2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Wearable EEG Entropy and Spectral Measures for Classification of Consumer Reward-based Evaluation of Odor Stimuli

Manuel S. Seet, Amritha V. Devarajan, Jazreel J. L. Low,
Junji Hamano, Mariana Saba, Nitish V. Thakor, Andrei Dragomir*

***Abstract*— Consumer neuroscience is a rapidly emerging field, with the ability to detect consumer attitudes and states via real-time passive technologies being highly valuable. While many studies have attempted to classify consumer emotions and perceived pleasantness of olfactory products, no known machine learning approach has yet been developed to directly predict consumer reward-based decision-making, which has greater behavioral relevance. In this proof-of-concept study, participants indicated their decision to have fragrance products repeated after fixed exposures to them. Single-trial power spectral density (PSD) and approximate entropy (ApEn) features were extracted from EEG signals recorded using a wearable device during fragrance exposures, and served as subject-independent inputs for 4 supervised learning algorithms (kNN, Linear-SVM, RBF-SVM, XGBoost). Using a cross-validation procedure, kNN yielded the best classification accuracy (77.6%) using both PSD and ApEn features. Acknowledging the challenging prospects of single-trial classification of high-order cognitive states especially with wearable EEG devices, this study is the first to demonstrate the viability of using sensor-level features towards practical objective prediction of consumer reward experience.**

***Index Terms*—Consumer Neuroscience, Reward Decision-Making, EEG, Olfaction, Machine Learning**

I. INTRODUCTION

Consumer neuroscience is an emerging interdisciplinary field that taps on neuroscientific knowledge and methods to glean deeper insights into the needs, preferences and choices of consumers. One frontier is the development of real-time passive technologies which can objectively detect consumers' attitudes and states that influence their product evaluations and behaviors. Interest for such technologies has been gaining traction, since it is known that classical rating scales widely-used in market research may fail to capture consumer's true preferences and intentions [1].

One popular approach to studying consumer's evaluation of the product is to measure emotions and perceived pleasantness. Many studies have employed electroencephalography (EEG) [2], [3], electrocardiography (ECG) [3], electrodermal recordings [4] and facial expression monitoring [5] to objectively measure neural and autonomic responses that may predict subjective pleasantness and emotions. Among these, EEG has stood out as the most popular and reliable choice, as it is the most accurate in measuring the brain's hedonic and affective responses [6]. Other modalities depend on downstream effects that may be less reliable due to physiological noise (for periphery measures) [3], or socio-cultural influences (for facial emotion measures) [7].

However, it is known from past literature that positive hedonic valuation and emotions do not necessarily translate to strong purchase intention, alluding to complex interactions with other determinants that contribute to consumer decision-making [8]. Therefore, to better predict consumers' purchase behavior, it is more advantageous to detect cognitive states that directly relate to their product-based decision-making.

Such decisions are driven not simply by 'liking' of the product, but the feeling of 'wanting' it, which taps on brain networks closely associated with reward appraisal [9]. While these dopaminergic reward networks primarily encompass subcortical circuitries linking the ventral striatum, basal ganglia and amygdala, they also project out to cortical regions such as the orbitofrontal cortex and the cingulate cortices that process reward-driven weighing of decisions that precipitates behavioral action [9]. Hence, EEG is theoretically capable of tracking consumer reward-based decisions using signals originating from these cortical substrates during exposure to the product. No known study has yet developed a machine learning approach towards practically achieving this goal, and so the current research pursued this direction.

The aim of this study is therefore to assess the feasibility of EEG-based prediction of consumer reward-based decisions. Here, participants were exposed to various fragrance products, immediately after which they indicated their decision to want to have the sample repeated; this metric reflects reward-driven evaluations [10]. Sensor-level features such as power spectral density and approximate entropy were extracted from the EEG signals, and used to train 4 popular supervised learning algorithms (kNN, Linear-SVM, RBF-SVM, XGBoost). Classification performance was assessed via a cross-validation procedure.

II. DATA COLLECTION

A. Experiment Procedures

Fourteen right-handed female participants of Chinese ethnicity (aged 21-45) at the National University of Singapore (NUS) took part in the research. Experimental procedures have been approved by the NUS Institution Review Board. Demographic factors (gender, handedness and ethnic background) were kept constant to minimize extraneous affective and neural differences across participants [11], [12].

This work was supported by BMRC SPF grant number APG2013/110. from A*STAR, Singapore and Procter & Gamble, Singapore. M. S. Seet, A. V. Devarajan, J. J. L. Low, N. V. Thakor and A. Dragomir are with the N.1 Institute for Health, National University of Singapore, 28 Medical drive, 05-COR, Singapore 117456. M. Saba and J. Hamano are with Procter & Gamble, International Operations SA Singapore Branch.
* Corresponding Author: andrei.drag@gmail.com

In this study, the participants were presented with fragrance product samples in 8-seconds trials while fixating on a fixation cross, immediately after which they rated sample pleasantness (11-point scale), and indicated their decision to want to have the sample repeated, on a 5-point balanced scale (i.e. a reward-repetition decision scale: Strongly No, No, Neutral, Yes, Strongly Yes). This scale is a measure of reward-based decision-making, adapted from neuroimaging studies on reward, e.g. [10]. Participants' responses did not actually change the schedule of sample presentations. To mitigate olfactory fatigue, a coffee odorant was presented to subjects after they made their responses and short breaks were administered between all trials.

Testing occurred over three consecutive testing days, with forty trials of sample presentations per day. Four samples were used, which were identified to be samples with the highest and lowest familiarity ratings out of 6 commercial products in a pre-screening procedure. Familiarity was chosen as the variable for stimulus selection, because it relates to how accepting individuals will be of the stimulus in question [13], thus providing inter-stimulus variability that can give rise to a distribution in reward-based decision responses.

B. EEG Acquisition and Analysis

1) Apparatus: EEG signals were recorded with the CGX Quick-20 headset (Cognionics; CA, USA), which has a 19 dry-sensor montage conforming to the International 10-20 System, in addition to forehead grounds and a bilateral earlobe reference. Impedances were maintained under 2,500 kΩ. EEG data were recorded at a 500-Hz sampling rate and digitized for wireless transmission to the CGX Data Acquisition software on the host computer for storage.

2) Preprocessing: EEG signals were passed through an automated preprocessing pipeline with the following components: Chevbyshev band-pass filter (Order 6; 0.3 to 40 Hz), resampling to 250 Hz, common average referencing, epoch extraction (8-second trials), and independent component analysis (ICA). The cleaned EEG signal was reconstructed after automated detection and rejection of artifact-related components using the ADJUST algorithm [14].

3) Feature Extraction: All features were extracted from the first 4 seconds of the trial, corresponding to the first sniff cycle. The preprocessed EEG signal underwent band-pass filtering to obtain signals in the theta (θ: 4- 8 Hz), alpha (α: 8 - 12 Hz), beta (β: 12 - 30 Hz) and gamma (γ: 30- 40 Hz) bands. Relative power spectral density (PSD) across the 4 frequency bands were estimated using Welch's method. Approximate entropy (ApEn)–which quantifies the unpredictability of signal fluctuations–was calculated for each frequency band, according to ref. [15], using window size, $m = 2$, and noise filtering level, $r = 0.2 \times SD$ of signal amplitude values.

III. Supervised Learning

A. Supervised Learning Algorithms

1) k-Nearest Neighbors: k-Nearest Neighbors (kNN) is a model-free classification method, in which a test instance is assigned the label belonging to the majority of k closest instances in terms of proximity within the feature space. In this study, the distance metric d is Euclidean distance, and label voting is inversely distance-weighted. k was subjected to tuning.

2) SVM with Linear Kernel: The classical support vector machine (SVM) with the linear kernel finds the linear hyperplane in the feature space that best separates classes, with support vectors that define the margin. The C regularization parameter controls the tolerance of training misclassifcation to prevent overfitting, and was subjected to tuning.

3) SVM with Radial Basis Function Kernel: SVM is often implemented with the radial basis function (RBF) kernel that can map onto a infinitely multidimensional space. In this study, we set $\gamma = 1/(N \times \sigma^2)$, where N is the total number of features and σ^2 is the variance of the feature dataset. C was subjected to tuning.

4) XGBoost: Extreme Gradient Boosting (XGBoost) is an ensemble learning algorithm that relies on gradient tree boosting, along with additional regularizing components (gain threshold γ and L2 regularization parameter λ) that control tree building and minimize over-fitting [16]. Many optimization features (e.g. approximate greedy split-finding, cache-aware access) help XGBoost learn efficiently with large datasets. In this study, we set $\gamma = 1$ and learning rate $\epsilon = 0.1$. λ was subjected to tuning.

B. Input Preparation

From the 1680 trial responses collected, 450 trials (26.8%) with "Neutral" responses were removed as the participants did not have a clear decision. Of the remaining 1230 responses, 478 (38.9%) were labeled as Yes-decisions ("Yes" + "Strongly Yes") and 752 (61.1%) as No-decisions ("No" + "Strongly No"). These Yes/No-Decisions served as labels, with either (i) PSD features (76 features), (ii) ApEn features (76 features) or (iii) PSD+ApEn features (152 features) of the corresponding trials fed as subject-independent inputs to the supervised learning algorithms. All feature inputs have been mean-centred and scaled to the unit variance.

C. Cross-Validation and Hyperparameter Tuning

Hyperparameters were optimized for weighted F1-score within a cross-validation grid search framework. For kNN, candidate $k \in \{1, 2, 3, ..., 10\}$, for Linear-SVM and RBF-SVM, candidate $C \in \{0.1, 1, 5, 10, 20, 50, 100\}$ and for XGBoost, candidate $\lambda \in \{1, 2, 3, ..., 10\}$. Classifier performance was assessed via a 20-fold cross-validation procedure, which trains on a 95% dataset subsample and tests on the remaining 5% subsample in each iteration. Performance was evaluated using (i) classification accuracy and (ii) weighted F1-score, which is the weighted average of the F1-score for each label to account for label imbalance.

IV. Results

A. Behavioral Analysis

The positive rank correlation between reward-repetition decision (all 5-point responses) and pleasantness ratings is

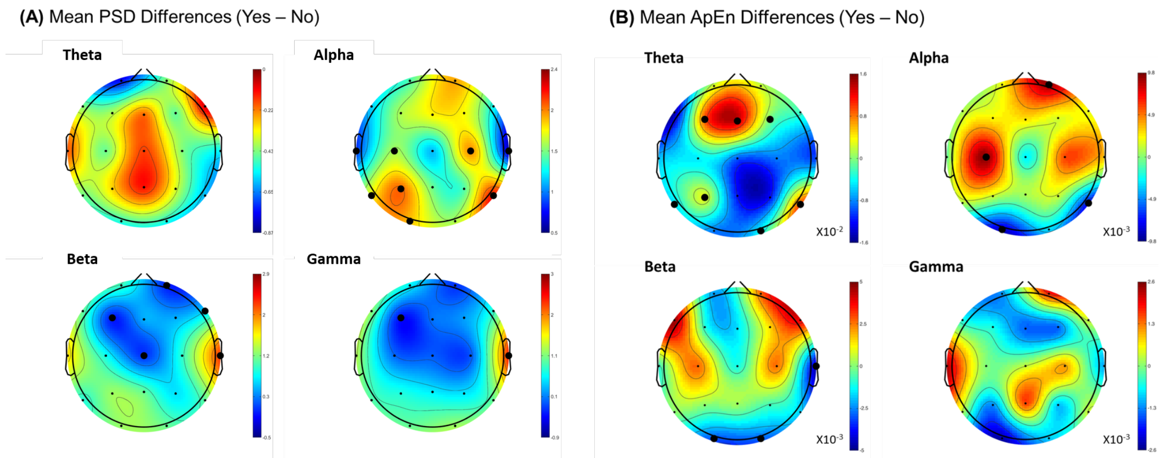

Fig. 1. Topographic scalp maps illustrating the mean differences in (A) PSD in dB and (B) ApEn between Yes-decision and No-decision trials, for each of the 4 frequency bands. All plot scales are zero-centred at green. Statistically significant differences (after FDR corrections) are marked with bold circles. For PSD, these differences are found in Alpha: C3, C4, T3, T4, P3, P7, P8, O1; Beta: Fp2, F3, F8, Cz, T4; Gamma: F3, T4. For ApEn, these differences are found in Theta: Fz, F3, F4, P3, P7, P8, O2; Alpha: Fp2, C3, P8, O1; Beta: T4, O1, O2. *Note that in the beta and gamma PSD difference plots, the right temporal activity localized to T4 is most likely attributable to effects of reward-repetition processing, evidenced by high PSD in Yes-decision trials (beta: $M = 0.214$; gamma: $M = 1.55\text{E-}03$) and relatively low PSD in No-decision trials (beta: $M = 0.124$; gamma: $M = 8.68\text{E-}04$).

strong (Spearman's $\rho = 0.63$, $p < .001$), and the coefficient of determination is $R^2 = 0.407$, meaning that pleasantness explains 40.7% of the variance of reward-repetition decision.

B. *EEG PSD and ApEn Analysis*

Figure 1 displays the topographic scalp map of the mean differences in PSD and ApEn between Yes- and No-decision trials. These differences were tested using t-tests, with the Benjamini–Yekutieli false discovery rate (FDR) procedure [17] applied to correct for multiple comparisons for $\alpha = .05$. Statistically significant differences in PSD were observed mainly in the alpha band amongst centro-posterior channels (raw ps$<$ 1.4E-5), as well as in the beta (ps$<$ .0003) and gamma bands (ps$<$ .0003). Statistically significant differences in ApEn were observed mainly in the theta band in frontal and posterior channels (ps$<$ .006), as well as in the alpha (ps$<$ .007) and beta bands (ps$<$ .0001).

C. *Classifier Performance*

According to the mean classification accuracy and weighted F1-score (Figure 2), merging PSD and ApEn features produced better classification performance across all classifiers, compared to using either only PSD or only ApEn features. However, the advantage of feature merging appears to be marginal, especially in the case of kNN and XGBoost.

kNN achieved the best classification accuracy at 77.6% (weighted F1-score = 0.763) when using all PSD and ApEn features, with $k = 10$. Other nonlinear classifiers (RBF-SVM with $C = 10$; XGBoost with $\lambda = 1$) also had comparable performance, with accuracy scores $> 75\%$ and weighted F1-scores $> 74\%$. Linear-SVM had the least accurate classification with 66.6% accuracy at best ($C = 100$).

V. Discussion

The current study set out to assess the feasibility of using wearable EEG features to predict consumer reward-based evaluations. First, it was found that subjective pleasantness was a significant correlate of reward-repetition decisions, though it only accounts for less than half (40.7%) of its variance. This finding is in agreement with the idea that there are other factors apart from products' hedonic value that contribute to reward-based evaluations [8]. Therefore, EEG is needed to detect neurocognitive signals that can better predict consumers' reward decision-making.

Second, statistical analyses reveal that PSD and ApEn are modulated with respect to reward-repetition decisions. PSD, which captures activation information in the frequency domain, showed Yes-No differences most prominently in the alpha, beta and gamma band. ApEn, which captures nonlinear dynamic information in the time domain, showed Yes-No differences most prominently in the theta, alpha and beta bands. Alpha-, theta-, and gamma-band modulations observed here are consistent with prior olfactory studies, e.g. [2], which have shown that activity within these bands relate to pleasantness evaluations. The additional presence of substantial beta-band modulation is noteworthy, as beta brain oscillations are associated with reward-based processing [18]. Overall, these patterns suggest that in making a decision concerning repetition of the stimuli, neural mechanisms underpinning hedonic judgements as well as reward-driven processing (amongst others) are engaged.

Most critically, both PSD and ApEn features can support reliable classification of reward-repetition, especially when using nonlinear classifiers (e.g. kNN, XGBoost). Merging the two brings marginal benefit in most cases. This indicates that both frequency-domain and time-domain features extracted at the sensor level from single trials have informative value relating to reward-based evaluation and decision-making. These results are notable in view of the challenges typically associated with single-trial EEG discrimination [19].

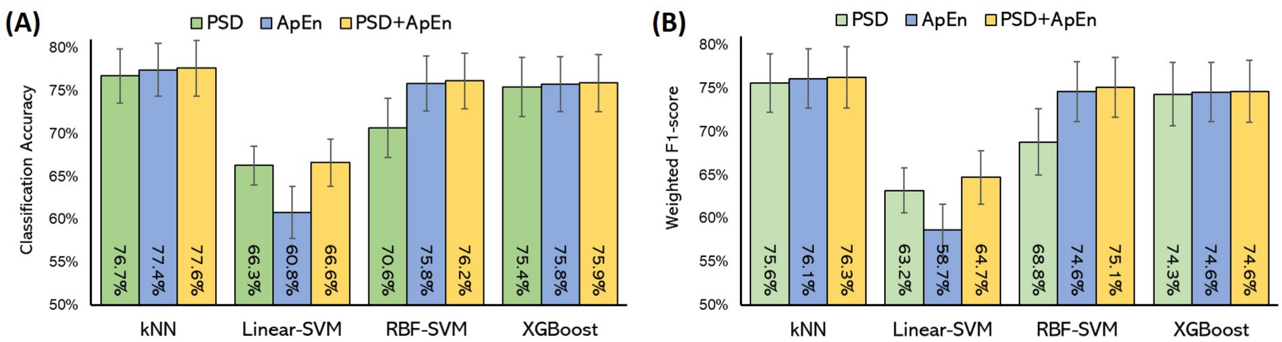

Fig. 2. (A) Mean classification accuracy (chance accuracy = 61%) and (B) Mean weighted F1-score of the 4 supervised learning models, using (i) PSD features, (ii) ApEn features, and (iii) the combination of PSD and ApEn features. Error bars represent ±1 standard error of the mean (SEM).

VI. Conclusions

This proof-of-concept study demonstrates the viability of using features estimated from wearable EEG signals to predict consumer reward-based evaluations of olfactory stimuli, with the best trial-based classification performance (77.6% accuracy) achieved with kNN using PSD and ApEn features. Our findings also illustrate the significant role played by brain oscillations spanning the theta, alpha, beta and gamma bands in reward-related decision-making, suggesting recruitment of a wide array of neurocognitive mechanisms.

It must be noted that the current study focused on participants of Chinese ethnicity to avoid extraneous inter-subject affective and neural differences. Future research will study how demographic and cognitive factors influence consumers' reward-based decision-making and to what degree machine learning models can be trained to account for these factors.

Considering the significant challenges raised by single-trial classification of high-order cognitive states especially with wearable EEG devices, we believe the present study constitutes a significant breakthrough in consumer neuroscience research, by demonstrating for the first time the feasibility of using sensor-level features towards practical objective prediction of consumer reward experience. Future work can focus on finer-tuned, multi-class classification of reward-repetition responses, further paving the way for naturalistic, non-verbal testing in consumer research.

References

[1] X. Villegas-Ruiz, O. Angulo, and M. O'Mahony, "Hidden and false "preferences" on the structured 9-point hedonic scale," *J. Sens. Stud.*, vol. 23, no. 6, pp. 780–790, 2008.

[2] N. I. Abbasi, R. Bose, A. Bezerianos, N. V. Thakor, and A. Dragomir, "Eeg-based classification of olfactory response to pleasant stimuli," in *2019 41st Conf. Proc. IEEE Eng. Med. Biol. Soc. (EMBC)*. IEEE, 2019, pp. 5160–5163.

[3] E. Kroupi, J.-M. Vesin, and T. Ebrahimi, "Subject-independent odor pleasantness classification using brain and peripheral signals," *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 422–434, 2015.

[4] M. S. Seet, M. R. Amin, N. I. Abbasi, J. Hamano, A. Chaudhury, A. Bezerianos, R. T. Faghih, N. V. Thakor, and A. Dragomir, "Olfactory-induced positive affect and autonomic response as a function of hedonic and intensity attributes of fragrances," in *2020 42nd Conf. Proc. IEEE Eng. Med. Biol. Soc. (EMBC)*. IEEE, 2020, pp. 3170–3173.

[5] B. Mahieu, M. Visalli, P. Schlich, and A. Thomas, "Eating chocolate, smelling perfume or watching video advertisement: Does it make any difference on emotional states measured at home using facial expressions?" *Food Qual. Prefer.*, vol. 77, pp. 102–108, 2019.

[6] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using eeg signals: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 374–393, 2017.

[7] F. A. M. da Silva and H. Pedrini, "Effects of cultural characteristics on building an emotion classifier through facial expression analysis," *J. Electron. Imaging*, vol. 24, no. 2, p. 023015, 2015.

[8] E. Kytö, M. Virtanen, and S. Mustonen, "From intention to action: Predicting purchase behavior with consumers' product expectations and perceptions, and their individual properties," *Food Qual. Prefer.*, vol. 75, pp. 1 – 9, 2019.

[9] E. T. Rolls, "Taste, olfactory, and food reward value processing in the brain," *Prog. Neurobiol.*, vol. 127, pp. 64–90, 2015.

[10] F. Grabenhorst, E. T. Rolls, and B. A. Parris, "From affective value to decision-making in the prefrontal cortex," *Eur. J. Neurosci.*, vol. 28, no. 9, pp. 1930–1939, 2008.

[11] C. Ferdenzi, S. C. Roberts, A. Schirmer, S. Delplanque, S. Cekic, C. Porcherot, I. Cayeux, D. Sander, and D. Grandjean, "Variability of affective responses to odors: culture, gender, and olfactory knowledge," *Chem. Senses*, vol. 38, no. 2, pp. 175–186, 2013.

[12] J.-P. Royet, J. Plailly, C. Delon-Martin, D. A. Kareken, and C. Segebarth, "fmri of emotional responses to odors:: influence of hedonic valence and judgment, handedness, and gender," *Neuroimage*, vol. 20, no. 2, pp. 713–728, 2003.

[13] S. Delplanque, D. Grandjean, C. Chrea, L. Aymard, I. Cayeux, B. Le Calve, M. I. Velazco, K. R. Scherer, and D. Sander, "Emotional processing of odors: evidence for a nonlinear relation between pleasantness and familiarity evaluations," *Chem. Senses*, vol. 33, no. 5, pp. 469–479, 2008.

[14] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "Adjust: An automatic eeg artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2011.

[15] J. Bruhn, H. Röpcke, and A. Hoeft, "Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane anesthesia," *J. Am. Soc. Anesthesiologists*, vol. 92, no. 3, pp. 715–726, 2000.

[16] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Conf. Proc. 22nd ACM Knowl. Discov. Data Min. (KDD '16)*, 2016, pp. 785–794.

[17] Y. Benjamini and D. Yekutieli, "False discovery rate–adjusted multiple confidence intervals for selected parameters," *J. Am. Stat. Assoc.*, vol. 100, no. 469, pp. 71–81, 2005.

[18] E. Mas-Herrero, P. Ripollés, A. HajiHosseini, A. Rodríguez-Fornells, and J. Marco-Pallarés, "Beta oscillations and reward processing: coupling oscillatory activity and hemodynamic responses," *Neuroimage*, vol. 119, pp. 13–19, 2015.

[19] S. Lemm, B. Blankertz, G. Curio, and K.-R. Muller, "Spatio-spectral filters for improving the classification of single trial eeg," *IEEE. Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1541–1548, 2005.