MDPI

You Look like You'll Buy It! Purchase Intent Prediction Based on Facially Detected Emotions in Social Media Campaigns for Food Products

Katerina Tzafilkou *, Anastasios A. Economides [ID] and Foteini-Rafailia Panavou

SMILE Lab, University of Macedonia, 54006 Thessaloniki, Greece; economid@uom.edu.gr (A.A.E.); fotini138@hotmail.com (F.-R.P.)
* Correspondence: tzafilkou@uom.edu.gr

Abstract: Understanding the online behavior and purchase intent of online consumers in social media can bring significant benefits to the ecommerce business and consumer research community. Despite the tight links between consumer emotions and purchase decisions, previous studies focused primarily on predicting purchase intent through web analytics and sales historical data. Here, the use of facially expressed emotions is suggested to infer the purchase intent of online consumers while watching social media video campaigns for food products (yogurt and nut butters). A FaceReader Online™ multi-stage experiment was set, collecting data from 154 valid sessions of 74 participants. A set of different classification models was deployed, and the performance evaluation metrics were compared. The models included Neural Networks (NNs), Logistic Regression (LR), Decision Trees (DTs), Random Forest (RF,) and Support Vector Machine (SVM). The NNs proved highly accurate (90–91%) in predicting the consumers' intention to buy or try the product, while RF showed promising results (75%). The expressions of sadness and surprise indicated the highest levels of relative importance in RF and DTs correspondingly. Despite the low activation scores in arousal, micro expressions of emotions proved to be sufficient input in predicting purchase intent based on instances of facially decoded emotions.

Keywords: consumer emotions; emotional artificial intelligence; face tracking; FaceReader Online™; prediction of purchase intent; social media food marketing

1. Introduction

Predicting customer purchase intent has been considered the holy grail of e-commerce business and digital marketing research. Several studies attempted to predict purchase intent through various models, mainly by utilizing the customer's history purchasing behavior [1], sales rank [2], or other web analytics data [3]. Previous studies applied neuroscience techniques, such as real-time EEG [4], others focused on measuring the platform's usability and engagement attributes [5–8], while others attempted to predict purchase intent through posts and user-generated content on social media [9].

Despite the close links between consumer emotions and purchase decisions, the field of measuring emotions in digital and social media campaigns to predict purchase behaviors is under-researched. Measuring emotions can be highly beneficial in the field of food market research since psychological and social factors can elicit food-related consumption behaviors. Emotions are also crucial determinants of food preferences and acceptability [10]. Hedonic foods (e.g., ice cream) primarily serve to provide sensory pleasure, whereas utilitarian foods (e.g., vegetables, bread) typically have more instrumental value [11]. Since exposure to food advertisements can influence a viewer's food choices, this associates food advertisements with food consumption decisions [11]. Face tracking is a popular methodology for emotion detection in the field because facial expressions are essential to the expression of emotions. Researchers agree that changes in facial expressions can

lead to new objective measures of the affective responses to foods, since they frequently accompany internal feelings [10]. Recent research has shown that most self-reported sensory and hedonic responses cannot efficiently capture a complete consumer response to foods because many responses are unconscious and/or physiological [12]. Facial expressions might provide additional information on fast-changing emotions during food consumption. Several previous studies examined food-evoked emotions or preferences through explicit dynamic approaches (e.g., Temporal Dominance of Emotions) and implicit face-tracking methodologies [13–18]. However, most of these studies included sensory attributes in physical human–product experiences (e.g., while tasting the product) and have not been conducted in the context of digital campaigns. In a face-tracking study, Mahieu et al. [19] confirmed that strong implicit emotions can be elicited in exposure to food advertisements, similarly to sensory-based/embodied experience. In the context of graphic-style elicited emotions, Yu and Ko [20] examined the FaceReader emotions that can be recognized while processing different stimuli of multimedia content such as static images. The findings revealed significant emotional differences between different graphic styles (e.g., colors) and content types. McDuff et al. [21] conducted a large-scale experiment using Affectiva face-tracking software, proving that ad liking and purchase intent can be predicted with quite high accuracy. However, their examined ad context included a wide range of markets and product categories and was not focused on food products. Several other studies examined the associations between ad preferences and facial expressions in either Web content [22] or product label features [23], utilizing a broad range of ads and market products. Recently, Tzafilkou et al. [24] showed that FacerReader Online™ can efficiently capture the differences in emotional responses elicited by different types of food and media posts in social media campaigns. Although there have been studies in each stream, a concerted effort to link and examine facially expressed emotions in relation to the desirability of the food product and purchase intention is scant. Moreover, to the best of our knowledge, there is no other attempt to predict purchase intent through face-tracking emotional intelligence data on food-related YouTube campaigns. This study addresses the above gap by utilizing the consumers' facially detected emotions through a face-tracking Artificial Intelligence (AI) tool to predict their intent to buy or taste the promoted food product.

To this end, the main research objective of the current study is to explore whether facially expressed emotions can reveal consumers' intention to buy or taste (try) the promoted utilitarian food products, while watching food-related video campaigns displayed on social media channels, such as YouTube, Facebook, and Instagram. The study focused on healthy utilitarian products since they are considered less likely to provoke strong food-elicited emotions compared to unhealthy and hedonic products [10,11]. The study also examines and compares a set of Neural Networks (NNs) and machine learning (ML) classification models. Finally, the relative importance of the variables is calculated and discussed.

The results are expected to contribute to the design and deployment of AI models to automatically predict consumer attitudes and intents through their facially decoded emotional states, even in stimuli that provoke low emotional responses. Researchers and marketers can utilize the models to examine consumer behavior and explore the associations between emotions and food product selection in social media image posts and videos more deeply. The findings might also be useful to campaign designers to efficiently measure the success of their creatives in the food market.

This paper is structured as follows: In the second section, the methodology is described, presenting all experimental design and data collection/processing steps. It also presents the classification methods that were deployed. Then, the results are presented, as regards the descriptive statistics and the performance metric scores of the classification models. The fourth section discusses the main findings of the study, presents the research and practical implications, and outlines the main limitations. The last section summarizes the conclusions of the study.

2. Materials and Methods

This section describes all the methodological steps of the research, including the selection of the content stimuli and the experimental design, the participants and procedure of the experimental tasks, the data collection/processing, and the classification models that were deployed.

2.1. Content Stimuli and Experimental Design

A multi-stage approach was applied to collect the required dataset through FaceReader Online™ and self-reported responses. Four different experimental tasks were designed and allocated to the audience. Commercially available products of branded milk yogurt, almond yogurt, and nut butters (hazelnuts, almond, sesame, etc.), were used as stimuli.

The first experiment included a YouTube video ad of a regular Greek yogurt, while the second one included a YouTube video ad of an almond yogurt and related almond milk products of the same brand. The third experiment included a set of Facebook image posts on a branded nut butter product, presenting only the product and its package. The fourth experiment presented the same nut butter, highlighting its use in a hedonic approach, e.g., through combinations with hedonic food (muffins and pancakes). Due to the FaceReader Online™ restriction to upload stimuli in image formats, the tasks in experiments 3 and 4 were conducted through a video slideshow presenting the selected social media image posts. The characteristics of the selected products and media stimuli are presented in Table 1.

Table 1. Product and media characteristics.

Experiment#	Stimulus Characteristics		
	Product	**Media Type**	**Duration**
	Product Type: Yoghurt | Stimuli Type: Video Ads		
Experiment #1	Milk Yogurt	Video (YouTube)	15 s
Experiment #2	Almond Yogurt	Video (YouTube)	17 s
	Product Type: Nut butters | Stimuli Type: Image Posts		
Experiment #3	Nut butters, showing only the product	Video Slideshow (Facebook and Instagram image posts)	15 s
Experiment #4	Nut butters, combined with hedonic food	Video Slideshow (Facebook and Instagram image posts)	16 s

2.2. Participants and Procedure

Seventy-six participants successfully completed at least one of the four experimental tasks within a period of 60 days. There was an approximately equal gender distribution in the sample (51% female, 49% male). The majority (64.5%) was in the 25–34 age group, 18.4% were in the 35–50 age group, 14.5% were between 18 and 24 years old, while 2.6% were above 51 years old. Only 5 participants reported food allergies, 2 of which concerned dairy products; hence, these participants were excluded from the yogurt experimental data. None of the participants reported being vegan, and 14.5% of the participants reported that they do not consume sugar.

An online invitation page including detailed instructions was distributed to all participants through email, describing the experimental steps and providing the corresponding FaceReader Online™ and self-reported questionnaire links. Participants completed each experiment separately on their desktop or laptop devices. After watching each stimulus, the participants were asked to reply to a self-reported questionnaire about their intention to buy the product. The questionnaire was based on a single-item five-point Likert scale: I intend to buy or taste the product (1 = Extremely Disagree; . . . ; 5 = Extremely Agree). Single-item questionnaires are acceptable regarding reliability of the model and can be

applied for particular and non-complex constructs that can be clearly and homogeneously perceived, as applied in other studies [25,26].

All participants provided their formal consent to be facially tracked by FaceReader Online™ for emotion detection purposes. The face-tracking and self-reported data collection methods were approved by the University's Ethical Committee.

2.3. Data Collection and Processing

The FaceReader Online™ software detects the six basic emotions of Ekman [27] (happy, sad, angry, surprised, scared, disgusted), a neutral state (neutral), as well as valence and arousal [28]. The emotional features are calculated in a range between 0 and 1, where a value closer to 1 denotes stronger expression of the emotion. Valence indicates whether the emotional state of the subject is positive or negative, and its value ranges between −1 and 1. Arousal indicates whether the test participant is active (+1) or not active (0). Studies have indicated that the software is an efficient tool for analyzing emotions, with an accuracy rate of 90% [24].

Recording and analysis were completed successfully in 203 sessions. All analyzed results were exported by FaceReader Online™ in csv format to be cleaned and further processed in Python scripts. After removing all records of low recording quality (less than 4.0/10.0), there were 154 valid records composing the final dataset (Experiment 1: 41 records, Experiment 2:36 records, Experiment 3:38 records, Experiment 4: 39 records). The quality score indicates the quality of the total recording, ranging from 0 to 10, where 0 indicates very poor quality and 10 high quality. It is based on the analysis quality and the amount of successfully analyzed frames. FaceReader Online™ suggests the exclusion of recordings that achieved a quality score below 4.0. The final records of the analyzed dataset achieved an average recording quality score of 9.0. Finally, no significant levels of outliers were observed in the generated boxplots.

The classification model input variables were the FaceReader™-detected average emotions [0, 1] from the recorded sessions. The outcome variable was the binary category (Low, High) of the self-reported 'intention to buy or taste the promoted product'. The five choice responses were collected through the Likert scale and were divided into two classes (0 = "Low" and 1 = "High"). This approach was followed to slightly increase the limited amount of data for each class to be trained [29,30]. Thus, each classifier would have two outputs, representing low (1, 2, 3) and high (4, 5) intention, respectively. In the sample dataset, there was an equal representation (almost 50%) of the responses in the two classes.

Table 2 summarizes the features that were processed and applied in the classification models. It also describes Ekman's [27] Facial Action Coding System's (FACS) muscular action units that are used by FaceReader Online™ to detect every emotional state. FACS is a comprehensive, anatomically based system for describing all visually discernible facial movement and breaks down facial expressions into individual components of muscle movements, called Action Units (AUs) (https://www.paulekman.com/facial-action-coding-system/ accessed on 10 March 2023).

2.4. Classification

Six machine learning techniques were applied for the classification of the collected data, including a simple sequential layered Neural Network (NN), a Multi-layer Perceptron classifier connecting to a NN (MLP NN), Support Vector Machine (SVM), Logistic Regression (LR), Decision Trees (DTs), and Random Forest (RF), and 20–80 percent of training and testing datasets were applied to all classification models to avoid overfitting. Only in RF, a 15–85 analogy was applied to achieve the highest accuracy. Since the dataset could be augmented, no other overfitting avoidance methods were applied in the datasets. The models were executed through the Python language, utilizing the Keras API (Tensorflow) for the NN models and the scikit-learn library for the other ML models of classification. A description of the main characteristics of the deployed models and a justification for their selection are presented below.

Table 2. Summary of extracted features.

Feature	Value Range-Data Type	AUs Description *
Neutral	[0, 1]-154 non-null float	Inactivity of facial muscles
Happy	[0, 1]-154 non-null float	Cheek Raiser, Lip Corner Puller
Sad	[0, 1]-154 non-null float	Inner Brow Raiser, Brow Lowerer, Lip Corner Depressor
Angry	[0, 1]-154 non-null float	Brow Lowerer, Upper Lid Raiser, Lid Tightener, Lip Tightener
Scared	[0, 1]-154 non-null float	Inner Brow Raiser, Outer Brow Raiser, Brow Lowerer, Upper Lid Raiser, Lid Tightener, Lip Stretcher, Jaw Drop
Surprised	[0, 1]-154 non-null float	Inner Brow Raiser, Outer Brow Raiser, Upper Lid Raiser, Jaw Drop
Disgusted	[0, 1]-154 non-null float	Nose Wrinkler, Lip Corner Depressor, Lower Lip Depressor
Valence	[−1, 1]-154 non-null float	It is calculated as the intensity of "happy" minus the intensity of the negative expression with the highest intensity
Arousal	[0, 1]-154 non-null binary	It is based on the activation of 20 Action Units (AUs) of the Facial Action Coding System (FACS)
Class	154 non-null binary	The binary class of perceived intention to buy or taste the promoted food product ("0 = Low", "1 = High")

Note. RangeIndex: 154 entries, 0 to 153. Data columns (total 10 columns). * The description of the emotions is according to the Facial Action Coding System, which is universally applied to [27,28].

Neural Networks (NNs): Neural Networks can perform several regression and classification tasks. NN efficiency is primarily based on the weights of the net to be trained. The weights are initially set to random values, and then instances of the training set are repeatedly generated. The values for the input of an instance are placed on the input units, and the output of the net is compared with the desired output for this instance [31]. NNs were selected in this study because of their high-performance scores in accuracy in relevant studies [8]. In this study, a simple NN was deployed based on a Keras sequential model composed of successive layers.

In a multi-layer perceptron (MLP) classifier, the weights of the network are found by solving a quadratic programming problem with linear constraints. In this study, the deployed MLP NN classifier applied 4 hidden layers, and the learning rate was set to 0.01. The alpha parameter was regulated at 1×10^5 to resolve overfitting.

Support Vector Machine (SVM): Support Vector Machine is an effective supervised learning algorithm, which is used for classification and regression problems, for both linear and nonlinear data. SVM was selected since it is an effective classifier. The idea of SVM is that the algorithm creates a hyper plane, which segregates the data points into classes, ensuring that the separation between the classes is maximum [32]. This study deployed a linear classifier, predicting each input's member class between two possible classifications.

Logistic Regression (LR): Logistic Regression is the iterative display of the most powerful linear combination of variables most likely to determine the observed outcome. LR aids in determining the likelihood that a new instance belongs to a specific class. Since it is a probability, the result is between 0 and 1. To employ the LR as a binary classifier, a threshold must be specified to distinguish between two classes. The LR model can be used to model a categorical variable with more than two values. Studying the relationships between a set of labeled data, data can be categorized into discrete classes. Logistic regression is one of the most frequently used methods in statistics and discrete data analysis [31].

Decision Trees (DTs): The Decision Tree algorithm is one of the most frequently and widely used supervised machine learning algorithms that can be used for both classification and regression tasks. Decision Trees are trees that classify instances by sorting them based on feature values. Each node in a decision tree represents a feature in an instance to be classified. Decision tree classifiers usually employ post-pruning techniques that evaluate

the performance of decision trees, as they are pruned by using a validation set [30]. In the deployed DT model, the criterion Gini index was applied to create split points. The accuracy scores of the training and testing datasets were comparable, implying the absence of overfitting.

Random Forest (RF): Random Forest is an ensemble learning method, which constructs multiple decision trees through different data subsets, voting on the results of multiple decision trees to obtain the prediction as output of the model. Random Forest algorithm can be used to solve both classification and regression problems. It is also considered to be a very accurate and robust model because it uses many decision trees to make predictions. Random Forest was chosen since the default hyperparameters of RF provide great results, and it is incredible at avoiding overfitting [33]. The Random Forest classifier was applied from the sklearn Python library, and random search was applied to narrow down the range for each hyperparameter.

3. Results

This section presents the results of the study, categorized in two subsections: (i) the results of descriptive statistics and (ii) the performance results of the classification models.

3.1. Descriptive Statistics

Table 3 presents the descriptive statistics for the FaceReader Online™-detected emotional states (ranging from 0 to 1) and the 5-level self-reported scores of the variable "intention to buy or taste". As depicted, neutrality reported significantly higher values, while sadness was the most prevailing emotion detected by FaceReader Online™, followed by anger. Anger was the second-most-intensive emotion (after sadness), followed by happiness, disgust, and surprise. Fear (scared) revealed the lowest scores of detections in the participants' facial expressions. The perceived intention to buy or taste the product was almost 3.5/5.0, and no significant differences emerged between the different products ($p < 0.05$, Kruskal–Wallis Test).

Table 3. Descriptive statistics of input and output data.

	Minimum	Maximum	Mean	Std. Error	Std. Deviation
Input data * (FaceReader Online ™ emotion values)					
Neutral	0.1529	0.9888	0.717859	0.0159904	0.1984357
Happy	0.0000	0.4690	0.035942	0.0066883	0.0829995
Sad	0.0006	0.9294	0.103791	0.0107723	0.1336808
Angry	0.0001	0.7478	0.075247	0.0094886	0.1177505
Surprised	0.0000	0.2900	0.026289	0.0034500	0.0428128
Scared	0.0000	0.1363	0.016016	0.0017813	0.0221054
Disgusted	0.0002	0.4042	0.026873	0.0041572	0.0515900
Valence	−0.9279	0.4029	−0.128644	0.0148014	0.1836807
Arousal	0.1644	0.5854	0.315235	0.0058099	0.0720988
Output data ** (self-reported values)					
Intention to Buy/Taste	1.00	5.00	3.38	0.093	1.149

* The mean value of the FaceReader Online™ recording quality was 8.56 (Min = 4 Max = 10 Std. Dev. = 1.08).
** The output data refers to the average value of the self-reported variable "intention to buy or taste the promoted product".

The histograms in Figure 1 reveal the equal distribution (50%) of the responses in the two classes of intention to buy or taste the product, as well as the distributions in the FaceReader Online™-detected emotional states. Valence and arousal were at relatively low levels, revealing a state of inactivity. Additionally, a Shapiro–Wilk test showed that the detected emotions did not follow a normal distribution in the sample dataset ($p < 0.01$).
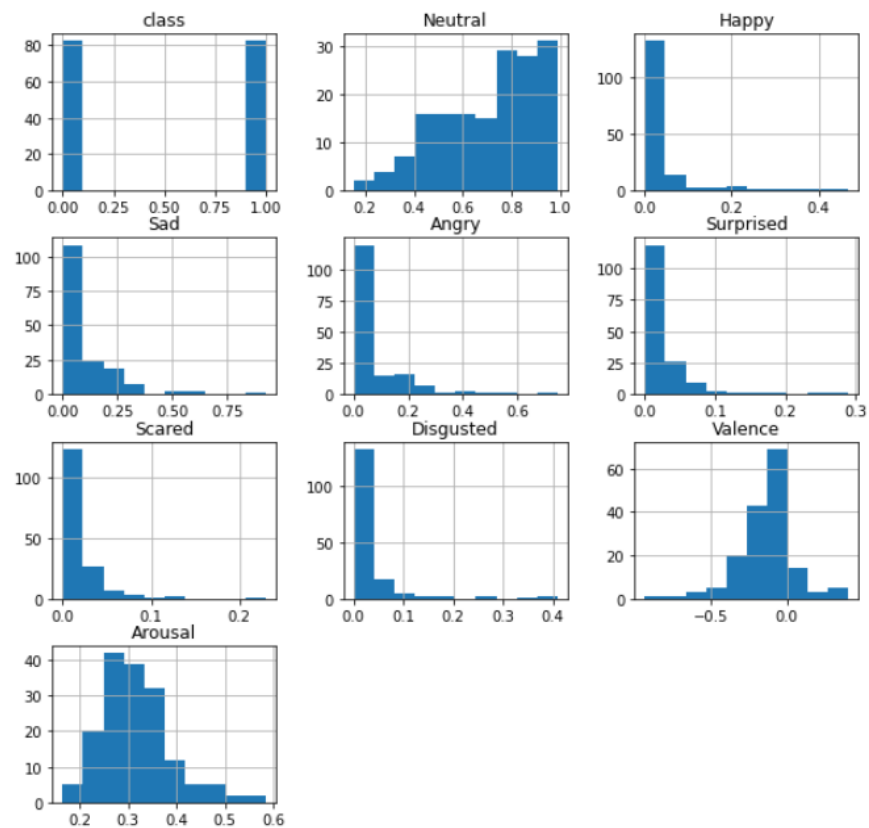
Figure 1. Histograms of examined emotional features and perceived purchase intent (class).

Despite the low-intensity scores of the detected emotions, their peak values were considered in the analysis. Moreover, FaceReader Online™ was able to detect parallel and mixed emotions. Figures 2–5 depict the reported time charts on the temporal average expressions of all participants in every campaign. The time-series data in the charts highlight the average levels (between 0 and 1 in axis y) of every measured emotion throughout the duration of every video campaign, separated in timeframes of 2 and 5 s (in axis x). The emotion "neutral" was excluded from the reports since its levels were significantly higher than all measured emotions (around 0.7/1.0), revealing low levels of emotional responses.
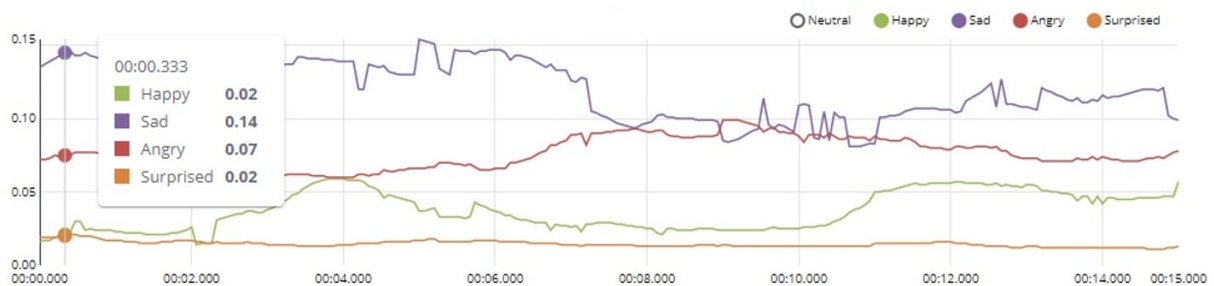
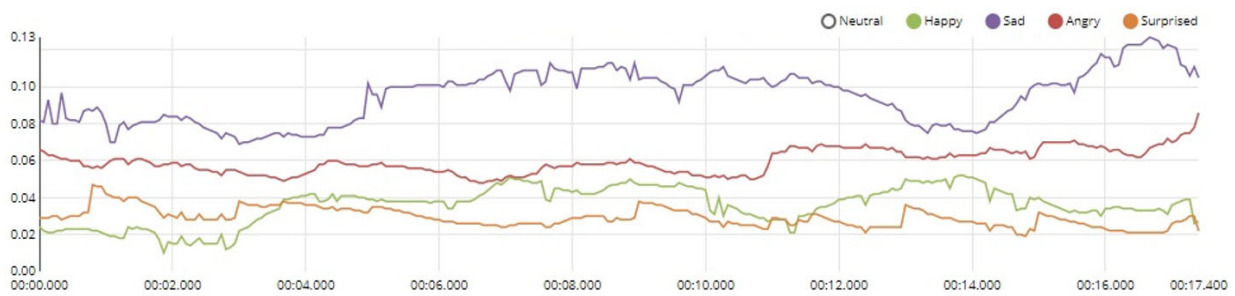**Figure 2.** Time chart of temporal average expressions—Experiment #1.

Figure 3. Time chart of temporal average expressions—Experiment #2.
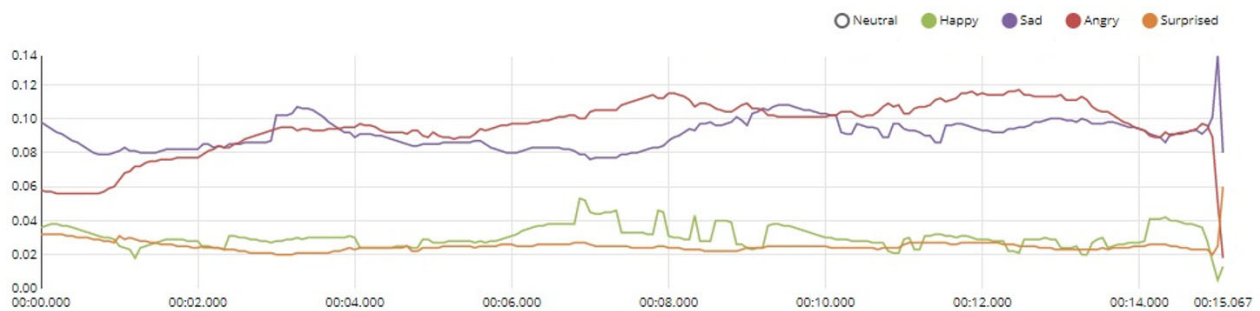

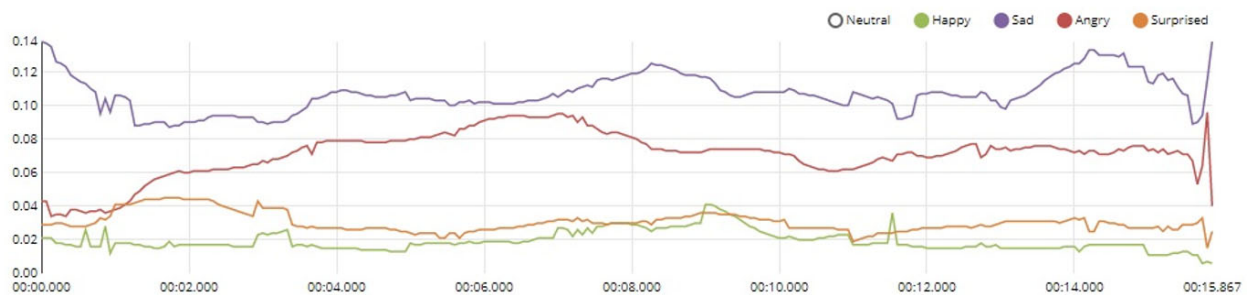**Figure 4.** Time chart of temporal average expressions—Experiment #3.

Figure 5. Time chart of temporal average expressions—Experiment #4.

3.2. Classification Results

Table 4 shows an overview of the different classification techniques undertaken on the collected dataset. The performance of the models is expressed in terms of accuracy and F-scores (considering precision and recall rates).

Table 4. Performance of classification models.

Model	Accuracy Score	F-Score
Neural Network–Sequential (NN-S)	0.90	0.81
Neural Network-Multi-Layer Preceptor (MLP NN)	0.91	0.83
Random Forest (RF)	0.75	0.75
Decision Trees (DTs)	0.70	0.70
Logistic Regression (LR)	0.58	0.60
Support Vector Machine -Linear (SVM)	0.55	0.68

As depicted, the NN approaches achieved the highest performance scores (90–91%), followed by RF (75%) and DTs (70%), while the models of LR, RF, and SVM revealed

lower scores (<60%). It should be noted that RF achieved a lower score (70%) when assigned 20–80 training and testing dataset percentages; the results are achieved through 15–85 percentages. Figure 6 visualizes the confusion matrix for the RF classification. The normalized percentages in the off-diagonal cells denote the correct predicted classes in the dataset.
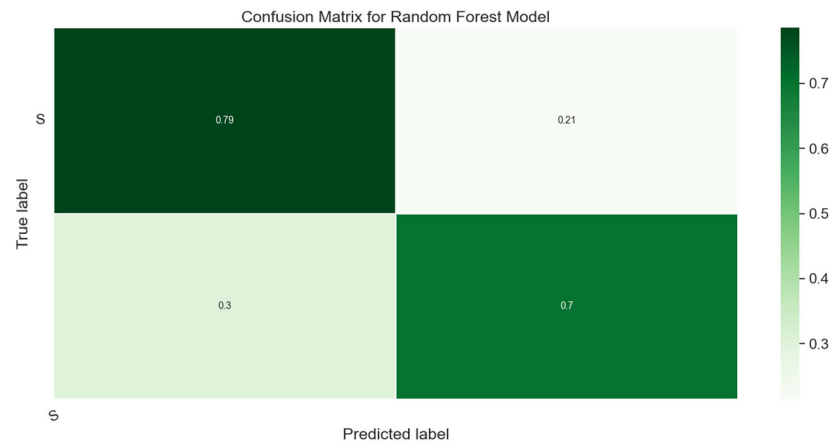

Figure 6. Confusion matrix for the Random Forest classification.

To investigate the emotional feature significance more deeply in the classification model performance, the relative importance was explored. Figure 7 depicts the relative importance of the extracted features for every measured emotion, calculated for the RF (a) and DT (b) classification models. As depicted, the strongest indicators for intention to buy or taste were the FACS expressions of "sadness" and "surprise", while the weakest indicators were "happiness" and "anger" correspondingly.
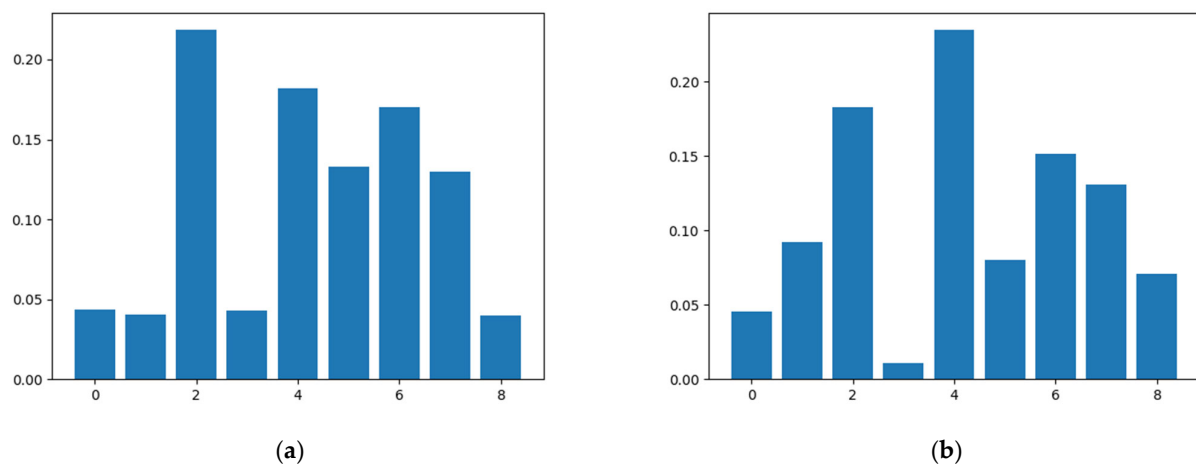

(**a**)   (**b**)

Figure 7. Features' relative importance (**a**) Random Forest classification; (**b**) Decision Tree classification. Note. 0 = Neutral, 1 = Happy, 2 = Sad, 3 = Angry, 4 = Surprised, 5 = Scared, 6 = Disgusted, 7 = Valence, 8 = Arousal.

4. Discussion, Implications, and Limitations

The study results are encouraging towards the detection of consumer purchase intent of utilitarian food products through facially detected emotional states. The study concludes that micro expressions captured by face-tracking AI tools are sufficient to feed classification models to predict consumer intent to buy or taste the food products. According to Ekman, micro expressions occur involuntarily within a fraction of a second and can expose a person's true emotions. Since emotions are associated with purchase intent of foods [34], and facial micro expressions can reveal emotional states, this study showed that it is

possible to predict purchase intent through facial emotions, even in low-emotion-expression activities (watching food video stimuli). Despite the low levels of arousal and valence amongst the participants and the prevalence of neutral emotional expressions, this study showed that slight facial expressions, which denote one of Ekman's basic emotional states, can be used as input to predict online customer purchase intent while watching digital promotions of utilitarian food products. Based on this, FaceReader Online™ proved to be a useful and valid means of assessing the consumers' attitudes towards a video-promoted product and their cognitive decision/desire to buy or try the product.

This study also showed that Neural Network models achieved quite high accuracy rates in classifying low or high levels of intention to buy/taste the product, through the FaceReader Online™ emotional data inputs. These findings are in accordance with previous studies in the field, confirming that deep learning and NN models outperform other models in predicting customer purchase intent [8]. The accuracy score is close to the one of Chaudhuri et al. [8], who achieved almost 90% accuracy in purchase intent on online platforms, through a deep learning NN of 128 neurons per layer. Their models' input data included platform engagement attributes, such as start of the session, price of product added to cart, day of the week, price of product clicked on, etc., as well as the time that passed since the last visit and the customer's account score to the online retailer. An accuracy score of 78% was achieved in [21] through SVM, in predicting changes in purchase intent, in a wide context of ads and products. This study revealed that NN approaches might achieve higher performance scores in detecting purchase intent for food products from facial data input. A similarly high accuracy score was achieved in [9], where the authors applied several ML models to predict potential buyers from social media user-generated content. Their highest score was achieved through the XGBoost classifier, reaching a score of 86%. The current study achieved similar results without scarping social media datasets, accessing e-shop web sessions, cart, and customer account data, but through a different dataset, solely relying on emotions detected through facial expressions.

Based on the FaceReader Online™ outcomes, sadness was the most prevailing emotion (after neutrality) amongst the participants. Sadness expressions were also denoted as the most significant factor for purchase intent in the RF classification model. However, this does not mean that participants felt sad while watching the campaigns; rather, their facial expressions looked like the FACS pattern of sadness, triggered by the muscles in particular face areas. According to FACS [35], sadness expression includes three muscular action units: inner brow raiser, brow lowered, and lip corner depressor. However, these muscles do not only contribute to a sad expression but also to other voluntary facial expressions, for instance when the person experiences cognitive difficulty or confusion [36]. Similar studies that applied FaceReader Online™ to examine emotions on multimedia content [20] concluded that, several times, the software perceived neutrality as sadness, while the concentration of participants on static images was recognized as anger. Additionally, happiness recognition is based on the units of cheek riser (orbicularis oculi) and lip corner pillar (zygomaticus major). That means that a person needs to smile to be recognized as happy [36]. However, happiness is a complex emotion, and researchers suggest the co-consideration of other modalities, such as head and gaze [37].

There are three major academic and business implications in this study. First, this study bridges the acknowledged need to predict online consumer intent to purchase a promoted product in social media campaigns. This can be achieved through facially decoded emotions that can be retrieved from face-tracking AI tools and applications. FaceReader emotional reports can be used as input to classify the viewers' intention to buy or taste a video-promoted product on social media. Researchers and marketers can use our findings to extend research in the field and estimate the consumer's purchase intent in different video creatives and product types.

Secondly, the findings of this study can be applied in the context of designing efficient campaigns to promote healthy and utilitarian products and enhance the consumers' well-being. Researchers and marketeers can use similar classification models to compare the

effects of different video campaigns or products and optimize their creatives to positively affect consumers towards adopting a healthier lifestyle.

Finally, this study compares the usage of multiple classification techniques for a prediction of the intended purchases from watching social media campaigns. A comparative approach was adopted to examine the efficacy of NNs and ML for the dataset. This examination extends previous approaches that applied SVM models to achieve high accuracy scores [21] and supports other studies that have demonstrated the higher predictive power of DL for consumer behavioral datasets [8]. As a result, it provides further empirical evidence to support the superior predictive capabilities of NNs in such a context. Decision Trees were also found to be efficient in terms of accuracy. Future research can use these insights when applying multiple techniques for research in this and other similar domains.

One limitation of this study is the laboratory setup of the experiment, where participants knowing that they are recorded might have affected their facial expressions [38]. For this, more research should be conducted outside the laboratory to receive more spontaneous and original expressional feedback. Similarly, the self-reported method to collect responses on the participants' intent to buy the product might have been affected by social desirability bias [39], where respondents might have rated higher than the original levels of their intentions towards a utilitarian (non-hedonic) food product.

Another main limitation of the study is the relatively small dataset, for this larger dataset should be used in future research to foster and generate the findings. Further, the type of the promoted stimuli could possibly differentiate the results; since utilitarian food products do not easily induce strong feelings, the FaceReader Online^{TM} results might be different for different products [24], affecting the performance scores of the models. The study also considers that there might be other mediating factors between emotional states and intent to buy a product. Such factors can include personal characteristics, likeness of the product, and attitude towards the brand. However, research showed that these factors are strongly associated with the final intention to buy the product; hence, it is assumed that they do not significantly affect results and the usefulness of the classification models.

Finally, the study is conducted on the analysis of facial responses on a group of respondents to produce generalized outcomes. FaceReader Online^{TM} also provides individual reports that can be applied in individual experimental settings.

Future work could include the deployment of more classification models and a comparison of more performance metrics, such as sensitivity and specificity of the models.

5. Conclusions

The study examined the prediction of purchase intent based on facially expressed emotions that were detected through FaceReader Online^{TM}, a face-tracking emotional AI tool. A four-stage experiment was designed, where participants were exposed to a set of videos and other multimedia-based social media campaigns promoting utilitarian food products, including yogurt and nut butters. While watching the campaigns, the participants' facial emotions were analyzed through Facereader Online^{TM}, and in the end, they self-reported their intention to buy or taste the promoted product. A set of classification methods was applied and compared in terms of accuracy and F-scores. Neural Networks achieved a relatively high level of accuracy (90–91%), while Random Forest and Decision Trees showed promising results (75%, 70%). The expressions of sadness and surprise indicated the highest levels of relative importance in predicting purchase intent in the context, followed by disgust and valence.

Some main limitations of the study regard the laboratory experimental approach and the stimuli characteristics, which might affect the results. Future research shall be conducted outside the laboratory, on various stimuli, and on larger datasets to reinforce the validity of our findings.

Overall, the results are promising towards the future application of face tracking to detect purchase intention and other decision-related cognitive processes. The findings can contribute to the deployment of new predictive models and their real-time application

in social-media-promoted content. Researchers and marketers can use such knowledge to predict (intended) sales and efficiently design social media campaigns for promoting utilitarian products and overall well-being.

References

1. Martínez, A.; Schmuck, C.; Pereverzyev, S.; Pirker, C.; Haltmeier, M. A Machine Learning Framework for Customer Purchase Prediction in the Non-Contractual Setting. *Eur. J. Oper. Res.* **2020**, *281*, 588–596. [CrossRef]
2. Jacobs, B.J.D.; Donkers, B.; Fok, D. Model-based purchase predictions for large assortments. *Mark. Sci.* **2016**, *35*, 389–404. [CrossRef]
3. Hu, X.; Huang, Q.; Zhong, X.; Davison, R.M.; Zhao, D. The Influence of Peer Characteristics and Technical Features of a Social Shopping Website on a Consumer's Purchase Intention. *Int. J. Inf. Manag.* **2016**, *36*, 1218–1230. [CrossRef]
4. Ravaja, N.; Somervuori, O.; Salminen, M. Predicting Purchase Decision: The Role of Hemispheric Asymmetry over the Frontal Cortex. *J. Neurosci. Psychol. Econ.* **2013**, *6*, a0029949. [CrossRef]
5. Venkatesh, V.; Agarwal, R. Turning Visitors into Customers: A Usability-centric Perspective on Purchase Behavior in Electronic Channels. *Manag. Sci.* **2006**, *52*, 367–382. [CrossRef]
6. Sismeiro, C.; Bucklin, R.E. Modeling Purchase Behavior at an E-Commerce Web Site: A Task Completion Approach Revised June 2003 Revised November 2003. *SAGE J.* **2004**, *41*, 35985.
7. Zhu, G.; Wu, Z.; Wang, Y.; Cao, S.; Cao, J. Online Purchase Decisions for Tourism E-Commerce. *Electron. Commer. Res. Appl.* **2019**, *38*, 100887. [CrossRef]
8. Chaudhuri, N.; Gupta, G.; Vamsi, V.; Bose, I. On the Platform but Will They Buy? Predicting Customers' Purchase Behavior Using Deep Learning. *Decis. Support Syst.* **2021**, *149*, 113622. [CrossRef]
9. Xu, Z.; Dang, Y.; Wang, Q. Potential Buyer Identification and Purchase Likelihood Quantification by Mining User-Generated Content on Social Media. *Expert Syst. Appl.* **2022**, *187*, 115899. [CrossRef]
10. Juodeikiene, G.; Zadeike, D.; Klupsaite, D.; Cernauskas, D.; Bartkiene, E.; Lele, V.; Steibliene, V.; Adomaitiene, V. Effects of Emotional Responses to Certain Foods on the Prediction of Consumer Acceptance. *Food Res. Int.* **2018**, *112*, 361–368. [CrossRef]
11. Otterbring, T.; Folwarczny, M.; Gidlöf, K. Hunger Effects on Option Quality for Hedonic and Utilitarian Food Products. *Food Qual. Prefer.* **2023**, *103*, 104693. [CrossRef]
12. Talen, L.; den Uyl, T.E. Complex Website Tasks Increase the Expression Anger Measured with FaceReader Online. *Int. J. Hum. Comput. Interact.* **2022**, *38*, 282–288. [CrossRef]
13. Danner, L.; Haindl, S.; Joechl, M.; Duerrschmid, K. Facial Expressions and Autonomous Nervous System Responses Elicited by Tasting Different Juices. *Food Res. Int.* **2014**, *64*, 81–90. [CrossRef]
14. Garcia-Burgos, D.; Zamora, M.C. Facial Affective Reactions to Bitter-Tasting Foods and Body Mass Index in Adults. *Appetite* **2013**, *71*, 178–186. [CrossRef] [PubMed]
15. He, W.; Boesveldt, S.; de Graaf, C.; de Wijk, R.A. The Relation between Continuous and Discrete Emotional Responses to Food Odors with Facial Expressions and Non-Verbal Reports. *Food Qual. Prefer.* **2016**, *48*, 130–137. [CrossRef]
16. Leitch, K.A.; Duncan, S.E.; Keefe, S.O.; Rudd, R.; Gallagher, D.L. Characterizing Consumer Emotional Response to Sweeteners Using an Emotion Terminology Questionnaire and Facial Expression Analysis. *FRIN* **2015**, *76*, 283–292. [CrossRef]
17. Van Bommel, R.; Stieger, M.; Visalli, M.; De Wijk, R.; Jager, G. Does the Face Show What the Mind Tells? A Comparison between Dynamic Emotions Obtained from Facial Expressions and Temporal Dominance of Emotions (TDE). *Food Qual. Prefer.* **2020**, *85*, 103976. [CrossRef]
18. Mena, B.; Torrico, D.D.; Hutchings, S.; Ha, M.; Ashman, H.; Warner, R.D. Understanding Consumer Liking of Beef Patties with Different Firmness among Younger and Older Adults Using FaceReader™ and Biometrics. *Meat Sci.* **2023**, *199*, 109124. [CrossRef]
19. Mahieu, B.; Visalli, M.; Schlich, P.; Thomas, A. Eating Chocolate, Smelling Perfume or Watching Video Advertisement: Does It Make Any Di Ff Erence on Emotional States Measured at Home Using Facial Expressions ? *Food Qual. Prefer.* **2019**, *77*, 102–108. [CrossRef]
20. Yu, C.; Ko, C. Applying FaceReader to Recognize Consumer Emotions in Graphic Styles. *Procedia CIRP* **2017**, *60*, 104–109. [CrossRef]

21. McDuff, D.; El Kaliouby, R.; Cohn, J.F.; Picard, R.W. Predicting Ad Liking and Purchase Intent: Large-Scale Analysis of Facial Responses to Ads. *IEEE Trans. Affect. Comput.* **2015**, *6*, 223–235. [CrossRef]
22. McDuff, D.; El Kaliouby, R.; Senechal, T.; Demirdjian, D.; Picard, R. Automatic Measurement of Ad Preferences from Facial Responses Gathered over the Internet. *Image Vis. Comput.* **2014**, *32*, 630–640. [CrossRef]
23. Pichierri, M.; Peluso, A.M.; Pino, G.; Guido, G. Health Claims' Text Clarity, Perceived Healthiness of Extra-Virgin Olive Oil, and Arousal: An Experiment Using FaceReader. *Trends Food Sci. Technol.* **2021**, *116*, 1186–1194. [CrossRef]
24. Tzafilkou, K.; Panavou, R.; Economides, A.A. Facially Expressed Emotions and Hedonic Liking on Social Media Food Marketing Campaigns: Comparing Different Types of Products and Media Posts. In Proceedings of the 2022 17th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP), Corfu, Greece, 3–4 November 2022; pp. 1–6. [CrossRef]
25. Ding, Y.; Zhao, T. Emotions, Engagement, and Self-Perceived Achievement in a Small Private Online Course. *J. Comput. Assist. Learn.* **2019**, *36*, 449–457. [CrossRef]
26. Tzafilkou, K.; Economides, A.A. Mobile Game-Based Learning in Distance Education: A Mixed Analysis of Learners' Emotions and Gaming Features. In Proceedings of the Learning and Collaboration Technologies: Games and Virtual Environments for Learning: 8th International Conference, LCT 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, 24–29 July 2021; pp. 115–132. [CrossRef]
27. Ekman, P. Are There Basic Emotions? *Psychol. Rev.* **1992**, *99*, 550–553. [CrossRef]
28. Loijens, L.; Krips, O. *Facereader Methodology*; Noldus Information Technology: Wageningen, The Netherlands, 2013.
29. Olsen, A.F.; Torresen, J. Smartphone Accelerometer Data Used for Detecting Human Emotions. In Proceedings of the 2016 3rd International Conference on Systems and Informatics (ICSAI), Shanghai, China, 19–21 November 2016; pp. 410–415.
30. Piskioulis, O.; Tzafilkou, K.; Economides, A.A. Emotion Detection through Smartphone's Accelerometer and Gyroscope Sensors. In Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21). Association for Computing Machinery, New York, NY, USA, 21–25 June 2021; pp. 130–137. [CrossRef]
31. Osisanwo, F.Y.; Akinsola, J.E.T.; Awodele, O.; Hinmikaiye, J.O.; Olakanmi, O.; Akinjobi, J. Supervised Machine Learning Algorithms: Classification and Comparison. *Int. J. Comput. Trends Technol.* **2017**, *48*, 128–138. [CrossRef]
32. Gove, R.; Faytong, J. *Machine Learning and Event-Based Software Testing: Classifiers for Identifying Infeasible GUI Event Sequences*; Elsevier Inc.: Amsterdam, The Netherlands, 2012; Volume 86, ISBN 9780123965356.
33. Pretorius, A.; Bierman, S.; Steel, S.J. A Meta-Analysis of Research in Random Forests for Classification. In Proceedings of the 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA)-RobMech 2016, Stellenbosch, South Africa, 30 November–2 December 2016. [CrossRef]
34. Jiang, Y.; King, J.M.; Prinyawiwatkul, W. A Review of Measurement and Relationships between Food, Eating Behavior and Emotion. *Trends Food Sci. Technol.* **2014**, *36*, 15–28. [CrossRef]
35. Ekman, P. *Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists Press: Palo Alto, CA, USA, 1978.
36. Root, A.A.; Stephens, J.A. Organization of the Central Control of Muscles of Facial Expression in Man. *J. Physiol.* **2003**, *549*, 289–298. [CrossRef]
37. Yudiarso, A.; Liando, W.; Zhao, J.; Ni, R.; Zhao, Z. Validation of Facial Action Unit for Happy Emotion Detection. In Proceedings of the 3rd International Conference on Psychology in Health, Educational, Social, and Organizational Settings (ICP-HESOS 2018)—Improving Mental Health and Harmony in Global Community, Surabaya, Indonesia, 16–18 November 2018; SCITEPRESS—Science and Technology Publications, Lda.: Surabaya, Indonesia; pp. 360–363. [CrossRef]
38. Wichchukit, S.; Mahony, M.O. The 9-Point Hedonic Scale and Hedonic Ranking in Food Science: Some Reappraisals and Alternatives. *J. Sci. Food Agric.* **2014**, *95*, 2167–2178. [CrossRef]
39. Krumpal, I. Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review. *Qual. Quant.* **2013**, *47*, 2025–2047. [CrossRef]