those accessing the study via the desktop website were included to ensure proper display of visual stimuli, excluding mobile and tablet users. This study adhered to the principles outlined in the Declaration of Helsinki. Participants provided informed consent by confirming their understanding that their responses would remain confidential, and they agreed to participate in the survey. They retained the option to withdraw from the survey at any point without providing a justification. Ethics clearance for this study was obtained through a subcommittee of the University of Oxford Central University Research Ethics Committee [R85145/RE001].

2.3. Statistical analysis

In Studies 1A and 1B, the rate of accurately identifying each food item as AI-generated was compared to the chance of a random guess (50

%) using a *t*-test. Additionally, a multiple comparison *t*-test for dependent samples was used to assess differences in mean recognition rates across the three levels of food processing, with significance thresholds adjusted for multiple comparisons using Bonferroni correction ($\alpha = 0.05$ prior adjustment). Furthermore, a fractional logistic regression analysis was performed to evaluate the influence of age on the proportion of AI-generated images that were accurately distinguished. Fractional response models are commonly used for dependent variables ranging from 0 to 1, such as rates and proportions (Papke & Wooldridge, 1996). We used a logit distribution for the conditional mean to ensure interpretability similar to that of conventional logistic regression analysis.

In Study 1C, a within-subjects repeated measures ANOVA was conducted to assess the effects of processing level (unprocessed, processed, or ultra-processed) and the photo's nature (real or AI-generated) on participants' confidence in the genuineness of the photo. Pairwise post-hoc comparisons (Bonferroni-corrected) followed the analysis. Finally, two linear regression analyses were performed to assess the impact of participant age on the confidence scores regarding the authenticity of the images viewed, both for actual real images and AI-generated ones. All analyses were conducted using Stata 18.

2.4. Results of Study 1A (Joint evaluation)

The proportion of food items accurately identified as AI-generated was very high (Table 4), with a range extending from 62 % for processed carrots (i.e., carrot juice) to 91 % for ultra-processed peanuts (i.e., chocolate peanut treats).

A multiple comparison *t*-test for dependent samples was used to assess the differences in mean recognition rates across the three levels of food processing. The average recognition rate of AI-generated images for ultra-processed foods ($M = 0.83$, $SD = 0.24$) was significantly greater than that for both unprocessed ($M = 0.78$, $SD = 0.25$), $t(99) = 2.36$, $p = .022$, and processed foods ($M = 0.78$, $SD = 0.24$), $t(99) = 2.45$, $p = .016$. Nonetheless, no significant difference was observed between the unprocessed and processed foods. Subsequently, a fractional logistic regression analysis was performed to evaluate the influence of age on the proportion of accurately distinguished AI-generated images. As depicted in Fig. 1, older participants were less likely to correctly identifying AI-generated images (OR = 0.97, $p < .001$).

Table 4
Percentages of images correctly identified as AI-generated in Study 1A.

Food	Processing	%
Milk	Unprocessed	77
	Processed	83
	Ultra-processed	72
Potato	Unprocessed	78
	Processed	87
	Ultra-processed	82
Apple	Unprocessed	83
	Processed	78
	Ultra-processed	88
Carrot	Unprocessed	76
	Processed	62
	Ultra-processed	84
Peanut	Unprocessed	74
	Processed	78
	Ultra-processed	91

Note: All the rates statistically differed from random guessing (50%) at the 5% level.

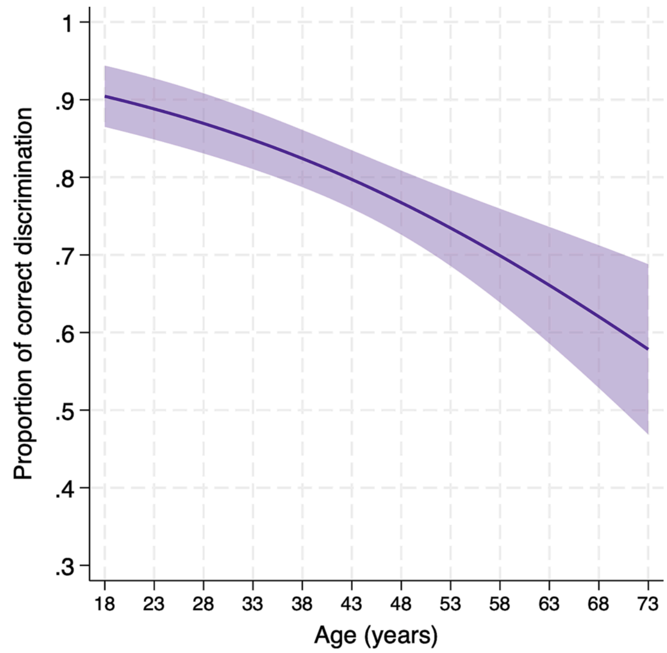

Fig. 1. Proportion of correct discrimination of AI-generated foods as a function of participants' age (line) in Study 1A, with 95% CIs (shaded area).

2.5. Results of Study 1B (Single evaluation)

The percentage of food items correctly identified was again generally high, though there were some exceptions perhaps attributable to the heightened complexity of the task (Table 5). The rate of correctly identified images ranged from 26 % for AI-generated processed apple (i.e., apple sauce) to 93 % for real ultra-processed peanut (i.e., chocolate peanut treats).

A multiple comparison *t*-test for dependent samples was conducted to evaluate the mean rates of correct identification across three levels of food processing. Consistent with the findings of Task B, the mean rate for ultra-processed foods ($M = 0.84$, $SD = 0.14$) was significantly greater than that for unprocessed ($M = 0.74$, $SD = 0.15$), $t(99) = 5.94$, $p < .001$, and processed foods ($M = 0.73$, $SD = 0.15$), $t(99) = 6.50$, $p < .001$. No significant difference was observed between the latter two categories. This pattern reinforces the observation that the ultra-processed foods selected for inclusion in this study were more readily identified as real or AI-generated. Furthermore, the experiment facilitated a comparison between correct identifications of real versus AI-generated images (see Fig. 2). A dependent samples *t*-test indicated that real images ($M = 0.81$, $SD = 0.14$) were identified more accurately compared to their AI-generated counterparts ($M = 0.73$, $SD = 0.16$), $t(99) = 4.31$, $p < .001$. Additionally, the correlation between the rates of correctly identified real and AI-generated images was not significant ($r = .11$, $p = .296$).

Furthermore, a fractional logistic regression analysis was used to evaluate the influence of age on the proportion of both real and AI-generated images correctly discriminated. As depicted in Fig. 3, older participants were less likely to correctly identifying AI-generated images (OR = 0.98, $p < .001$), while age had no influence on the likelihood of correctly identifying real food photos (OR = 0.99, $p = .181$).

2.6. Results of Study 1C (Single evaluation with confidence scores)

A within-subjects repeated measures ANOVA was conducted to assess the effects of processing level (unprocessed, processed, or ultra-processed) and photo's nature (real or AI-generated) on participants' confidence in the genuineness of the photo. The analysis indicated significant main effects for both processing level, $F(2, 490) = 6.65$, $p = .001$,

Table 5
Percentages of images correctly identified as real or AI-generated in Study 1B.

Food	Processing	Photo's nature	%
Milk	Unprocessed	Real	65
		AI	71
	Processed	Real	**49**
		AI	98
	Ultra-processed	Real	**57**
		AI	84
Potato	Unprocessed	Real	90
		AI	**52**
	Processed	Real	92
		AI	60
	Ultra-processed	Real	92
		AI	70
Apple	Unprocessed	Real	92
		AI	88
	Processed	Real	86
		AI	***26***
	Ultra-processed	Real	77
		AI	94
Carrot	Unprocessed	Real	92
		AI	**54**
	Processed	Real	85
		AI	69
	Ultra-processed	Real	90
		AI	90
Peanut	Unprocessed	Real	70
		AI	67
	Processed	Real	90
		AI	77
	Ultra-processed	Real	93
		AI	88

Note: In bold, the rates that do not statistically differ from a random guess (50 %, p >.05). In bold and italic, the rates that statistically differ from a random guess (p <.05) for a wrong answer.

and the nature of the photo, $F(1, 490) = 685.85$, p <.001. Additionally, the interaction effect of processing level and the image's nature was significant, $F(2, 490) = 21.54$, p <.001. Post-hoc comparisons depicted in Fig. 4 elucidate that participants' confidence was consistently higher for real photos, with a notably greater disparity observed for ultra-processed food images, aligning with the findings from Studies 1A and 1B.

Finally, two linear regression analyses were conducted to assess the impact of participant age on the confidence scores regarding the authenticity of the images viewed, both for actual real images and AI-generated ones. The results, illustrated in Fig. 5, revealed that as age increased, participants exhibited a tendency to be less confident in the authenticity of real photos ($\beta = -0.554$, p <.001) while becoming more confident in the authenticity of AI-generated images ($\beta = 0.316$, p =.001). This estimated trend converged towards the middle of the confidence scale, indicating uncertainty for both types of photos, after

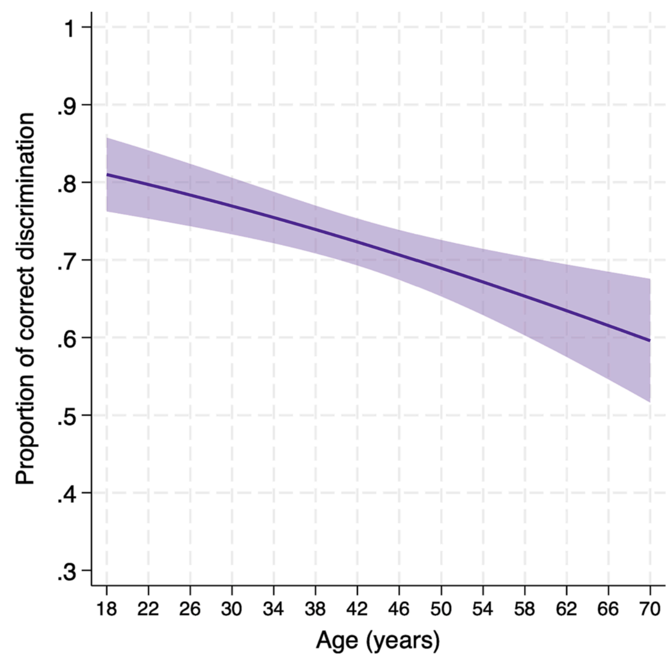

Fig. 3. Proportion of correct discrimination of AI-generated food images as a function of participants' age (line) in Study 1B, with 95% CIs (shaded area).

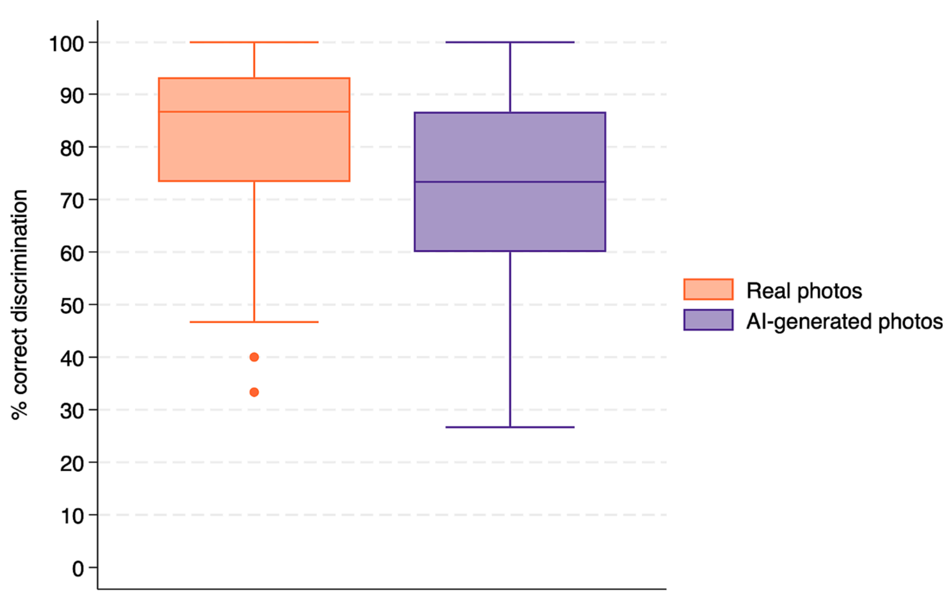

Fig. 2. Boxplots depicting the distribution of correct discrimination percentages for real versus AI-generated images in Study 1B. The plot includes the lowest and highest data points (whiskers), the first and third quartiles (box extremities), the median (line within the box), and any outliers (represented as dots).

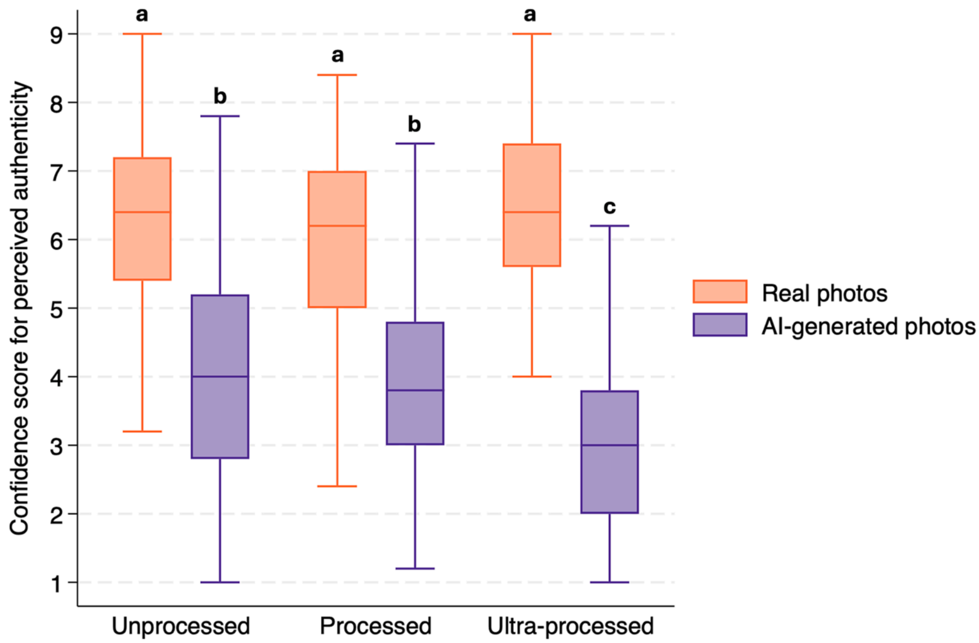

Fig. 4. Comparison of confidence scores for perceived authenticity of real versus AI-generated food images across processing levels in Study 1C. The boxplot of the distribution includes the lowest and highest data points (whiskers), the first and third quartiles (box extremities), and the median (line within the box). Letters indicate statistical differences at the 5 % level (p <.003 after Bonferroni correction).

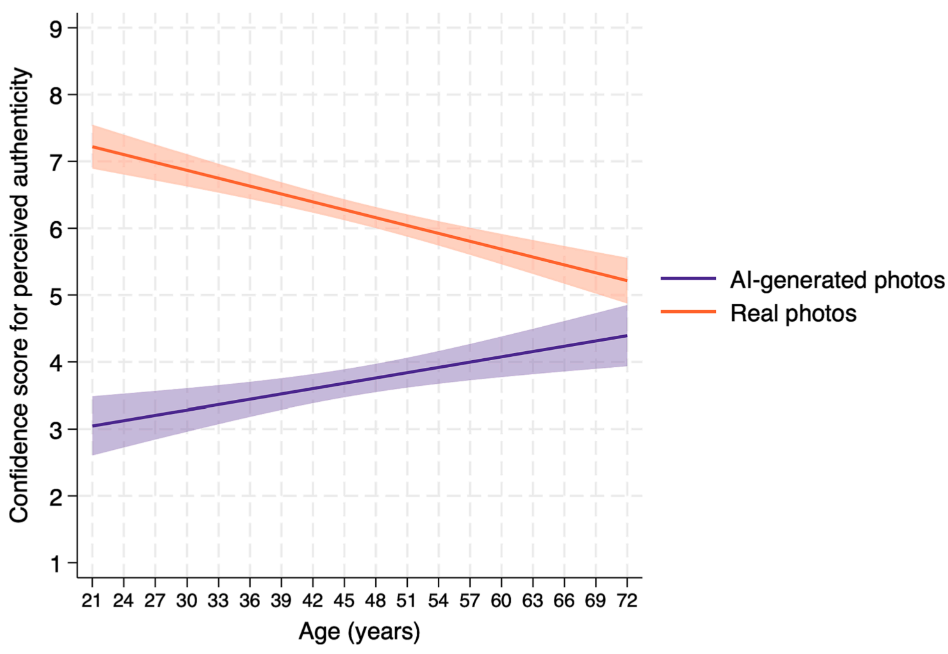

Fig. 5. Confidence scores for perceived authenticity of AI-generated food images (blue line) and real food images (orange line) as a function of participants' age in Study 1C, with 95% CIs (shaded areas).

the age of 70 years.

The findings from the three experiments conducted as part of Study 1 suggest that participants were largely able to distinguish between authentic and AI-generated food items, with this ability exhibiting some variations in accordance with the food's level of processing.

3. Study 2

3.1. Stimuli

The same set of 30 photos from Study 1 was also used for Study 2.

3.2. Participants and procedure

Study 2 explored whether knowledge of the image's nature/origin would influence how appetizing the various foods appeared. Similar to Study 1C, the participants assessed a single food image per trial (either real or AI-generated), in randomized order, and rated its appeal on a 9-point scale, from 1 ("Not at all appetizing") to 9 ("Extremely appetizing"). Thus, 297 participants (M_{age} = 39.9 years, SD_{age} = 13.2) were randomly assigned to one of three conditions: Unlabeled (N = 90; M_{age} = 38.2 years, SD_{age} = 12.8): Participants evaluated the stimuli with no explicit information provided about the nature of each photo; Labeled

($N = 87$; $M_{age} = 40.1$ years, $SD_{age} = 13.2$): The participants evaluated the stimuli with labels indicating whether the photo was real or generated by AI; Mislabeled ($N = 120$; $M_{age} = 41$ years, $SD_{age} = 13.5$): Participants evaluated the stimuli with labels that were reversed—the real images were tagged as AI-generated and vice versa. No differences in terms of sociodemographic characteristics were observed across label disclosure conditions, confirming the success of the randomization (Table 6).

Data collection took place in November 2023 via Google Forms, with UK participants recruited through Prolific Academic. Only participants with normal or corrected-to-normal vision were invited, and those accessing the study via the desktop website were included to ensure the proper display of the visual stimuli, excluding mobile and tablet users. This study adhered to the principles outlined in the Declaration of Helsinki. All of the participants provided informed consent prior to their participation. Subsequently, those in the unlabeled and mislabeled conditions received comprehensive debriefing post-task and were re-requested for informed consent. Ethics clearance for this study was obtained through a subcommittee of the University of Oxford Central University Research Ethics Committee [R85145/RE001].

3.3. Statistical analysis

To assess the influence of label disclosure (a between-subjects factor) and attributes of photos, including processing level and photo's nature (within-subjects factors), on participants' evaluations of food palatability, a linear mixed-effects model (LMM) was used. This model, estimated via maximum likelihood, integrated fixed effects for the main factors and interaction effects. These main effects and interactions were considered fixed effects in the analysis. Random intercepts for subjects and food items (15 combinations determined by food category and processing level) were incorporated to address variability in baseline palatability ratings between individuals (by-subjects intercept) and for specific foods/dishes (by-items intercept). Formally, the estimated model can be described as follows:

$$palatability_{ij} = \beta_0 + \boldsymbol{\beta} \boldsymbol{X}_{ij} + u_i + u_j + \epsilon_{ij} \quad (1)$$

where $i = 1,\ldots,297$ individuals and $j = 1,\ldots,15$ food items evaluated. β_0 represents the intercept, and $\boldsymbol{\beta}$ is a vector of parameters corresponding to the variables in $\boldsymbol{X}$ (i.e., label disclosure, processing level, photo's nature, and their interactions). $\beta_0 + \boldsymbol{\beta} \boldsymbol{X}_{ij}$ constitutes the fixed portion of the model, akin to any linear regression estimated via ordinary least squares. The random effects u_i and u_j shift this regression line based on participant and food item, respectively. ϵ_{ij} represents the error term, and it captures the unexplained variability in the palatability ratings of individual i evaluating food item j, after accounting for the fixed effects and the random effects. All analyses were conducted using Stata 18.

Table 6
Participants' characteristics in Study 2 across label disclosure conditions.

	Unlabeled ($N = 90$)	Labeled ($N = 87$)	Mislabeled ($N = 120$)	Total ($N = 297$)	UK Census
Age					
≤ 15 years	0 %	0 %	0 %	0 %	19 %
16–24 years	17 %	10 %	9 %	12 %	11 %
25–34 years	27 %	30 %	25 %	27 %	13 %
35–49 years	35 %	33 %	39 %	36 %	19 %
50–64 years	19 %	21 %	20 %	20 %	19 %
≥ 65 years	2 %	6 %	7 %	5 %	19 %
Sex assigned at birth					
Male	36 %	33 %	33 %	34 %	49 %
Female	64 %	67 %	67 %	66 %	51 %

Note: No significant difference was observed between disclosure conditions regarding both age ($F(2, 294) = 1.11, p = .332$) and sex ($\chi^2(2) = 0.14, p = .933$). Data for the latest UK census (2021) were retrieved from: https://www.ons.gov.uk/census.

3.4. Results

The LMM results highlighted the significance of all main effects, and interactions as well (except for processing level and photo's nature), in influencing the palatability score (Table 7), taking into account the significant variability between subjects and between food items (Table 8).

The interaction between label disclosure and the nature of the photo, as depicted in Fig. 6, reveals several insights: AI-generated images appeared more appetizing than real photos when the participants were unaware of their origin. However, when correctly labeled, there was no significant difference between real and AI-generated photos in terms of perceived palatability. Moreover, in the condition where the photos were mislabeled, participants rated those they believed to be real as more appetizing, even though they were actually AI-generated. Additionally, AI-generated photos were perceived as the most appetizing in the mislabeled condition (where participants were told that AI-photos were actually real), whereas real photos were rated as more appetizing when labeled as real, though there was no significant difference from the mislabeled condition. Indeed, while knowing that a photo was real increased its appetitive score, knowing that it was AI-generated did not significantly affect the score. Overall, these results suggest a positive influence on the palatability score due to the superior aesthetic appeal of AI-generated images compared to real ones, although counterbalanced by the negative effect of being aware that a food photo was generated by AI.

The analysis of the interaction between the level of food processing and the nature of the photos under different label disclosure scenarios (Fig. 7) revealed several key findings: 1) The ultra-processed foods were consistently rated as more appetizing than the unprocessed and processed foods, irrespective of the image's origin; 2) In those scenarios

Table 7
Fixed effects parameters of LMM for Study 2.

	Estimate	*SE*	*z*	*p*
(Intercept)	4.838***	0.137	35.36	< 0.001
Label condition				
Labeled	0.925***	0.195	4.74	< 0.001
Mislabeled	0.511**	0.181	2.82	0.005
Processing level				
Processed	0.151	0.134	1.13	0.260
Ultra-processed	1.391***	0.134	10.38	< 0.001
Photo's nature				
AI	0.333***	0.091	3.64	< 0.001
Label condition × Processing level				
Labeled × Processed	−0.310	0.191	−1.62	0.105
Labeled × Ultra-processed	−0.501**	0.191	−2.62	0.009
Mislabeled × Processed	−0.359*	0.177	−2.03	0.043
Mislabeled × Ultra-processed	−0.348	0.177	−1.96	0.050
Label condition × Photo's nature				
Labeled × AI	−0.474***	0.130	−3.63	< 0.001
Mislabeled × AI	0.375**	0.121	3.10	0.002
Processing level × Photo's nature				
Processed × AI	0.029	0.129	0.22	0.823
Ultra-processed × AI	−0.04	0.129	−0.31	0.757
Label condition × Processing level × Photo's nature				
Labeled × Processed × AI	0.134	0.185	0.73	0.467
Labeled × Ultra-processed × AI	−0.394*	0.185	−2.14	0.033
Mislabeled × Processed × AI	0.021	0.171	0.12	0.902
Mislabeled × Ultra-processed × AI	−0.312	0.171	−1.82	0.069

Note: *** $p < .001$; ** $p < .01$; * $p < .05$.

Table 8
Random effects parameters of LMM for Study 2.

	σ^2	*SE*	95 % CI	
			LL	*UL*
Subjects				
(Intercept)	0.876	0.089	0.718	1.069
Food items				
(Intercept)	2.160	0.071	2.025	2.303
Residual	1.883	0.040	1.806	1.962

Note: Likelihood-ratio test vs. linear model: $\chi^2(2) = 2753.96$, $p < .001$.

where the nature of the images was not disclosed to participants (unlabeled condition), AI-generated images were preferred over their real counterparts; 3) When the nature of the photos was disclosed to participants (labeled condition), there were no significant differences in appetizing ratings between real and AI-generated images, except for ultra-processed foods, which were found to be more appetizing in their real forms; 4) In those situations where the labels were reversed (mislabeled condition), participants consistently rated the AI-generated versions as more appetizing across all processing levels.

The results of Study 2 suggest that AI-generated food images were perceived differently in terms of their appeal when compared to their real counterparts, with the level of food processing impacting these ratings. Furthermore, the disclosure of the photograph's origin—whether real or AI-generated—was found to significantly influence these assessments.

4. Discussion

The aim of the two studies reported here was to investigate people's ability to distinguish between AI-generated and authentic food images and also to understand their perceptions of the appeal of AI-generated food items under various disclosure conditions. The participants in Study 1 demonstrated a high level of performance in terms of identifying AI-generated food photographs, especially in Study 1A, which used a joint evaluation mode. According to the General Evaluability Theory (Hsee, 1996; Hsee & Zhang, 2010), in joint evaluations, participants can use the attributes of one image as a reference to evaluate another, thereby enhancing its evaluability and participants' sensitivity to value. In our case, this value was not monetary but was represented by the authenticity of the photos. The joint evaluation mode, as opposed to the separate evaluation in Studies 1B and 1C, likely facilitated participants' ability to accurately discern between AI-generated and authentic food images.

Recognition was also higher in the case of the ultra-processed foods. This could be attributed to the fact that these foods are already highly manipulated, thus making the AI alterations more conspicuous and perhaps too artificial. Additionally, the recognition rate for real photos was higher than for AI-generated ones, aligning with the findings of other studies (e.g., Lu et al., 2023). This is likely due to participants' lifelong exposure to real food stimuli (Herrmann, 2023; Lu et al., 2023). Notably, there was no correlation between the ability to identify real and AI-generated images, suggesting that these are two different domains of perceptual visual expertise (see also Ivy et al., 2023). The ability to distinguish between real and AI-generated images decreased with age, suggesting a potential generational divide in terms of recognition skills when interacting with such new technical possibilities in terms of food image generation. This aligns somewhat with Herrmann's (2023) findings, where older adults were shown to be less adept at distinguishing modern AI-synthesized speech from human speech when compared to younger adults.

Study 2 explored how labeling influenced the perceived appeal of food images. Without disclosure, the AI-generated images were consistently rated as appearing more appetizing than their real counterparts. This aligns with research on the attractiveness of AI-generated human faces (Miller et al., 2023; Tucciarelli et al., 2022). Such a preference with regards to food photos may reflect the tendency of AI models like DALL-E to enhance image desirability, paralleling discoveries in gastrophysics (Spence et al., 2022). For instance, the AI-generated stimuli appear somewhat glossier (refer to Fig. S3 in Supplementary Material) and with warmer and more uniform lighting (as shown in Fig. S4, Supplementary Material) compared to their original counterparts. Note that glossiness, which has been associated with the perception of freshness (Bailey,

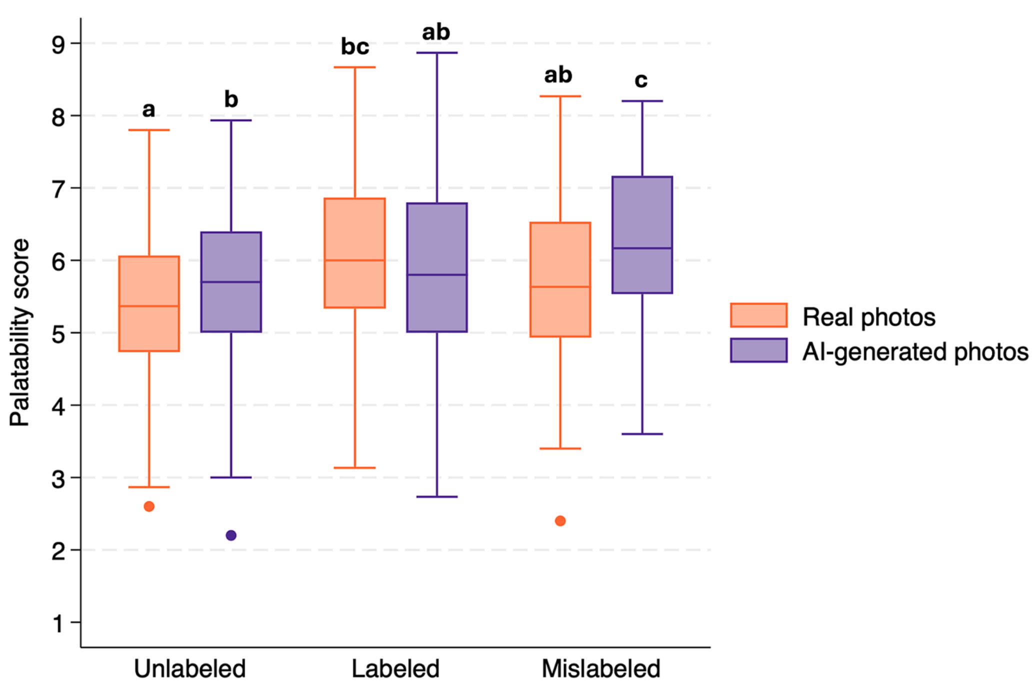

Fig. 6. Comparison of palatability scores for real and AI food photos across label disclosure conditions in Study 2. The boxplot of the distribution includes the lowest and highest data points (whiskers), the first and third quartiles (box extremities), the median (line within the box), and any outliers (represented as dots). Letters indicate statistical differences at the 5 % level ($p < .003$ after Bonferroni correction).

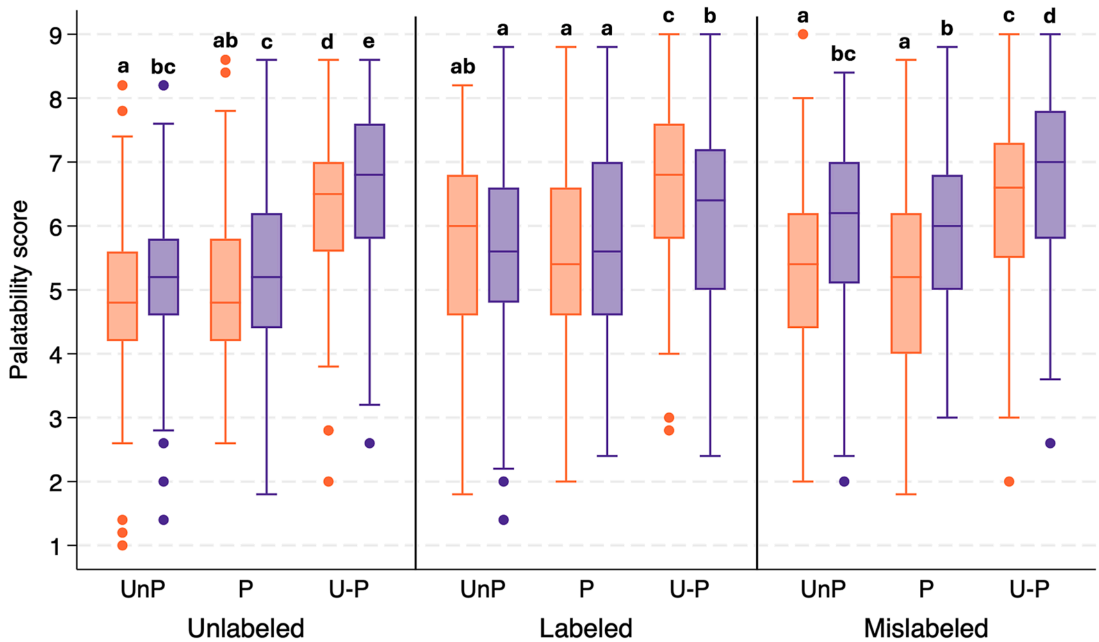

Fig. 7. Comparison of palatability scores for real and AI food photos across processing levels for each label disclosure condition in Study 2. The boxplot of the distribution includes the lowest and highest data points (whiskers), the first and third quartiles (box extremities), the median (line within the box), and any outliers (represented as dots). Letters indicate statistical differences at the 5 % level ($p < .003$ after Bonferroni correction) within each label disclosure condition. UnP = Unprocessed foods; P = Processed foods; U-P = Ultra-processed foods.

2017; De Kerpel et al., 2020; Murakoshi et al., 2013; Spence, 2021), along with warm and uniform lighting, significantly enhances the visual appeal of food (Rodriguez et al., 2022; Spence et al., 2022). Moreover, although direct comparisons might be challenging, several AI images would appear to present a version of the foods that appears more energy-dense compared to the original, as exemplified by the increased number of potato fries in Fig. S2 (Supplementary Material). This raises concerns about the widespread distribution of such idealized food images potentially exacerbating the phenomenon of visual hunger (Spence et al., 2016), especially given the susceptibility of individuals to cue-induced eating (Spence, 2017a; Versace et al., 2019). Future studies could provide a more comprehensive assessment of these issues by delving into aspects such as the actual differences in color parameters through detailed image analysis (for reference, see Motoki et al., 2021, on such an approach).

Despite the superior aesthetic appeal of AI-generated food images when compared to real ones, when labels indicating whether a photo was real or AI-generated were provided, people's preferences tended to shift towards those images that had been labeled as real, irrespective of their actual nature. This aligns with the findings of Nozawa et al. (2022), suggesting that consumers evaluated restaurants more negatively when informed of their use of AI (vs. human) providers. It further underscores the psychological impact of perceived naturalness and authenticity in enhancing the palatability of food (Califano et al., 2023; Roman et al., 2017). Notably, the positive effect of AI-generated images' aesthetic appeal on the palatability score appeared to outweigh the negative effect of being aware that the image was AI-generated. This finding should encourage transparent disclosure policies regarding the nature of food images shown to consumers.

Furthermore, the influence of the level of food processing was notable in the study's findings. For instance, in both the unlabeled and mislabeled conditions, where participants were unaware of, or deceived about, the nature of the image, unprocessed foods, such as apple and carrots, were deemed more appealing in their AI rendition. Considering the movement towards more sustainable patterns of consumption (Caso et al., 2023; Gallagher et al., 2022), especially regarding fruits and vegetables that are visually suboptimal (De Hooge et al., 2017), there is a concern that if these models consistently produce less "ugly" food images, this might inadvertently end up nudging people towards an unsustainable ideal of what natural food should look like.

Nonetheless, in the labeled condition, ultra-processed AI-generated images were judged to be less appetizing than their real counterparts. This outcome is intriguing, especially considering the higher recognition rate of ultra-processed foods in Study 1. The 'uncanny valley' effect (Mori et al., 2012) may be relevant here. The additional layer of AI manipulation, combined with the high degree of processing inherent in these foods, could make the images appear overly perfect and thus artificial. This may reduce their appeal compared to real versions, which could retain a sense of organic authenticity, particularly when the nature of the photos is disclosed. Although further studies are needed to disentangle these complex relationships, many of the participants in Study 1 reported identifying AI-generated images by looking for those images that somehow appeared "too perfect" or "too good to be real". Finally, when the labels were reversed, AI versions of foods across all levels of processing were preferred to their real counterparts, underscoring AI's general superiority in terms of stimulating visual hunger, compounded by the (misleading) belief in their authenticity.

The present study carries implications for both marketers and policymakers. Our findings suggest that AI-generated food images are generally well-received by consumers, often even better than real photos when individuals are unaware of the photo's nature. While this may present an opportunity for marketers and the industry (e.g., to reduce the costs associated with food photoshoots), there is a potential risk of exacerbating 'visual hunger', which could influence unhealthy eating behaviors or create unrealistic expectations about food among consumers. To address this concern, clear disclosure of the content's origin is essential. We argue that transparent disclosure may represent a win–win solution, benefiting both the industry and the well-being of individuals and society. Indeed, although we observed a negative effect on the appeal of these photos when consumers were informed about their AI origin, they were still deemed just as appealing as real ones.

However, in interpreting these findings, it is essential to consider certain limitations. First, the study's sample may not be fully representative of the general population. Men and individuals over the age of 65 years were somewhat underrepresented in our studies, potentially limiting the generalizability of the findings (though see Woods et al., 2015). Additionally, the experiment was based on specific stimuli generated by a specific AI model, in this case, DALL-E 3. This specificity means that the results might not universally apply to all AI-generated food imagery, as different models, such as Midjourney, are likely to yield varying degrees of realism. Furthermore, the rapid advancement of generative models must be acknowledged. These results should therefore be considered as representing a snapshot in time and should be interpreted within such limitations. Consequently, further research is essential to validate and build upon these preliminary findings.

5. Conclusions

In exploring the influence of AI-generated food imagery on consumer perception, the present study investigates the complex interplay between technological advancement and human responses, particularly within the sphere of digital food marketing. The findings reported here reveal an intricate picture concerning the way in which consumers discern and react to such images, highlighting several facets of AI's role in shaping consumer perception. Looking to the future, the study opens up several avenues for future research. For instance, one might consider conducting similar studies on comfort foods, where the emotional connection with food might influence the consumer acceptance of digitally generated content. Such a study would, of course, need to consider variations in the definition of comfort food (see Spence, 2017b) across different genders and geographical locations. The intriguing realm of molecular gastronomy or modernist cuisine, known for its deliberate manipulation of both taste/flavor and the appearance of food also presents an interesting area for future exploration (Spence & Youssef, 2018; Velasco et al., 2016). Additionally, extending the research to include perception of those foods purported to be generated by 3D food-printing technologies offers another promising direction for future research (Ledford, 2015). Lastly, the potential influence of food aromas on the perception of the naturalness of food images, no matter whether AI-generated or real represents an intriguing research direction that could further enrich this rapidly developing field of study (Barwich, 2019; Zhang & Spence, 2023).

In summary, this research represents a first step in understanding the nuanced relationship between AI-generated food imagery and consumer perceptions. As AI technology continues to evolve, it is essential to keep exploring these dynamics, especially considering the implications for ethical marketing practices and the field of consumer psychology (see Spence, 2020).

Ethical Statement

Participants gave informed consent via the statement "I am aware that my responses are confidential, and I agree to participate in this survey" where an affirmative reply was required to enter the survey. They were able to withdraw from the survey at any time without giving a reason.

This study has been reviewed by, and received ethics clearance through, a subcommittee of the University of Oxford Central University Research Ethics Committee [R85145/RE001].

CRediT authorship contribution statement

Giovanbattista Califano: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Charles Spence:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodqual.2024.105149.

References

Astrup, A., & Monteiro, C. A. (2022). Does the concept of "ultra-processed foods" help inform dietary guidelines, beyond conventional classification systems? NO. *The American Journal of Clinical Nutrition, 116*(6), 1482–1488.

Bailey, R. L. (2017). Influencing eating choices: Biological food cues in advertising and packaging alter trajectories of decision making and behavior. *Health Communication, 32*(10), 1183–1191.

Barwich, A. S. (2019). A critique of olfactory objects. *Frontiers in Psychology, 10*, 1337.

Califano, G., Furno, M., & Caracciolo, F. (2023). Beyond one-size-fits-all: Consumers react differently to packaging colors and names of cultured meat in Italy. *Appetite, 182*.

Caso, D., Canova, L., Capasso, M., & Bianchi, M. (2023). Integrating the theory of planned behavior and the self-determination theory to promote Mediterranean diet adherence: A randomized controlled trial. *Applied Psychology: Health and Well-Being, 2023*, 1–22.

Cunningham, C. A., & Egeth, H. E. (2018). The capture of attention by entirely irrelevant pictures of calorie-dense foods. *Psychonomic Bulletin & Review, 25*, 586–595.

De Hooge, I. E., Oostindjer, M., Aschemann-Witzel, J., Normann, A., Loose, S. M., & Almli, V. L. (2017). This apple is too ugly for me!: Consumer preferences for suboptimal food products in the supermarket and at home. *Food Quality and Preference, 56*, 80–92.

De Kerpel, L., Volles, B. K., & Van Kerckhove, A. (2020). Fats are glossy, but does glossiness imply fatness? the influence of packaging glossiness on food perceptions. *Foods, 9*, 90. https://doi.org/10.3390/foods9010090

Gallagher, R., Raimondo, M., & Caracciolo, F. (2022). Eating the "inedible": How to improve the consumption of the perceived inedible parts of fruits and vegetables in Ireland and Italy? *Food Quality and Preference, 99*.

Hässig, A., Hartmann, C., Sanchez-Siles, L., & Siegrist, M. (2023). Perceived degree of food processing as a cue for perceived healthiness: The NOVA system mirrors consumers' perceptions. *Food Quality and Preference, 110*, Article 104944.

Herrmann, B. (2023). The perception of artificial-intelligence (AI) based synthesized speech in younger and older adults. *International Journal of Speech Technology*, 1–21.

Hou, Y., Yang, W., & Sun, Y. (2017). Do pictures help? the effects of pictures and food names on menu evaluations. *International Journal of Hospitality Management, 60*, 94–103.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*(3), 247–257.

Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science, 5*(4), 343–355.

Ivy, S., Rohovit, T., Stefanucci, J., Stokes, D., Mills, M., & Drew, T. (2023). Visual expertise is more than meets the eye: An examination of holistic visual processing in radiologists and architects. *Journal of Medical Imaging, 10*(1).

Jackson, S. (2023). Is AI a good alternative for photography? *Slerp.* https://www.slerp.com/blog/2023/05/16/is-ai-a-good-alternative-for-photography/.

Lazard, A. J., Mackert, M. S., Bock, M. A., Love, B., Dudo, A., & Atkinson, L. (2018). Visual assertions: Effects of photo manipulation and dual processing for food advertisements. *Visual Communication Quarterly, 25*(1), 16–30.

Ledford, H. (2015). Foodies embrace 3D-printed cuisine: Printers unleash creative cookery, but will consumers bite? *Nature News, April 20th*.

Lu, Z., Huang, D., BAI, L., Qu, J., Wu, C., Liu, X., & Ouyang, W. (2023, November). Seeing is not always believing: Benchmarking human and model perception of AI-generated images. In Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track.

Miller, E. J., Steward, B. A., Witkower, Z., Sutherland, C. A., Krumhuber, E. G., & Dawel, A. (2023). AI hyperrealism: why AI faces are perceived as more real than human ones. *Psychological Science, 09567976231207095*.

Monteiro, C. A., Cannon, G., Levy, R. B., Moubarac, J. C., Louzada, M. L., Rauber, F., & Jaime, P. C. (2019). Ultra-processed foods: what they are and how to identify them. *Public Health Nutrition, 22*(5), 936–941.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100.

Motoki, K., Saito, T., Suzuki, S., & Sugiura, M. (2021). Evaluation of energy density and macronutrients after extremely brief time exposure. *Appetite, 162*. https://doi.org/10.1016/j.appet.2021.105143

Motoki, K., Takahashi, A., & Spence, C. (2021). Tasting atmospherics: Taste associations with colour parameters of coffee shop interiors. *Food Quality and Preference, 94*.

Murakoshi, T., Masuda, T., Utsumi, K., Tsubota, K., & Wada, Y. (2013). Glossiness and perishable food quality: visual freshness judgment of fish eyes based on luminance distribution. *PLoS ONE, 8*(3).

Nozawa, C., Togawa, T., Velasco, C., & Motoki, K. (2022). Consumer responses to the use of artificial intelligence in luxury and non-luxury restaurants. *Food Quality and Preference, 96*.

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics, 11*(6), 619–632.

Petit, O., Velasco, C., & Spence, C. (2018). Are large portions always bad? using the delboeuf illusion on food packaging to nudge consumer behaviour. *Marketing Letters, 29*(4), 435–449. https://doi.org/10.1007/s11002-018-9473-6

Prem, E. (2023). From ethical AI frameworks to tools: A review of approaches. *AI and Ethics, 3*, 699–716.

Rodriguez, O. E. S., Masuko, S., & Yamanaka, T. (2022). The impression of deliciousness through food photography-a photographer's approach to test the commonly manipulated factors and the interactions on the food image creation process. *International Journal of Affective Engineering, 21*(1), 67–76.

Roman, S., Sánchez-Siles, L. M., & Siegrist, M. (2017). The importance of food naturalness for consumers: results of a systematic review. *Trends in Food Science & Technology, 67*, 44–57.

Sætra, H. S. (2023). Generative AI: here to stay, but for good? *Technology in Society, 75*.

Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare, 11* (6), 887.

Spence, C. (2017a). *Gastrophysics: The new science of eating*. Penguin UK.

Spence, C. (2017b). Comfort food: A review. *International Journal of Gastronomy and Food Science, 9*, 105–109. https://doi.org/10.1016/j.ijgfs.2017.07.001

Spence, C. (2020). On the ethics of neuromarketing & sensory marketing. In J. Trempe-Martineau & E. Racine (Eds.), Organizational neuroethics: Reflections on the contributions of neuroscience to management theories and business practice (pp. 9-30). Cham, Switzerland: Springer Nature.

Spence, C. (2021). Glossy packaging: On the questionable appeal of glossy/shiny food packaging. *Foods, 10*, 959. https://doi.org/10.3390/foods10050959

Spence, C., & Youssef, J. (2018). Assessing the long-term impact of the molecular gastronomy movement on haute cuisine. *International Journal of Gastronomy & Food Science, 14*, 35–44. https://doi.org/10.1016/j.ijgfs.2018.10.001

Spence, C., Motoki, K., & Petit, O. (2022). Factors influencing the visual appeal of food: From energy-density to the aesthetics of gastroporn. *Food Quality & Preference, 102*.

Spence, C., Okajima, K., Cheok, A. D., Petit, O., & Michel, C. (2016). Eating with our eyes: From visual hunger to digital satiation. *Brain and Cognition, 110*, 53–63.

Tucciarelli, R., Vehar, N., Chandaria, S., & Tsakiris, M. (2022). On the realness of people who do not exist: the social processing of artificial faces. *IScience, 25*(12).

Van Esch, P., & Stewart Black, J. (2021). Artificial intelligence (AI): revolutionizing digital marketing. *Australasian Marketing Journal, 29*(3), 199–203.

Van Tullen, C. (2023). *Ultra-processed people: Why do we all eat stuff that isn't food…and why can't we stop*. London, UK: Cornerstone Press.

Velasco, C., Michel, C., & Spence, C. (2021). Gastrophysics: Current approaches and future directions. *International Journal of Food Design, 6*(2), 137–152.

Velasco, C., Michel, C., Youssef, J., Gamez, X., Cheok, A. D., & Spence, C. (2016). Colour-taste correspondences: Designing food experiences to meet expectations or to surprise. *International Journal of Food Design, 1*, 83–102. https://doi.org/10.1386/ijfd.1.2.83_1

Versace, F., Frank, D. W., Stevens, E. M., Deweese, M. M., Guindani, M., & Schembre, S. M. (2019). The reality of "food porn": Larger brain responses to food-related cues than to erotic images predict cue-induced eating. *Psychophysiology, 56* (4).

Walhout, H. (2023, August 1). Why AI food photography is so unappetizing. Eater. https://www.eater.com/23814463/ai-image-generator-food-photography-unappetizing.

Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. PeerJ, 3:e1058. doi: 10.7717/peerj.1058.

Zhang, T., & Spence, C. (2023). Orthonasal olfactory influences on consumer food behaviour. Appetite, 190:107023.