Food Quality and Preference 93 (2021) 104290

Contents lists available at ScienceDirect

Food Quality and Preference

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

Investigating visual attention toward foods in a salad buffet with mobile eye tracking

Marjaana Puurtinen[a,*], Ulla Hoppu[b,1], Sari Puputti[b,1], Saila Mattila[b,1], Mari Sandell[b,c,1,2]

[a] Department of Teacher Education, University of Turku, Finland
[b] Functional Foods Forum, University of Turku, Finland
[c] Department of Food and Nutrition, University of Helsinki, Finland

ARTICLE INFO

Keywords:
Fixation
Food selection
Eye tracking
Pupil size
Salad buffet
Scanpath

ABSTRACT

Mobile eye tracking (MET) enables the recording of gaze data in less-controlled research environments, but best practices for its use in studies about visual attention to foods are yet undetermined. This study supports the building of a coherent framework for this methodological approach by discussing current eye-tracking trends in the field, applying MET in an experiment with real foods, and proposing methodological approaches for future studies. In the experiment, 32 female participants' gaze data were recorded while they inspected a salad buffet for 20 s and then assembled a self-choice salad. The functionality of fixation, scanpath, and pupil size measures was investigated, focusing on associations between eye movements and food item color and position, eye movements and food item preference, and pupil size and selected measures. Dish placement affected the relative amount of visits to a single food item, whereas food item color and preference were not associated with the examined measures. The pupil-size measure did not function with the elderly participants. Importantly, a simple cluster analysis, based on a scanpath and a food selection measure, helped to illustrate different profiles of food view and selection. It was determined that food item position should be carefully considered in MET studies involving real foods, and scanpath measures could be useful in bringing forth behavioral differences that are not revealed by fixation parameters alone. Importantly, identifying "attention-action" profiles by combining eye-tracking and other measures seems to be a fruitful way of approaching individual differences in food viewing and selection.

1. Introduction

Visual attention toward food has typically been examined by recording participants' eye movements while they look at food images (e.g., Garcia-Burgos et al., 2017; Graham et al., 2011; Hummel et al., 2017, 2018; Nijs et al., 2010; Nummenmaa et al., 2011; Peng-Li et al., 2020). With eye tracking, we can accurately record the durations and locations of individuals' *fixations*, while they look at a given visual stimulus. Fixations are the short moments when we target our gaze to a certain location in our visual array and process information from it before moving our gaze to another location to process the information there (Holmqvist et al., 2011; Rayner, 1998, 2009). Fixation durations vary according to and during a visual task; typically they average around 200–300 ms, but in text reading, for instance, the shortest fixation durations may last only for 50–75 ms (Rayner, 2009). Eye-tracking studies are traditionally conducted with static stimuli and remote eye trackers in controlled research environments, and this has been the custom in food-related studies, too. Undoubtedly, these studies have tapped into several important aspects of visual perception and cognition, but, at the same time, they are quite distant from real-life encounters with foods.

Mobile eye tracking (MET) enables the easy recording of gaze data, even in less-controlled research environments (e.g., Pérez-Edgar et al., 2020), and modern head-mounted eye trackers are as comfortable as a pair of eyeglasses. So far, they have rarely been applied in studies with authentic food items. A notable exception is the study by Wang et al. (2018), where MET was applied when males viewed and selected foods from a real buffet with high- and low-calorie food items. However, the

* Corresponding author at: Dept. of Teacher Education, 20014 University of Turku, Finland.
E-mail address: marjaana.puurtinen@utu.fi (M. Puurtinen).
[1] Postal address: Functional Food Forum, 20014 University of Turku, Finland.
[2] Postal address: P.O 66 (Agnes Sjöbergin katu 2), 00014 University of Helsinki, Finland.

https://doi.org/10.1016/j.foodqual.2021.104290
Received 15 December 2020; Received in revised form 10 May 2021; Accepted 10 May 2021
Available online 17 May 2021

shift from controlled laboratory experiments to real-life ones presents researchers with some challenges and necessarily affects the formulation of research questions and analytical choices (see Pérez-Edgar et al., 2020): measures and practices typical to laboratory settings do not always translate easily to real-life contexts. This study bridged this gap by recording participants' gaze data with MET while they inspected a salad buffet in controlled conditions but with real foods, during lunchtime, and eating the salad they assembled for themselves. The study aims to support the building of a coherent methodological framework for MET applications in research about visual attention toward foods by discussing current trends in eye-tracking studies in the field, applying MET in an experiment with real foods, and, based on the results, proposing next steps for MET studies about food viewing and selection.

1.1. Visual attention to food

Studies applying a passive picture-presentation paradigm (Hendrikse et al., 2015) present a participant with picture pairs or picture arrays of food and non-food items and measure with eye-tracking data and/or visual probe tasks how the participant's attention is targeting and divided between these visual items. In general, human visual attention appears to be biased toward food items over non-food ones (Nijs et al., 2010; Nummenmaa et al., 2011). However, when the visual items look similar, the bias may disappear (Nummenmaa et al., 2011).

Several studies have examined whether the bias toward food items is associated with weight-gain and obesity (for reviews, see Hendrikse et al., 2015; Werthmann et al., 2015). Also eating disorders, such as binge eating, have been suggested to have an association with visual attention toward food (Schmidt et al., 2016). Some studied factors perhaps mediating the food bias are, among others, subjective satiety or hunger (Nijs et al., 2010), gender (Doolan et al., 2014), and food item preference or pleasantness (Garcia-Burgos et al., 2017; Motoki et al., 2018; Wang et al., 2018; Nummenmaa et al., 2011).

The findings in studies addressing these issues are somewhat mixed, however. Nijs et al. (2010) suggested that the more hunger overweight or obese women reported, the more their attention was drawn toward food items, as opposed to non-food items. Similarly, Werthmann et al. (2011) reported that overweight females, who reported food cravings, directed their gaze more often toward food pictures, compared to healthy-weight participants, but overweight females also showed reduced maintenance of attention toward these pictures. On the contrary, Nummenmaa et al. (2011) noticed a *negative* association between body mass index (BMI) and food detection. Schmidt et al. (2016) observed that adolescents with a binge-eating disorder showed delayed disengagement from food items, and the effect was stronger for preferred food items—the latter finding is, however, described as exploratory. Doolan et al. (2014) did not observe *any* BMI effects in their study, only a bias toward high-energy food items compared to low-energy food items. Hummel et al. (2017), in turn, did not observe this same bias, and Motoki et al. (2018) reported only effects of tastiness (instead of healthiness) for visual attention toward food items placed among non-food distractors.

Thus, these characteristics' exact roles and interplay are yet unclear, even though they have been addressed in highly controlled research conditions, where as many confounding factors as possible are controlled for. With real foods, the situation becomes even more complex: food perception is, in fact, multisensory, including all five senses (Spence, 2015). The sense of sight as well as smell give the first impression of food before tasting, and visual cues, including the color and shape of the food, also become important determinants of food choice (Jantathai et al., 2013; Wadhera and Capaldi-Phillips, 2014). For example, the visual layout of the food elements on a plate is associated with a diner's experience of the dish (Michel et al., 2015), and colorful salad portions with high color contrasts between salad components are considered more attractive than less colorful ones (Paakki et al., 2019). In sum, when observing visual attention towards real foods, several intertwined factors might be at play: some are stimulus-oriented (e.g. food item vs. non-food item, color) and others related to participant characteristics (e.g., BMI).

Of the aforementioned studies, only Nijs et al. (2010) allowed their participants to eat in their study session, although their bogus taste test was conducted separately from a passive picture presentation task. Nevertheless, hungry overweight women ate more than normal-weight ones. Indeed, being aware of the possibility of eating what one is looking at could affect attentional and selection processes. Taking a step toward real-life settings, where food is not just viewed but also consumed, requires the refinement of the research methodology applied in studies thus far.

1.2. Eye tracking in studies about food viewing and selection

The eye-tracking method enables the study of factors influencing gazing behavior and decision time (Vu et al., 2016). In practice, depending on the experimental set-up, the method provides the possibility of examining both stimulus-driven and goal-driven attention—or, in other words, bottom-up and top-down processes—in visual tasks (see Vu et al., 2016; Werthmann et al., 2015). In eye-tracking research, bottom-up and top-down processes are considered to have different temporal characteristics: early attention components reflect the automated and involuntary (and immediate) bottom-up processes, whereas later attention components tap into the more controlled, and even strategic, top-down processes (Werthmann et al., 2015).

In the passive picture-presentation paradigm, the typical way of using eye-tracking data is to calculate gaze direction and gaze duration biases (e.g. Doolan et al., 2014; Graham et al., 2011; Nijs et al., 2010; Nummenmaa et al., 2011; Schmidt et al., 2016). The former is the percentage of times the target item (e.g. a food item) is fixated on prior to a distractor (e.g. a non-food item); the latter is the percentage of fixation time spent on a target item, compared to time spent looking at the distractor(s) (out of the total fixation time). Depending on their exact use, these measures can tap into both early attention processes (gaze direction) and later attention processes (gaze duration); note that even in the "later" attention processes, the observed intervals may lie within just hundreds of milliseconds. In food choice tasks, Jantathai et al. (2013) and Wang et al. (2018) reported that increases in the selected fixation measures (fixation count for the former and fixation duration for the latter) were associated with selecting a particular food item, suggesting a relationship between some later attention measures and food item selection.

These fixation measures focus on time spent on selected targets or the number of gaze visits on them. Food selection and decision making are, however, temporal processes and in real-life settings, the studied intervals necessarily become longer than the brief presentation of images typical to the passive picture-presentation paradigm. Since fixation parameters do not take into account *when* events occur, in real-life tasks parameters reflecting the course of the viewing process seem an important addition. A typical approach in, for instance, consumer research is to analyze *scanpaths*, traces of "a participant's eye-movements in space and time" (Holmqvist et al., 2011, p. 253). In its simplest form, a scanpath can be a list of fixated targets in their correct order, but scanpaths can be complemented with information about fixation durations at these locations, too. In addition, new parameters, representing selected aspects of the individual's scanpaths, can be calculated based on the sequential information from the course of the viewing process (this approach was used in the present study). There are several techniques of varying degrees of sophistication for conducting scanpath analyses, and their appropriateness depends on the research question and characteristics of the data set (see Eraslan et al., 2016; Le Meur & Baccino, 2013). Many of these measures are, however, designed for controlled experimental set-ups. So far, it appears that in real-life, food-related studies, the most suitable scanpath parameters have yet to be determined.

In addition to eye-movement measures, modern eye trackers record participants' pupil sizes. Pupil size is in constant fluctuation and affected by existing conditions, such as luminance, as well as participant characteristics; for instance, pupil diameters get smaller with age (Guillon et al., 2016). When the research environment is controlled to an adequate extent, pupil size can be used as an arousal measure. In practice, cognitive and/or emotional arousal (e.g. when facing difficulties in a reading task or looking at emotionally captivating images) may result in pupils gradually getting larger than their original states. (Holmqvist et al., 2011, pp. 393–394.) However, this effect occurs with some delay and varies greatly between individuals and tasks (Holmqvist et al., 2011, p. 435).

With regard to visual attention toward foods, pupil size data has been reported only rarely. Graham et al. (2011) reported a decreased pupil diameter in their high-BMI participant group, when they looked at images of high-calorie sweet foods, compared to low-calorie images. Wang et al. (2018) reported smaller average pupil sizes during the viewing of a dessert buffet among overweight rather than lean participants. However, since pupil size is in constant fluctuation and affected by visual stimuli, environmental conditions, and cognitive and emotional arousal, in the latter study the 1-minute length of the recording session was perhaps long enough to allow several confounding factors to intervene with the measurement. Averaging constantly fluctuating pupil diameter over long periods should be performed with caution, so that the actual cause of the effect can be reliably determined. Nevertheless, these prior studies suggest that this arousal measure is also worth exploring when studying visual attention toward foods.

1.3. Aims

This study has a main methodological goal: to investigate the functionality of MET in an experimental setting that includes real foods and that allows the participants to select and eat the foods of their choice. Food viewing and selection is examined from four perspectives, all of which build on previous studies. The study examines associations between (a) eye movements and food item color and position, (b) eye movements and food item preference, and (c) pupil size during food viewing, selected background measures, and later food consumption. It also seeks ways to identify (d) participant profiles illustrating associations between initial food viewing and later food selection. Based on the results, the study discusses the possibilities of MET in a real-life food viewing and selection task and suggests potential methodological traits for future work on the topic.

2. Methods

2.1. Participants

Data was collected in fall 2019. Sixty-two adult female participants, who had taken part in extensive sensory tests first in 2015–2016 (as part of a larger participant group; see Puputti et al., 2018) and again in fall 2018 or spring 2019, and were thus familiar with the laboratory facilities, were invited to participate; finally, 32 volunteers took part in the present study. During the invitation process, pregnant or breastfeeding women and those with coeliac disease were excluded and invitees with scent hypersensitivity advised not to participate. Participants were also required to have normal vision (below −1.0 diopter) or apply contact lenses during the study. Food allergies and intolerances were enquired into before the study visit and just before the meal. The number of participants, though not extensive, was comparable to prior studies applying MET in real-life settings (e.g., there were 32 participants in Wang et al. (2018)) and considered appropriate with respect to the laborious research protocol, where a salad buffet was prepared and served individually twice for each participant (see Sections 2.2 and 2.4, below).

Although all 32 participants were female and from the city of Turku or nearby, these volunteers did vary in terms of age and calculated BMI. The participants were between 24 and 76 years old ($M = 53$ years, $SD = 14$ years, $MED = 50$ years). During the lunch buffet visit, they also reported their height and weight, and their calculated BMI varied from 19 to 55 kg/m^2 ($M = 26.8$ kg/m^2, $SD = 6.9$ kg/m^2, $MED = 26$ kg/m^2; two participants' BMI information was missing). This variability is taken into account in data analyses. The participants could eat a self-assembled lunch salad during both visits. No financial compensation was offered. The study protocol was approved by the ethics committee of the University of Turku (statement 56/2017), participation was voluntary, and all participants gave written informed consent. The use of the mobile eye-tracker was explained to each participant at the beginning of the test session, as was the type of data collected with the equipment, and participants were able to view their own eye-movement recordings after their second visit.

2.2. Food item selection, preparation, and presentation

The lunch buffet included 14 different fresh food items typical to lunch salad buffets in Finland. Foods (see Table 1) were selected based on three criteria. First, the same participants had previously taken part in sensory tests and had ranked the favorability of a number of food items, and some of these food items (olives, broccoli, feta cheese, and salted nuts) were chosen also for the buffet items. From these, olives were selected as the target food item for this study due to large differences in individual liking for it: the whole hedonic scale from 1 (dislike extremely) to 9 (like extremely, $MED = 7$) was represented in the group's ratings. In addition, food items' visual appearance was taken into account so that two food items formed color-matched pairs (see Fig. 1; with the exception of pasta food items), and nutrient contents were considered so that the participants could assemble a healthy and versatile lunch salad.

The same local supermarket delivered food items weekly and the salad buffet was prepared for each participant separately in the laboratory kitchen and, when needed, stored in the refrigerator. The preparation of the food items (e.g., rinsing, defrosting, chopping, peeling and/or cutting; cooking the pasta, cooling it and mixing it with pesto or aioli sauce) followed a pre-planned procedure, and the serving size (g) for each food item was the same for all participants. Foods were served in glass bowls (15 cm × 15 cm) and arranged on a serving trolley in three different rows (see Fig. 1). The order of the serving bowls was randomized between each participant. Serving tools were ordinary tablespoons, except for a salad server for lettuce. A soup plate of white porcelain, with a diameter of 22 cm, was used as the food plate. Water, rye and oat bread, margarine, olive oil with lemon flavor, and French dressing were also offered. After the meal, each participant was served coffee or tea and cookies.

2.3. Eye tracker

Tobii Pro Glasses 2 (wireless; with a 50 Hz recording frequency) were used to record participants' gaze data and pupil size. A one-point

Table 1
Food item pairs used in the study. Food items in each pair were chopped into similar-sized pieces (apart from green food items).

Pair	Food items	
Black items	Kalamata olives	Grapes
Beige items	Salted nuts	Chickpeas
Green items	Broccoli	Iceberg lettuce
White items	Feta cheese	Mozzarella cheese
Red items	Cherry tomatoes	Sweet pepper
Orange items	Oranges	Cantaloupe melon
Pasta items[1]	Aioli pasta	Pesto pasta

[1] Fusilli pasta mixed with white aioli sauce and fusilli pasta mixed with green pesto sauce differed slightly in coloring.

Fig. 1. A view of the laboratory, the food trolley with the 14 food items in glass bowls, and the black dot applied in the study for preventing a preview of the food items (see section 2.4). Picture was taken by one of the authors and is published with permission.

calibration procedure was applied twice: first when introducing the participant to the equipment and practicing the procedure, and the second time prior to the actual recording.

2.4. Procedure

The test session was organized at a multisensory laboratory (see Fig. 1). The data set was collected jointly with a study about the effect of multisensory environments on food intake and emotions (Hoppu et al., 2020), and all participants lunched in the laboratory twice. Unlike the study by Hoppu et al. (2020) that analyzed within-subject differences between two conditions, the present study focused on visual attention toward foods during the participants' first visit, as that was when they first saw what was being served. During the second visit the viewing of the foods may have become a recognition or a memory task; if so, this would have added more intervening factors into interpreting the eye-movement data. The test session consisted of a pre-questionnaire, introductory/practice stage (getting familiar with the MET device and procedure), initial view of the buffet (20 s), food selection, eating, and dessert/post-questionnaire. The protocol was practiced beforehand with four of the five authors and a research assistant, and piloted with one volunteer adult.

Study participants were asked to attend two study sessions at the same time of the day at either 10:45 a.m. or 1:30 p.m. at least 1 week apart. Participants were advised to choose a slot closer to their normal lunchtime, not eat anything during the one hour before their session, and avoid the use of scented cosmetic products during the study visits. When entering the laboratory, the participants filled out a pre-questionnaire and stated their perceived hunger state on a scale, with (1) being "very hungry," (2) "quite hungry," (3) "somewhat hungry," and (4) "not at all hungry." Each session lasted approximately 45 min.

In eye-tracking experiments, it is important to control for the location of the first fixation; this is, however, much more complex with MET than with remote trackers. Hence, the following procedure was developed for the study and practiced with each participant prior to the actual recording. In the practice stage outside the laboratory, after a calibration, the participant was informed that she could choose cookies for her dessert. She was instructed to look at a large black dot printed on paper and placed up on the wall above eye level and then close her eyes. Next, the researcher placed three bowls of cookies on the table below the black dot. The participant was instructed to open her eyes, look at the black dot and count aloud to two, and then look down at the cookie bowls for five seconds. The researcher started timing when the participant said "two," and mentioned when the time was up. After this, the participant named the cookies she would have for dessert.

This same procedure, just with a 20-s viewing time, was later performed in the laboratory. After a new calibration, the participant was informed that she would be able to see the salad buffet for 20 s. The participant was tasked to look at the food trolley "as if you were in a lunch restaurant checking what is for offer." Next, the participant was directed to stand behind a mark on the floor, instructed to look up at a black dot on the wall opposite to her, and close her eyes (see Fig. 1). The food trolley was brought in and placed in its marked position between the participant and the wall (the closest edge of the trolley being 60 cm away from the participant). The participant was advised to open her eyes and count aloud to two, and after counting, view the food items for 20 s from her marked position, and according to the given instructions (the researcher again mentioning when the time was up). Both with the help of the live replay of the eye-tracking recording, visible on the researcher's laptop, and the actual eye-movement recording inspected later, it was ensured that all of the participants kept their eyes closed, there was no preview of the foods, and all participants started viewing the food items from the same location (black dot). Room luminance in the multisensory laboratory was kept constant for all participants.

After this 20-s view, the participant exited the room and a smaller food trolley (see Fig. 1) with bread, spread and salad dressings was brought in. After that, the participant re-entered and was instructed that she could now freely assemble a salad. She was advised to take as much food and spend as much time eating as she wanted, and knock on the door when she was finished. The food trolley was taken out of the room and the laboratory door was closed, so that the participant could eat in private. The experimenter monitored the lunch session from the next room via the live replay of the MET recording. When the participant let the experimenter know she was ready by knocking on the door, the experimenter stopped the eye-tracking recording, served the participant coffee/tea and cookies, and gave her a tablet with the final online questionnaire.

2.5. Data handling and analysis

2.5.1. Food selection, food consumption, and time spent eating

Two participants' data about food selection, consumption and time spent eating was coded as missing data due to unsuccessful recordings for one participant and a very small amount of food selected by the other participant. (For the latter participant, her data was nevertheless included in analyses about the effect of food color or dish position for the selected eye-movement measures.) A parameter gauging food selection, *food item take*, was calculated by simply counting the number of times the participant took up a spoon and put food stuff on her plate (range 8–15, $MED = 11.5$, $n = 30$. Note that the same food item was sometimes taken twice). The weight of the served and *consumed* amount of each food was measured with a scale (Mettler Toledo PB3002-S) to 1 g accuracy (range 123–514 g, $MED = 366$ g, $n = 30$; two participants' data were missing). Each participant ate all the food she took. *Time spent eating* was calculated based on the eye-tracking recording and defined as the time from when the participant sat down to eat to the moment when she knocked on the door and let the experimenter know she was finished (note that the parameter was subject to deviations due to manual coding). Participants' eating time varied from 7 to 19 min ($MED = 13$ min, $n = 30$; two participants' data were missing).

2.5.2. Eye-movement measures

The eye-movement data was pre-processed using Tobii Pro Lab 1.123 software. The Tobii I-VT attention filter (see Appendix A) was applied in defining a fixation, discarding fixations shorter than 60 ms. Two participants' eye-movement data had a significant amount of missing data points, and one participant spoke during the inspection of the food trolley (despite the instructions); these three participants were removed from the eye-movement analyses.

Fixations from each participant's 20-s initial food viewing were manually mapped on an image of that specific food trolley with the snapshot and mapping functions of Tobii Pro Lab 1.123 by the first author (with considerable experience in handling eye-tracking data). Manual mapping was selected after first experimenting with the automated mapping of the software, but noticing that a considerable amount of fixations was not registered by the mapping tool and a manual checking would nevertheless be necessary. Thus, manual mapping, although performed with only one researcher, was in this case considered the most reliable tool for, for example, identifying the first fixation toward a specific food item. During mapping, fixations occasionally seemed to overshoot or undershoot the intended food items; in these cases, the landing position was interpreted based on the overall fixation pattern of the participant.

Square- or rectangle-shaped areas of interest (AOI) were drawn around each of the fourteen dishes for each participant. This allowed the matching of each fixation with a particular AOI and a food item, and eye-movement measures were calculated based on this information. During planning the set-up, it was taken into account that dishes on the back row, although elevated compared to the middle and first row, would, from the participant's standing point, cover a smaller visual area than the middle and first row. Similarly, the middle row dishes would cover a smaller visual area than the dishes on the first row (see Fig. 1). After considering the advantages and disadvantages of different types of set-ups, the 3-row set-up was selected on the basis that it allowed all of the food items to be inspected from the same marked standing point without extensive head movements. The dish placement was randomized across all participants, which compensated for between-participant differences in the absolute AOI size for a specific food item. The AOIs were drawn for each participant (i.e., each recording) separately, but their sizes were kept as constant as possible between participants: AOI width was 120–160 pixels and height 120–160 pixels for the dishes on the first row, 110–130 and 100–130 pixels for the dishes on the middle row, and 90–130 and 90–130 pixels for the dishes on the back row.

The 20-s interval contained, in practice, fixations from the first fixation toward the food trolley (after the participant opened her eyes and lowered her gaze) and until the participant moved her gaze toward the researcher, when hearing that the 20-s time had ended. In this real-life setting, variability still occurred, since the participants slightly varied in how quickly they moved their gaze toward the trolley after counting to two (with the researcher starting timing of the 20-s interval from the word "two"), and how quickly they responded to the researcher saying that the time was up. The temporal length of the analyzed interval varied from 18.9 s to 23.8 s ($MED = 20.8$ s). This variability was taken into account in the data analyses, as described below.

Considering the need to explore suitable eye-movement measures for MET studies, three eye-movement measures that are considered to signal different types of visuo-cognitive processes were selected for the analyses. The *first fixation duration for each food item* and for each participant was directly exported from Tobii Pro Lab 1.123. This measure reports the absolute time spent on a food item when it is first fixated, and is often considered an indicator of immediate and target-specific effects on visual attention. Second, the *total fixation time* for a food item is the sum duration of participant's fixations that targeted that food item, calculated as a percentage of their sum total fixation time for all food items (due to slight differences in the total viewing time). This measure addresses the distribution of fixation time across all food items, instead of focusing on absolute fixation durations, and overrides any possible individual differences in absolute fixation durations. A high value can, in practice, consist of either few but long fixations on a target item, or several but shorter ones. Third, the *visit count* of a food item is the number of gaze visits to that item, calculated as a percentage of all participant's visits to all food items (due to slight differences in the total viewing time). This simple measure thus only addresses the distribution of frequency of visits toward all food items and ignores the time spent on each of them. In practice, these three measures utilize different features of the same eye-tracking data set and address different aspects of visual processing. Knowing that fixation durations and locations are sensitive to multiple intervening factors, and that the eye-tracking method and the level of detail of eye-tracking measures were originally developed for highly controlled research settings, the selection of three different types of parameters allowed the exploration of what type of measures seem to function in a real-life MET study, and how.

Participants' scanpaths during the 20-s view were analyzed with three simple parameters, created for the purposes of this study and computed manually based on the AOI hit list exported from Tobii Pro Lab 1.123. First, *the order of fixation on the food items* was tracked for each participant. Second, *returns to food items* was calculated as the percentage of cases when a food item was (a) fixated on and (b) re-fixated on after visits to a maximum of 3 other food items, out of all gaze shifts from one dish to another made by the participant. This parameter was intended to indicate a participant's tendency to compare and re-fixate on food items. Third, each dish was given a running number, starting from the dish placed far left in the back row (back row: numbers 1 to 5; middle row: numbers 6 to 9; front row: numbers 10 to 14). *Stepwise visual processing* is the percentage of cases when a food item was fixated on after a participant fixated on a dish with a position number ± 1 (i.e. a nearby dish), out of all gaze shifts between dishes made by the participant. This parameter was intended to indicate a participant's tendency to visually inspect the food dishes in a systematic, row-based manner, moving from one dish to the next.

2.5.3. Pupil size data

Measuring pupil size in less controlled conditions such as applied in this study needs to be done with care. For instance, as with all time series data, averaging the constantly fluctuating pupil sizes during longer time periods easily ignores the rapid peaks and valleys that could indicate arousal effects. In less controlled environments, there are also likely individual differences in both the onset and duration of any changes in pupil sizes, even though caused by the same occurrence, which makes it difficult to reliably select one given time point at which to compare pupil sizes between participants. Furthermore, general individual differences in pupil sizes mean that any analyses that apply absolute pupil sizes (diameters) typically require defining a baseline for each participant prior to searching for arousal effects – this, too, if often challenging outside the laboratory.

A rough arousal measure was designed for this study in order to explore the measure's functionality in a way that takes into account the above-mentioned potential risks. The purpose of the measure was to grasp changes in pupil diameter irrespective of the participant-specific baseline pupil size or temporal features of the potential effects. It was simply assumed that a large pupil deviation at any point during the 20-s interval could signal increased arousal in the food-view situation. Thus, arousal during food view was measured by inspecting the *maximum deviation of a participant's pupil size during the 20-s interval*. The eye tracker recorded pupil size at a 50 Hz frequency for both eyes separately. Since the left and right eye do not react similarly to changes in arousal states (or stimuli), and the recordings contained intervals when only one of the pupils was detected, the difference between the maximum and minimum pupil size was calculated separately for the right and left eye of each participant. For each participant, the larger of these deviations was applied in the analyses. During the 20-second interval, lighting in the room was kept constant for all participants; however, slight head movements could have affected this measure to some extent.

2.5.4. Statistical analyses

In order to examine aim (a), Friedman tests and Dunn-Bonferroni post hoc tests were applied. Aim (b) was examined with Kruskal–Wallis analyses. For aim (c), a one-way ANOVA, Pearson correlations, and Kruskal-Wallis analyses were used, and for aim (d) Pearson correlations, independent samples t-tests, and K-means cluster analyses were conducted. All calculations were performed with IBM SPSS Statistics 26 software.

3. Results

3.1. Food item color and position and visual attention

The 14 food items consisted of seven pairs (see Table 1), ranging from bright red tomatoes and sweet peppers to beige nuts and chickpeas. The potential effect of food-item pair on participants' visual attention was examined by summing up the parameter values for pairs of food items, and running Friedman tests for these related samples. The distribution of total fixation time across food item pairs (see Fig. 2 for median and Appendix B for mean values) did not differ ($\chi^2(6) = 7.133$, $p = .309$; $n = 29$; three participant's information was missing); neither did visits to the food item pairs ($\chi^2(6) = 3.641$, $p = .725$; $n = 29$) or the sum duration of first fixations on them ($\chi^2(6) = 4.422$, p = .620; $n = 29$). Despite these non-significant overall findings, we mention one pairwise comparison for future studies: there was a statistically significant difference between the total fixation time for the black food items (which gained the longest median total fixation time) and pasta food items (which gained the shortest median total fixation time; $p = .023$, $n = 29$).

Since the dishes were placed on three rows (see Fig. 1), the effect of dish position was then examined with Friedman tests. The distribution of total fixation time across all 14 dishes did not differ ($\chi^2(13) = 18.470$, $p = .140$, $n = 29$; for means values, see Appendix C), and neither did the first fixation durations ($\chi^2(13) = 18.078$, $p = .155$, $n = 29$). However, the distribution of visits to a food item was affected by the dish position ($\chi^2(13) = 85.208$, $p = .000$, $n = 29$). Fig. 3 presents the median percentages of visits to each of the 14 dishes. A visual inspection of the figure suggests a general trend of relatively more visits to the center of the back, middle and front rows (Fig. 3: dishes 2–4, 7–8, and 11–13) than to the extreme ends of these rows (Fig. 3: dishes 1 & 5, 6 & 9, and 10 & 14), and this is especially so in the case of the first row, dishes 10 and 14 being the least visited ones, and the middle row, dish 7 being the most visited one. This observation was supported by a series of pairwise comparisons: visits to these three dishes differed significantly from visits to several other dishes (p-values for all pairwise comparisons are presented Table 2).

In summary, the number of visits to a food item was affected by the placement of the food item on the trolley, but the distribution of total fixation time across all food items, and first fixation duration toward each food item, were not. The smallish sample size does not allow for the inclusion of dish position as a factor in our subsequent analyses; thus, although the placement of each food item was randomized for each participant and this procedure should diminish potential location effects, we acknowledge that it might be a potential confounding factor and therefore focus only on fixation measures in the following sections.

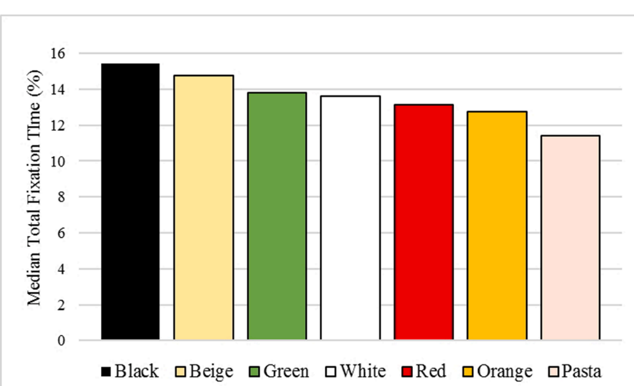

Fig. 2. Sum total fixation times for pairs of food items (% of all total fixation time, median values, $n = 29$). Apart from pasta food items, the pairs were color-matched. No statistically significant overall color effect.

3.2. Food item preference and visual attention

Participants were divided into three groups based on their self-reported olive preference.[3] On a scale from 1 to 9, those who chose 1–2 were considered to find olives "unpleasant," those who chose 3–6 "somewhat unpleasant to pleasant," and those who chose 8–9 "pleasant." Despite the small group sizes, this categorization was believed to best describe the participants, compared to, for example, creating groups based on median split, and these groups were examined with a series of Kruskal-Wallis analyses. Not surprisingly, these three groups differed in terms of olive consumption ($\chi^2(2) = 10.217$, $p = .006$, $n = 29$; three participant's information was missing); none of the participants who considered olives "unpleasant" took them at all (see Table 3). Pairwise comparisons showed the difference to lie between the two extreme groups ($p = .004$).

Due to the study design, where food items formed color-matched pairs, potential eye-movement effects related to olive preference could appear on one or both of two levels: on the level of the food item itself, or on the level of the color-matched pair of food items. First, for the food-item effect, the duration of first fixation on the olive dish was, however, similar across the three groups ($\chi^2(2) = 1.557$, $p = .459$, $n = 29$; see Table 3). There were also no significant group-based differences in the total fixation time for the olive dish ($\chi^2(2) = 2.964$, $p = .227$, $n = 29$). Second, thinking that black grapes acted as distractors and that the participant might have visually compared the similar-looking black food items (olives and grapes) to one another, olive preference could be associated with increased attention toward *both* of the food items (the target and its distractor with similar appearance). For this reason, the sum duration of the first fixations on olive and grape dishes was also calculated. However, the groups did not differ in this respect, either ($\chi^2(2) = 0.399$, $p = .819$, $n = 29$), or regarding the sum total fixation time for the two black food items ($\chi^2(2) = 0.399$, $p = .819$, $n = 29$). In sum, although the groups differed greatly in their liking of the target food item, and treated it differently when composing their salads, the two applied fixation parameters showed no obvious differences in how the target food item was fixated on prior to actual food consumption.

3.3. Pupil size and food consumption

The deviation of pupil size correlated negatively with age ($r = -0.558$, $p = .002$; $n = 29$). Participants were therefore divided into three data-driven, age-based groups, titled "adults" (24–45 years, $n = 10$), "middle-aged adults" (46–65 years, $n = 12$), and "seniors" (66–76 years, $n = 7$), with mean pupil size deviations of 2.01 ($SD = 0.58$), 1.82 ($SD = 0.52$), and 1.12 ($SD = 0.33$), respectively. According to one-way ANOVA, there were significant differences ($F(2,26) = 6.749$, $p = .004$) between the three groups, and based on pairwise comparisons (Tukey), the senior group differed from the adult ($p = .004$) and middle-aged adult group ($p = .020$), but the adult group did not differ from the middle-aged participants ($p = .654$). Due to the obvious effect of age regarding this arousal measure, the senior group was omitted from the following pupil size analyses.

For the remaining 22 participants, the pupil size deviation was not statistically significantly associated with BMI ($r = -0.157$, $p = .509$; $n = 20$; one outlier [BMI > 50] was omitted), amount of food later consumed (g; $r = 0.231$, $p = .509$; $n = 21$; one participant's information was missing), or time spent eating the salad lunch ($r = 0.366$, $p = .103$, $n =$

[3] It was pre-checked that there were no statistically significant correlations between (a) olive preference, olive use, food item take (freq.), food amount (g), and time spent eating and age, or (b) olive preference, olive use, food item take (freq.), food amount (g), and time spent eating and BMI. (One outlier [BMI > 50] was omitted from the BMI correlation analyses.) In sum, there were no overall age- or BMI-based differences in the participants' eating behavior or prior preferences toward our food item of interest.

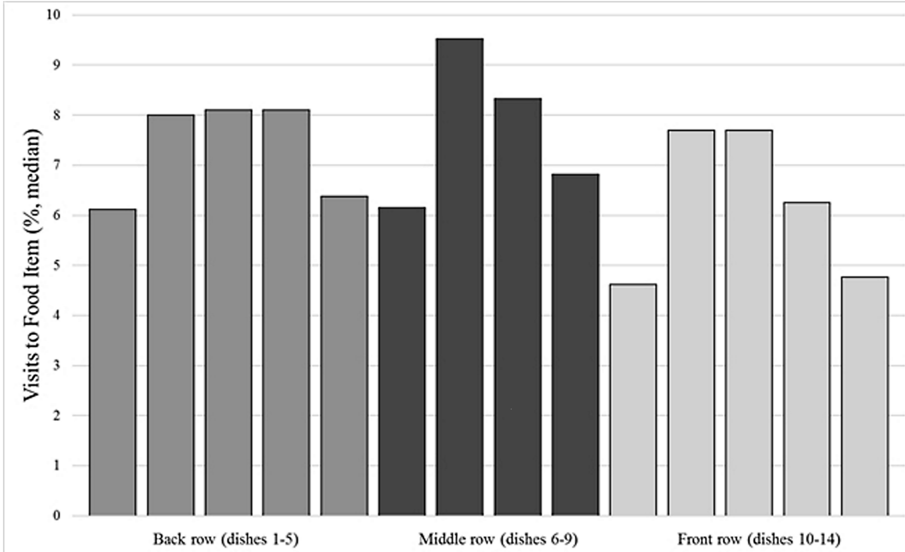

Fig. 3. Distribution of (gaze) visits to the fourteen dishes (% of all visits, median values, $n = 29$).

Table 2
Significance values for a series of Dunn-Bonferroni post hoc tests (pairwise comparisons) regarding gaze visits to the 14 dishes (n = 29) during the initial 20-s viewing of the food items. Significance values were adjusted with the Bonferroni correction for multiple tests.

	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.248	0.836	0.371	1.000	1.000	0.004**	0.089	1.000	1.000	1.000	1.000	1.000	1.000
2		1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.001**	1.000	1.000	1.000	0.001**
3			1.000	1.000	1.000	1.000	1.000	1.000	0.006*	1.000	1.000	1.000	0.005*
4				1.000	1.000	1.000	1.000	1.000	0.002**	1.000	1.000	1.000	0.002**
5					1.000	0.031*	0.498	1.000	1.000	1.000	1.000	1.000	1.000
6						0.036*	0.548	1.000	1.000	1.000	1.000	1.000	1.000
7							1.000	0.498	0.000***	1.000	1.000	0.051	0.000***
8								1.000	0.000***	1.000	1.000	0.727	0.000***
9									0.603	1.000	1.000	1.000	0.575
10										0.042*	0.010*	1.000	1.000
11											1.000	1.000	0.040*
12												1.000	0.010*
13													1.000

* $p < .05$, ** $p < .005$, *** $p < .001$

Table 3
Means and standard deviations (in parentheses) for olive take (g) and selected fixation parameters for the three olive preference groups.

Olive preference	*n*	Olive take (g)	First fixation duration (ms)		Total fixation time (%)	
			Olive	Black food items (sum)	Olive	Black food items (sum)
Unpleasant	6	0.00 (0.00)	396 (235)	810 (350)	7.49 (2.69)	16.43 (2.65)
Somewhat unpleasant to pleasant	10	14.10 (11.44)	364 (282)	750 (558)	6.17 (2.93)	15.76 (6.24)
Very pleasant	13	18.54 (9.85)	457 (228)	758 (327)	8.53 (3.11)	14.76 (3.64)

21; one participant's information was missing). Participants had also reported their hunger prior to seeing the buffet on a four-step scale. Of the 22 participants, six reported being "very hungry" (score 1), twelve reported being "quite hungry" (score 2) and four reported being "somewhat hungry" (score 3). None of these participants selected option "not at all hungry" (score 4). This distribution of selected categories did now allow re-categorizing the participants into fewer groups with more subjects in each; thus, despite the small group sizes, deviation of pupil size for these three groups was compared with a Kruskal-Wallis analysis. Experienced hunger was not associated with pupil size deviation ($\chi^2(2) = 4.595$, $p = .101$, $n = 22$). In sum, variability in pupil size deviations during initial food viewing was not associated with BMI, perceived hunger (prior to eating), food consumption (g) or time spent eating the self-choice salad.

3.4. Profiles of food view and selection

To search for associations between food viewing and food selection and select appropriate measures for a cluster analysis, correlations between two scanpath measures, selected background measures, and food selection measures were examined. *Returns to food items* (please see section 2.5.2 for details) during the 20-s first viewing were not statistically significantly associated with age ($r = -0.010$, $p = .958$, $n = 29$), BMI ($r = -0.243$, $p = .232$, $n = 26$; one outlier [BMI > 50] was omitted), food item take ($r = 0.198$, $p = .321$, $n = 27$), food consumption (g; $r = 0.020$, $p = .922$, $n = 27$), or time spent eating ($r = 0.209$, $p = .296$, $n = 27$). *Stepwise visual processing* (please see section 2.5.2 for details) of the food items during the 20-s first viewing was not statistically significantly associated with age ($r = -0.238$, $p = .214$, $n = 29$), BMI ($r = -0.330$, $p =$

.099, $n = 26$; one outlier [BMI > 50] was omitted), food consumption (g; $r = 0.128$, $p = .526$, $n = 27$), or time spent eating ($r = 0.006$, $p = .976$, $n = 27$). However, there was a statistically significant negative correlation between stepwise visual processing during the initial 20-s viewing of the food trolley and later food item take ($r = -0.393$, $p = .043$, $n = 27$). These two parameters were chosen for the K-means clustering. Considering the small sample size ($n = 27$; 5 participants data was missing), a two-cluster solution ($n = 15$ for cluster 1 and $n = 12$ for cluster 2) was selected as the basis for the following descriptive analysis. Fig. 4 illustrates how the participants were positioned with respect to the two applied parameters.

Cluster 1 included participants with a tendency to visually inspect the 14 dishes in a more row-based manner (cluster center for stepwise processing index = 60%), and take fewer food items, when assembling their salad (cluster center for food item take = 10.27). Participants in cluster 2 typically inspected the food items in a less systematic manner (cluster center for stepwise processing index = 48%) and assembled their salad with more food item takes (cluster center for food item take = 12.83). Participant groups created based on the clustering differed significantly both in terms of stepwise visual processing ($t(25) = 3.688$, $p = .001$) and food item take ($t(25) = -6.681$, $p = .000$).

Fig. 5 illustrates, via two case examples (participants 79 and 15), these associations between visual attention toward the food items during the initial 20-s viewing and later food take. In these cases, the final salad sizes were very similar (413 g for participant 79 and 452 g for participant 15), though they were the results of very different visual and food selection processes.

Participant 79 (Fig. 5, left) exemplifies the first cluster. This participant's scanpath was systematic: she inspected food items typically one row at a time, moving both from left to right and right to left (grey dotted lines in Fig. 5 separate the three rows of dishes on the food trolley). When assembling her salad after the initial viewing, she took each food item only once and tended to move from one dish to another that was either horizontally or vertically close by (the three first taken food items were from dishes 7, 11, and 12; then dishes 5, 9 and 14; and finally dishes 1, 2 and 3). In addition, she took fruits last, placed them on the bread plate, and had them as a dessert. In conclusion, this participant studied the buffet systematically, appeared to have a clear idea of how much to take of each of the food items, and even personalized the use of the bread plate. The presentation of the food items appeared to affect not only her visual processing but also her food take, since she selected her food items in three sequences of nearby food items.

Participant 15 (see Fig. 5, right) demonstrated a very different behavioral pattern. Her initial viewing process was marked by skips between dishes and rows. Unlike participant 79, her food selection did not proceed from one dish to a nearby dish: there were only two sequences where she selected food from two or three nearby dishes (dish 1 and 2; dishes 8, 4 and 3). She also took both iceberg lettuce and pesto pasta twice, and the latter re-take took place after already getting bread and a spread from the side table. In sum, this participant appeared to be more guided by the food items themselves and not their positioning on the food trolley both in terms of visual processing and food selection. She also seemed less decisive in how much food to take, as signaled by the re-takes of lettuce and pasta.

4. Discussion

This study investigated visual attention toward food items in a salad buffet setting, where participants not only viewed the foods while their viewing processes were recorded with a mobile eye tracker, but also assembled and ate a salad of their choosing. Fixation, scanpath, and arousal measures were used in searching for associations between eye-movement, stimulus-related, and food-selection measures, with the ultimate goal of developing a systematic use of MET in real-life studies about food viewing and selection.

Aim (a) focused on the stimulus-driven effects of food item color and position. In this study, food item color did not affect visual processing, but foods in central positions on the trolley gained relatively more gaze visits than those at the corners. However, dish placement did not affect the time spent on fixating on the dishes (cf. Attila et al., 2020). Aim (b) targeted the possible associations between food item preference and visual attention when a specific target food item was placed among several other food items. Unlike in other food choice studies (e.g.,

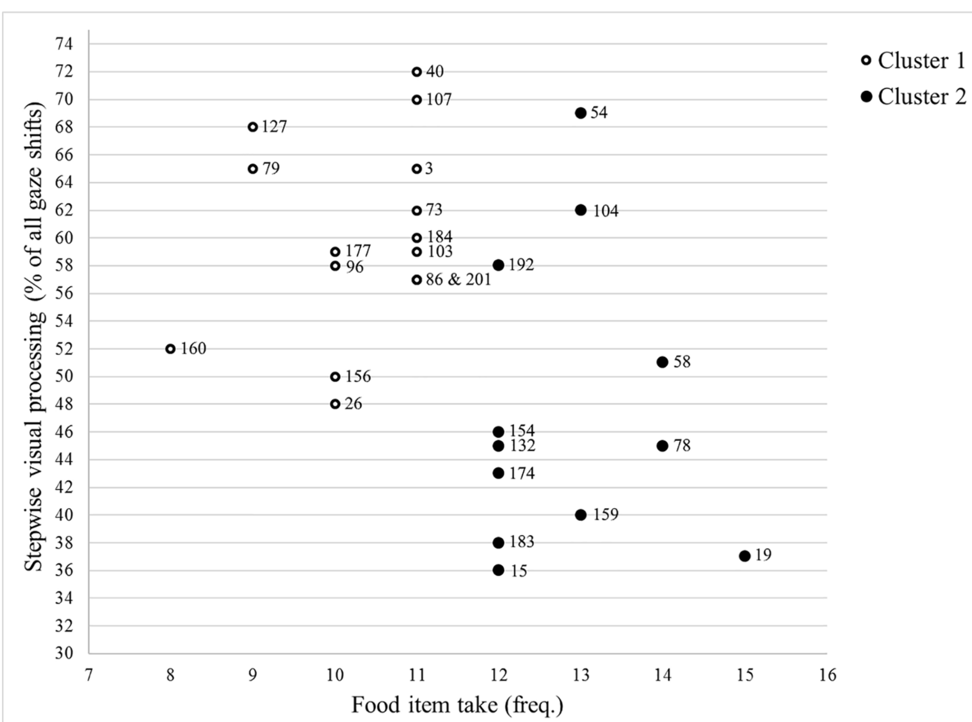

Fig. 4. Participants, their ID numbers, cluster number, and values for their stepwise visual processing (%) during the 20-s initial viewing of the food items, and later food item take (freq.). Fourteen food items were served; note that the same food item was sometimes taken twice.

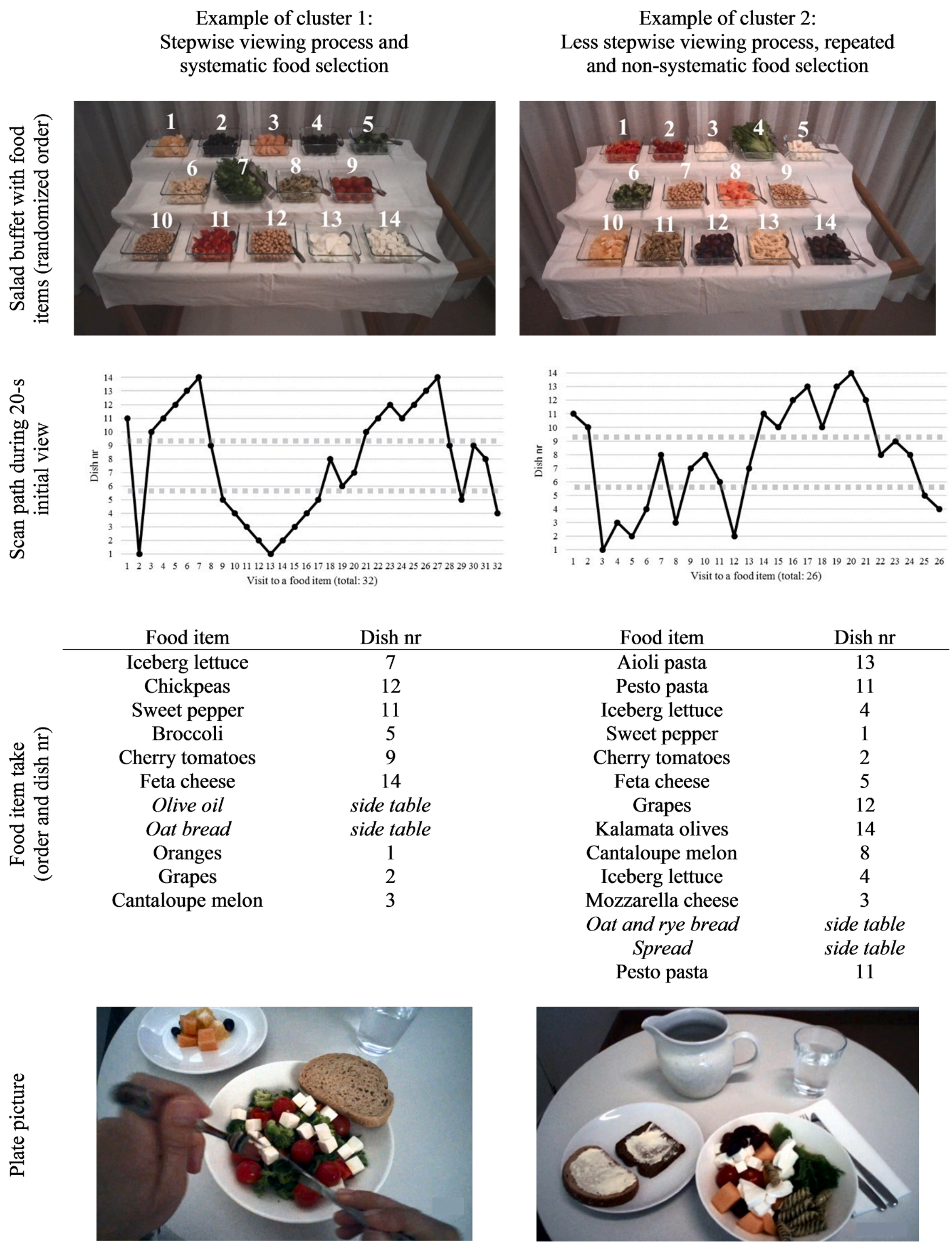

Fig. 5. Case examples from the two clusters: salad buffets, scanpath visualizations, food item take, and plate picture for participants 79 (left) and 15 (right). Pictures were obtained from Tobii Pro Glasses 2 recordings (scene camera view).

Jantathai et al., 2013; Wang et al., 2018), in this design, the participants not only looked at the preferred and non-preferred food item and chose between them, but were able to actually take the food item as part of their salad. Indeed, participants with dislike toward the target item did not later include it in their self-choice salad. Nevertheless, similarly to Motoki et al. (2018) but contrary to Attila et al. (2020), there were no significant effects with respect to the applied fixation measures on the preferred or non-preferred food item.

With respect to aim (c), it was assumed that changes in pupil size might be associated with BMI (Graham et al., 2011; Wang et al., 2018). It was also explored whether looking at a tempting lunch buffet might result in larger arousal effects for participants that were hungry, who later assembled larger salads, or who spend a longer time eating their salad. First, this arousal measure was not reliable in the case of elderly participants (see Guillon et al., 2016), something that future studies should take into account when recruiting participants. Second, for the remaining (smallish) participant group, none of the explored measures showed any associations with changes in pupil sizes. Due to difficulties in pinpointing factors affecting changes in pupil size in free-viewing tasks, arousal was measured only as the deviation of the pupil diameter during the whole 20-s viewing. However, only general situational arousal could be explored in this manner and changes in pupil size could not be linked to any specific food item. Should one wish to study the associations of pupil-size and a specific visual target, a much more controlled experimental set-up is needed. All in all, if pupil size is applied in MET studies about food viewing and selection, a very careful use of the measure (avoiding averaging and reflecting on the necessity of defining participant-specific baselines) is advised.

Aim (d) was approached in a more exploratory manner. The degree of stepwise visual processing of the salad buffet during the initial food viewing and the number of taken food items during later food selection formed a basis for a cluster analysis, and two case descriptions exemplified thus formed profiles of food viewing and selection in this relatively freely defined food-view task. In sum, a simple combination of a scanpath measure and the number of selected food items helped to pinpoint the key differences between two “attention-action“ profiles. When able to assemble a salad of her own choice, the participant in cluster 1 tended to be more affected by the placement of the dishes in terms of visual processing and food take. She moved her gaze to nearby food items and also selected food items in this manner, and demonstrated decisiveness by taking each food item only once. The participant in cluster 2, however, appeared to be more guided by the dish contents, moving between dishes that were placed far apart (both when viewing them and when selecting the food items) and returning to some dishes when assembling her salad. (These profiles bear resemblance to what have been dubbed impulsive and reflective systems; see Motoki et al., 2018.) In terms of visual attention towards food in real-life contexts, acknowledging these types of (situation-specific) profiles could be important when planning future studies. For instance, if one is aiming to influence participants’ choices by nudging them towards healthy products, it may be that people differ in their sensitivity to these attempts according to their situation-specific food viewing and selection profiles.

It needs to be noted that with MET, steps in data handling are subject to researcher-originated errors, and even more so than data handling with remote eye trackers. In this study, for instance, fixations from the viewing of the salad buffet were manually mapped onto still images of the buffet, as this was required for conducting the further analyses. In the presented setup, the food items were placed quite distinctly from one another and the mapping was considered reliable enough for the performed analyses, though it was still admittedly subject to errors; especially so as the mapping was conducted by only one researcher (though with considerable experience in handling eye-tracking data). Other MET studies with real foods should also take into account that the setting should enable as reliable data pre-processing as possible. We also encourage a research team to invite members with hands-on experience in handling eye-tracking data, or train their coders properly beforehand. All in all, careful pretesting of the setup as well as data handling is highly advisable.

Another remark one needs to make relates to the small sample size in this study. In the presented type of set-up the individual testing, preparation of real foods, and laborious handling of eye-tracking data all hinder the collection of large data sets. This prevents from using more sophisticated statistical analyses that could answer slightly more complex research questions. In order to increase data set size, a practical alternative is to collaborate between research teams: one could systematically repeat set-ups previously reported by others, providing that enough method details are presented in the manuscripts, or even share eye-movement data between teams. Similar suggestions have already been made in other less-researched domains that apply the eye-tracking method and often struggle with collecting large data sets (see, for example, Puurtinen, 2018).

The eye-movement data for this study was collected during the 20-s initial viewing of the foods. Thus, the set-up, including dish position and viewer’s distance to these visual targets were created with that 20-s task in mind. The participants also wore the eye-tracker during food selection and eating, as this was a way for the experimenter to keep track of what occurred in the multisensory room. Eye-movement data from these intervals is, however, in most cases, unusable; the MET device applied in this study easily lost track of the pupils of the viewer when the viewer looked down. If one wishes to collect eye-movement data also during food selection, visual targets need to be placed high enough and also possibly tilted toward the viewer. Recording the actual eating requires even more innovative solutions for the eye-tracking to succeed. Future studies should continue work on designing and testing set-ups that come closer to how food is served in restaurants, while still enabling systematic study of food selection and eating behavior.

All in all, the need for careful checking of data quality and potential manual data processing in MET studies make the collection of large quantitative data sets difficult for a single research team, and this was also the case in this study (combined with a laborious research protocol). Since the reliability of statistical testing is, on occasion, questionable, the results should be read as pointing toward some promising methodological traits, instead of describing visual attention toward foods, or food choice behavior, as is. These types of preliminary and descriptive studies are still needed in order to develop the best practices for larger data collections that could be performed jointly by several research teams.

5. Conclusion

The results of this study suggest that in real-life food viewing tasks, it might be particularly useful to focus on measures such as visit counts and scanpath measures, though these are typically not the measures researchers beginning to use eye tracking start with. These measures might show behavioral differences between participants, which are not revealed by fixation parameters, since in real-life settings the latter are easily confounded by several uncontrolled (or even unknown) factors. Overall, the interplay between visual attention and selected actions seems an intriguing avenue for future research about food viewing and selection, and identifying attention-action profiles and creating experiments around them, instead of only comparing groups based on more traditional background measures, might help in explaining some of the so-far mixed results concerning visual attention toward food.

Overall, methodological discussions are still much needed. Important contributions have been made by, for example, Doolan et al. (2014), who discuss the benefit of direct measures in addition to indirect ones, and Vu et al. (2016), who tested the effect of selected design variables for some eye-tracking measures. The gradual and systematic building of a coherent methodological framework becomes especially important when shifting from static and remote eye trackers, placed in carefully controlled environments and with a long research tradition to build on, to mobile eye tracking in less-controlled research environments and tasks. With MET, ensuring data quality, understanding significant intervening factors, and developing the best tools for data analyses are still underway (e.g., Pérez-Edgar et al., 2020). Steps toward reconciling laboratory and real-life conditions have recently been taken in, for instance, consumer research (Bialkova et al., 2020), and food-related studies could do this type of careful experimenting with research conditions, too. With systematic method testing, MET could enable food perception research to take into account food’s multisensory nature, and help in combining the currently separate scientific discussions about visual perception and other food-related sensory experiences.

Funding sources

This work was supported by the Academy of Finland [grant numbers 309408 and 318894].

CRediT authorship contribution statement

Marjaana Puurtinen: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing - review & editing, Writing - original draft. **Ulla Hoppu:** Conceptualization, Formal analysis, Investigation, Methodology, Writing - original draft. **Sari Puputti:** Conceptualization, Investigation, Methodology. **Saila Mattila:** Investigation. **Mari Sandell:** Conceptualization, Funding acquisition, Project administration, Writing - review & editing.

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.

Acknowledgments

We thank Nora Logren for her work as a research assistant during the data collection and the participants of the study for their time and effort.

Appendix A. Fixation filter settings applied in data pre-processing.

Software	Tobii Pro Lab 1.123
Gaze filter	Tobii I-VT (Attention)
Gap fill-in	Max. gap length 75 ms
Noice reduction	Moving median
	Window size 3 samples
Velocity calculator	Window length 20 ms
I-VT classifier	Threshold 100/s
Merge adjacent fixations (yes)	Max. time between fixations 75 ms
	Max. angle between fixations 0.5◦
Discard short fixations (yes)	Min. fixation duration 60 ms

Appendix B. Mean values (with *SD*s in parenthesis) for the three eye-movement parameters and for each food item pair ($n = 29$).

Food item pair	Total fixation time (%)*	Visits to food items (%)*	First fixation duration (ms)**
Black food items	15.45 (4.46)	14.59 (3.32)	770 (410)
Yellow food items	14.65 (6.11)	14.43 (5.29)	710 (310)
Green food items	14.21 (4.19)	14.50 (3.89)	700 (400)
Red food items	13.31 (4.64)	14.73 (3.49)	650 (280)
White food items	14.73 (4.97)	14.57 (4.03)	600 (320)
Beige food items	14.76 (4.64)	14.03 (2.92)	740 (430)
Pasta food items	12.90 (4.95)	13.14 (3.47)	650 (360)

* sum for a pair of food items.
** sum for a pair of food items.

Appendix C. Mean values (with *SD*s in parenthesis) for the three eye-movement parameters and for the 14 dishes ($n = 29$).

Dish nr	Total fixation time (%)	Visits to food items (%)	First fixation duration (ms)
1	7.10 (3.50)	5.84 (1.97)	410 (285)
2	6.97 (2.64)	8.13 (2.13)	256 (134)
3	7.73 (3.84)	8.10 (3.13)	324 (211)
4	6.87 (2.06)	7.88 (2.48)	360 (237)
5	8.11 (4.90)	6.24 (2.21)	424 (290)
6	6.65 (3.46)	6.35 (2.82)	430 (377)
7	8.42 (3.40)	9.24 (2.41)	333 (225)
8	8.45 (3.79)	9.28 (4.01)	396 (267)
9	6.53 (3.65)	6.89 (2.00)	362 (249)
10	6.35 (3.64)	5.14 (2.13)	356 (228)
11	6.93 (3.32)	7.61 (2.23)	274 (140)
12	7.27 (3.28)	7.80 (2.61)	332 (270)
13	5.96 (3.20)	6.45 (2.73)	263 (121)
14	6.67 (3.04)	5.04 (1.98)	302 (249)

References

Attila, G., Danner, L., Dürrschmid, K., Kókai, Z., Sipos, L., Huzsvai, L., & Kovács, S. (2020). Structure of presented stimuli influences gazing behavior and choice. *Food Quality and Preference, 83*, Article 103915. https://doi.org/10.1015/j.goodqual.2020.103915

Bialkova, S., Grunert, K. G., & van Trijp, H. (2020). From desktop to supermarket shelf: Eye-tracking exploration on consumer attention and choice. *Food Quality and Preference, 81*, 103839. https://doi.org/10.1016/j.foodqual.2019.103839

Doolan, K. J., Breslin, G., Hanna, D., Murphy, K., & Gallagher, A. M. (2014). Visual attention to food cues in obesity: An eye-tracking study. *Obesity, 22*, 2501–2507. https://doi.org/10.1002/oby.20884

Eraslan, S., Yesilada, Y., & Harper, S. (2016). Eye tracking scanpath analysis techniques on web pages: A survey, evaluation and comparison. *Journal of Eye Movement Research, 9*. https://doi.org/10.16910/jemr.9.1.2

Garcia-Burgos, D., Lao, J., Munsch, S., & Caldara, R. (2017). Visual attention to food cues is differentially modulated by gustatory-hedonic and post-ingestive attributes. *Food Research International, 97*, 199–208. https://doi.org/10.1016/j.foodres.2017.04.011

Graham, R., Hoover, A., Ceballos, N. A., & Komogortsev, O. (2011). Body mass index moderates gaze orienting biases and pupil diameter to high and low calorie food images. *Appetite, 56*(3), 577–586. https://doi.org/10.1016/j.appet.2011.01.029

Guillon, M., Dumbleton, K., Theodoratos, P., Gobbe, M., Wooley, C. B., & Moody, K. (2016). The effects of age, refractive status, and luminance of pupil size. *Optometry and Vision Science, 93*, 1093–1100. https://doi.org/10.1097/OPX.0000000000000893

Hendrikse, J. J., Cachia, R. L., Kothe, E. J., McPhie, S., Skouteris, H., & Hayden, M. J. (2015). Attentional biases for food cues in overweight and individuals with obesity: A systematic review of the literature. *Obesity Reviews, 16*(5), 424–432. https://doi.org/10.1111/obr.2015.16.issue-510.1111/obr.12265

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Hoppu, U., Puputti, S., Mattila, S., Puurtinen, M., & Sandell, M. (2020). Food consumption and emotions at a salad lunch buffet in a multisensory environment. *Foods, 9*(10), 1349. https://doi.org/10.3390/foods9101349

Hummel, G., Ehret, J., Zerweck, I., Winter, S. S., & Stroebele-Benschop, N. (2018). How eating behavior, food stimuli and gender may affect visual attention: An eye tracking study. *Eating Behaviors, 31*, 60–67. https://doi.org/10.1016/j.eatbeh.2018.08.002

Hummel, G., Zerweck, I., Ehret, J., Winter, S. S., & Stroebele-Benschop, N. (2017). The influence of the arrangement of different food images on participants' attention: An experimental eye-tracking study. *Food Quality and Preference, 62*, 111–119. https://doi.org/10.1016/j.foodqual.2017.07.003

Jantathai, S., Danner, L., Joechl, M., & Dürrschmid, K. (2013). Gazing behavior, choice and color of food: Does gazing behavior predict choice? *Food Research International, 54*(2), 1621–1626. https://doi.org/10.1016/j.foodres.2013.09.050

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: Strengths and weaknesses. *Behavior Research Methods, 45*(1), 251–266. https://doi.org/10.3758/s13428-012-0226-9

Michel, C., Velasco, C., Fraemohs, P., & Spence, C. (2015). Studying the impact of plating on ratings of the food served in a naturalistic dining context. *Appetite, 90*, 45–50. https://doi.org/10.1016/j.appet.2015.02.030

Motoki, K., Saito, T., Nouchi, R., Kawashima, R., & Sugiura, M. (2018). Tastiness but not healthfulness captures automatic visual attention: Preliminary evidence from an eye-tracking study. *Food Quality and Preference, 64*, 148–153. https://doi.org/10.1016/j.foodqual.2017.09.014

Nijs, I. M. T., Muris, P., Euser, A. S., & Franken, H. A. (2010). Differences in attention to food and food intake between overweight/obese and normal-weight females under conditions of hunger and satiety. *Appetite, 54*, 243–254. https://doi.org/10.1015/j.appet.2009.11.004

Nummenmaa, L., Hietanen, J. K., Calvo, M. G., Hyönä, J., & Greenlee, M. W. (2011). Food catches the eye but not for everyone: A BMI–contingent attentional bias in rapid detection of nutriments. *PLoS ONE, 6*(5), e19215. https://doi.org/10.1371/journal.pone.0019215

Paakki, M., Sandell, M., & Hopia, A. (2019). Visual attractiveness depends on colorfulness and color contrasts in mixed salads. *Food Quality & Preference, 76*, 81–90. https://doi.org/10.1016/j.foodqual.2019.04.004

Peng-Li, D., Byrne, D. V., Chan, R. C. K., & Wang, Q. J. (2020). The influence of taste-congruent soundtracks on visual attention and food choice: A cross-cultural eye-tracking study in Chinese and Danish consumers. *Food Quality and Preference, 85*, 103962. https://doi.org/10.1016/j.foodqual.2020.103962

Pérez-Edgar, K., MacNeill, L. A., & Fu, X. (2020). Navigating through the experienced environment: Insights from mobile eye tracking. *Current Directions in Psychological Science, 29*(3), 286–292. https://doi.org/10.1177/0963721420915880

Puputti, S., Aisala, H., Hoppu, U., & Sandell, M. (2018). Multidimensional measurement of individual differences in taste perception. *Food Quality and Preference, 65*, 10–17. https://doi.org/10.1016/j.foodqual.2017.12.006

Puurtinen, M. (2018). Eye on music reading: A methodological review of studies from 1994 to 2017. *Journal of Eye Movement Research, 11*, 1–16. https://doi.org/10.16910/jemr.11.2.2

Rayner, K. (1998). Eye movement in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506. https://doi.org/10.1080/17470210902816461

Schmidt, R., Lüthold, P., Kittel, R., Tetzlaff, A., & Hilberg, A. (2016). Visual attentional bias for food in adolescents with binge-eating disorder. *Journal of Psychiatric Research, 80*, 22–29. https://doi.org/10.1016/j.jpsychires.2016.05-016

Spence, C. (2015). Multisensory flavor perception. *Cell, 161*(1), 24–35. https://doi.org/10.1016/j.cell.2015.03.007

Vu, T. M. H., Tu, V. P., & Duerrschmid, K. (2016). Design factors influence consumers' gazing behaviour and decision time in an eye-tracking test: A study on food images. *Food Quality and Preference, 47*, 130–138. https://doi.org/10.1016/j.foodqual.2015.05.008

Wadhera, D., & Capaldi-Phillips, E. D. (2014). A review of visual cues associated with food on food acceptance and consumption. *Eating Behaviors, 15*(1), 132–143. https://doi.org/10.1016/j.eatbeh.2013.11.003

Wang, E., Cakmak, Y. O., & Peng, M. (2018). Eating with eyes – Comparing eye movements and food choices between overweight and lean individuals in a real-life buffet setting. *Appetite, 125*, 152–159. https://doi.org/10.1016/j.appet.2018.02.003

Werthmann, J., Jansen, A., & Roefs, A. (2015). Worry or craving? A selective review of evidence for food-related attention biases in obese individuals, eating-disorder patients, restrained eaters and healthy samples. *Proceedings of the Nutrition Society, 74*(2), 99–114. https://doi.org/10.1017/S0029665114001451

Werthmann, J., Roefs, A., Nederkoorn, C., Mogg, K., Bradley, B. P., & Jansen, A. (2011). Can(not) take my eyes off it: Attention bias for food in overweight participants. *Health Psychology, 30*, 561–569. https://doi.org/10.1037/a0024291