Categorical vs Coordinate Relationships do not reduce to spatial frequency differences.

Vishaal Prasad (v2prasad@ucsd.edu)
UCSD Department of Computer Science, 9500 Gilman Dr.
La Jolla, CA 92093 USA

Ben Cipollini (ben.cipollini@gmail.com)
Classy, Diamond View, 350 Tenth Ave
San Diego, CA 92101 USA

Garrison Cottrell (gary@ucsd.edu)
UCSD Department of Computer Science, 9500 Gilman Dr.
La Jolla, CA 92093 USA

Abstract
Categorical and coordinate stimuli were proposed by Kosslyn (1987) as a set of lateralized visual tasks with a left hemisphere advantage for categorical and right hemisphere advantage for coordinate. He hypothesized that categorical tasks depended on low spatial frequencies and coordinate were preferentially processed in higher spatial frequencies (e.g. Baker et al. 1999); however, the literature in subsequent years was inconclusive on this hypothesis (Jager and Postma 2003). Slotnick et al. (2001) directly tested Kosslyn’s hypothesis and also arrived at conflicting results. By stratifying by difficulty, they showed that Kosslyn's initial hypothesis only when tasks are difficult enough. Our differential encoding model is a three layer neural network that accounts for lateralization of visual processing via the biologically and developmentally plausible mechanism of differences in the connection spread of long-range lateral neural connections. We show that our model accounts for Slotnick's psychological data, and show Slotnick's analysis does not convincingly explain the conflicting results. Instead, we propose that Kosslyn's initial hypothesis was incorrect: categorical and coordinate stimuli are not differentiated solely by spatial frequencies. These results therefore cannot be captured by models such as Ivry and Robertson’s “Double Filtering by Frequency” model, which is driven directly by lateralization in spatial frequency processing.

Keywords: Differential Encoding; spatial frequency processing; categorical vs. coordinate

Introduction

The human brain is composed of two relatively independent hemispheres that communicate via a bridge of neurons known as the corpus callosum. The level of independence suggests that for at least some cognitive processes, it may be advantageous for each hemisphere to specialize and reduce redundancy. This is known as lateralization, and it occurs for many diverse cognitive facilities (Stephan et al. 2003). Prominently, language processing is strongly left hemisphere dominant in right-handed people (Knecht et al. 2000). Of particular interest to us is visual lateralization. Past studies have shown visual lateralization in processing stimuli ranging from frequency gratings to facial recognition (e.g. Sergent 1985, Ivry & Robertson 1998).

One such stimulus, from Navon (1977), is the hierarchical letter, such as a large, “global” T composed of small, “local” Ls. Sergent (1982) showed that the left hemisphere (LH) has an advantage at identifying the local level target (the letter L in the above example), whereas the right hemisphere (RH) has an advantage at identifying the global level target. She concluded that the LH performs better at high-frequency information, whereas the RH does better at low-frequency information. Kitterle et al. (1990) directly tested this hypothesis with frequency gratings and showed that this frequency lateralization was driven by task demands, rather than purely by stimulus properties.

Based on Sergent’s (1982) theory, Ivry & Robertson (1998) proposed their Double Filter by Frequency (DFF) theory to explain this asymmetric processing. DFF theory proposes that the hemispheres both process images via spatial frequency information, but that they preferentially favor information from different frequency bands. In addition to Sergent’s and Kitterle’s results, DFF theory also accounts data from Christman et al. (1991) by proposing that the frequency processing differences between the hemispheres are not absolute, but instead relative to the frequency band relevant for solving the current task. Finally, the model accounts for the categorical and coordinate spatial relations proposed by Kosslyn (1987) and Kosslyn et al. (1989). However, there is no neurological basis for the core mechanisms of the DFF, nor is there a developmental explanation of how or why this phenomenon would emerge (see Cipollini, 2014 for a further discussion).

Competing with DFF theory is our Differential Encoding (DE) theory (Hsiao et al. 2008, Hsiao et al. 2013), which is
described in detail below. This model has accounted for many of the same data, but with specific biological mechanisms with a plausible neurodevelopmental cause. However, it is yet to be tested on the categorical and coordinate dichotomy presented by Kosslyn and his colleagues.

Kosslyn (1987) and also Kosslyn et al. (1989) argued that humans process visual stimuli using two distinct types of spatial relations. Coordinate relations rely on an absolute, metric basis; for example, the statement “the glass of water is 3 inches from my hand” defines a coordinate judgment of one’s hand and the glass of water. In contrast, categorical relations rely on an abstract, relative terms. The statement “the glass of water is on top of the table” does not tell us exactly where the glass is, only its relative position to a table. In his 1987 paper, Kosslyn observed a LH advantage on categorical relations and a RH advantage on coordinate relations. He proposed two different subsystems governed these two relations and that they lateralized based on a snowball effect deriving from the LH’s known advantage in speech (Kosslyn 1987). Other work (e.g. Hellige & Michimata, 1989) provided further support for this hypothesis with more varied types of stimuli (e.g. a bar and dot stimulus).

However, in 1991, Sergent found that lateralization presented only when stimuli were degraded, and several analyses have noted that the lateralization only presents in right-handed people (Slotnick et al. 2001). Other researchers likewise found weak or inconclusive evidence for lateralization of categorical and coordinate stimuli, especially for the LH advantage on categorical (Okubo & Michimata 2002). Nevertheless, researchers generally agree that a distinction exists, even if it is weaker than originally thought (see Jager & Postma 2003 for a review, but also van der Ham & Postma 2010, Baumann et al. 2012). In light of these conflicting data, Kosslyn and colleagues have refined their hypothesis and now argue that lateralization in categorical and coordinate stimuli exists due to a difference in frequency processing, potentially based on hemispheric differences in neuronal receptive fields (see Kosslyn et al. 1992, Kosslyn et al. 1998, Baker et al. 1999, Chapter 9 of Hugdahl et al. 2003).

Of particular interest here are the experiments and results of Slotnick et al. (2001). The authors sought to test Kosslyn’s hypothesis directly. Using all of the same stimuli used in Kosslyn et al. (1989), they conducted a series of five experiments on 134 subjects, each with one hemisphere physically deactivated. Though the results were generally in line with Kosslyn’s hypothesis, one coordinate experiment did not show the expected RH dominance, instead showing the opposite lateralization as in the original paper. Noting that their task was easier due to larger reference differences, Slotnick et al.stratified along difficulty and suggested that only when a task is made more difficult does lateralization arise on coordinate experiments. They thereby concluded that task difficulty is an important factor in the lateralized processing of categorical and coordinate stimuli.

Figure 1: Taken from Hsiao et al. (2013), this diagram shows the autoencoder models with varying connection spreads and symmetric connections. Notice the left hemisphere’s hidden units connect to a wider set of neurons on average.

**Differential Encoding Model**

The Differential Encoding (DE) model is another explanation for human visual lateralization (Hsiao et al. 2008, Hsiao et al. 2013). It is inspired from the auditory system’s long range lateral connections (LRLCs), in which a LH neuron connects to neighbors generally closer to itself than the RH’s on average (Galuske et al. 2000).
The DE model hypothesizes these LRLCs as the driving factor behind visual lateralization as well. Over the DFF theory, the DE model has the advantages of having neuro-developmental and neuro-physical plausibility (Cipollini 2014).

Computationally, the Differential Encoding model is a standard 3-layer neural network which can be thought of as a recurrent neural network unrolled one step in time (Hsiao et al. 2013). The first set of connections is a sparse, noisy autoencoder, trained on natural images, to represent how a stimulus might be represented in the early stages of the brain using low level processing such as Gabor filters. Each neuron in the autoencoder corresponds to a spatial location, and it connects to 5-20 other neurons generated randomly from a Gaussian centered around the neuron itself. The LH and RH networks vary merely by the standard deviation, or “sigma” parameter, of the Gaussian, to mimic the lateralized connection spread of the LRLCs, shown in Figure 1. Each output neuron of the autoencoder then connects to a single output unit, a perceptron that then learns to classify a task via the standard procedure of backpropagation.

This model has accounted for Sergent (1982)’s data (Hsiao et al. 2013), as well as Kitterle’s data showing task dependence of lateralization. Moreover, we’ve shown via spectral breakdowns that the RH encodes more information than the LH at higher frequencies, and vice-versa (Cipollini et al. in prep). This suggests the DE model has the very frequency encoding properties that Sergent (1982) hypothesized. As Kosslyn and colleagues have suggested (e.g. Baker et al. 1999), the distinction between categorical and coordinate stimuli ostensibly stems from a difference in frequency information. Therefore, we test the network on Slotnick et al. (2001)’s stimuli to further establish the relationship between our model and frequency lateralization, as well attempt to reach parity with DFF on these stimuli.

Materials and Methods

Slotnick et al. (2001) Stimuli

The stimuli used in the 2001 study can be seen in Figure 2. There are three types of stimuli: blob/dot, plus/minus, and paired squares. All three stimulus types involve coordinate tasks. Blob/dot requires an evaluation of how far apart the blob and dot are, and plus/minus likewise require an evaluation of how far apart the plus and minus are. The paired squares task, in contrast, requires judging whether the two sets of paired squares are equidistant or not. The former two stimulus types also have categorical tasks. The blob/dot categorical task requires evaluating whether the dot is on the blob or off of it, and the plus/minus categorical task requires evaluating whether the plus is on the right or the left. There is no categorical task for paired squares.

Procedure

The simulation was implemented in MATLAB. All code is open source\(^1\). Input images were implemented as bitmaps, following the images published in the original paper as best possible. To accomplish this, plus/minus and paired squares images were 34x25 pixels and the blob/dot images were 68x50 pixels. Due to the varying resolutions, the experiment sizes had different hyper-parameters. Of particular interest, the 34x25 images had a LH and RH standard deviation (sigma) of 4 and 10 pixels.

\(^1\)https://github.com/guruucsd/DifferentialEncoding/releases/tag/slotnick
respectively; the 68x50 had 4 and 15. In both cases, each neuron had five connections per hidden unit, with a total of 3400 hidden units. For the 34x25 experiments, this meant

The human experiment used 100 LH subjects and 124 RH subjects; however, 54 hemispheres were excluded. We followed the same analyses done in the human experiment, and in an attempt to roughly match statistical power, we instantiated each hemisphere in our computational model 100 times.

Results and Discussion

Figure 3 shows that the model follows the task x hemisphere interaction of the human data across all five experiments. These results are stable and statistically significant. The two categorical experiments show RH dominance, as expected. Coordinate paired squares is also the only

![Figure 2](image)

**Figure 2.** For each task, mean percent error of all left hemispheres is plotted on the left and mean percent error of all right hemispheres is plotted on the left.

![Figure 3](image)

**Figure 3:** A comparison of the DE network’s performance versus the Slotnick results. The DE model’s results closely follow the relative performance difference of the human data. This includes the coordinate blob/dot task that Slotnick would later try to explain.

stimulus type with a RH advantage. As with the human data, the coordinate blob/dot violates Kosslyn’s hypothesis and shows a significant RH advantage. There is a slight deviation in our results for categorical plus/minus, which is more strongly lateralized in the human results as compared to in our model.

Figure 4 shows the human data versus the DE network’s when stratified across difficulty. The hemisphere x task
interaction follows Slotnick once again, but \textit{hemisphere x task x difficulty} does not follow as well. We take a closer look into Slotnick’s difficulty stratification, in order to understand the model’s deviation from the reported human results.

**Figure 4:** A comparison of DE versus the Slotnick data on easy versus hard. Note that for a task such as coordinate plus/minus, we would expect LH lateralization to decrease in hard compared to easy tasks, but that sort of interaction does not follow. Likewise for coordinate plus/minus and categorical blob/dot.

### Revisiting Slotnick

Slotnick et al. (2001) directly measured lateralization in a subject by having them perform the task after one hemisphere had been temporarily deactivated. The deactivation procedure was conducted as part of a routine presurgical evaluation for treatment of intractable epilepsy. Consequently, this meant the authors could only run their experiments once, as the subjects underwent surgery afterwards. Their data did not align with Kosslyn’s hypothesis, and so they conducted a post hoc analysis of the data based on difficulty. The crucial conclusion of their analysis is that lateralization follows prevailing wisdom only when stratified to be difficult enough.

We find reasons to doubt Slotnick et al.’s conclusions. First, when we reexamine Slotnick et al.’s data, we find critical inconsistencies in their figures. Their results shown here in Figures 3 and 4 show that though coordinate plus/minus (2nd plot from right) overall did not lateralize, the easy and hard samples both did. We would expect, based on the easy and hard plots, for the overall coordinate plus/minus to be LH-dominant, yet that is not the case. Coordinate blob/dot has a similar inconsistency with the LH-injected numbers, though this is less apparent and can explained with there being far more “easy” samples. It is hard to state why such inconsistencies exist. Nevertheless, it calls into question difficulty stratification as an explanation for coordinate blob/dot lateralizing contrary to expectation.

In addition, the stratifications between easy and hard were not built in a principled manner and therefore lack validity. Slotnick et al. (2001) state in their paper that the stratification of the paired square coordinate task was simply a heuristic based on the empirical results, as there was no neat way of differentiating easy and hard stimuli. Similarly, they note that no analogous concept of difficulty really exists for categorical stimuli, so they simply used the same division as their coordinate counterparts.

Coupled with the methodological issues, the diverging results from our model lead us to question difficulty stratification as the sole explanation of coordinate blob dot’s divergence. We therefore turn to examine the actual task itself.

**Frequency preferences of the DE model: revisiting Kosslyn’s original assumption**
Kosslyn and colleagues (e.g. Kosslyn et al. 1998) concluded that categorical and coordinate processing lateralized according to preferential frequency processing. As originally hypothesized in Sergent (1982), the LH is thought to outperform the RH in processing high spatial frequencies (HSFs), with the RH outperforming the LH for low spatial frequencies (LSFs). The Differential Encoding model has also shown this same differential frequency processing (Hsiao et al. 2013).

To examine whether Slotnick’s tasks were driven by spatial frequency differences, we tested the model with different bandpass filters for each task. This was done by running the same networks on the bandpass-filtered stimuli. Specifically, all networks trained on the same, unchanged image patches to learn the same features, simulating typical visual experience. However, the perceptron was trained on stimuli run through a bandpass filter window of size four and eight CPI, for 34x25 and 68x50 images respectively. This would allow us to empirically deduce which frequencies the network best responded to. Results were agnostic to a host of parameter choices, including sigma, dropout, and bandpass width (within reason), so we believe the results are general to the task, and not specific to anything about our network setup.

Per Kosslyn’s hypothesis, we expect to see lateralization in accordance to task type: categorical tasks should show a preference for higher spatial frequencies and coordinate tasks a preference for low spatial frequencies. Figure 5 shows the results. Coordinate paired squares performs best at the window centered around 10 CPI and performs worse at higher frequencies. Coordinate plus/minus shows a slight frequency preference at the highest frequencies, whereas categorical shows a bimodal preference, a weaker one at lower spatial frequencies and a stronger one at higher spatial frequencies. Blob/dot stimuli for both categorical and coordinate showed a preference for the lower CPIs. However, categorical performed equally well at the windows centered from 10 to 18 CPIs, whereas coordinate performed best in the window centered at 10 CPIs and was locally parabolic around that area. Outside of these frequency ranges, critical image features are lost and both networks perform similarly poorly.

These results support the idea that model (and human) performance is driven by spatial frequency differences in task-specific information, as the frequency breakdown above matches the lateralization results shown in Figure 3. In Figure 6, we show that with five connections, the RH encodes more information about spatial gratings than the LH starting at about 8 CPI and 14 CPI for 34x25 and 68x50 images respectively for the sigmas we have chosen. Results can be explained by a general LH advantage. However, if enough information is encoded at higher frequencies, then it can be nearly offset (plus/minus coordinate) or reversed (paired squares). Therefore, we conclude that the spatial frequency analysis of our network is internally consistent with the overall results.

These results contradict the hypothesis that categorical and HSF information are intertwined, and likewise for coordinate and LSF. Since the proposal in 1987, many researchers have failed to provide significant evidence the null hypothesis via different stimuli (see Slotnick et al. 2001 for a summary) in various lab conditions. Researchers have generally attributed these results to methodological differences including stimulus type, lab luminescence, and other factors (Slotnick et al. 2001, Kosslyn et al. 1992, Christman et al. 1991, among others). Slotnick et al. (2001) proposed that only a sufficiently difficult task

<table>
<thead>
<tr>
<th>Task</th>
<th>Coordinate</th>
<th>Categorical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired Square</td>
<td><img src="image" alt="Graph" /></td>
<td>Not Applicable</td>
</tr>
</tbody>
</table>
Figure 5: A comparison of the Differential Encoding network’s frequency preferences, based on network accuracy on a task using bandpasses centered around a certain frequency. We would expect, if categorical and coordinate processing were based on spatial frequency profiles, for there to be a consistent pattern between categorical and coordinate that would show increased performance on categorical on the higher spatial frequencies and likewise for coordinate and lower spatial frequencies.
lateralizes. However, if categorical and coordinate stimuli are inherently HSF and LSF tasks respectively, and the RH and LH respond preferentially, it is unclear exactly what changes when the task is difficult.

Our data leads us to a simpler conclusion: categorical and coordinate are not completely tied to spatial frequencies. The reason why blob/dot coordinate does not show RH-dominance is not because it is too easy of a task, but because the because it does not encode more information at higher spatial frequencies like the other two coordinate tasks do. This simple explanation explains the aforementioned suggestions about methodologies – if categorical and coordinate are not inherently beholden to encoding information at differing frequencies, then differences such as luminescence, task difficulty, or even the choice of stimulus itself can change the frequency profiles to make coordinate tasks look like categorical tasks, or vice-versa.

We hope that this work leads to a new discussion on the dependence between categorical/coordinate stimuli and spatial frequencies. We plan to look into the effects of contrast balancing, as highlighted in Michimata & Okubo (2002), to further the analysis about the relation of spatial frequencies and coordinate and categorical tasks.

References


Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. (1989). Evidence for two types of spatial representations: hemispheric


