

## Article

# Predicting Temporal Liking of Food Pairings from Temporal Dominance of Sensations Data via Reservoir Computing on Crackers and Spreads

Hiroharu Natsume <sup>†</sup> and Shogo Okamoto <sup>\*,†</sup> 

Department of Computer Science, Tokyo Metropolitan University, Hino 191-0065, Japan;  
natsume-hiroharu@ed.tmu.ac.jp

\* Correspondence: okamoto@tmu.ac.jp

<sup>†</sup> These authors contributed equally to this work.

## Abstract

The temporal dominance of sensations (TDS) and temporal liking (TL) methods offer complementary insights into the evolution of sensory and hedonic responses during food consumption. This study investigates the feasibility of predicting TL curves for food pairings from their TDS profiles using reservoir computing, a type of recurrent neural network. Participants evaluated eight samples—two crackers (plain, sesame), two spreads (peanut butter, strawberry jam), and their four binary combinations—performing both TDS and TL evaluations. This process yielded paired time-series data of TDS and TL curves. We trained various reservoir models under different conditions, including varying reservoir sizes (64, 128, 192, or 256 neurons) and the inclusion of auxiliary input dimensions, such as flags indicating the types of foods tasted. Our results show that models with minimal auxiliary inputs achieved the lowest root mean squared errors (RMSEs), with the best performance being an RMSE of 0.44 points on a 9-point liking scale between the observed and predicted TL curves. The ability to predict TL curves for food pairings holds some promise for reducing the need for extensive sensory evaluation, especially when a large number of food combinations are targeted.

**Keywords:** food combination; reservoir computing; cracker, spread



Academic Editor: Reza Farhoosh

Received: 9 August 2025

Revised: 20 September 2025

Accepted: 27 September 2025

Published: 29 September 2025

**Citation:** Natsume, H.; Okamoto, S. Predicting Temporal Liking of Food Pairings from Temporal Dominance of Sensations Data via Reservoir Computing on Crackers and Spreads. *Foods* **2025**, *14*, 3373. <https://doi.org/10.3390/foods14193373>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The temporal dominance of sensations (TDS) [1–3] and temporal liking (TL) [4,5] methods allow recording of sensory and hedonic dynamics during food consumption. By combining the temporal evolution of dominant sensations captured by the TDS method with the numerical hedonic evaluations provided by the TL method, researchers have attempted to elucidate the temporal relationship between perception and liking [5–19]. These approaches aim to identify which sensory attributes—such as taste, aroma, and texture—determine liking at each moment or interval during consumption. For example, a pioneering study proposed using the interval mean of liking scores during periods when a specific attribute is dominant as an indicator of the sensory drivers of liking [5]. This approach was later refined by weighting each attribute according to its dominance duration [11]. Kuesten and Bi [16] repeatedly conducted time intensity methods for different sensory attributes instead of the TDS method, and investigated how those intensities related to the temporal liking of fruit chews. They used a non-additive model to connect the sensory and liking data, and then employed relative importance measures used in

machine learning and statistic learning techniques as evolving temporal drivers of liking. Okamoto et al. demonstrated the validity of vector autoregression [17] and state-space [18] models that determine the current liking value based on the current and past sensory properties of food items. Long before the advent of the TDS method, numerous studies had investigated the relationship between sensory attributes such as taste, aroma, and texture and consumer liking. Classic review articles and textbook chapters (e.g., [20–24]) comprehensively summarize analytical approaches such as preference mapping, regression, and multivariate analysis for linking sensory profiles to hedonic responses.

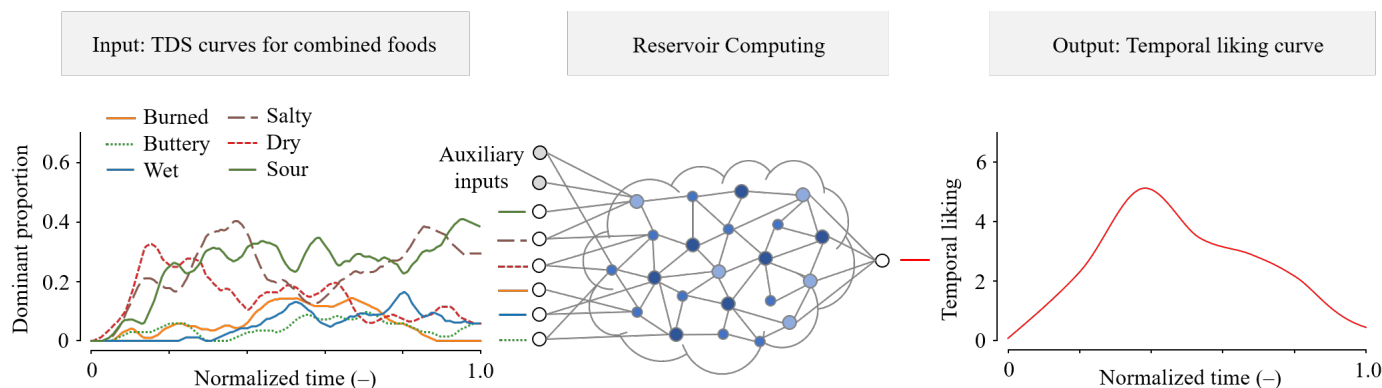
In connection with these techniques, our group has developed methods for predicting TL outcomes based on TDS data. For example, we have used neural networks to estimate corresponding TL curves from TDS curves for strawberries and coffee [8,25]. A primary motivation for this work is to reduce the cost of sensory evaluation or consumer tests and to streamline the food development process—a goal widely shared among food industry practitioners [26–28]. Sensory evaluation is employed in nearly 70% of food development cases, involving both expert panels and consumers, and can represent a major cost driver in the process [29]. Consequently, the development and adoption of rapid sensory profiling methods, consumer-oriented approaches, and predictive modeling, including machine learning approaches, have become important topics for reducing the time and expense of sensory-driven innovation.

A scenario in which cost efficiency becomes particularly important is the evaluation of food pairings, which have long been of great interest in sensory science [30–33]. For example, when considering combinations such as drinks and foods, even with only 10 varieties of each, the number of possible pairings reaches 100. Conducting sensory tests on all these combinations would substantially increase costs. Therefore, there is value in developing methods that can identify favorable food pairings while minimizing the cost of sensory evaluation. To this end, for example, Kustos et al. [33] efficiently investigated compatible food and wine pairings by combining descriptive sensory analysis with hedonic evaluation.

Several research groups have investigated methods for food pairing using machine learning [34]. For example, Ahn et al. [35] analyzed a global recipe database to identify preferred ingredient combinations. Similarly, Park et al. [36] estimated the compatibility of ingredient combinations using neural networks trained on recipe data. Furthermore, Park et al. [37] developed a method based on a knowledge graph to assess ingredient pairing compatibility. These approaches aim to discover favorable ingredient pairs from large-scale databases. In contrast, our study evaluates the quality of ingredient pairings by applying machine learning to TDS and TL data—a fundamentally different approach. To our knowledge, there have been no prior studies that have used TDS and TL methods to estimate the quality of ingredient pairings. While Galmarini et al. [19] used TDS and TL methods to investigate the temporal drivers of liking for wine and cheese pairings, their objective was not to predict pairing compatibility. Considering the growing popularity of TDS methods in food science and their adoption in industry [38], developing approaches that can predict liking from TDS data could provide an effective tool for supporting product innovation and reducing the cost of sensory testing in the food industry.

In this study, we develop a method for estimating the compatibility of food pairings using TDS and TL curves. As a prediction method, we employ reservoir computing [39–41], a type of recurrent neural network well suited for handling time-series data. We focus on two types of crackers and two types of spreads. TDS and TL curves are collected for each item when tasted alone and in combination. As shown in Figure 1, the reservoir computing model trained on these data predicts the TL curve for a given food pairing based on its TDS curve. Although predicting both the TDS and TL curves for pairings directly from the TDS

curves of individual items remains an aspirational research direction, this study represents an initial attempt to explore feasibility within a limited context.



**Figure 1.** Estimation of a TL curve from TDS curves obtained when two foods are tasted in combination.

## 2. Methods: Sensory Evaluation and Data Processing

### 2.1. Food Samples

We used two types of crackers (plain, sesame) and two types of spreads (peanut, strawberry). The plain crackers were Premium Saltine Crackers (Nabisco, Inc., East Hanover, NJ, USA; lot number: 24 OCA00433D2). The sesame crackers were Black Sesame Soft Crackers (Maeda Confectionery Co., Ltd., Sakai, Osaka, Japan; best-before date: 14 October 2025). The peanut butter was Verde Peanut Whip (Aohata Corporation, Takehara, Hiroshima, Japan; production number: 2026.3/DA2; best-before date: March 2026). The strawberry jam was Spoon Free Strawberry (Aohata Corporation, Takehara, Hiroshima, Japan; production number: 2026.3.4/C; best before date: March 2026). The names and manufacturers of the individual products are listed in Table 1. These foods were selected due to their ease of handling, consistent quality and availability, and suitability for combination consumption.

**Table 1.** Food products.

Category	Name	Original Name	Manufacturer
Cracker	Plain	Premium Saltine Crackers	Nabisco, Inc.
	Sesame	Black Sesame Soft Crackers	Maeda Confectionery Co., Ltd.
Spread	Peanut	Verde Peanut Whip	Aohata Corporation
	Strawberry	Spoon Free Strawberry	Aohata Corporation

Crackers were served as bite-sized pieces, and spreads were served using a spoon (3.5 g). In addition to these individual items, we prepared four cracker–spread combinations: plain–peanut, plain–strawberry, sesame–peanut, and sesame–strawberry. In the combination task, bite-sized crackers were served with 3.5 g of spread applied on top.

In this study, two types of crackers and two types of spreads were selected, resulting in four pairings and, together with the single-item conditions, eight tasting conditions in total. This number of samples was selected to balance the variety of foods examined with the practical feasibility of the experiment (see Section 2.3).

### 2.2. Panel

Sixteen university students (age: 20–25; 5 females and 11 males) participated as the panel in the experiment. Prior to the experiment, participants confirmed that at least two hours had passed since their last meal and that they did not dislike the food being served. While this panel size satisfies the minimum recommendation for semi-trained panels

specified in ISO 13299 [2], the homogeneity of the panel in terms of age and background is a limitation that should be considered when generalizing the findings.

The panel received training focused on the identification of sensory attributes. Participants first reviewed and agreed upon definitions of the TDS attributes. Two training sessions were conducted to familiarize them with the TDS method, including practice in selecting dominant sensations and becoming accustomed to the computer interface used for data collection. During these sessions, participants practiced with sample foods not included in the main experiment. In addition, all participants were familiarized with the TL task and the use of the nine-point liking scale. They completed two practice sessions in which they provided real-time liking responses to sample foods using the interface.

### 2.3. TDS and TL Tasks

The TDS and TL experiments were conducted in a quiet room maintained at approximately 23 °C. Each participant evaluated two types of crackers, two types of spreads, and all four possible combinations of these foods. After tasting each food sample, participants rinsed their mouths with soft water. The order of food sample presentation was randomized, as was the sequence of TDS and TL tasks for each food. The TDS and TL tasks were conducted separately, with each task repeated three times per food item. In total, each participant completed 48 trials (8 food conditions  $\times$  2 tasks  $\times$  3 repetitions), which required approximately two and a half hours.

We used custom-made software written in Python 3 (version 3.13.3) to collect the TDS and TL data. For the TDS task, participants assessed changes in their dominant sensations using a graphical interface with buttons labeled by sensory attributes. At any moment, participants pressed the button corresponding to the attribute they perceived as most dominant, switching buttons when their perception changed. Only one attribute could be selected at a time.

Twelve sensory attributes were used in the TDS tasks: aromatic, burned, buttery, dry, nutty, salty, sesame-like, smooth, sour, sweet, wet, and wheat-like. These attributes were selected by consensus among the authors and their three colleagues from a list compiled via web sources and other references [42–47]. The authors and three colleagues independently evaluated this list using a check-all-that-apply procedure to judge which attributes were applicable to the food samples. Attributes selected by at least four of the five evaluators were retained, yielding the final set of twelve. The same set of attributes was used for crackers, spreads, and their combinations. Definitions of these attributes are provided in Table 2.

**Table 2.** Definitions of attributes used in the TDS tasks.

Attribute	Definition
Aromatic	Complex but pleasant fragrant
Burned	Well-baked, toasty
Buttery	Butter-like flavor
Dry	Moisture-free, crispy
Nutty	Reminiscent of dried nuts
Salty	Salty (basic taste)
Sesame-like	Sesame-like flavor
Smooth	Smooth mouthfeel, no graininess texture
Sour	Sour (basic taste)
Sweet	Sweet (basic taste)
Wet	Moist, no dryness
Wheat-like	Wheat-like flavor

For the TL task, participants also used a graphical interface with nine buttons labeled 1 through 9. While tasting each food, participants indicated their liking in real time by pressing the button that best represented their current preference—selecting 9 for most favorable and 1 for least favorable.

#### 2.4. Dataset Generation with Bootstrap Resampling

The TDS task yielded TDS curves representing the temporal changes in dominance proportion for each sensory attribute. The dominance proportion for an attribute is defined as the proportion of trials in which that attribute was selected at each time point.

Similarly, the TL method produced TL curves, which show the temporal changes in the average liking score. Although the minimum possible score in the experiment was 1, the score was initialized to 0 at the beginning of each trial and remained 0 until the first button was pressed.

The TDS and TL tasks provide average time-series curves that represent the collective experiences of all participants. However, this results in a dataset that is insufficient for training machine learning or statistical models. To address this, we applied bootstrap resampling—a data augmentation technique—to the TDS and TL curves [48,49]. Initially, there were 16 pairs of TDS and TL curves, one pair for each participant. Bootstrap resampling was performed by randomly sampling 16 pairs with replacement from the original set of 16 panels, then averaging the sampled data to produce a single pair of TDS and TL curves. By repeating this process, we generated 100 curve pairs for each of the eight food samples.

### 3. Prediction of TL Curves Based on TDS Curves

#### 3.1. Reservoir Model Architecture

To predict TL curves, we employed a reservoir model [39–41], a type of recurrent neural network well suited for time-series data. The model outputs a TL curve when provided with TDS curves as input. For the implementation, ReservoirPy [50] (version 0.3.12), a Python library for reservoir computing, was used.

The architecture consisted of three layers: input, reservoir, and output. The input layer comprised 12 nodes, each corresponding to a sensory attribute. The output layer was a single node, representing the scalar liking score. The number of reservoir neurons was varied across four configurations: 64, 128, 192, and 256. Based on previous studies [8,25], we considered it reasonable to expect that the optimal number would fall within this range.

In addition to the 12-dimensional attribute input, we explored seven alternative configurations that incorporated auxiliary information. These auxiliary inputs were encoded as flag vectors, where each dimension was set to 1 if the corresponding condition was satisfied and 0 otherwise. The purpose of the auxiliary input flags was to provide contextual information—such as whether the sensory evaluation was for a cracker or spread alone, or a combination—in an effort to improve prediction performance. We compared various architectures, differing in both the number of auxiliary input nodes and the number of reservoir neurons, with respect to prediction accuracy.

Table 3 summarizes the total input dimensions, the number of auxiliary dimensions added, and provides a description of each configuration.

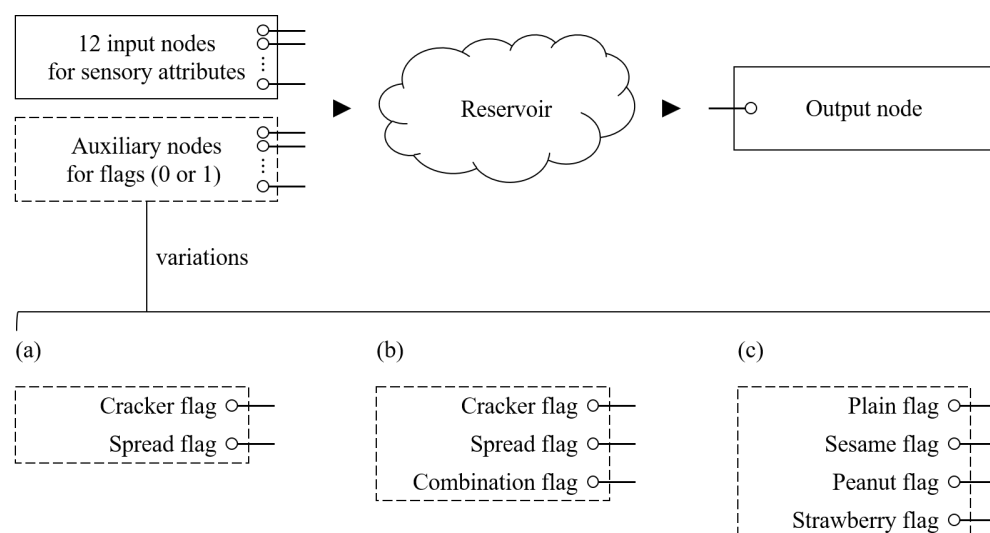
For example, as illustrated in Figure 2a, when cracker and spread flags were included, the input layer consisted of 14 nodes: 12 representing sensory attributes and 2 representing the flags. For single foods, only the corresponding flag was set to 1, while for paired samples, both flags were set to 1. For example, during the training phase, when the TDS curves for crackers were provided to the model, the cracker flag was set to 1 and the spread flag was set to 0.

**Table 3.** Auxiliary input dimensions and their descriptions.

Total Input Dimensions	Auxiliary Dimensions	Description
12	0	No auxiliary information (only TDS curves provided)
13	1	Combination flag (0: single item; 1: paired cracker and spread)
14	2	Cracker flag and spread flag (both set to 1 for paired samples)
15	3	Cracker, spread, and combination flags (exclusive, only one set to 1)
16	4	Brand flags (premium, sesame, peanut, strawberry)
17	5	Brand flags and combination flag
18	6	Brand flags, cracker flag, and spread flag
19	7	Brand flags, cracker, spread, and combination flags

As shown in Figure 2b, when cracker, spread, and combination flags were used, the input layer comprised 15 nodes. For paired foods, only the combination flag was set to 1, with the others set to 0.

In Figure 2c, when four brand flags were used, each flag indicated the presence of a specific ingredient (e.g., premium cracker, sesame, peanut, or strawberry). For instance, for a sample of premium cracker with strawberry jam, the premium and strawberry flags were set to 1, while sesame and peanut flags remained 0.



**Figure 2.** Structure of the input layer. (a) Two auxiliary nodes. (b) Three auxiliary nodes. (c) Four auxiliary nodes.

### 3.2. Training Dataset

The training data included TDS and TL curves collected from both single food items and pairings. The dataset for pairings was used in a cross-validation framework. For example, when predicting the TL curve for the plain–strawberry combination, the model was trained on the other seven food samples: plain, sesame, peanut, strawberry, plain–peanut, sesame–peanut, and sesame–strawberry. There were 100 resampled curve sets for each tasting condition; hence, each model was trained on a total of 700 curve sets.

### 3.3. Model Evaluation Metrics

The trained models were evaluated using a set of 100 curve pairs that were not included in the training data. For each curve set, the model predicted the TL curve by inputting the corresponding TDS curves, and the root mean squared error (RMSE) between the predicted and actual TL curves was calculated. This procedure was repeated for all

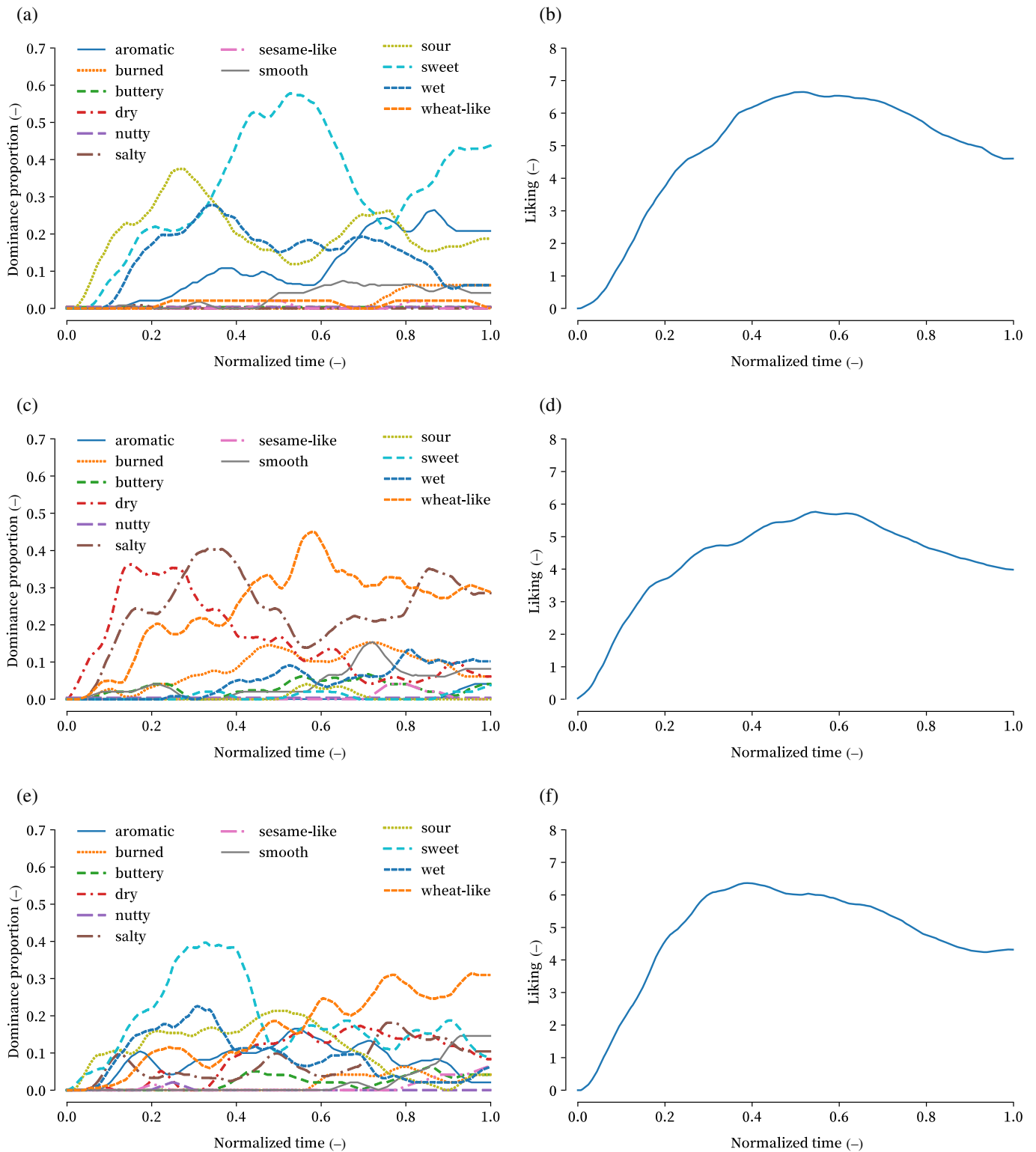
100 curve sets. Finally, the median and interquartile range (IQR) of the 100 RMSE values were calculated.

#### 4. Results

As an example, Figure 3 shows the TDS and TL curves for strawberry jam, plain cracker, and their pairing. For strawberry jam in (a), “sour” initially appeared as the dominant attribute, followed by “wet”. In the middle phase, approximately 60% of panels selected “sweet”. The jam received its highest liking scores during this middle phase, exceeding 6, as shown in (b). For the plain cracker in (c), “dry”, “salty”, and “wheat-like” attributes were each perceived as dominant at different times. The TL curve in (d) resembles that of the strawberry jam, although its maximum value was below 6. For the combination of strawberry jam and plain cracker in (e) and (f), a wider range of attributes was selected than for the individual items. Attributes characteristic of the jam, such as “sour”, “wet”, and “sweet”, appeared more frequently in the early phase, while cracker-related attributes like “dry” and “wheat-like” appeared later. In the curves, “salty” and “dry” were selected less frequently than in the cracker-alone condition. In the TL curve, the peak liking score was reached earlier than for the individual items, around 0.4 on the normalized time scale, with the maximum score exceeding 6.

The observed TL curves can be interpreted in light of the sensory attributes identified in the TDS tasks. For example, plain crackers were predominantly characterized by dry, salty, and wheat-like attributes, which corresponded with lower liking scores that did not exceed six points. In contrast, strawberry jam was described as sour, sweet, and wet, attributes that generally align with higher hedonic responses, and indeed its TL curve reached higher values. When the two foods were paired, the TDS curves revealed a complementary pattern: sour, sweet, and wet attributes dominated early in the tasting, while dry and wheat-like attributes appeared later. This temporal interplay likely contributed to the higher and earlier peak in the TL curve compared to crackers alone. Such results suggest that hedonic enhancement in pairings may arise from the combination of contrasting and complementary attributes, with crackers balancing the stickiness and sweetness of the jam and the jam mitigating the dryness of the crackers. Although these interpretations remain qualitative, they provide a plausible link between ingredient properties, sensory attributes, and temporal liking outcomes. Future studies involving a wider variety of food categories and ingredient processing methods may further clarify how intrinsic product properties shape both TDS profiles and hedonic dynamics.

The emergence of wheat-like and dry attributes in crackers may be linked to structural protein elements (such as gluten network) and processing steps like baking that affect water content, crumb/crust structure, and flavor release in cereal products. Links between protein aggregation, thermal processing, and sensory attributes in flour-based goods are discussed in recent reviews and studies [51–54]. For example, it was demonstrated how increasing protein additions change dough viscoelasticity and affect hardness and mouthfeel in crackers [53]. However, a detailed discussion of ingredient composition or processing effects is beyond the scope of this study, which primarily aims to develop and evaluate predictive models of temporal liking from TDS data.



**Figure 3.** Examples of TDS and TL curves. (a) TDS curves of strawberry jam. (b) TL curve of strawberry jam. (c) TDS curves of plain cracker. (d) TL curve of plain cracker. (e) TDS curves of plain-strawberry. (f) TL curve of plain-strawberry.

Table 4 presents the RMSEs of the 32 models examined. The ten models with the smallest (best) RMSEs and the ten with the largest (worst) RMSEs are listed in ascending order of RMSE. Ranks were determined based on the median RMSE values across the four food combinations, while the table also reports RMSEs for each individual combination.

The model performances varied substantially across the models. The top-ranked model achieved a mean RMSE of 0.44, with an IQR of 0.34–0.54. This model employed 128 neurons and a single auxiliary flag indicating whether the sample was a single food item or a cracker–spread pairing. In contrast, the lowest-ranked model, which used 64 neurons and five auxiliary flags, recorded a mean RMSE of 3.58.

**Table 4.** RMSEs for each model architecture. “Jam” refers to strawberry jam. Models are listed in ascending order of RMSE.

Rank	Neurons	Flags	Median RMSE (IQR)				
			All Pairs	Plain–Jam	Plain–Peanut	Sesame–Jam	Sesame–Peanut
1	128	1	0.44 (0.34–0.54)	0.41 (0.31–0.49)	0.46 (0.35–0.58)	0.42 (0.34–0.50)	0.51 (0.39–0.62)
2	192	2	0.44 (0.34–0.55)	0.39 (0.30–0.50)	0.42 (0.30–0.57)	0.42 (0.36–0.50)	0.52 (0.39–0.67)
3	192	0	0.46 (0.35–0.58)	0.42 (0.33–0.54)	0.40 (0.33–0.52)	0.45 (0.35–0.53)	0.59 (0.46–0.73)
4	64	1	0.46 (0.36–0.59)	0.44 (0.37–0.52)	0.48 (0.34–0.61)	0.42 (0.32–0.53)	0.59 (0.44–0.70)
5	192	3	0.46 (0.36–0.60)	0.46 (0.35–0.59)	0.49 (0.35–0.61)	0.41 (0.34–0.49)	0.57 (0.45–0.75)
6	256	1	0.47 (0.35–0.57)	0.44 (0.33–0.56)	0.46 (0.35–0.57)	0.45 (0.36–0.54)	0.51 (0.40–0.62)
7	192	1	0.47 (0.37–0.59)	0.44 (0.33–0.52)	0.47 (0.34–0.57)	0.46 (0.40–0.54)	0.57 (0.43–0.70)
8	256	0	0.47 (0.37–0.59)	0.44 (0.34–0.56)	0.50 (0.39–0.59)	0.43 (0.34–0.54)	0.50 (0.43–0.64)
9	128	2	0.48 (0.37–0.62)	0.60 (0.45–0.73)	0.47 (0.35–0.57)	0.41 (0.32–0.48)	0.52 (0.42–0.68)
10	256	3	0.48 (0.37–0.63)	0.41 (0.32–0.53)	0.44 (0.33–0.58)	0.47 (0.40–0.60)	0.62 (0.50–0.78)
	⋮				⋮		
23	192	6	1.49 (1.13–1.71)	1.66 (1.46–1.83)	1.47 (1.39–1.61)	0.88 (0.71–0.99)	1.71 (1.56–1.89)
24	256	7	1.68 (1.05–3.76)	2.55 (2.30–2.83)	1.08 (0.86–1.26)	1.00 (0.85–1.22)	4.65 (4.47–4.90)
25	256	5	1.91 (1.08–2.25)	1.60 (1.40–1.87)	2.53 (2.34–2.75)	2.10 (1.95–2.21)	0.68 (0.56–0.86)
26	128	7	2.38 (0.82–4.72)	5.60 (5.31–5.86)	0.60 (0.44–0.79)	3.92 (3.73–4.12)	1.01 (0.89–1.23)
27	192	5	2.38 (1.65–4.08)	1.80 (1.61–1.97)	1.55 (1.43–1.67)	3.20 (3.01–3.38)	5.55 (5.41–5.68)
28	192	4	2.58 (1.14–4.53)	1.59 (1.42–1.77)	6.58 (6.34–6.77)	3.73 (3.56–3.88)	0.77 (0.60–0.96)
29	64	7	2.71 (2.05–3.72)	2.23 (2.04–2.49)	2.95 (2.83–3.11)	1.95 (1.73–2.08)	4.30 (4.12–4.47)
30	256	6	3.01 (2.00–4.74)	6.33 (6.16–6.52)	3.60 (3.37–3.81)	1.65 (1.40–1.93)	2.50 (2.24–2.68)
31	192	7	3.42 (1.61–5.25)	1.96 (1.62–2.20)	5.07 (4.89–5.21)	5.51 (5.30–5.70)	1.39 (1.23–1.56)
32	64	5	3.58 (2.28–4.14)	4.61 (4.44–4.79)	3.38 (3.22–3.57)	0.58 (0.44–0.74)	3.81 (3.63–3.95)

## 5. Discussion

The median RMSE and IQR for the best-performing model were 0.44 and 0.34–0.54, respectively. Given that the liking score is based on a 9-point scale, this represents a small prediction error. In a previous study [25], the RMSE for predicting the liking of four types of coffee, when the model was trained only on data from the same brand, ranged from 0.46 to 0.59. In a study on strawberries [8], the RMSE ranged from 0.68 to 0.99. The top ten models used in the present study outperformed these previous results in terms of RMSE.

According to Table 4, the model with 128 neurons and a single auxiliary input flag exhibited the highest prediction accuracy. The top five models achieved comparable performance, with the number of auxiliary inputs ranging from 0 to 3. In contrast, models with four or more auxiliary input flags performed poorly, regardless of the number of neurons. A possible explanation is overfitting: when brand-level flags were provided, the model may have relied too heavily on this specific information, thereby improving the fit to the training data but diminishing generalization to unseen food items or pairings. By contrast, auxiliary inputs that only indicate broad categories (e.g., cracker, spread, or combination) appear to strike a better balance by providing useful context without overly constraining the model to brand-level information. These results suggest that the inclusion of brand flags negatively affected prediction accuracy, and that using at most three auxiliary input flags may be effective for improving estimation accuracy. However, since two of

the top ten models did not utilize any auxiliary input flags, the overall effect of auxiliary inputs may be limited. While the findings indicate that employing a very small number of auxiliary input flags is likely beneficial, it would be premature to draw firm conclusions based solely on this limited set of food samples (crackers and spreads).

For problems of the scale addressed in this study, a reservoir size of approximately 128 to 192 neurons appears to be appropriate. In general, increasing the number of neurons improves the model's capacity to address complex input–output relationships, but also increases the risk of overfitting [39,55]. Conversely, using too few neurons may result in an insufficient number of state variables to adequately explain the liking response. Although there are few studies investigating the relationship between problem or domain size and reservoir size for TL curve prediction, previous work has indicated a similar optimal range [8,25]. It should be noted, however, that only four neuron sizes (64, 128, 192, and 256) were tested in this study, and thus it cannot be excluded that optimal values might exist outside this range.

Considering these findings together, we were able to identify, to some extent, appropriate reservoir model configurations for food pairing prediction. The number of neurons should be in the range of approximately 128–196, and the use of auxiliary inputs, if any, should be limited to about one to three types. However, these characteristics may change if the number of food items to be processed increases.

Several approaches could be considered to further improve estimation accuracy. First, optimizing the number of participants is important. Although the optimal panel size for the TDS method itself has not been fully determined [56], the same holds true for the prediction of TL curves from TDS data. Additionally, it is not necessarily the case that binary flags are the only suitable auxiliary inputs; exploring flags with three levels or continuous values may also be beneficial. That said, as mentioned earlier, the top-performing models investigated in this study likely already achieve sufficiently high accuracy among approaches using reservoir computing. Looking ahead, it may be more important to devise strategies to prevent unexpectedly large estimation errors.

There are still some limitations and unresolved issues with the proposed method.

A limitation of this study is the restricted number of food items examined: two types of crackers and two types of spreads, yielding four pairings. While this limited scope constrains the generalizability of the findings to broader categories of food pairings, it was chosen as a practical compromise between experimental feasibility and participant burden. Expanding the number of food types would require alternative designs, such as incomplete block designs or reducing the number of pairings tested per participant. Future research should therefore investigate whether the proposed approach can be extended to a wider range of food items under such designs.

Another major limitation of the present approach concerns the normalization of time. The method assumes that all curves are temporally normalized. While this normalization increases the consistency of the analysis, it also results in the loss of information regarding individual differences, such as the rate and duration of flavor changes in the mouth. For example, the average tasting times for crackers and spreads were approximately 20 s and 10 s, respectively, but time normalization eliminates this information. At present, it is unclear how this normalization may have influenced the results.

We acknowledge that bootstrap resampling does not reproduce the true variability of an independent consumer population, as it relies on repeated sampling from the same limited panels. In this study, it was used solely as a data augmentation method to enable model training. Future work with larger and more diverse participant groups will be needed to validate the generalizability of the proposed approach.

In the current approach, training the model to predict the liking of food pairings requires data from all other pairings, which does not reduce the overall cost of sensory evaluation. Ideally, the model would generate the TL curve for a pairing solely from the TDS curves of the two individual foods. Furthermore, it would be advantageous if the training data consisted only of the TDS and TL curves for single items and a limited number of key pairings. For instance, with 10 types of crackers and 10 types of spreads, training the model using only the single-item data (20 foods) and a few critical pairings would substantially reduce the sensory evaluation workload. Although developing such a method presents considerable challenges, we aim to pursue this direction in future work.

While the present study does not directly quantify reductions in sensory evaluation costs, our approach is motivated by the industrial reality that sensory testing is widely conducted and often resource intensive. A recent survey [29] reported that sensory evaluations are used in nearly 70% of food development cases, frequently involving large consumer groups and expert panels, and that cost was among the most cited reasons for omitting expert evaluations. Against this backdrop, the present findings represent an initial methodological step toward predictive approaches that may, in the long term, contribute to reducing the burden of sensory evaluation in food development.

## 6. Conclusions

This study demonstrated the feasibility of predicting TL curves for food pairings based on TDS data using reservoir computing. By collecting TDS and TL data for two types of crackers, two types of spreads, and their combinations, we developed and evaluated several reservoir computing models with varying numbers of neurons and auxiliary input flags. The most accurate models achieved a median RMSE of 0.44 on a 9-point liking scale, which is a notably small prediction error when compared to previous research on temporal liking prediction. Models with a moderate number of neurons (128–192) and a limited number of auxiliary inputs (up to three) performed best, while models with excessive auxiliary input flags tended to perform poorly.

These results suggest that reservoir computing can provide robust predictions of hedonic dynamics in food pairings, with potential to reduce the scale and cost of sensory evaluations. Nevertheless, key limitations remain: the present models require data from all food pairings during training, which currently limits their practical benefit in reducing sensory evaluation workload. Furthermore, the use of temporal normalization, while enhancing analytical consistency, results in the loss of potentially informative dynamics.

Future research should address these limitations by developing methods that can generalize from data on single foods and a limited set of key pairings, ideally enabling the prediction of TL curves for any pairing from the TDS curves of its components. Despite these challenges, the present findings constitute an important step toward more efficient, data-driven approaches for understanding and predicting food-pairing preferences.

**Author Contributions:** Conceptualization, H.N. and S.O.; methodology, H.N. and S.O.; software, H.N.; validation, H.N. and S.O.; formal analysis, H.N. and S.O.; investigation, H.N. and S.O.; resources, H.N.; data curation, H.N.; writing—original draft preparation, H.N. and S.O.; writing—review and editing, H.N. and S.O.; visualization, H.N.; supervision, S.O.; project administration, S.O.; funding acquisition, S.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Urakami Foundation for Food and Food Culture Promotion (2024-16).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Hino Campus, Tokyo Metropolitan University (Approval Number: R7-076).

**Informed Consent Statement:** Written informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** The research data are available from the corresponding author upon direct request, accompanied by a clear explanation of the intended purpose.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

TDS Temporal dominance of sensations  
TL Temporal liking

## References

1. Pineau, N.; Schlich, P.; Cordelle, S.; Mathonnière, C.; Issanchou, S.; Imbert, A.; Rogeaux, M.; Etiévant, P.; Köster, E. Temporal Dominance of Sensations: Construction of the TDS curves and comparison with time-intensity. *Food Qual. Prefer.* **2009**, *20*, 450–455. [\[CrossRef\]](#)
2. ISO 13299; Sensory Analysis—Methodology—General Guidance for Establishing a Sensory Profile. International Organization for Standardization: Geneva, Switzerland, 2016.
3. Schlich, P. Temporal Dominance of Sensations (TDS): A new deal for temporal sensory analysis. *Curr. Opin. Food Sci.* **2017**, *15*, 38–42. [\[CrossRef\]](#)
4. Taylor, D.E.; Pangborn, R.M. Temporal aspects of hedonic responses. *J. Sens. Stud.* **1990**, *4*, 241–247. [\[CrossRef\]](#)
5. Thomas, A.; Visalli, M.; Cordelle, S.; Schlich, P. Temporal Drivers of Liking. *Food Qual. Prefer.* **2015**, *40*, 365–375. [\[CrossRef\]](#)
6. Castura, J.C.; Antúnez, L.; Giménez, A.; Ares, G. Temporal Check-All-That-Apply (TCATA): A novel dynamic method for characterizing products. *Food Qual. Prefer.* **2016**, *47*, 79–90. [\[CrossRef\]](#)
7. Meyners, M. Temporal liking and CATA analysis of TDS data on flavored fresh cheese. *Food Qual. Prefer.* **2016**, *47*, 101–108. [\[CrossRef\]](#)
8. Natsume, H.; Okamoto, S. Prediction of Temporal Liking from Temporal Dominance of Sensations by using reservoir computing and its sensitivity analysis. *Foods* **2024**, *13*, 3755. [\[CrossRef\]](#)
9. Nguyen, Q.C.; Varela, P. Identifying temporal drivers of liking and satiation based on temporal sensory descriptions and consumer ratings. *Food Qual. Prefer.* **2021**, *89*, 104143. [\[CrossRef\]](#)
10. Thomas, A.; Chambault, M.; Dreyfuss, L.; Gilbert, C.; Hegyi, A.; Henneberg, S.; Knippertz, A.; Kostyra, E.; Kremer, S.; Silva, A.; et al. Measuring temporal liking simultaneously to Temporal Dominance of Sensations in several intakes. An application to Gouda cheeses in 6 European countries. *Food Res. Int.* **2017**, *99*, 426–434. [\[CrossRef\]](#)
11. Thomas Carr, B.; Lesniasukas, R.O. Analysis of variance for identifying temporal drivers of liking. *Food Qual. Prefer.* **2016**, *47*, 97–100. [\[CrossRef\]](#)
12. Ares, G.; Alcaire, F.; Antúnez, L.; Vidal, L.; Giménez, A.; Castura, J.C. Identification of drivers of (dis)liking based on dynamic sensory profiles: Comparison of Temporal Dominance of Sensations and Temporal Check-all-that-apply. *Food Res. Int.* **2017**, *92*, 79–87. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Silva, A.P.; Voss, H.P.; van Zyl, H.; Hogg, T.; de Graaf, C.; Pintado, M.; Jager, G. Temporal dominance of sensations, emotions, and temporal liking measured in a bar for two similar wines using a multi-sip approach. *J. Sens. Stud.* **2018**, *33*, e12459. [\[CrossRef\]](#)
14. van Bommel, R.; Stieger, M.; Boelee, N.; Schlich, P.; Jager, G. From first to last bite: Temporal dynamics of sensory and hedonic perceptions using a multiple-intake approach. *Food Qual. Prefer.* **2019**, *78*, 103748. [\[CrossRef\]](#)
15. Okamoto, S.; Natsume, H.; Watanabe, H. Synergy analysis between the Temporal Dominance of Sensations and Temporal Liking curves of strawberries. *Foods* **2025**, *14*, 992. [\[CrossRef\]](#)
16. Kuesten, C.; Bi, J. Temporal Drivers of liking based on Functional Data Analysis and non-additive models for multi-attribute Time-Intensity data of fruit chews. *Foods* **2018**, *7*, 84. [\[CrossRef\]](#)
17. Okada, T.; Okamoto, S.; Yamada, Y. Affective Dynamics: Causality Modeling of Temporally Evolving Perceptual and Affective Responses. *IEEE Trans. Affect. Comput.* **2022**, *13*, 628–639. [\[CrossRef\]](#)
18. Tachi, K.; Okamoto, S. Dynamic state-space modeling with factorial memories in temporal dominance of sensations, emotions and temporal liking. *IEEE Trans. Affect. Comput.* **2023**, *14*, 2266–2275. [\[CrossRef\]](#)

19. Galmarini, M.; Dufau, L.; Loiseau, A.L.; Visalli, M.; Schlich, P. Wine and cheese: Two products or one association? A new method for assessing wine-cheese pairing. *Beverages* **2018**, *4*, 13. [[CrossRef](#)]
20. Lawless, H.T.; Heymann, H. *Sensory Evaluation of Food: Principles and Practices*; Food Science Text Series; Springer: New York, NY, USA, 2010. [[CrossRef](#)]
21. Drake, M.; Watson, M.; Liu, Y. Sensory analysis and consumer preference: Best practices. *Annu. Rev. Food Sci. Technol.* **2023**, *14*, 427–448. [[CrossRef](#)]
22. Miller, R. Drivers of consumer liking for beef, pork, and lamb: A review. *Foods* **2020**, *9*, 428. [[CrossRef](#)]
23. García-Barón, S.E.; Carmona-Escutia, R.P.; Herrera-López, E.J.; Leyva-Trinidad, D.A.; Gschaedler-Mathis, A. Consumers' drivers of perception and preference of fermented food products and beverages: A systematic review. *Foods* **2025**, *14*, 713. [[CrossRef](#)]
24. De Pelsmaeker, S.; Schouteten, J.J.; Lagast, S.; Dewettinck, K.; Gellynck, X. Is taste the key driver for consumer preference? A conjoint analysis study. *Food Qual. Prefer.* **2017**, *62*, 323–331. [[CrossRef](#)]
25. Natsume, H.; Okamoto, S. Cross-brand machine learning of coffee's Temporal Liking from Temporal Dominance of Sensations curves. *Appl. Sci.* **2025**, *15*, 948. [[CrossRef](#)]
26. Delarue, J.; Lawlor, J.B.; Rogeaux, M. (Eds.) *Rapid Sensory Profiling Techniques*; Woodhead Publishing: Cambridge, UK, 2015. [[CrossRef](#)]
27. Varela, P.; Ares, G. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Res. Int.* **2012**, *48*, 893–908. [[CrossRef](#)]
28. Torrico, D.D. The potential use of ChatGPT as a sensory evaluator of chocolate brownies: A brief case study. *Foods* **2025**, *14*, 464. [[CrossRef](#)] [[PubMed](#)]
29. Świąder, K.; Marczewska, M. Trends of using sensory evaluation in new product development in the food industry in countries that belong to the EIT regional innovation scheme. *Foods* **2021**, *10*, 446. [[CrossRef](#)] [[PubMed](#)]
30. Eschevins, A.; Giboreau, A.; Julien, P.; Dacremont, C. From expert knowledge and sensory science to a general model of food and beverage pairing with wine and beer. *Int. J. Gastron. Food Sci.* **2019**, *17*, 100144. [[CrossRef](#)]
31. Harrington, R.J. The wine and food pairing process: Using culinary and sensory perspectives. *J. Culin. Sci. Technol.* **2005**, *4*, 101–112. [[CrossRef](#)]
32. Rune, C.J.B.; Münchow, M.; Perez-Cueto, F.J.A. Systematic review of methods used for food pairing with coffee, tea, wine, and beer. *Beverages* **2021**, *7*, 40. [[CrossRef](#)]
33. Kustos, M.; Heymann, H.; Jeffery, D.W.; Goodman, S.; Bastian, S.E. Intertwined: What makes food and wine pairings appropriate? *Food Res. Int.* **2020**, *136*, 109463. [[CrossRef](#)]
34. Tseng, Y.J.; Chuang, P.J.; Appell, M. When machine learning and deep learning come to the big data in food chemistry. *ACS Omega* **2023**, *8*, 15854–15864. [[CrossRef](#)] [[PubMed](#)]
35. Ahn, Y.Y.; Ahnert, S.E.; Bagrow, J.P.; Barabási, A.L. Flavor network and the principles of food pairing. *Sci. Rep.* **2011**, *1*, 196. [[CrossRef](#)]
36. Park, D.; Kim, K.; Park, Y.; Shin, J.; Kang, J. KitcheNette: Predicting and ranking food ingredient pairings using Siamese Neural Network. In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization IJCAI-2019, Macao, China, 10–16 August 2019; pp. 5930–5936. [[CrossRef](#)]
37. Park, D.; Kim, K.; Kim, S.; Spranger, M.; Kang, J. FlavorGraph: A large-scale food-chemical graph for generating food representations and recommending food pairings. *Sci. Rep.* **2021**, *11*, 931. [[CrossRef](#)] [[PubMed](#)]
38. Visalli, M.; Galmarini, M.V. Multi-attribute temporal descriptive methods in sensory analysis applied in food science: A systematic scoping review. *Compr. Rev. Food Sci. Food Saf.* **2024**, *23*, e13294. [[CrossRef](#)] [[PubMed](#)]
39. Verstraeten, D.; Schrauwen, B.; D'Haene, M.; Stroobandt, D. An experimental unification of reservoir computing methods. *Neural Netw.* **2007**, *20*, 391–403. [[CrossRef](#)]
40. Tanaka, G.; Yamane, T.; Héroux, J.B.; Nakane, R.; Kanazawa, N.; Takeda, S.; Numata, H.; Nakano, D.; Hirose, A. Recent advances in physical reservoir computing: A review. *Neural Netw.* **2019**, *115*, 100–123. [[CrossRef](#)]
41. Abbas, A.H.; Abdel-Ghani, H.; Maksymov, I.S. Classical and quantum physical Reservoir Computing for onboard artificial intelligence systems: A perspective. *Dynamics* **2024**, *4*, 643–670. [[CrossRef](#)]
42. Brar, J. Sensory and Consumer Evaluation of Commercially Available Gluten-Free Crackers. Master's Thesis, College of Health and Human Sciences, Kansas State University, Manhattan, KS, USA, 2024.
43. Ikegaya, A.; Toyozumi, T.; Kosugi, T.; Arai, E. Taste and palatability of strawberry jam as affected by organic acid content. *Int. J. Food Prop.* **2020**, *23*, 2087–2096. [[CrossRef](#)]
44. Oliver, P.; Cicerale, S.; Pang, E.; Keast, R. Developing a strawberry lexicon to describe cultivars at two maturation stages. *J. Sens. Stud.* **2018**, *33*, e12312. [[CrossRef](#)]
45. Teixeira, F.; Santos, B.A.D.; Nunes, G.; Soares, J.M.; Amaral, L.A.D.; Souza, G.H.O.D.; Resende, J.T.V.D.; Menegassi, B.; Rafacho, B.P.M.; Schwarz, K.; et al. Addition of orange peel in orange jam: Evaluation of sensory, physicochemical, and nutritional characteristics. *Molecules* **2020**, *25*, 1670. [[CrossRef](#)]

46. Kurotobi, T.; Hoshino, T.; Kazami, Y.; Hayakawa, F.; Hagura, Y. Relationship between sensory analysis for texture and instrument measurements in model strawberry jam. *J. Texture Stud.* **49**, 359–369. [[CrossRef](#)]
47. Sithole, T.R.; Ma, Y.X.; Qin, Z.; Liu, H.M.; Wang, X.D. Influence of peanut varieties on the sensory quality of peanut butter. *Foods* **2022**, *11*, 3499. [[CrossRef](#)]
48. Okamoto, S. Bootstrap Resampling of Temporal Dominance of Sensations Curves to Compute Uncertainties. *Foods* **2021**, *10*, 2472. [[CrossRef](#)] [[PubMed](#)]
49. Castura, J.; Rutledge, D.; Ross, C.; Næs, T. Discriminability and uncertainty in principal component analysis (PCA) of temporal check-all-that-apply (TCATA) data. *Food Qual. Prefer.* **2022**, *96*, 104370. [[CrossRef](#)]
50. Trouvain, N.; Pedrelli, L.; Dinh, T.T.; Hinaut, X. ReservoirPy: An efficient and user-friendly library to design Echo State Networks. In Proceedings of the Artificial Neural Networks and Machine Learning–ICANN 2020, Bratislava, Slovakia, 15–18 September 2020; pp. 494–505. [[CrossRef](#)]
51. Lund, M.N.; Ray, C.A. Control of maillard reactions in foods: Strategies and chemical mechanisms. *J. Agric. Food Chem.* **2017**, *65*, 4537–4552. [[CrossRef](#)] [[PubMed](#)]
52. Grafenauer, S.; Miglioretto, C.; Solah, V.; Curtain, F. Review of the sensory and physico-chemical properties of red and white wheat: Which makes the best whole grain? *Foods* **2020**, *9*, 136. [[CrossRef](#)]
53. En Tay, R.R.; Agatha, T.; Somang, G.; Yulianti, O.; Lin Tan, E.L. Structuring wheat flour-based crackers using whey protein isolate. *Int. Dairy J.* **2022**, *128*, 105314. [[CrossRef](#)]
54. Madhavan, S.; Sreekantaiah, H.M. Influence of high protein on the rheological properties, microstructure and X-ray diffraction of crackers formulations. *Curr. Res. Nutr. Food Sci. J.* **2025**, *13*, 132–144. [[CrossRef](#)]
55. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)]
56. Visalli, M.; Galmarini, M.V.; Schlich, P. Use of temporal sensory evaluation methods with consumers: A position paper. *Curr. Opin. Food Sci.* **2023**, *54*, 101102. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.