

Early differential processing of material images: Evidence from ERP classification

Christiane B. Wiebel

Abteilung Allgemeine Psychologie, Universität Gießen,
Giessen, Germany



Matteo Valsecchi

Abteilung Allgemeine Psychologie, Universität Gießen,
Giessen, Germany



Karl R. Gegenfurtner

Abteilung Allgemeine Psychologie, Universität Gießen,
Giessen, Germany



Investigating the temporal dynamics of natural image processing using event-related potentials (ERPs) has a long tradition in object recognition research. In a classical Go-NoGo task two characteristic effects have been emphasized: an early task independent category effect and a later task-dependent target effect. Here, we set out to use this well-established Go-NoGo paradigm to study the time course of material categorization. Material perception has gained more and more interest over the years as its importance in natural viewing conditions has been ignored for a long time. In addition to analyzing standard ERPs, we conducted a single trial ERP pattern analysis. To validate this procedure, we also measured ERPs in two object categories (people and animals). Our linear classification procedure was able to largely capture the overall pattern of results from the canonical analysis of the ERPs and even extend it. We replicate the known target effect (differential Go-NoGo potential at frontal sites) for the material images. Furthermore, we observe task-independent differential activity between the two material categories as early as 140 ms after stimulus onset. Using our linear classification approach, we show that material categories can be differentiated consistently based on the ERP pattern in single trials around 100 ms after stimulus onset, independent of the target-related status. This strengthens the idea of early differential visual processing of material categories independent of the task, probably due to differences in low-level image properties and suggests pattern classification of ERP topographies as a strong instrument for investigating electrophysiological brain activity.

Introduction

Material perception research has received increasing interest over the last decade. The recognition of materials and material properties is important as it helps us to interact properly with our environment. For instance when we plan to grasp a slippery object, we have to visually assess not only the object's geometry but also its surface properties. Despite the ecological importance of material class recognition and of the assessment of material properties, the influence of complex natural surfaces on various visual tasks remained uninvestigated for a long time (Maloney & Brainard, 2010). Most studies on visual perception have used simple flat and matte stimuli, ignoring the complexity of real world surfaces (Brainard & Maloney, 2004; Maloney & Brainard, 2010). Thus, it is necessary to examine how we process information about materials and how this information interacts with other factors that form our visual perception of the world around us. Recently, the topic has gained more attention and a number of studies have investigated some of these issues. Examples include color perception for real objects made out of different materials (Giesel & Gegenfurtner, 2010); color categorization using real world surfaces and real illuminants (Olkkonen, Witzel, Hansen, & Gegenfurtner, 2010); the influence of different light fields on the perception of different material properties like gloss or roughness (Fleming, Dror, & Adelson, 2003); interactions between different surface material properties, for example gloss and three-dimensional (3-D) texture (Ho, Landy, & Maloney, 2008); and interactions between material classification and judgments of material qualities in the visual and semantic domain (Fleming, Wiebel, &

Citation: Wiebel, C. B., Valsecchi, M., & Gegenfurtner, K. R. (2014). Early differential processing of material images: Evidence from ERP classification. *Journal of Vision*, 14(7):10, 1–13, <http://www.journalofvision.org/content/14/7/10>, doi:10.1167/14.7.10.

doi: 10.1167/14.7.10

Received October 28, 2013; published June 24, 2014

ISSN 1534-7362 © 2014 ARVO

Gegenfurtner, 2013). Moreover, a number of studies have dealt with the question what kind of information in an image is related to the perception of certain material qualities, for example gloss (Kim & Anderson, 2010; Motoyoshi, Nishida, Sharan, & Adelson, 2007), or with building computational material classifiers (Liu, Sharan, Adelson, & Rosenholtz, 2010).

Within this context, one fundamental question is whether the time course of material processing within the visual system is fast enough to determine our interactions with our dynamical environment. In a behavioral study (Wiebel, Valsecchi, & Gegenfurtner, 2013), we investigated the speed and accuracy of material categorization. In agreement with Sharan (2009), we found material classification to be quite fast: Presentation durations of 30 ms or less were sufficient to produce above-chance classification performance. Here, we aimed at investigating the temporal characteristics of material perception in more detail, by taking advantage of the high temporal resolution of electrophysiological recordings. More precisely, we sought to determine a lower bound estimate of differential activity for the visual processing of different material categories. We also tried to find evidence for two different temporally distinct processing stages: an early sensory process and a later higher level processing stage (VanRullen & Thorpe, 2001). This imposes an interesting complement to prior work in material perception, especially to the neuroimaging literature that cannot provide an exact temporal framework of the underlying processes involved in material perception.

Electrophysiological recordings have a long-standing position in the investigation of the time course of perceptual and cognitive processes. Thorpe, Fize, and Marlot (1996) used electrophysiological recordings in a Go-NoGo animal detection task to explore the temporal processing of complex natural images. They contrasted the brain activity elicited on correct Go trials with that of correct NoGo trials to get a measure of how long it takes the visual system to reliably discriminate scenes containing an animal from scenes containing no animal. Activity significantly diverged approximately 150 ms poststimulus onset, leading the authors to the conclusion that enough information for solving the task must have been processed at this point. We will refer to this differential target related activity as “target effect” in the following.

Many studies replicated this result under varying conditions (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Rousselet, Fabre-Thorpe, & Thorpe, 2002; Rousselet, Mace, Thorpe, & Fabre-Thorpe, 2007; VanRullen & Thorpe, 2001). However, when interpreting differential electroencephalography (EEG) activity, it should be considered that there are many potential sources that might be responsible for the signal. Thus, in order to investigate the time course of

visual processing in natural images, one should take into account the role of different low-level image features which might contribute to the process of object recognition but not necessarily reflect its completion (Johnson & Olshausen, 2003).

VanRullen and Thorpe (2001) tried to account for this fact by using the same stimuli as targets and nontargets, eliminating the possibility that low-level image features only diagnostic for a subgroup of images were responsible for their results. In addition to that, they extended the previous findings by identifying two distinct undergoing processes in a similar Go-NoGo detection experiment using animal and vehicle pictures as target stimuli. They compared category-specific activity independent of the task, in addition to the replication of the known-target effect for both categories of visual stimuli (animals and vehicles). They found differential activity between the two categories emerging about 75 ms after stimulus presentation. The authors subsequently claimed that this early effect is related to a perceptual process, while the later one should index the decision-making process.

In contrast to earlier results, Johnson and Olshausen (2003) found that the target effect was modulated by reaction times (RTs) only if they controlled for low-level features between the images. The authors thus concluded that this response-dependent mechanism represents object recognition rather than the response-independent component reported before. However, it is still unclear, in the case that low-level features are controlled, whether the target effect reflects the result of sensory processing for target detection or whether this effect represents postsensory mechanisms (Johnson & Olshausen, 2005). Johnson and Olshausen (2005) dealt with this question, suggesting that the previously reported target effect might be related to the post-sensory P300 component instead of representing facilitated sensory processing. Comparing the results of two tasks, one where the target category was either introduced as a word and the target itself was an image and another where the presentation modality was swapped, they showed that the target effect was only related to the target status of a certain stimulus independent of its representation as a word or as an image. In line with this assumption, they also found the effect to show the same general pattern as the P300 and to depend upon RTs, consistent with previous observations (Johnson & Olshausen, 2003).

Taken together, earlier works on the temporal dynamics of objects in natural scenes have postulated two event-related potential (ERP) components of interest: the target effect and the category effect. The first could be interpreted as a measure of the minimum amount of time the visual system needs to produce a signal that reliably differentiates between a response-relevant image category and a response-irrelevant

image category. The category effect is the earliest measured differential activity between two different image categories independent of their behavioral relevance.

Those two measures can be used to shed light on the time course of the perceptual processing of natural images. In the present study we used a similar Go-NoGo paradigm to investigate the processing of material categories (wood and stone), calculating both the target and category effect so as to temporally characterize the process that allows human observers to categorize material surfaces. Besides looking at these two characteristic effects in average ERPs, we also applied pattern classification to the ERP topographies in order to investigate both effects in single trials.

Classifying brain activity on a single-trial basis has the advantage that, contrary to conventional analysis techniques, it does not require data to be aggregated and averaged over trials and possibly observers. Single trial pattern classifiers also provide a relatively direct measure for how informative specific brain activity patterns are for interpreting specific processes (Rieger et al., 2008). Pattern classification has been applied to multiple neurophysiological methods. For example, Rieger et al. (2008) showed that recognition success for natural scene images could be predicted based on single trial magnetoencephalographic (MEG) recordings using a linear support vector machine (SVM). Hiramatsu, Goda, and Komatsu (2011) showed that nine different material categories could be decoded from different brain activity patterns using a linear SVM classifier in a fMRI study. They presented real-world materials rendered on a controlled 3-D shape to the subjects who passively viewed the images in the scanner. Furthermore, they found activation dissimilarities in early visual areas (V1/V2) and higher visual areas (FG/CoS) to be significantly correlated with dissimilarities either based on differences in image statistics or with perceptual dissimilarities.

However, in ERP research this method has not yet gained the same popularity. Liu, Agam, Madsen, and Kreiman (2009) used a linear SVM classifier to predict object category information from intracranial field potentials. They showed that five different object categories could be discriminated on a single trial basis 100 ms after stimulus presentation.

In the present study we combined ERP measurements together with a single-trial linear classification procedure to study the temporal dynamics of material classification (wood vs. stone) in a Go-NoGo paradigm as used by VanRullen and Thorpe (2001).

We also validated our paradigm and data analysis approach in an object (animal vs. human) task. Category effects in ERPs have been shown mostly with object images, and being able to classify object

categories from ERP topographies is a necessary proof of concept for our linear classification approach.

Given the high temporal resolution of electrophysiological recordings, we were able to evidence the emergence of signals informative about the category of the materials as early as 100 ms after the onset of the image. This is an important extension to the previous work on the neural basis of material property perception that we summarized above.

Methods

Observers

Twenty-four observers participated in the study. The data of two observers were excluded from the analysis due to technical problems during the EEG recordings. The remaining 22 observers included eight men and 14 women with a mean age of 24.5 years (range: 19–30 years). Twenty-one observers were right-handed, while one observer was left-handed. All observers had normal or corrected-to-normal vision and were naive to the task and the images used in the experiment. All observers provided written informed consent in agreement with the Declaration of Helsinki. Methods and procedures were approved by the local ethics committee LEK FB06 at Giessen University (proposal number 2009-0008).

Stimuli

A set of 160 object and 150 material images were used in the study, representing two categories respectively: people and animals, as well as wood and stone. Object images were taken from the commercially available COREL database (Corel, 1996), whereas material images were photographed by ourselves using a Nikon D70 camera (Nikon, Tokyo, Japan) under varying indoor and outdoor illumination conditions. Images were normalized in terms of mean luminance and contrast (pixel standard deviation of luminance was normalized within each image category to yield 50% detection thresholds). Material images are available online under <http://www.allpsych.uni-giessen.de/MID>. Material images were consistently validated in a separate study. That is, images were correctly assigned to their respective material category by four independent observers. Details of the normalization and validation procedure can be read in (Wiebel et al., 2013).

In order to characterize the material images in greater detail, we calculated the spectral power distribution across spatial frequencies, averaged across

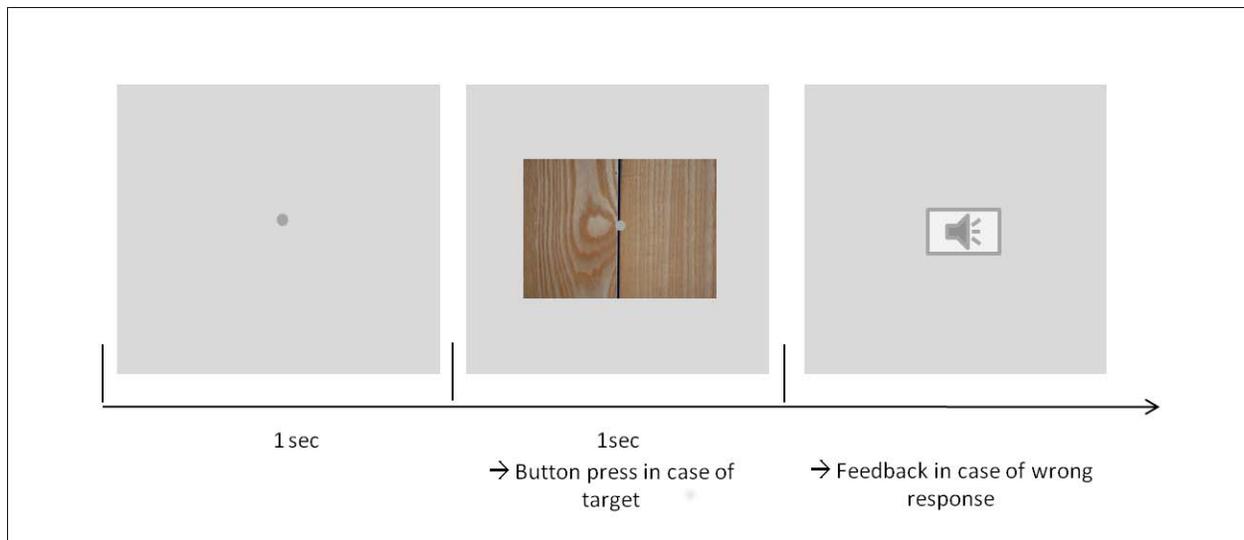


Figure 1. Schematic procedure of one trial.

all orientations. This typically results in a $1/f$ relationship with slopes of approximately -1 (Field, 1987). We used a robust regression fit to estimate that slope for each image. To look at orientation inhomogeneity, we calculated the circular variance over the average amplitude spectrum at each orientation.

Experimental setup

Stimuli were presented on a Samsung SyncMaster 2230R7 22-in. monitor (Samsung Group, Seoul, South Korea) with a refresh rate of 120 Hz. The stimuli had a spatial resolution of 512×768 pixels, corresponding to a viewing angle of $8.23^\circ \times 12.39^\circ$. The study took place in a dimly lit room, where subjects were seated approximately 100 cm in front of the screen.

Procedure

Observers completed a Go-NoGo task consisting of four blocks. In each block, either one of the material categories or one of the object categories served as a target, whereas the other served as a distractor. Target probability in each block was 50%. In the beginning of each trial a fixation dot was presented at the center of the screen for 1 s, followed by the stimulus. To assist observers to maintain their fixation during stimulus presentation, the fixation dot was left at the center of the screen, while presenting the stimulus image at the same time. Stimuli were presented until observers gave a response, but lasted no longer than 1 s. Observers were instructed to respond as fast as possible by pressing a button on a standard response box in case of a target and to suppress any response in case of a

nontarget. If a wrong response was given within the critical period, a short feedback tone was played (see Figure 1).

EEG recordings and analysis

Thirty-two electrodes from the 10-20 system were used to record electrical activity from the brain (FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz, Pz, FC1, FC2, CP1, CP2, FC5, FC6, CP5, CP6, TP9, TP10, Hleo, Veo, Hreo). The ground electrode was placed at the AFz position. Recording was conducted with the reference electrode on the left mastoid. Data were average-referenced offline. Recordings were sampled at 1000 Hz and low-pass filtered at 40 Hz. Artifact rejection was assessed over the following time interval $[-100 \text{ ms}; +400 \text{ ms}]$ with two criteria: eye movements recorded from two horizontally and one vertically placed electrodes around the eyes $[-80 \text{ Hz}; 80 \mu\text{V}]$ and accumulated α range activity $[-40; 40 \mu\text{V}]$ on parietal electrodes. A baseline correction was applied based on the 100 ms before stimulus presentation.

Results

Material task

Behavioral data

RTs were analyzed for every correct Go trial. The average of each observer's median RT in the wood condition was 426 ms. Median RTs in the stone condition were on average 25 ms slower compared to

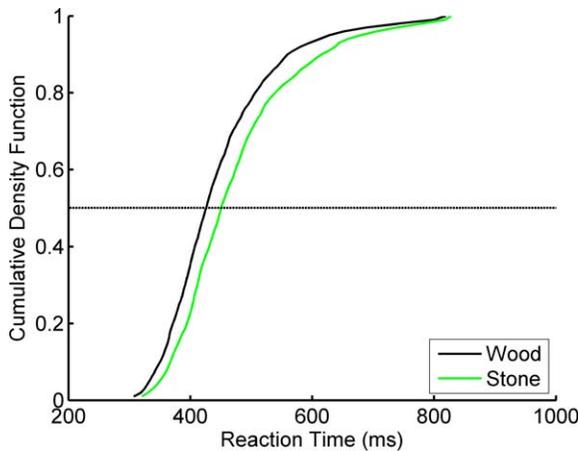


Figure 2. Empirical cumulative density function of the two material category (wood and stone) RT data. The crossing with the dashed line indicates the median RT averaged over observers.

the wood condition (Figure 2). This difference was statistically significant, $t(21) = 2.657$, $p < 0.05$.

Electrophysiology

ERPs: From the EEG recordings, differential activity between the different conditions was assessed by performing one-sample t tests at every time sample of the recordings (every ms). We used a criterion similar to VanRullen and Thorpe (2001) where an effect was defined as being significant if a number of 20 consecutive t tests were found to be significant on a $p < 0.01$ level.

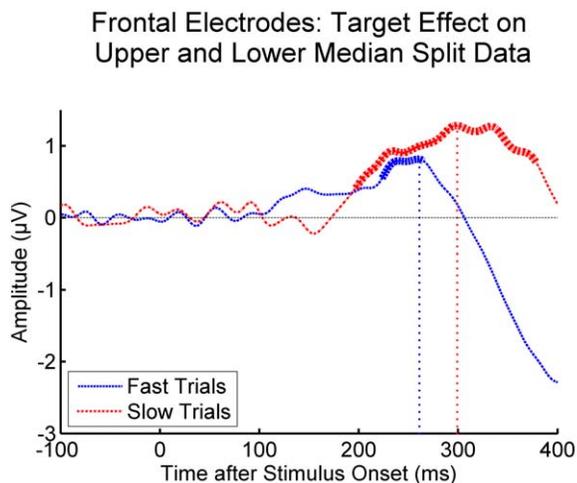


Figure 3. Target effect at frontal sites for fast and slow trials based on the median split on the RT data. Wood and stone trials were pooled together. The dashed blue curve represents fast trials. The dashed red curve represents slow trials. Thick parts of the curves illustrate periods of significant differential activity between Go and NoGo trials compared to baseline activity. Baseline is indicated by the black dashed line.

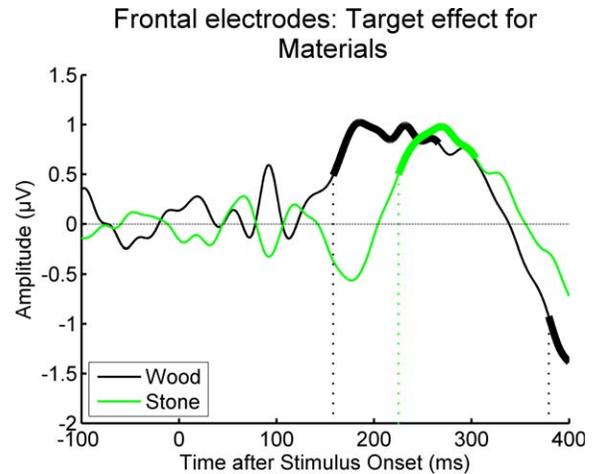


Figure 4. Target effect for wood and stone images at frontal electrodes. The green curve represents the stone condition. The black curve represents the wood condition. Thick parts of the curves illustrate periods of significant differential activity between Go and NoGo trials compared to baseline activity. Baseline is indicated by the black dashed line.

Onset latencies of a significant effect represent the first time sample of the consecutive t tests.

In a first step the differential activity between Go and NoGo trials (the target effect) within each category was investigated at frontal electrodes (FP1, FP2, F3, F4, F7, F8, Fz). In addition, we analyzed the data independent of their target related status as in the study by VanRullen and Thorpe (2001). That is, all trials in which a wood image was presented and all trials in which a stone image was presented were compared at parietal electrodes (P3, P4, P7, P8, Pz).

A frontal target effect was found for both material categories. We further refined our analysis by means of a median split on the RTs of Go trials. The peak differential activity between Go and NoGo trials was present later in time in the subset of trials with slower RTs compared to the subset of trials with higher RTs, an indication that the effect is tightly related to response generation (see Figure 3). A similar pattern was evident when comparing the two material categories. The target effect in case of the wood images emerged earlier (159 ms) than for the stone images (226 ms), which is consistent with the observation that subjects on average responded faster on the wood category than on the stone category (see Figure 4).

According to the analysis applied by VanRullen and Thorpe (2001), we averaged the potential in parietal electrodes in all trials from the wood category and the stone category, independent of their target-related status, in order to isolate the category specific signals in ERPs. Based on this analysis, we found one period of significant differences between the two material categories starting from 142 ms after stimulus presentation. The largest

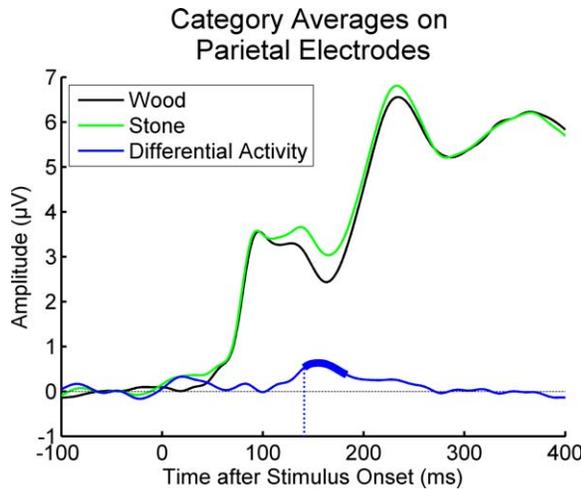


Figure 5. Category averages at parietal electrodes for wood and stone images. Go and NoGo trials were summed up for each category. The black curve indicates activity measured in wood trials. The green curve indicates activity measured in stone trials. The blue curve shows the average differential activity between wood and stone trials. Thicker parts of the curve show significant differences between the activity in the wood compared to the stone condition. The black dashed line indicates the baseline.

amplitude difference between the two categories within this significant time window was $0.63 \mu\text{V}$ (see Figure 5). In the following, we want to explore whether these activity differences can be exploited for classifying the two material image categories on a single trial basis.

Linear classifier

In addition to the canonical analysis of event-related potentials, we performed a single-trial leave-one-out classification analysis on the ERP data using a linear classifier on all recording sites except Cz. All data were aggregated over bins of 10 ms. In a leave-one-out classification procedure, the classifier is trained on all trials except for the one to be classified. That is, training and test data are never intermingled, guaranteeing that the classification is immune to overfitting (Rieger et al., 2008). The classification algorithm was based on the classify function implemented in Matlab (R2007b). This function performs a linear discriminant analysis on the data by fitting a multivariate normal density to each group, with a pooled estimate of covariance (TheMathWorks Inc., Natick, MA).

To evaluate the performance of the classifier, it is necessary to establish when the classification accuracy is significantly above the guessing level. As suggested by Rieger et al. (2008), we opted for a permutation procedure. This involved randomly assigning the class labels to all trials before performing the leave-one-out classification analysis. This procedure was repeated 200

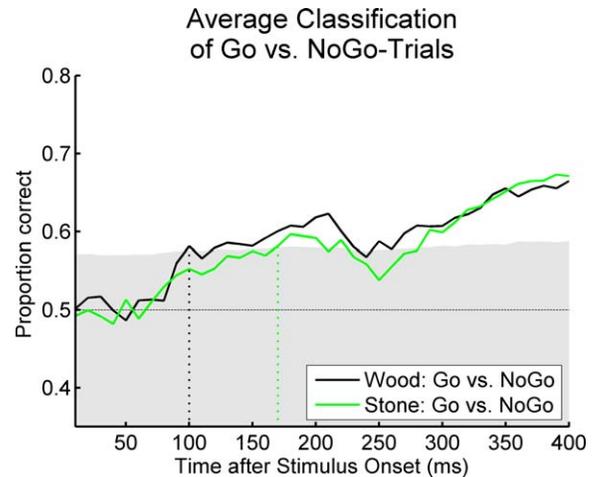


Figure 6. Results of the Go versus NoGo Classification for each single category at each time bin (10 ms) after stimulus onset. The black curve shows the classification results for wood Go versus NoGo trials, and the green curve shows the classification results for stone Go versus NoGo trials. The mean upper bound (averaged across observers and conditions) of the 95% confidence interval calculated for each observer and each time bin is indicated by the gray area.

times. Based on this, 95% confidence intervals for each observer's guessing rate were calculated.

First, analogously to the classical investigation of the target effect, Go versus NoGo trials were classified for each of the two material categories. Second, the two material categories were decoded within Go and NoGo trials respectively. Results are shown in Figures 6 and 7.

When the classifier was trained to decode trials within a category as Go and NoGo trials, classification accuracy reached above-chance levels as early as 100 ms after stimulus onset for wood images and around 170 ms after stimulus onset for stone images, with accuracies peaking between 150 ms and 200 ms. The largest gain in classification accuracy, however, was reached starting from 300 ms after picture onset. It can be assumed that potentials at this point in time were already linked to motor preparation and execution. Overall the classification pattern reflects the pattern of target effects in the canonical ERPs. Above-chance classification and ERP target effects appear with a comparable latency. Moreover, above-chance Go-NoGo classification was observed earlier for trials in the wood category compared to trials in the stone category. This result is consistent with the RT and ERP results, suggesting that the activity that was classified was also related to response production.

In line with the category-specific effects found on parietal electrodes, we could classify the individual trials into the single-material categories both in the Go and NoGo conditions very early in time (110 ms Go condition, 100 ms NoGo condition). In the canonical

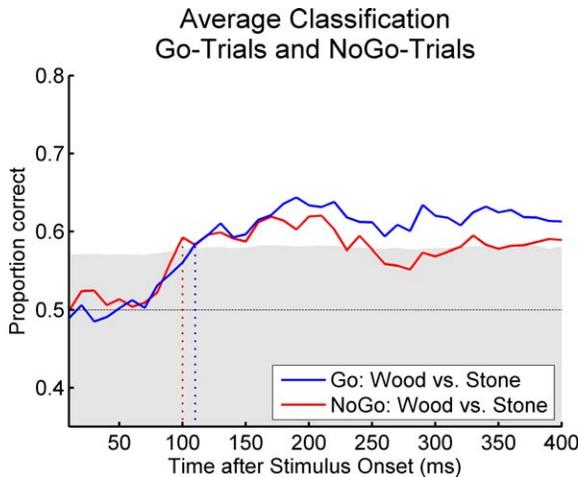


Figure 7. Average proportion correct of classifying wood versus stone in either Go trials or NoGo trials at each time bin (10 ms) after stimulus onset. The dashed blue curve illustrates classification accuracy of wood versus stone trials in the Go condition. The dashed red curve illustrates classification accuracy for wood versus stone trials in the NoGo condition. The mean upper bound of the 95% confidence interval calculated for each observer and each time bin individually is indicated by the gray plane and was here averaged across observers and conditions. The black dashed line indicates the 50% chance level.

ERP analysis this effect only became significant later in time, but started to build up around the same latency as observed here. The fact that the two material categories can be decoded within Go and NoGo trials at about the same point in time assures that this result cannot be due to motor-related activity in the data, but must be related to early differential processing of the visual stimuli (see Figure 7).

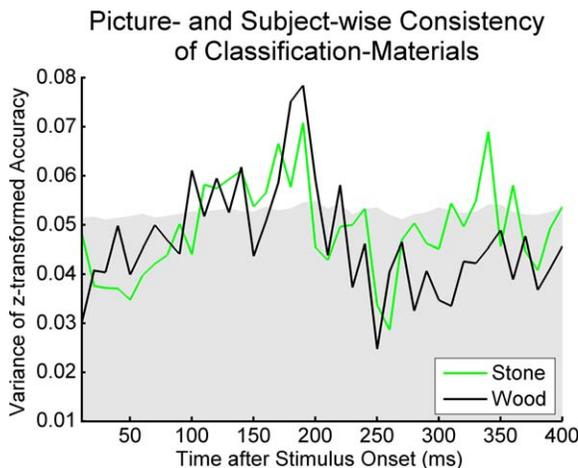


Figure 8. Picture and subject consistency of classification of material trials based on the variances of z-transformed accuracies over time. The gray area indicates the one-sided 95% confidence interval. The black curve indicates wood trials. The green curve indicates stone trials.

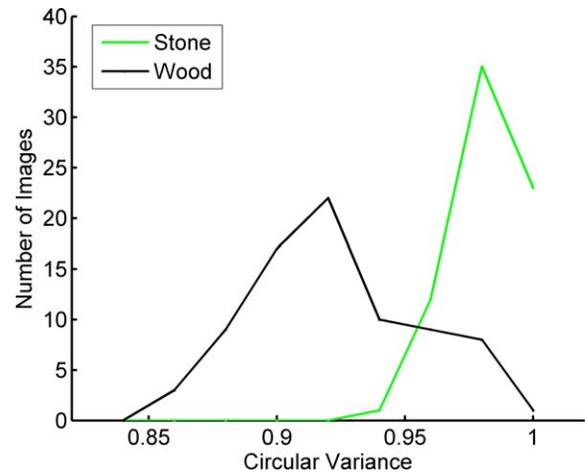


Figure 9. Histogram bin counts of circular variance calculated over the average amplitude spectrum at each orientation for wood and stone images.

As a measure of subject- and image-consistency of all wood and stone classifications we examined the variances of the z-transformed accuracies and compared them to a confidence interval based on 1,000 permutations of the image numbers. The reasoning behind this analysis was that if the same images were classified correctly and the same images were classified incorrectly across observers, the mean variance of the classification across subjects and images of one category should be higher than if the classification results would be not consistent. The case of inconsistent classification of images was simulated by permuting the image numbers and thus assigning them randomly to the classification results. We found evidence for consistency across images and subjects in both categories in the early phase of above-chance classification (see Figure 8).

Subsequently, we analyzed the spatial frequency and orientation content of our material images as potential low-level source of information that could have driven the early successful classification of the two image categories. As mentioned before, the spectral power distribution across spatial frequencies, averaged across all orientations was calculated. Slope parameters for wood compared to the stone images were quite similar and did not differ significantly, $t(148) = -1.545$, $p > 0.05$. To look at orientation inhomogeneities, we calculated the circular variance over the average amplitude spectrum at each orientation. Here distinct and significant differences, $t(148) = 14.779$, $p < 0.001$, in the orientation content of the two material categories were found (see Figure 9). For many of the wood images, there is a concentration of energy at one particular orientation. Therefore, the circular variance was smaller for wood images indicating less variance in the oriented contrast energy. This means that observers

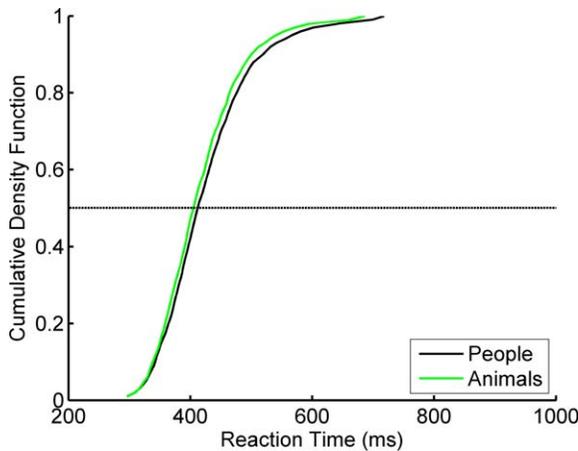


Figure 10. Empirical cumulative density function of the two material category (people and animals) RT data. The crossing with the dashed line indicates the median RT averaged over observers.

could have used oriented contrast to differentiate between wood and stone images.

Intuitively this makes a lot of sense. Wooden surfaces are very often naturally patterned, typically involving line-like structures or elements organized in a systematically oriented way. On the other hand, most stone surfaces rather have a more isotropic structure with less systematic variations or texture. Therefore, the differences in oriented contrast seem like a likely candidate feature to differentiate between wood and stone images.

Frontal Electrodes: Target Effect on Upper and Lower Median Split Data

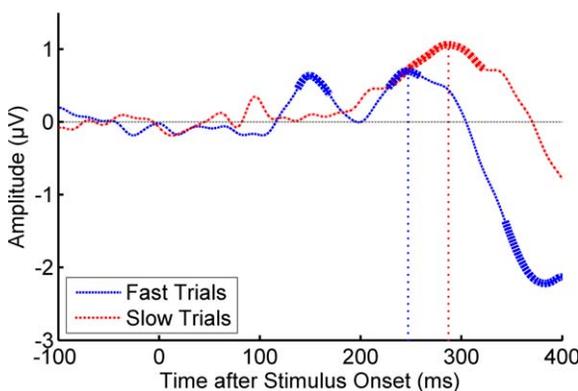


Figure 11. Target effect at frontal sites for fast and slow trials based on the median split on the RT data. People and animal trials were pooled together. The dashed blue curve represents fast trials. The dashed red curve represents slow trials. Thick parts of the curves illustrate periods of significant differential activity between Go and NoGo trials compared to baseline activity. Baseline is indicated by the black dashed line.

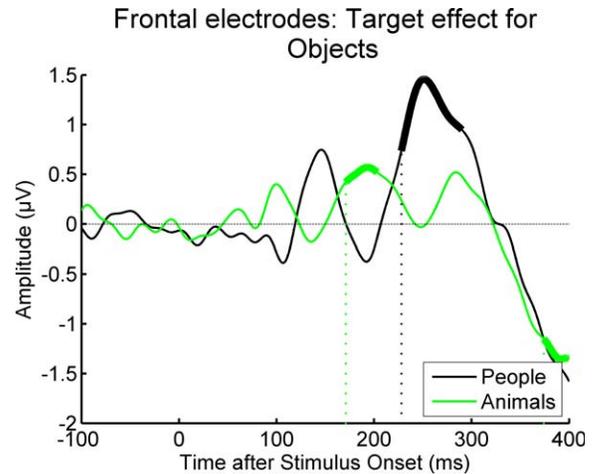


Figure 12. Target effect for people and animal images at frontal electrodes. The green curve represents the animal condition. The black curve represents the people condition. Thick parts of the curves illustrate periods of significant differential activity between Go and NoGo trials compared to baseline activity. Baseline is indicated by the black dashed line.

Object task

Behavioral data

As in the material task, average median RTs were compared between the two object categories: people and animals. Mean RTs were not significantly different (People: 412 ms; Animals: 405 ms). Data are shown in Figure 10. Again we performed a median split based on the RTs in Go trials for the further refinement of the analysis.

Electrophysiology

ERPs: In the canonical ERP analysis we replicated the target effect at frontal electrodes in both object categories. Splitting the data into fast and slow trials according to the RTs revealed again that the frontal target effect is RT-dependent. Both fast and slow object trials showed a period of differential activity between Go and NoGo trials starting around 230 ms, but an earlier period of differential activity starting at 138 ms after stimulus onset, was only evident for fast trials (see Figure 11).

Independent of RTs a significant target effect for people emerged at 229 ms after stimulus onset. For animals two significant periods were found. The first one started at 172 ms after stimulus onset and the second one started 375 ms after stimulus onset. Data are shown in Figure 12.

In contrast to the material task, we did not find differential activity between the two object categories people and animal at parietal electrodes, when both Go and NoGo trials were taken together (see Figure 13)

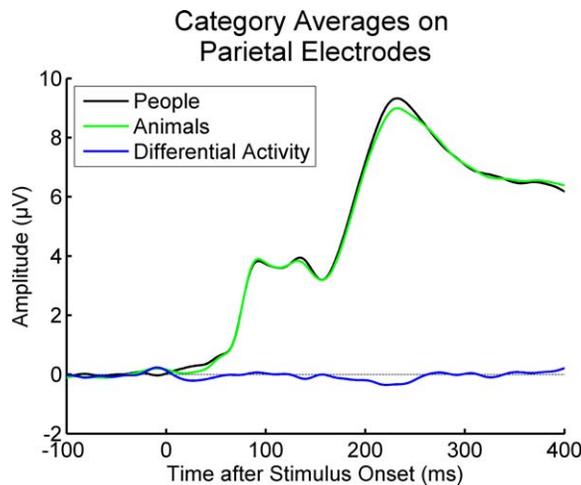


Figure 13. Category averages at parietal electrodes for people and animal images. Go and NoGo trials were summed up for each category. The black curve indicates activity measured in people trials. The green curve indicates activity measured in animal trials. The blue curve shows the average differential activity between people and animal trials. The black dashed line indicates the baseline.

Linear classifier

The linear classification analysis was conducted in the same way as for the material task. First, we tried to classify the target effect within each object category. In agreement with the canonical ERP analysis, Go trials could be decoded from NoGo trials slightly earlier for animal images (180 ms after stimulus onset) than for images of people (190 ms after stimulus onset). However, the general pattern of the classification performance was very similar in both object categories.

In a next step, we classified the two object categories within Go and NoGo trials. People and animals were earlier decoded in Go trials around 140 ms after stimulus onset, while in NoGo trials earliest significant classification of the two object categories was possible 170 ms after stimulus onset. The analysis of subject- and image-consistency of the classification showed that especially in the early phase around 150 ms, classification results for the decoding of the object categories are highly reliable (see Figure 16).

Discussion

VanRullen and Thorpe (2001) established the idea of finding evidence for two distinct processes of visual image processing in ERP recordings: an early sensory component and a later component related to higher-level evaluation of the sensory input. Within this framework, we set out to use the well-established Go-NoGo paradigm for ERP recordings to investigate the

temporal characteristics undergoing the processing of two different material image categories (wood and stone). Besides canonical ERP analysis, we used a pattern analysis approach to classify the material categories and the target relevance of the stimuli. For validation, we also conducted the well-established Go-NoGo task with two object categories (people and animals). We largely replicated the pattern of results that emerged in the data from the material classification task.

Material task

The target effect

We replicated the well-known target effect for two material categories: wood and stone images. In line with Johnson and Olshausen (2003, 2005), the target effect was related to RTs, meaning that it appeared earlier in trials with faster responses. This was reflected in the ERPs by the fact that peak latencies in the target effect were shifted to longer latencies for slow material trials compared to fast material trials, which were determined by means of a median split on the RT data (see Figure 3).

Johnson and Olshausen (2005) claimed that the target effect was postsensory in nature and rather related to a decisional process, thus only providing an upper temporal bound for the accomplishment of object recognition. The latency of the target effect could thus be determined by both the duration of the visual processing and the duration of the decision process. However, since we used the same images as targets and nontargets, we can assume that no differences in low-level image features have driven this effect.

This result was also reflected in the linear classification analysis of the target effect. The classifier was able to reliably distinguish Go from NoGo trials in both categories on a trial-by-trial basis, largely strengthening the result (see Figure 6).

Category differences

The task-independent analysis of categorical processing differences between the two material classes showed that differential activity between the wood and the stone category became evident approximately 140 ms after stimulus presentation in the canonical ERP analysis. In addition, we were able to decode the pattern of event-related electrophysiological activity into the two respective material categories on a single trial basis (see Figure 7). This was shown for Go as well as for NoGo trials starting from around 100 ms after stimulus presentation, ruling out the possibility that this classification result was based on motor-related

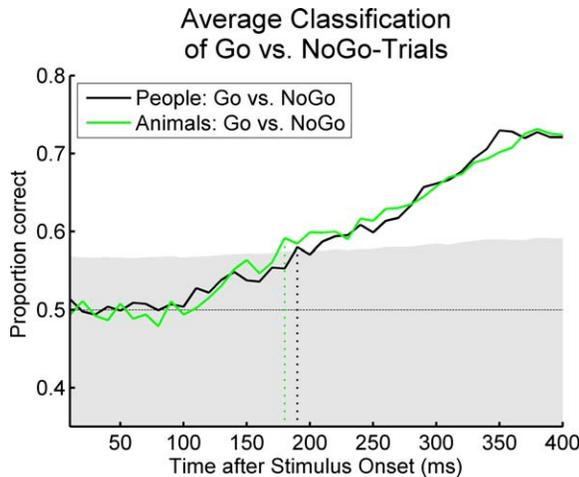


Figure 14. Results of the Go versus NoGo classification for each single category at each time bin (10 ms) after stimulus onset. The black curve shows the classification results for people Go versus NoGo trials, and the green curve shows the classification results for animals Go versus NoGo trials. The mean upper bound (averaged across observers and conditions) of the 95% confidence interval calculated for each observer and each time bin is indicated by the gray area.

activity. This finding, together with the fact that classification accuracy reached over chance performance at about the same latency for Go as well as NoGo trials, shows that these activity differences are related to the early differential processing of the visual stimuli itself and are response-independent. Interestingly, in the canonical ERP analysis, activity started to diverge between the wood and the stone category around 100 ms poststimulus and reached significance much later, but the classification approach was powerful enough to capture this differentiation in its very early stage. These results are also in line with earlier studies (Liu et al., 2009; Simanova, van Gerven, Oostenveld, & Hagoort, 2010) showing that different object categories could be decoded from intracranial field potentials as well as from event-related EEG. In addition, we showed that classification performance at this time point was statistically consistent across images and observers (see Figure 8). Our images were normalized in terms of luminance and contrast, but individual low-level differences between material images must have had a significant impact on their sensory processing as evidenced in the ERP topography. Computational studies have brought up a number of features beyond luminance and contrast which are important for both texture analysis and synthesis—e.g., Portilla and Simoncelli (2000)—as well as material image classification—e.g., Liu et al. (2010). Here analysis on the spatial frequency and orientation content of our images as one potential source of information revealed that wood images had less

variance in oriented contrast compared to the stone images, meaning that there was a concentration of energy at one particular orientation for many of these images. Considering a typical wooden or stone surface this appears to be a useful source of information. Wood usually has some kind of line-like pattern with a certain orientation, while stone most of the time looks more like an isotropic structure. Therefore, the oriented contrast seems like a likely candidate feature to differentiate between wood and stone images. Even though we cannot exclude other potential factors, this seems like an appropriate strategy observers might have used.

In general, the question remains whether the results reported here can be devolved to other material categories and other samples of material images. Behavioral experiments on the speed and accuracy of material categorization have shown highly consistent results based on different material categories and image databases (Sharan, 2009; Wiebel et al., 2013). Here, due to our study design, testing more than two material categories at once was not feasible. However, it would be important in the future to examine whether the results can be generalized to a broader range of material stimuli.

Object task

In line with earlier research (Thorpe et al., 1996; VanRullen & Thorpe, 2001), we replicated a frontal target effect for our two object categories: people and animals (see Figure 12). Furthermore, we validated our single trial linear classification analysis, by showing that this well-known effect was significantly predicted from about 200 ms poststimulus onwards in both object categories (see Figure 14).

On the contrary, no significant category effect on parietal electrodes was evident in the canonical ERPs. Despite this, the linear classification analysis was able to decode both categories in Go as well as in NoGo trials around 200 ms after image presentation (see Figure 15). This is later than what has been reported before (Liu et al., 2009; Simanova et al., 2010; VanRullen & Thorpe, 2001). There might be several reasons for this.

The early categorical effect must come from differences within the images that lead to very early differential processing. We can assume that the strength of this effect is highly conditioned by the specific images that are used. Our object categories (animals vs. people) might have been more similar in their physical structure compared to the categories (animals vs. vehicles) used by VanRullen and Thorpe (2001). Moreover, we took an effort to reduce the physical variability between the pictures from the different classes by applying

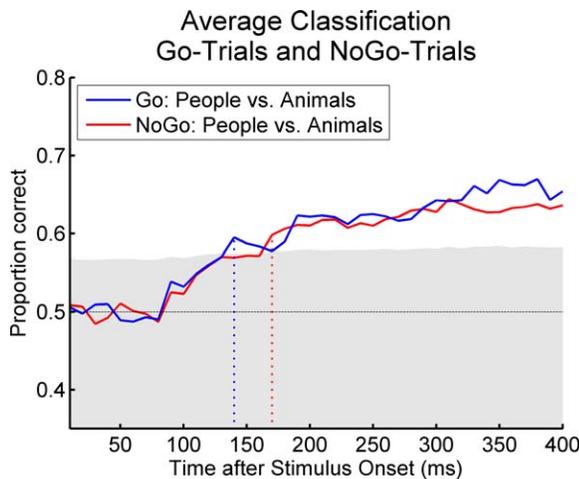


Figure 15. Average proportion correct of classifications of people versus animal in either Go trials or NoGo trials at each time bin (10 ms) after stimulus onset. The dashed blue curve illustrates classification accuracy of people versus animal trials in the Go condition. The dashed red curve illustrates classification accuracy for people versus animal trials in the NoGo condition. The mean upper bound of the 95% confidence interval calculated for each observer and each time bin individually is indicated by the gray plane and was here averaged across observers and conditions. The black dashed line indicates the 50% chance level.

normalization to contrast and average luminance. In the classification studies that we mentioned (Liu et al., 2009; Simanova et al., 2010), different methodologies and stimuli were used so that a direct comparison of the results is difficult. For instance, Simanova et al. used line drawings of objects, certainly a less complex stimulus compared to the images that were used here. The questions of what specific image statistics differ between our object categories and to what extent they promote the differential psychophysiological signature that is captured by the classification at the later stage are beyond the scope of our contribution.

General discussion

We have extended previous reports on material categorization (Hiramatsu et al., 2011; Sharan, 2009; Wiebel et al., 2013) by showing that differential visual processing of material categories in natural images can be decoded approximately 100 ms after stimulus onset based on electrophysiological brain activity patterns.

In general, our findings in the material task agree well with previous behavioral reports showing that materials can be recognized relatively fast (Sharan, 2009; Wiebel et al., 2013). Using fMRI, Hiramatsu et al. (2011) found evidence for material category specific

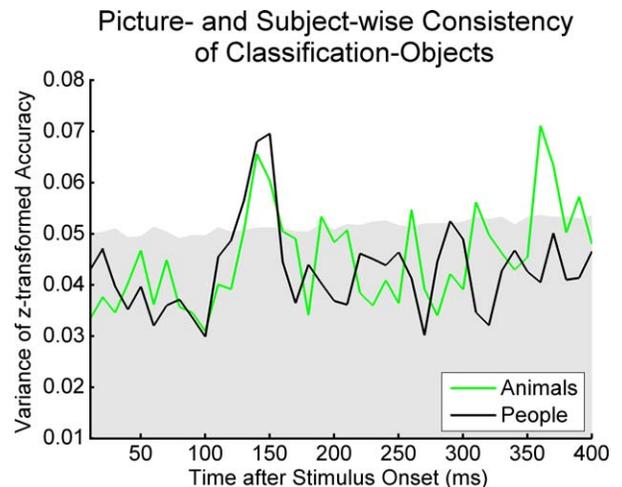


Figure 16. Picture- and subject-consistency of classification of object trials based on the variances of z-transformed accuracies over time. The gray area indicates the one-sided 95% confidence interval. The black curve indicates people trials. The green curve indicates animal trials.

activity in V1 based on a multivoxel pattern analysis. While fMRI does not allow conclusions about temporal aspects at a short time scale, localization of differences in V1 at least agrees with very early processing differences. These early differences went along with image statistics differing between categories. Our results extend these findings by giving a lower-bound estimate of the point in time where the differentiation of material categories begins, that is about 100 ms after image presentation.

In general, many recent neuroimaging studies considering material perception have shown evidence for separate processing of shape information and surface property information in extrastriate areas (Cant, Arnott, & Goodale, 2009; Cant & Goodale, 2007, 2011; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010a, 2010b). Hiramatsu et al. (2011) also found activation in areas higher up the ventral stream to reflect perceptual differences between different materials. Thus, it could be hypothesized that the image information leading to the early differential signal we observe here is later on used for more complex interpretations of the surface properties of the stimuli.

We also replicated a frontal target effect for the material image categories as well as for two object categories. In line with Johnson and Olshausen (2003, 2005) the effect was modulated by RTs. Despite the fact that RTs were on average slower in the material task compared to the object task, we did not observe a consequent latency shift in the target effect for material categories compared to the object target effect. Based on this fact, we suggest that the frontal target effect cannot be used as direct predictor for the behavioral output of a Go-NoGo task. The early differential

processing of the two material categories evidenced by our linear classification analysis demonstrates, however, that the initial processing of material categories seems to start astonishingly early (see Figure 7). We thus speculate that the manual RT disadvantage for materials must build up well within the decision process (see Figures 4 and 12).

Our linear classification procedure was largely able to capture the overall pattern of results from the canonical analysis of the ERPs and even extend it. The fact that the performance of the classifiers for the individual images was consistent across observers largely strengthens the results.

To conclude, by means of single trial linear classification of electrophysiological brain recordings we determined a lower bound estimate for the differential processing of natural material images. We propose that pattern classification of ERP topographies is a powerful tool to investigate electrophysiological brain activity.

Keywords: material categorization, Go-NoGo, ERP, linear classifier

Acknowledgments

This work was supported by grant DFG 879/9 to KRG. MV was supported by a postdoctoral fellowship from the Alexander von Humboldt Foundation and by the EU Marie Curie Initial Training Network “PRISM” (FP7 – PEOPLE-2012-ITN, Grant Agreement: 316746). CBW now has a new address at the Technische Universität Berlin, Modelling of Cognitive Processes Group, Department of Software Engineering and Theoretical Computer Science, christiane.wiebel@tu-berlin.de

Commercial relationships: none.

Corresponding author: Christiane B. Wiebel.

Email: Christiane.B.Wiebel@psychol.uni-giessen.de.

Address: Justus-Liebig-Universität Gießen, Abteilung Allgemeine Psychologie, Giessen, Germany.

References

- Brainard, D. H., & Maloney, L. T. (2004). Perception of color and material properties in complex scenes. *Journal of Vision*, 4(9):i, ii–iv, <http://www.journalofvision.org/content/4/9/i>, doi:10.1167/4.9.i. [PubMed] [Article]
- Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research*, 192(3), 391–405.
- Cant, J. S., & Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex*, 17(3), 713–731.
- Cant, J. S., & Goodale, M. A. (2011). Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. *Journal of Neuroscience*, 31(22), 8248–8258.
- Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010a). Separate channels for processing form, texture, and color: Evidence from fMRI adaptation and visual object agnosia. *Cerebral Cortex*, 20(10), 2319–2332.
- Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010b). Separate processing of texture and form in the ventral stream: Evidence from fMRI and visual agnosia. *Cerebral Cortex*, 20(2), 433–446.
- Corel. (1996). *Corel photo stock library* [CD-ROM]. Ottawa, Ontario: Author.
- Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13(2), 171–180.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12), 2379–2394.
- Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, 3(5):3, 347–368, <http://www.journalofvision.org/content/3/5/3>, doi:10.1167/3.5.3. [PubMed] [Article]
- Fleming, R. W., Wiebel, C. B., & Gegenfurtner, K. R. (2013). Perceptual qualities and material classes. *Journal of Vision*, 13(8):9, 1–20, <http://www.journalofvision.org/content/13/8/9>, doi:10.1167/13.8.9. [PubMed] [Article]
- Giesel, M., & Gegenfurtner, K. R. (2010). Color appearance of real objects varying in material, hue, and shape. *Journal of Vision*, 10(9):10, 1–21, <http://www.journalofvision.org/content/10/9/10>, doi:10.1167/10.9.10. [PubMed] [Article]
- Hiramatsu, C., Goda, N., & Komatsu, H. (2011). Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *Neuroimage*, 57(2), 482–494.
- Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008).

- Conjoint measurement of gloss and surface texture. *Psychological Science*, 19(2), 196–204.
- Johnson, J. S., & Olshausen, B. A. (2005). The earliest EEG signatures of object recognition in a cued-target task are postsensory. *Journal of Vision*, 5(4): 2, 299–312, <http://www.journalofvision.org/content/5/4/2>, doi:10.1167/5.4.2. [PubMed] [Article]
- Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, 3(7):4, 499–512, <http://www.journalofvision.org/content/3/7/4>, doi:10.1167/3.7.4. [PubMed] [Article]
- Kim, J., & Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness. *Journal of Vision*, 10(9):3, 1–17, <http://www.journalofvision.org/content/10/9/3>, doi:10.1167/10.9.3. [PubMed] [Article]
- Liu, C., Sharan, L., Adelson, E. H., & Rosenholtz, R. (2010, June). Exploring features in a Bayesian framework for material recognition. In *Computer Vision and Pattern Recognition*, 2010 IEEE Conference (pp. 239–246). San Francisco, CA: IEEE.
- Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, 62(2), 281–290.
- Maloney, L. T., & Brainard, D. H. (2010). Color and material perception: Achievements and challenges. *Journal of Vision*, 10(9):19, 1–6, <http://www.journalofvision.org/content/10/9/19>, doi:10.1167/10.9.19. [PubMed] [Article]
- Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, 447(7141), 206–209.
- Olkkonen, M., Witzel, C., Hansen, T., & Gegenfurtner, K. R. (2010). Categorical color constancy for real surfaces. *Journal of Vision*, 10(9):16, 1–22, <http://www.journalofvision.org/content/10/9/16>, doi:10.1167/10.9.16. [PubMed] [Article]
- Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1), 49–71.
- Rieger, J. W., Reichert, C., Gegenfurtner, K. R., Noesselt, T., Braun, C., Heinze, H. J., . . . Hinrichs, H. (2008). Predicting the recognition of natural scenes from single trial MEG recordings of brain activity. *Neuroimage*, 42(3), 1056–1068.
- Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience*, 5(7), 629–630.
- Rousselet, G. A., Mace, M. J., Thorpe, S. J., & Fabre-Thorpe, M. (2007). Limits of event-related potential differences in tracking object processing speed. *Journal of Cognitive Neuroscience*, 19(8), 1241–1258.
- Sharan, L. (2009). *The perception of material qualities in real-world images* (Ph.D. Thesis). Massachusetts Institute of Technology, <http://dspace.mit.edu/handle/1721.1/54644>.
- Simanova, I., van Gerven, M., Oostenveld, R., & Hagoort, P. (2010). Identifying object categories from event-related EEG: Toward decoding of conceptual representations. *PLoS One*, 5(12), e14465.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520–522.
- VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4), 454–461.
- Wiebel, C. B., Valsecchi, M., & Gegenfurtner, K. R. (2013). The speed and accuracy of material recognition in natural images. *Attention Perception & Psychophysics*, 75(5), 954–966.