Neural correlates of the food/non-food visual distinction

Kleovoulos Tsourides[1], Shahriar Shariat[2], Hossein Nejati[1,4], Tapan K. Gandhi[5], Annie Cardinaux[1], Christopher T. Simons[3], Ngai-Man Cheung[4], Vladimir Pavlovic[2] and Pawan Sinha[1]

[1]Massachusetts Institute of Technology
Cambridge, MA 02139

[2]Rutgers State University of New Jersey
Piscataway, NJ 08854

[3]Ohio State University
Columbus, OH 43218

[4]Singapore University of Technology and Design
Singapore, 138682

[5]Indian Institute of Technology
Delhi, New Delhi, 110016

Abstract

An evolutionarily ancient skill we possess is the ability to distinguish between food and non-food. Our goal here is to identify the neural correlates of visually driven ‘edible-inedible’ perceptual distinction. We also investigate correlates of the finer-grained likability assessment. Our stimuli depicted food or non-food items with sub-classes of appealing or unappealing exemplars. Using data-classification techniques drawn from machine-learning, as well as evoked-response analyses, we sought to determine whether these four classes of stimuli could be distinguished based on the patterns of brain activity they elicited. Subjects viewed 200 images while in a MEG scanner. Our analyses yielded two successes and a surprising failure. The food/non-food distinction had a robust neural counterpart and emerged as early as 85ms post-stimulus onset. The likable/non-likable distinction too was evident in the neural signals when food and non-food stimuli were grouped together, or when only the non-food stimuli were included in the analyses. However, we were unable to identify any neural correlates of this distinction when limiting the analyses only to food stimuli. Taken together, these positive and negative results further our understanding of the substrates of a set of ecologically important judgments and have clinical implications for conditions like eating-disorders and anhedonia.

Background and Significance

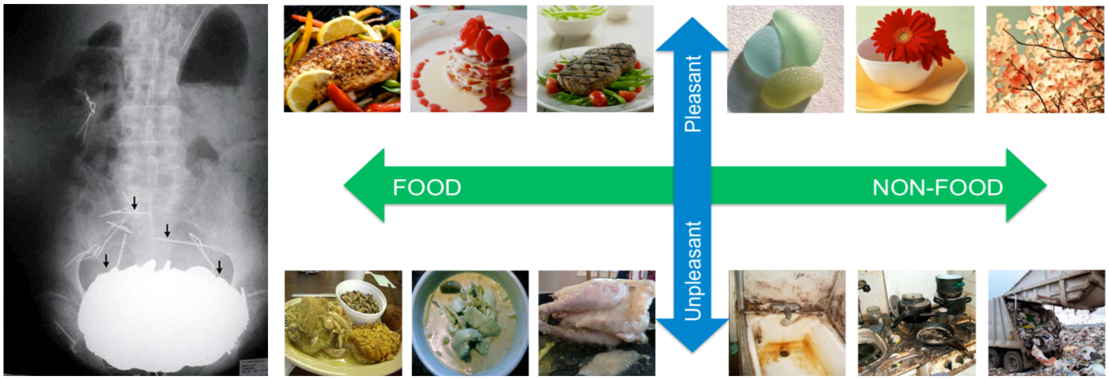
The ability to visually distinguish between food and non-food is critical for our survival. Disruption of this ability, as in cases of pica, an eating disorder characterized by persistent ingestion of nonnutritive substances (figure 1a), can have catastrophic consequences (Francois & Brenet, 2004). On a finer grain, we are also able to make rapid hedonic judgments about food. Given the ecological significance of these distinctions, identifying their neural correlates can yield important benefits. The spatial localization and time course of emergence of these distinctions in brain recordings can provide insights into the underlying processes involved in making the perceptual judgments (Thorpe, Fize, & Marlot, 1996), and also serve as biomarkers for neurological conditions involving anomalous responses to foods and non-foods.

Food selection is guided primarily by the visual, olfactory and taste systems. Past visual studies of hedonic perception in the context of food have focused largely on the relationship between a food’s caloric content and its perceived palatability. Palatability is found to be a useful cue for separating foods with high and low caloric contents, and also edible from non-edible items (Ohla, Toepel, le Coutre, & Hudry, 2012). Additional neuro-imaging studies have focused on the neural correlates in visual processing of food images within the context of rare syndromes such as Prader-Willi (Key & Dykens, 2008), anorexia and bulimia (Blechert, Feige, Joos, Zeeck, & Tuschen-Caffier, 2011). More broadly, several studies have examined the neural correlates of aesthetic and affective preferences, but the stimuli they have used do not typically involve the food versus non-food distinction (Amrhein, Mühlberger, Pauli, & Wiedemann, 2004; Jacobs, Renken, & Cornelissen, 2012; Kawabata & Zeki, 2004; Olofsson, Nordin, Sequeira, & Polich, 2008; Osaka, Ikeda, Rentschler, & Osaka, 2007; Schupp, Junghöfer, Weike, & Hamm, 2004) and hence cannot be used to infer the neural correlates of this specific distinction. FMRI studies as in (van der Laan, de Ridder, Viergever, & Smeets, 2011) focus on satiety and its modulation. To the best of our knowledge, no electrophysiological studies thus far have compared the classes of food and non-food imagery while controlling for affective dimensions. Additionally, there appear to be no studies that focus on food images without a confound of calorie and palatability contents or attentional bias (Bradley et al., 2003; P. A. Gable & Harmon-Jones, 2010; P. Gable & Harmon-Jones, 2008; Harmon-Jones, Gable, & Price, 2011). Given this background, no firm consensus has emerged regarding the neural markers corresponding to the perceptual distinction between pictures of foods and non-foods.

Our goal in this study is to employ computationally sophisticated pattern classification techniques to identify such correlates. A key question for us is whether our visual system exhibits responses to food stimuli that are different from non-food ones regardless of the level of pleasantness or affective valence. Furthermore, we examine whether the neural response differences across stimulus categories can be accounted for simply via systematic variations in low-level properties of an image such as color distributions and textural statistics.

Past research on neural correlates of visual categorization has focused on identifying components in electrophysiological data (EEG or MEG) corresponding to object classes such as faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996). These studies serve to contextualize our work and the methods we use. Specifically, although the correlates of the face/non-face distinction are generally accepted (although not without dissent, see Thierry, Martin, Downing, & Pegna, 2007), subtler perceptual distinctions (such as gender, age, familiarity) have been harder to identify in neural data. We believe that part of this difficulty may arise from the limitations of conventional data-analysis techniques. In particular, the evoked response field (ERF) type of analysis, which requires averaging of multiple temporally aligned signal fragments from one or a few sensors, is not well suited to picking up on distributed patterns of neural activity that may correspond to a perceptual judgment. A more 'agnostic' data classification approach drawn from the domain of machine learning may be better suited for this purpose. The dimension of like/dislike has also been examined by a few neuroimaging studies. For instance, Healey, Morgan, Musselman, Olino, & Forbes, 2014, have implicated activity in medial pre-frontal cortex in anhedonia in the social context. We have the opportunity to build on these results in two significant ways. First, we can explore the like/dislike dimension in a non-social setting and, second, through the use of electrophysiological recordings, we can obtain more precise temporal information about the onset of the neural distinction.

We used magneto-encephalography (MEG) to record brain activity elicited in response to two categories of visual stimuli: images depicting foods and non-foods. Each of these categories was further subdivided into two equal-sized classes, differing in their hedonic valence (positive and negative). Figure 1 shows examples of the stimuli we used. We recorded brain activity from 306 sensors distributed across the scalp while subjects passively viewed all 200 of these stimuli in random order. These continuous traces were subsequently segmented into 1 second epochs, temporally aligned to the onset of each stimulus. The collection of these segmented traces was then subjected to pattern classification analyses using techniques drawn from the domain of machine learning, as well as to conventional evoked response field (ERF) analyses common in the EEG domain (Niedermeyer & Silva, 2004; Vecchiato et al., 2011).



(a) (b)

Figure 1. *(a) X-ray of a 62 year old French man suffering from pica. The patient had ingested over 350 coins, needles and necklaces (from Francois, 2004). (b) Sample stimuli used in our experiment. The top row shows food stimuli with the left panel comprising images that were rated as being more palatable*

relative to those on the right. The lower row shows a few non-food stimuli segregated into pleasant (left) and unpleasant (right) subclasses.

Our pattern classification analyses used sparse logistic regression to classify raw MEG signals corresponding to the different image categories. The classifier was provided the first 1000 ms of all magnetometer signals, without any ad-hoc sensor selection. In order to determine information available for classification in different time epochs, we used a 10 ms sliding window over the signals, shifting this window 1 ms at a time. Our classifier therefore receives 10 ms worth of data from all sensors in each step. Furthermore, motivated by the use of resting state signals for reducing the signal noise, we used the first 100 ms baseline (the resting state and before the start of the trigger) as an additional source of training for the classifier, resulting in improvements in classification performance. Details of our classification approach are described in the Methods section.

Methods

Stimuli

Full-color images were chosen by 20 volunteers from multiple image repositories and culinary websites. The volunteers rated each image in terms of its hedonic valence. 50 images in each of the four classes (food: appetitive, food: non-appetitive, non-food: pleasant, non-food: unpleasant) that received high or low scores most consistently across the raters were then used to constitute the final stimulus set. Images were processed to all have the same mean luminance and size. Descriptions of all of the images we used are provided in the supplementary material. (Although they were not available when we commenced our study, it is worth pointing out that two sets of publicly accessible image databases have recently been created (Foroni et al., 2013; Blechert et al., 2014). Such databases can play an important role in providing a common base to be able to compare results from disparate studies.)

Subjects

14 healthy subjects including 6 men and 8 women (ages: 18-27, mean age 22) participated in this study. Subjects were all within normal BMI range (21-26). This healthy BMI range is in part due to constraints posed by the experimental equipment; our MEG machine cannot accommodate individuals weighing above 190 pounds. All subjects were undergraduate or graduate students. One subject was a professional culinary artist. All had normal or corrected-to-normal visual acuity and none had a history of neurological or psychiatric disorders. Subjects had consumed their last meal at least 3 hours before the experiment took place. Each subject gave written informed consent according to procedures approved by the Massachusetts Institute's of Technology Institutional Review Board (MIT IRB).

No behavioral tasks were required. Subjects were instructed simply to maintain gaze on the screen, while trying to minimize movements and eye blinks. After participating in the experiment, participants rated the images. Ratings were based on category of images (into food and non-food classes), edibility for the food class, and pleasantness for the non-food class. We also obtained additional ratings from individuals who did not undergo MEG scanning. Responses regarding the categorical labeling of images across make evident the fact that the classification of images into the two categories was robust and consistent across observers.

MEG Recordings

Recordings were made using a recumbent Elekta Magneto-encephalography scanner with 306 sensors (102 magnetometers and 204 gradiometers). Images were back-projected on a screen placed 5 feet in front of the subject. Each image was presented for 350 ms; inter-stimulus interval time varied from 1750-2500 ms. Pre-processing included Elekta's proprietary artifact rejection filter, band-pass filtering between 0.5-40Hz and z-score. Brainstorm software was used to export the preprocessed signals to MATLAB for further analysis using customized scripts for ERF and classification analyses. For ERF-significance analysis, for each subject, we selected the sensors exhibiting the maximum on the global ERF average of all image stimuli classes. We analyzed using the average of the absolute values of each ERF potential (ie: Both ERFs in and out of the cortex. We performed t-test on all subjects, for each ms, and corrected for multiple comparisons using False Discovery Rate (FDR). Additionally, we performed 2x2 ANOVA.

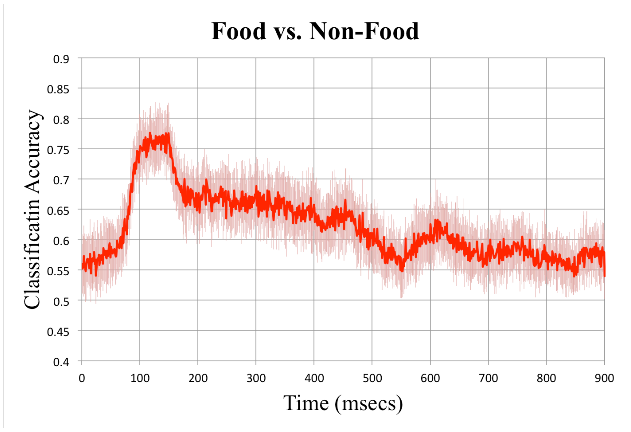
Classification analyses

We used Glmnet implementation of two-class logistic regressor (Friedman, Hastie, & Tibshirani, 2010). In each experiment, the classifier is trained on 10 ms worth of signal data from all MEG sensors. This window is then slid 1 ms at a time, producing classification performances for each time point. Note that at any given time point, the classifier only views the past 10ms of sensors to that time point. We divided the entire set of data to disjoint training and testing sets (60% and 40% respectively). In addition, we added 100 ms of the baseline to the training set to improve noise modeling. After training the classifier on the training set, we tested the resulting classification model on the testing set. We performed 4 folds cross-validation and reported the average performance. The performance of the classifier at each ms is then reported. For performance reporting we chose area under the curve (AUC). AUC was calculated by computing the area under the ROC curve created by changing the threshold of the regressor score from minimum to maximum (among all computed scores) and plotting true positive rate against false positive rate. The threshold for deciding whether the signal evoked by a particular image belongs to positive or negative class was found by using the ROC curve that resulted in the highest AUC and its best operating point. This point on the curve indicates the results in the least number of false positives and highest number of true positives. Reported classification accuracies are regarding this best operating point in each of the experiments.

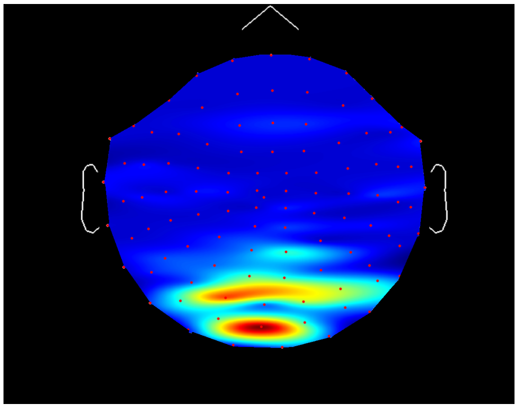
Results

Our results showed that accuracy of classification (food versus non-food) averaged across 14 subjects was at chance during the first 0-70 ms, but then rose to significantly above-chance levels. The discriminating information became evident first in the 70-130 ms epoch and persisted until at least 200 ms after the stimulus onset. Sensors in the occipital area were found to have the highest feature weights for the food vs. non-food distinction. We found similar results for the pleasant versus unpleasant distinction (without regard to the food and non-food categories). After hovering near chance for the first 0-70 ms, the classifier performance rises to well-above chance subsequently and accuracy remains high for the next 60 ms. However, the sensors receiving the highest weights from the classifier in this setting were right lateralized towards the anterior occipital and posterior temporal regions. Figures 2 summarizes the results. Plots at (a), (c), and (e) to (j) in figure 2 represent classifier's performance in discriminating between the two classes of MEG signals. For example, in figure 2a, the classifier is discriminating the class of signals evoked by viewing food images vs. the class of signals evoked by viewing non-food images. Having this performance plot, we can derive sensors that the classifier relies on for the classification. It should be noted that the higher the performance of the classifier, the more reliable sensor weights given by the classifier. These sensors weights can be projected back to the 3D sensor space, representing the cortical regions that are most significantly different between the two classes of signals. For example, the sensor weight map in figure 2b represents weights assigned by the classifier to MEG sensors, indicating their importance for discrimination between signals evoked by viewing food images vs. signals evoked by viewing non-food images, at the time of highest classification accuracy. Similarly, figure 2d illustrates the respective sensors weights for discrimination between signals evoked by pleasant vs. unpleasant images. Finally, it is important to note that dipolar patterns presented here are not an artifact of manual selection of sensors since the selection is based on classifier-derived weights.

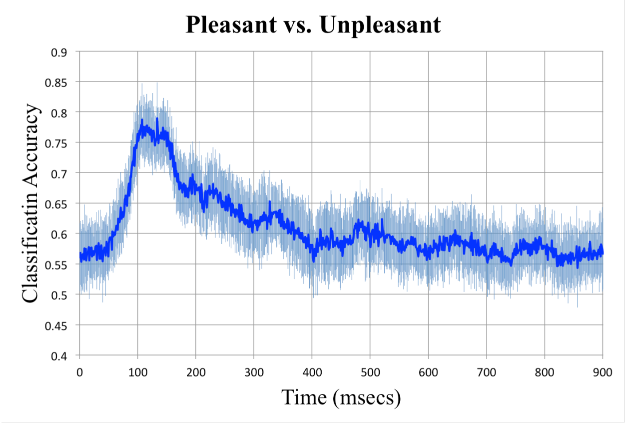
It is interesting to note that classifier-based analysis is able to find information in the neural signals capable of distinguishing food from non-food, pleasant from unpleasant items, and pleasant non-foods from *un*pleasant non-foods. However, it fails at the appetizing foods versus *un*appetizing foods distinction. Apparently, our scalp based neural measurements are insufficient to pick up information that would allow us to distinguish between the valences of different food stimuli. This failure could potentially also arise from our not using a powerful enough classifier for the task. However, it is worth noticing that the classifier does succeed at the other distinctions, attesting to the possibility that the within-food distinction might be an inherently harder classification task than the other ones.



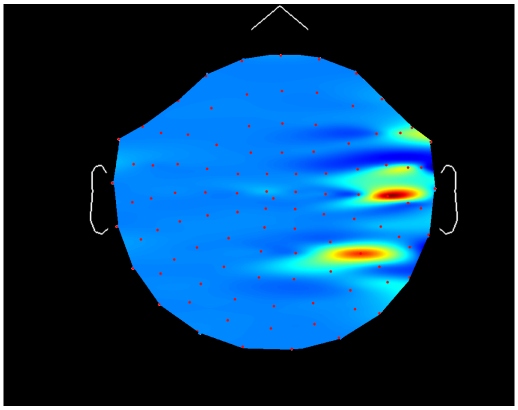
(a)



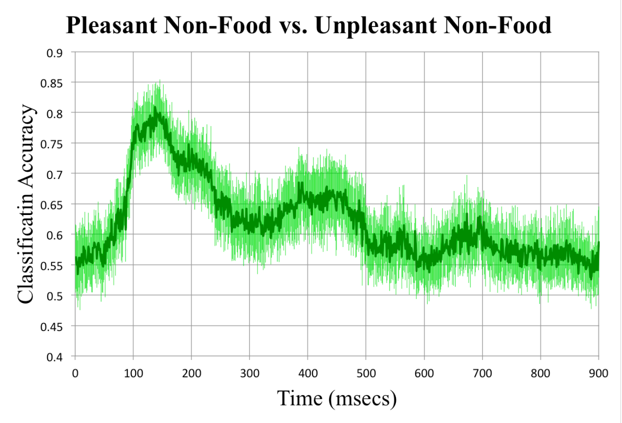
(b)



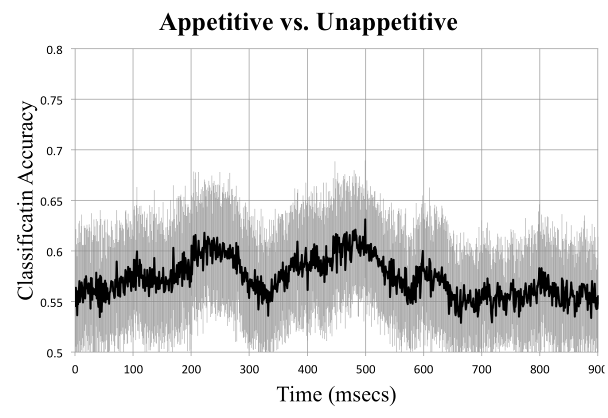
(c)



(d)



(e)



(f)

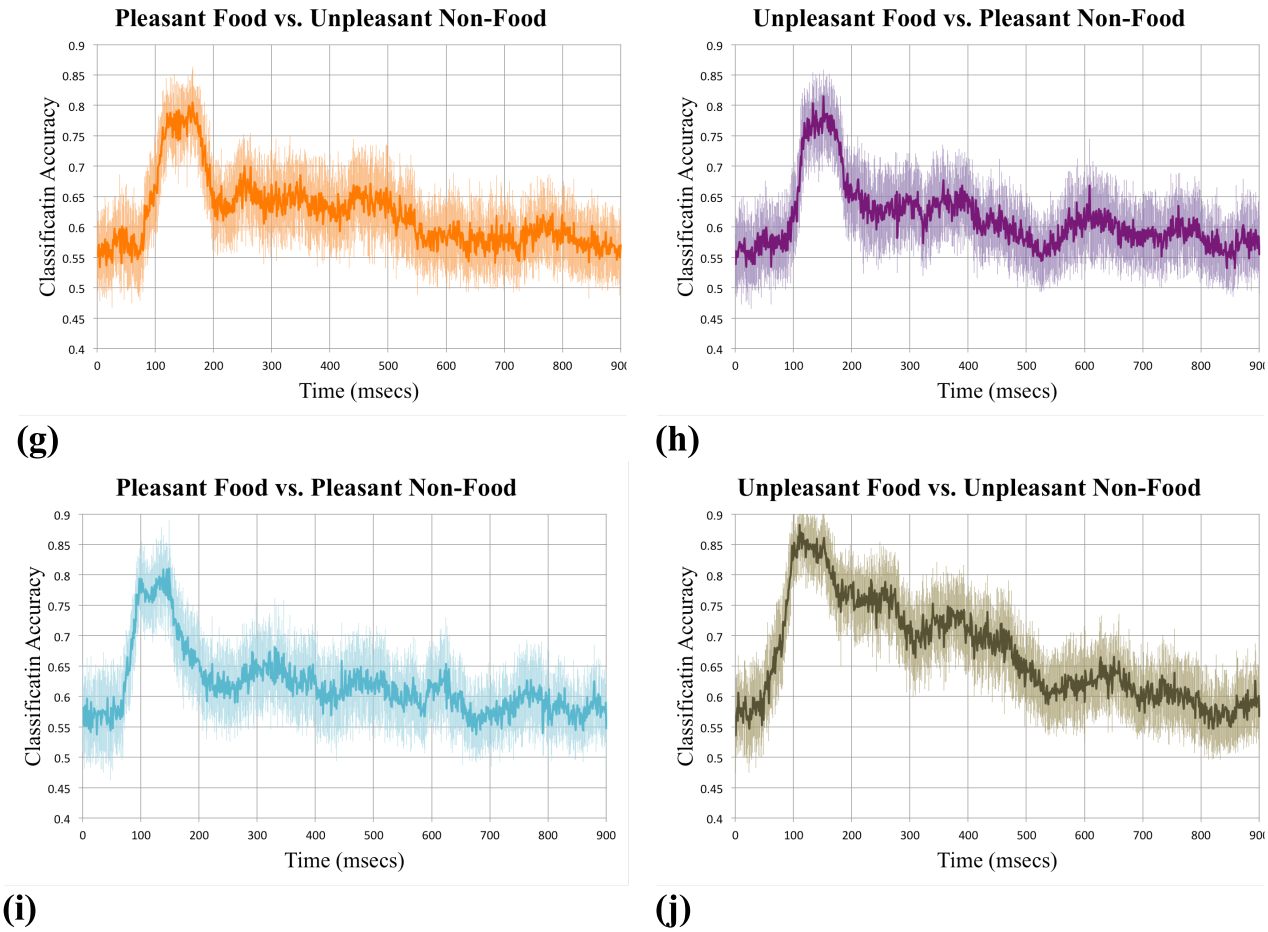
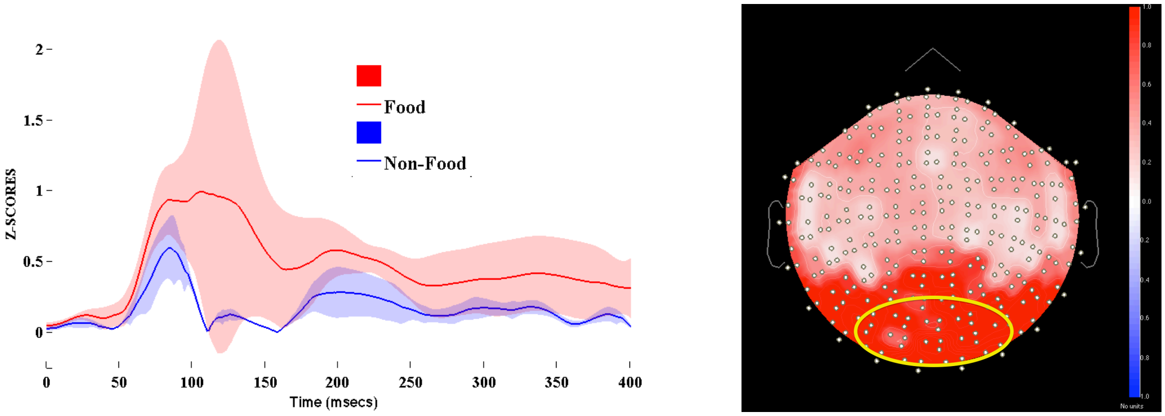


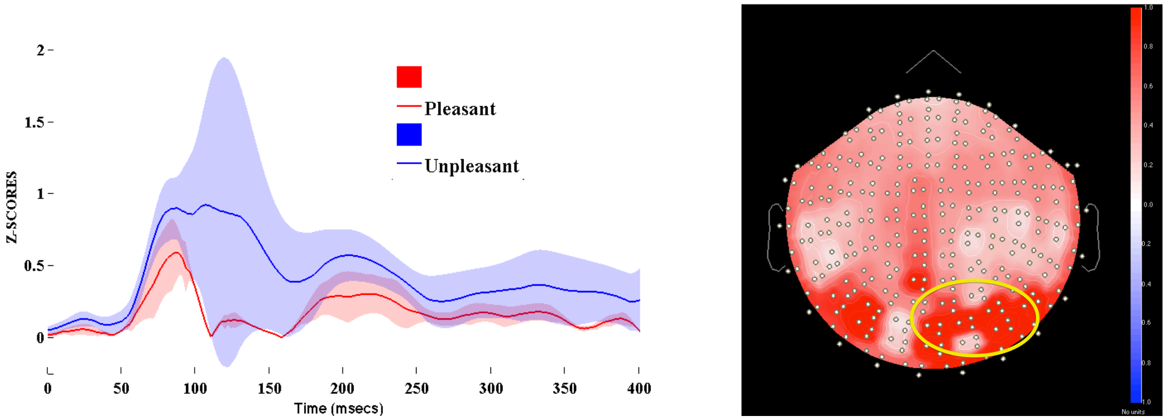
Figure 2. *Bar plots of recognition rates (averaged across 14 subjects) in different time-bins. The four bar-plots shown here are: (a) Classification accuracy of the food vs. non-food distinctions. Also shown (b) is the spatial distribution of feature weights as derived from the classifier analysis. Most of the high weights are concentrated in the occipital sensors. (c) Classification accuracy of the pleasant vs. unpleasant distinction. (d) The spatial distribution of feature weights as derived from the classifier analysis. Most of the high weights are concentrated in the right occipital and temporal sensors (e) Classification accuracy of the pleasant non-food vs. unpleasant non-food distinction. (f) Classification accuracy of the pleasant food vs. unpleasant food distinction. (g) Classification accuracy of the pleasant food vs. unpleasant non-food distinction. (h) Classification accuracy of the unpleasant food vs. pleasant non-food distinction. (i) Classification accuracy of the pleasant food vs. pleasant non-food distinction. (j) Classification accuracy of the unpleasant food vs. unpleasant non-food distinction. The inputs to the classifier are 1000ms of all magnetometer and gradiometer signals derived across the whole head.*

Guided by these classifier-based results, we sought to determine whether any components might be evident in conventional evoked response analyses that would be different for the food/non-food classes and the pleasant/unpleasant classes. One should note that the classification weights might be different from actual evoked responses, as the classifier may leave a highly active sensor that carries redundant information, for a relatively less active sensor that carries more discriminative information.

For the first distinction (food versus non-food), we focused on the 70 to 130 ms time-bin since that is when classification accuracy first reached above-chance levels. Also, our ‘region of interest’ was the occipital area since that is where the classifier assigned the highest feature weights (figure 2b). Our evoked response field (ERF) analyses did indeed reveal differences between the responses elicited by the group of food images on the one hand and non-food images on the other. As figure 3a shows, statistically significant differences ($p<0.05$) in the evoked responses become evident within the first 100 ms after the onset of the stimulus. Using similar guidance from the classifier results for the pleasant versus unpleasant distinction, we selected right occipital sensors and found significant differences ($p<0.05$) in evoked responses at approximately 100ms post-stimulus onset (figure 3b). Although classification performances in figure 2 and ERF plots in figure 3 are indicating the same basic phenomenon, is important to differentiate between these two; while the former illustrates success rate of the classifier in discriminating between the two classes of signals at a particular time, the later directly illustrates the (averaged) differences between the two signals. Interestingly, an ANOVA revealed no statistically significant differences across any of the stimulus groups. This contrast between the significantly above chance classification performance using machine learning techniques and the seeming indistinguishability of the groups with conventional ERF style analyses makes an important methodological point. T-Test ERF analyses (complemented with ANOVAs) are limited in terms of their ability to use signal information that may be distributed across space and time. Classification analyses, on the other hand are not subject to such a constraint and are able to pick up on distinctions that may be spatio-temporally non-local. Furthermore, the latter approach obviates the need for preconceived notions of ROIs; the classifier can examine all of the sensor data 'agnostically' and determine which sensors contribute the most to distinguishing between classes. We believe that the results we have presented here illustrate these distinctions between analytical approaches and highlight the potential benefits of machine learning tools for mining high-dimensional human electrophysiology data.



(a)



(b)

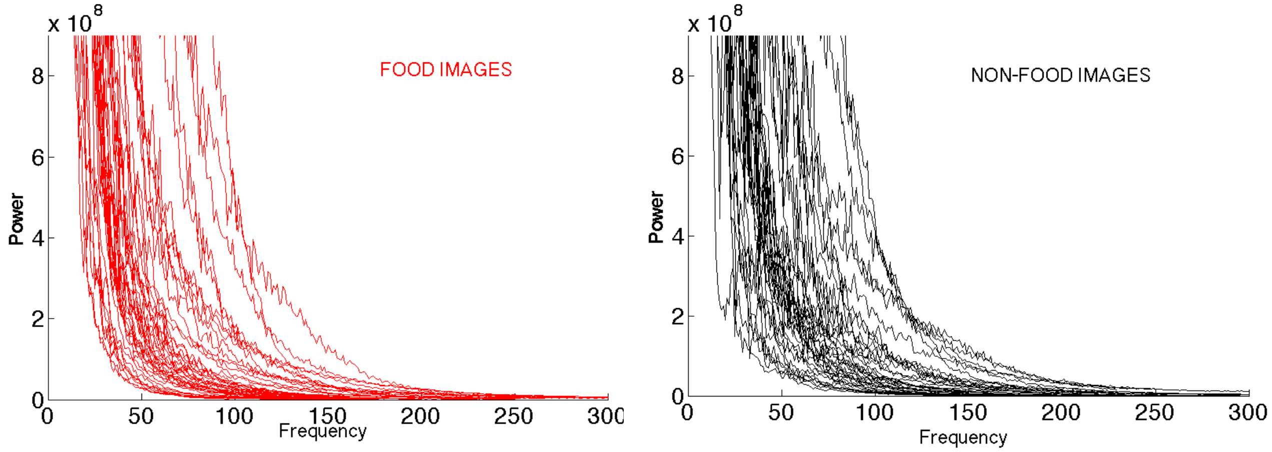
Figure 3. *ERF results for two distinctions: (a) Food versus non-food, and (b) Pleasant versus unpleasant. The left panels display evoked responses and the right panels indicate the subset of sensors used for their computation.*

An interesting aspect of the results, from both the classifier as well as evoked-response analyses, is the rapidity with which the perceptual distinctions become evident in neural signals. Given that the latency of primary visual cortical response is approximately 50-60 ms (Poghosyan & Ioannides, 2007), these findings suggest that the food versus non-food perceptual distinction can be made within 50 ms of the arrival of the information in V1. (Interestingly, similarly rapid responses have been reported in the domain of face-odor conditioning by Junghofer and colleagues (Steinberg et al., 2012)). The observed rapidity of neural discrimination raises questions about the underlying mechanisms responsible for the distinction. Various possibilities exist. For instance, One possibility is that rapid visual analysis may be performed sub-cortically (LeDoux, 1996); the early distinctions we see in our data could be manifestations of the output of such analyses. While plausible, we consider this account unlikely given the subtle visual signals that need to be analyzed to make the distinctions in this study (see also Cauchoix & Crouzet, 2013). Alternatively, the observed responses over visual cortex may be driven by feedback signals from reward related areas, such as mPFC (Tzschentke, 2000). This account is not very satisfying because it sidesteps the issue of how the initial visual analysis for driving the reward system is accomplished. Our data analysis should, in principle, have revealed such processing, as a precursor to activation in the mPFC and then in the visual cortex. Could the rapid processing of disgust account for our results (e.g. Smith, 2012)? While this is potentially applicable to accounting for the like/dislike results, it does not explain the rapid distinction observed between the food and non-food categories.

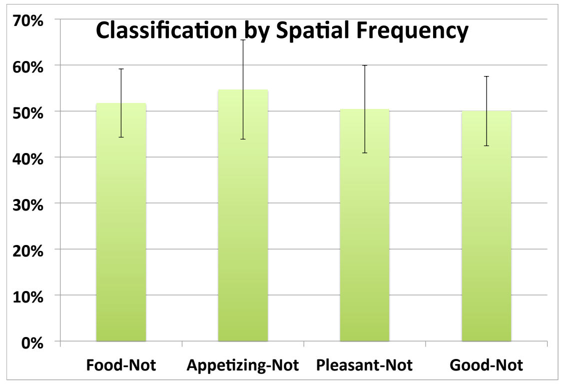
Another possibility is that the semantic distinction between food and non-food may in fact be based on some low-level image properties that can be analyzed rapidly. Two of the key perceptually salient low-level properties are spatial frequency content and color profile. Perhaps food images occupy a different section of the spatial-frequency and color spaces relative to non-food images, and hence can be rapidly distinguished from the latter. To test this possibility, we have characterized the spatial-frequency content and color distributions of the four stimulus categories and subjected the characterizations (power spectra and color histograms) to the same kinds of classifier analyses as those that we used for the MEG data. Figure 4 shows the results. The classifier is unable to distinguish between the different semantic groups based on the frequency and color information. This suggests that the stimulus categories do not have systematic differences along these dimensions that would permit a trivial low-level distinction between them.

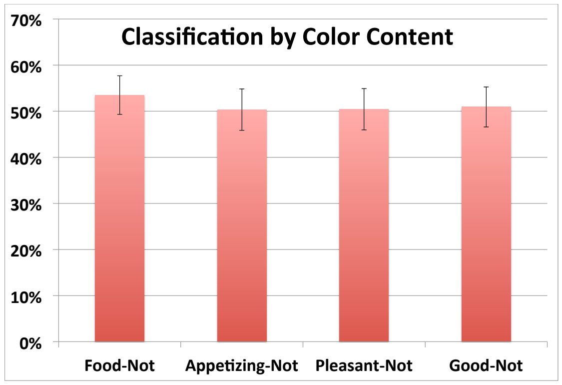
In order to examine the hypothesis of low-level factors further, we investigated whether popular image-descriptors drawn from the domain of computer vision could provide different signatures for the image classes we have used here. Histograms of oriented Gradients around keypoint detectors and related features have been amongst the most popular and powerful representational strategies for encoding spatial content and arrangement (Bay, Ess, Tuytelaars, & Van Gool, 2008; Csurka, Dance, Fan, Willamowski, & Bray, 2004; Dalal & Triggs, 2005; Lowe, 2004). We first extracted SIFT and SURF features from each image in our stimulus set. We then proceeded by populating the features in a Bag of Words Model (BOW) (Csurka et al., 2004; Van De Sande, Gevers, & Snoek, 2010) to create a 'visual' dictionary that could hypothetically characterize the 'lexicon' of an image containing edible or inedible content. We then trained an SVM classifier based on the L1 histogram distance of these visual lexicons for the images. In addition to SIFT and SURF, we also tested the possible discriminative power of image edges, by testing classification performance on Gabor features (Daugman, 1985; Jain, Ratha, & Lakshmanan, 1997). Simple cells in the visual cortex of mammalian brains can be modeled by Gabor functions (Jain et al., 1997), thus it is thought that image analysis by Gabor functions might mimic some aspects of perception in the human visual system. We extracted features using a Gabor bank of 8 scales x 8 directions and used the VLFeat implementation of SVM (Vedaldi & Fulkerson, 2008). Classifier analysis based on these three sets of features was unable to distinguish between image classes based on this encoding of image content (figures 4 (d),(e), and (f)).

These results suggest that rapidly emerging distinctions between different stimulus categories we have observed in neural signals are likely driven by a sophisticated analysis of image content, rather than by simple low-level features of the kind we have described above. An important open question for future studies concerns the 'signal substrate' for such semantic analyses. Of particular relevance here is past work showing that low spatial frequencies in an image may undergo rapid analysis and thereby generate hypotheses that guide and constrain the interpretation of higher spatial frequency content (Bar et al., 2006; Ahlfors et al., 2015). In the context of our work, this points to the need for determining whether category distinctions can be made even with low spatial frequencies. If psychophysical tests demonstrate that this is possible, then_coupled with results from past studies, we would have a candidate mechanism for explaining the rapid neural distinctions between stimulus classes.



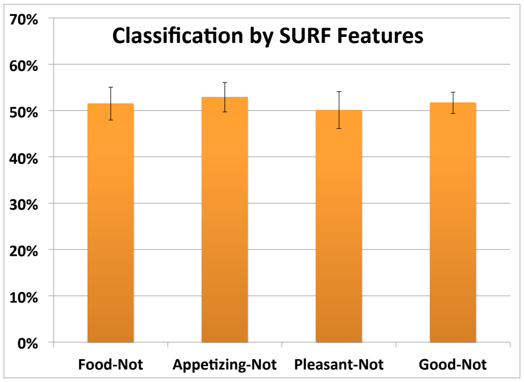
(a)



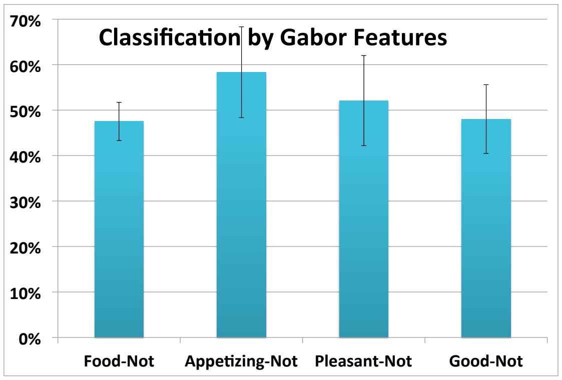


(b)

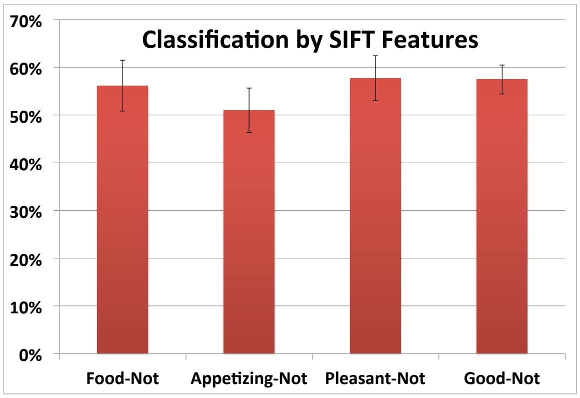
(c)



(d)



(e)



(f)

Figure 4. *Classification results using image content. Overall spatial frequency content, shown in (a), does not explain the observed differences. As illustrated in (b), classification accuracy is indistinguishable from chance using spectral content as the input. (c) Classification based on color distributions. (d),(e) and (f) Classification based on SURF, Gabor and SIFT features. All image classes were indistinguishable on these dimensions.*

Taken together, these results present a robust neural correlate of the food versus non-food distinction and also demonstrate how classification-based analytical techniques can complement conventional evoked response ones. The neuronal responses appear to encode this distinction rapidly and robustly. This poses an interesting challenge of explaining the genesis of the rapid discrimination. Intuitively appealing low-level factors, such as spatial frequency content, color distributions and basic characterizations of spatial features, seem not to be adequate to account for the observed differences. While this does not definitively rule out a low-level factor as underlying the observed neural differences, it makes the hypothesis less likely. This leads us to favor an explanation that involves the extraction of basic semantic information (whether the stimulus depicts an edible or non-edible item) based on a combination of multiple low-level cues, none of which is independently capable of supporting the judgment. Even as we search for the underlying factors, the neural correlates that we have described here can prove useful for probing classification strategies used by the brain, their differences across sub-groups of healthy participants and the changes brought about by neurological conditions that manifest as anhedonia on the one extreme (that can in serious cases lead to a failure to thrive due to a lack of interest in food) and an eating disorder marked by excessive consumption and resulting obesity, on the other.

References

Ahlfors, S. P., Jones, S. R., Ahveninen, J., Hämäläinen, M. S., Belliveau, J. W., Bar, M. (2015). Direction of magnetoencephalography sources associated with feedback and feedforward contributions in a visual object recognition task. Neurosci Lett. 2015 Jan 12;585:149-154.

Amrhein, C., Mühlberger, A., Pauli, P., & Wiedemann, G. (2004). Modulation of event-related brain potentials during affective picture processing: A complement to startle reflex and skin conductance response? *International Journal of Psychophysiology*, *54*(3), 231–240. doi:10.1016/j.ijpsycho.2004.05.009

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, *110*(3), 346–359. doi:10.1016/j.cviu.2007.09.014

Bar, M., Kassam, L. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., Hamalainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., Halgren, E. (2006). Top–down facilitation of visual recognition, *Proc. Natl. Acad. Sci. U. S. A.* 103, 449–454.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*, 551–565. doi:10.1162/jocn.1996.8.6.551

Blechert, J., Feige, B., Joos, A., Zeeck, A., & Tuschen-Caffier, B. (2011). Electrocortical processing of food and emotional pictures in anorexia nervosa and bulimia nervosa. *Psychosomatic Medicine*, *73*(5), 415–421. doi:10.1097/PSY.0b013e318211b871

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: an image database for experimental research on eating and appetite. *Frontiers in Psychology,* 5. doi: 10.3389/fpsyg.2014.00617.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience*, *117*, 369–380. doi:10.1037/0735-7044.117.2.369

Cauchoix, M., & Crouzet, S. M. (2013). How plausible is a subcortical account of rapid visual recognition? *Front. Hum. Neurosci.*, *7*, 1–4. doi:10.3389/fnhum.2013.00039

Csurka, G., Dance, C., Fan, L. X., Willamowski, J., & Bray, C. (2004). isual categorization with bags of keypoints. In *Proc. of ECCV International Workshop on Statistical Learning in Computer Vision* (pp. 1–22).

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, *1*. doi:10.1109/CVPR.2005.177

Daugman, J. . (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*(7), 1160–1169.

Foroni, F., Pergola, G., Argiris, G., & Rumiati, R. I. (2013). The FoodCast research image database (FRIDa). *Front Hum Neurosci*, 7, 51. doi: 10.3389/fnhum.2013.00051.

Francois, B., & Brenet, O. (2004). A medical mystery. *The New England Journal of Medicine*, *350*, 59–59.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1), 1–22. doi:10.1359/JBMR.0301229

Gable, P. A., & Harmon-Jones, E. (2010). Late positive potential to appetitive stimuli and local attentional bias. *Emotion (Washington, D.C.)*, *10*, 441–446. doi:10.1037/a0018425

Gable, P., & Harmon-Jones, E. (2008). Relative left frontal activation to appetitive stimuli: Considering the role of individual differences. *Psychophysiology*, *45*, 275–278. doi:10.1111/j.1469-8986.2007.00627.x

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2011). Leaning embodies desire: Evidence that leaning forward increases relative left frontal cortical activation to appetitive stimuli. *Biological Psychology*, *87*, 311–313. doi:10.1016/j.biopsycho.2011.03.009

Healey, K. L., Morgan, J., Musselman, S. C., Olino, T. M., & Forbes, E. E. (2014). Social anhedonia and medial prefrontal response to mutual liking in late adolescents. *Brain and Cognition*, (89), 39–50. doi:10.1016/j.bandc.2013.12.004

Jacobs, R. H. a H., Renken, R., & Cornelissen, F. W. (2012). Neural correlates of visual aesthetics--beauty as the coalescence of stimulus and internal state. *PloS One*, *7*, e31248. doi:10.1371/journal.pone.0031248

Jain, A. K., Ratha, N. K., & Lakshmanan, S. (1997). Object detection using gabor filters. *Pattern Recognition*, *30*, 295–309. doi:10.1016/S0031-3203(96)00068-4

Kawabata, H., & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, *91*, 1699–1705. doi:10.1152/jn.00696.2003

Key, A. P. F., & Dykens, E. M. (2008). "Hungry Eyes": Visual processing of food images in adults with Prader-Willi syndrome. *Journal of Intellectual Disability Research*, *52*, 536–546. doi:10.1111/j.1365-2788.2008.01062.x

LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* (1st Editio.). New York, NY: Simon and Schuster.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*, 91–110. doi:10.1023/B:VISI.0000029664.99615.94

Niedermeyer, E., & Silva, F. H. L. Da. (2004). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. *Lippincott Williams and Wilkins* (Vol. 1, p. 1309). Retrieved from http://books.google.com/books?id=tndqYGPHQdEC&pgis=1

Ohla, K., Toepel, U., le Coutre, J., & Hudry, J. (2012). Visual-gustatory interaction: orbitofrontal and insular cortices mediate the effect of high-calorie visual food cues on taste pleasantness. *Plos One*, *7*(3), e32434–e32434. doi:10.1371/journal.pone.0032434

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247–265. doi:10.1016/j.biopsycho.2007.11.006

Osaka, N., Ikeda, T., Rentschler, I., & Osaka, M. (2007). PPA and OFC correlates of beauty and ugliness: An event – related fMRI study. *Perception*, *36*. doi:10.1068/v070067

Poghosyan, V., & Ioannides, A. A. (2007). Precise mapping of early visual responses in space and time. *NeuroImage*, *35*, 759–770. doi:10.1016/j.neuroimage.2006.11.052

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology*, *41*, 441–449. doi:10.1111/j.1469-8986.2004.00174.x

Smith, M. L. (2012). Rapid processing of emotional expressions without conscious awareness. *Cerebral Cortex*, *22*, 1748–1760. doi:10.1093/cercor/bhr250

Steinberg, C., Dobel, C., Schupp, H. T., Kissler, J., Elling, L., Pantev, C., & Junghöfer, M. (2012). Rapid and Highly Resolving: Affective Evaluation of Olfactorily Conditioned Faces. *Journal of Cognitive Neuroscience*, *24*(1), 17–27. doi:10.1162/jocn_a_00067

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, *10*, 505–511. doi:10.1038/nn1864

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. doi:10.1038/381520a0

Tzschentke, T. M. (2000). The medial prefrontal cortex as a part of the brain reward system. *Amino Acids*, *19*(1), 211–219. doi:10.1007/s007260070051

Van De Sande, K., Gevers, T., & Snoek, C. (2010). Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1582–1596.

Van der Laan, L. N., de Ridder, D. T. D., Viergever, M. A., & Smeets, P. A. M. (2011). The first taste is always with the eyes: A meta-analysis on the neural correlates of processing visual food cues. *NeuroImage*, *55*(1), 296–303.

Vecchiato, G., Astolfi, L., De Vico Fallani, F., Toppi, J., Aloise, F., Bez, F., … Babiloni, F. (2011). On the Use of EEG or MEG brain imaging tools in neuromarketing research. *Computational Intelligence and Neuroscience*.

Vedaldi, A., & Fulkerson, B. (2008). VLFeat: An Open and Portable Library of Computer Vision Algorithms.