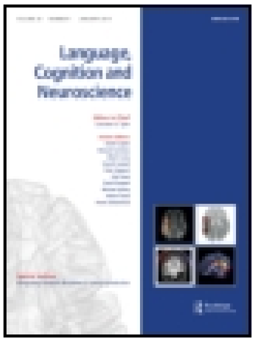

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

# Recognising a zebra from its stripes and the stripes from “zebra”: the role of verbal labels in selecting category relevant information

Lynn K. Perry & Gary Lupyan

LANGUAGE, COGNITION AND NEUROSCIENCE, 2016
http://dx.doi.org/10.1080/23273798.2016.1154974

Routledge
Taylor & Francis Group

# Recognising a zebra from its stripes and the stripes from "zebra": the role of verbal labels in selecting category relevant information

Lynn K. Perry[a] and Gary Lupyan[b]

[a]Department of Psychology, University of Miami, Coral Gables, FL, USA; [b]Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

**ABSTRACT**
Distinguishing members from non-members of some categories can be accomplished by identifying one or several diagnostic features (e.g. zebra-stripes are diagnostic of zebras). Other categories lack diagnostic features (e.g. dogs). Consequently, distinguishing members from non-members requires attending to many correlated dimensions. Interestingly, children and non-human animals are less adept at using diagnostic features compared to adults – possibly due to adults' more developed verbal labelling abilities. We examined whether recognition of categories with diagnostic features ("sparse" categories) is (1) linked to better abilities to selectively attend to relevant information and (2) aided by labelling. In Experiments 1–2, we quantify and validate a measure of category sparsity. Experiment 3 demonstrates that sparse categorisation, assessed by an implicit naming task, correlates with performance in the flanker task, a measure of selective attention. Experiment 4 demonstrates up-regulating activity over Wernicke's area via transcranial direct current stimulation – hypothesised to enhance labelling – selectively improves sparse categorisation.

**ARTICLE HISTORY**
Received 13 July 2014
Accepted 10 February 2016

**KEYWORDS**
Categorisation; labelling; transcranial direct current stimulation; selective representation; object recognition



Because no two perceptual inputs are ever exactly the same, recognising an input as being the "same" as a previous input, is an act of categorisation. Categorisation can be thought of as the process by which an input is aligned in some way with that of previously encountered members of the same category. Properties relevant for the category become highlighted, while properties irrelevant are abstracted over. For example, recognising an object as a cup involves representing it so that aspects of its shape, size, and material are highlighted, while colour – a property uninformative of cup-ness – is (partly) abstracted over. This process has often been called "selection" and parallels have been drawn between selection of task-relevant information in perception (e.g. attending to horizontal vs. vertical forms, or to colour and not to shape), and selection in the conceptual domain (e.g. thinking of a knife as something that is usually made of metal rather than as something that is sharp (e.g. Kan & Thompson-Schill, 2004). In both cases, "selection" (indeed, selective attention itself) can be thought of more generally as a warping of a representation (a distributed neural activation pattern) into a task-relevant form.[1] Here, we focus on a subset of categorisation – categorisation in the service of naming. Naming an object, for example, naming a chair as a "chair" requires categorising it at what is often called a *basic* level, selectively representing those features relevant to being a chair and (temporarily) down-weighting features irrelevant to chair-ness.

## The role of labelling in categorisation

It is easy to see how naming depends on categorisation. But there is accumulating evidence that categorisation also depends on naming. For example, named categories are easier to learn (Balaban & Waxman, 1997; Lupyan, Rakison, & McClelland, 2007; Nazzi & Gopnik, 2001; Perry & Samuelson, 2013; Plunkett, Hu, & Cohen, 2008). Once a category is learned, the knowledge of its attributes is more effectively activated by a verbal label than other highly associated cues, such as nonlinguistic sounds (Boutonnet & Lupyan, 2015; Lupyan & Thompson-Schill, 2012). Consistent with the possibility of a causal involvement of labels in categorisation, language impairments (in particular, naming impairments such as aphasia) produce categorisation impairments (Gainotti, 2014; Lupyan & Mirman, 2013 for review), and interfering with language in healthy adults impairs categorisation (Lupyan, 2009). Taken together, these findings suggest that labels help to *reify* categories by selectively

**CONTACT** Lynn K. Perry lkperry@miami.edu

activating critical features and abstracting over irrelevant ones (see Lupyan, 2012b for discussion and a computational model).

Importantly, the influence of verbal labels on categorisation appears to interact with the structure of the category. While there are many ways of measuring category structure, we focus here on the number of features/dimensions that category members have in common, what we will refer to as *category sparsity*. The construct of category sparsity is well-summarised by Sloutsky (2010, pp. 1250–1251):

> Categories that are statistically dense have multiple intercorrelated (or covarying) features relevant for category membership, with only a few features being irrelevant. Good examples of statistically dense categories are basic-level animal categories such as cat or dog. Category members have particular distributions of values on a number of dimensions (e.g. shape, size, color, texture, number of parts, type of locomotion, type of sounds they produce, etc.). These distributions are jointly predictive, thus yielding a dense (albeit probabilistic) category. Categories that are statistically sparse have very few relevant features, with the rest of the features varying independently. Good examples of sparse categories are dimensional groupings (e.g. "round things"), relational concepts (e.g. "more"), scientific concepts (e.g. "accelerated motion"), or role-governed concepts (e.g. cardinal number).

Deciding what should count as a feature or dimension is not at all straightforward,[2] and so we will be fairly noncommittal about the exact meaning of these terms. An example of what we mean by a *dense* category is something like BIRDS[3] – categories whose members share many properties but cannot be distinguished based on any single feature (what Lupyan, Mirman, Hamilton, & Thompson-Schill, 2012 called high-dimensional categories; see also Pothos, 2005). A paradigmatic example of a *sparse* category is RED-THINGS.

Such sparse categories (sometimes also called "rule-based" categories) have been shown to present substantial learning difficulties for non-human animals (e.g. Couchman, Coutinho, & Smith, 2010) and young children (Kloos & Sloutsky, 2008; Minda, Desroches, & Church, 2008a) who, when given a choice, tend to default to a more dense similarity-based category structure (e.g. Smith & Kemler, 1977). Ashby, Maddox, and colleagues have argued that learning categories organised by overall similarity (i.e. dense categories) utilises more automatic processes while learning rule-based categories (i.e. sparse categories) places more demands on abilities to selectively represent relevant features (see e.g. Ashby & Maddox, 2011; Minda & Miles, 2010). Thus, selective representation demands, or the need to represent a small amount of task-relevant information to the exclusion of irrelevant information should be greater for sparse categorisation.

It is interesting then that verbal labelling has been linked specifically to sparse categorisation. For example, verbal interference impacts the ability to categorise based on a specific feature (e.g. size), but not on more global properties (Lupyan, 2009). Similarly, people with naming impairments show problems with categorising based on specific features such as ANIMALS WITH STRIPES or OBJECTS THAT ARE ROUND, but are similar to controls on dense such as BIRDS (e.g. Lupyan & Mirman, 2013). Young children struggle with sparse categorisation, but providing them with novel category labels can facilitate their learning of such categories (Perry & Samuelson, 2013). In contrast, under normal circumstance, adults, are strongly biased to use sparse categorisation (e.g. Couchman et al., 2010; Minda et al., 2008a). Down-regulation of neural activity over Wernicke's area using transcranial direct current stimulation (tDCS) (as a means of inhibiting neural processing involved in naming (see e.g. Price, 2000)) tends to promote forming a more dense category structure (Perry & Lupyan, 2014). Taken together, evidence from studies of children, typical adults, and adults with aphasia, suggest that labelling supports the ability to represent sparse categories.

However, it has remained unclear *why* such a link might exist between labelling and sparse categorisation and what processes underlie recognition of sparse categories. The current studies are designed to begin answering these two questions. First, are selective representation demands greater for sparser as compared to denser categories? Second, are category labels more helpful for recognising members of sparse categories (perhaps due to the greater need for selective representation such categories require)?

Rather than using contrived categories like those in much of past work (Couchman et al., 2010; Kloos & Sloutsky, 2008; Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010; Minda, Desroches, & Church, 2008b; Perry & Lupyan, 2014), we used common animal and artefact categories varying in sparsity. For example, DOG is a fairly dense category – there is no feature that is shared by all dogs and that distinguishes dogs from non-dogs. In contrast, ZEBRA is a sparser category in that members have a salient feature (zebra-stripes) that is shared by all the members and can be often used to distinguish zebra from non-zebras.

We address the first question by examining whether individual differences in performance on a nonlinguistic selective attention task relate to individual differences in performance on a picture-word verification task that measures speed of implicit naming. We address the second question by modulating processes involved in labelling[4] using tDCS and examining how this modulation affects performance on (implicitly) naming of members of sparser vs. denser categories.

## Rationale for present work

The ability to represent task-relevant information is critical to a variety of tasks (often studied within the disciplines of "cognitive control" and "executive function") (see Banich, 2009; Kan & Thompson-Schill, 2004) and varies considerably across individuals even within a typical population (e.g. Vogel, McCollough, & Machizawa, 2005). We reasoned that if categorising members of sparser categories like ZEBRA requires representing items in a more selective way relative to categorising members of denser categories like DOG, then people who perform better on nonverbal tasks requiring selection may be relatively better at categorising items from sparser categories. In Experiments 1 and 2, we determine the sparsity of 12 familiar categories to use as stimuli and validate our measure of category sparsity. In Experiment 3, we related individuals' performance on the flanker congruity task (see Figure 1) – commonly used to study selection processes (see Eriksen, 1995) – to individual differences in a picture-word verification task (see Figure 2), a measure of implicit naming.

If labelling helps to support selective representation of category-relevant information, and if selective representations are more critical for recognising members of sparse categories like ZEBRA than members of dense categories like DOG, then manipulating the ease with which people can label may impact the recognition of members from sparse categories compared to members from dense categories. In Experiment 4, we manipulated the labelling process by using tDCS (Perry & Lupyan, 2013). Specifically, we examined effects of up- and down-regulating activity over Wernicke's area via tDCS on verification of picture names. We predicted that if labelling supports selective representation, then up-regulating activity over Wernicke's area should facilitate recognition of sparse categories while down-regulating activity should impair it.

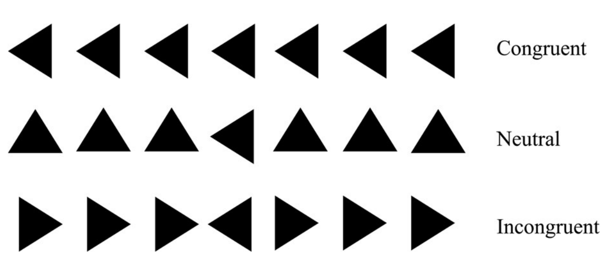


**Figure 1.** Sample trial types in the flanker task. Participants' task is to report the direction (left, right) of the central triangle.

## Experiment 1: stimulus selection and norming

The primary goal of Experiment 1 was to quantify category sparsity for the materials used in our picture-word verification tasks (Experiments 2–4). The pictures were photographs obtained from online image collections. We selected 12 picture categories which had basic-level names of approximately equal frequencies (based on American National Corpus written frequency norms, Reppen, Ide, & Suderman, 2005), and which had approximately equal concreteness (based on Medical Research Council psycholinguistic database norms, Wilson, 1988) (see Table 1 for stimulus characteristics). We next recruited 59 English-speaking adults from Amazon Mechanical Turk and asked them to do one of three tasks: a commonality-listing task, a name agreement task, or a typicality-rating task.

### *Measuring category sparsity*

One way to quantify category sparsity is by asking participants to list common attributes of category members. We reasoned that insofar as a given feature is central to a category, it will be readily identified by participants tasked with telling us what all the members of a given category have in common. We presented 15 participants with 10 pictures of items from each of the 12 categories and asked them to list the features that all the items from a given category had in common. Category sparsity was defined as the

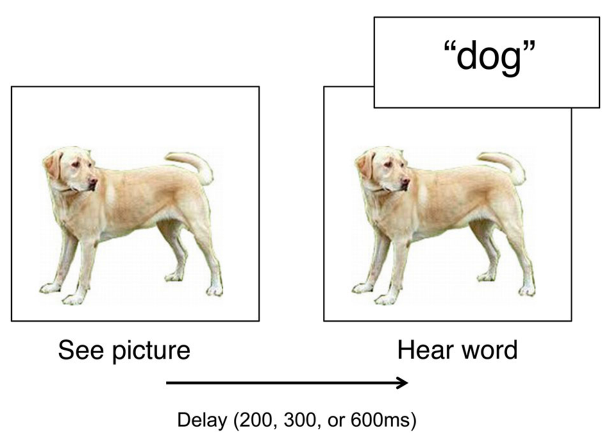


**Figure 2.** Schematic of stimuli presentation in picture-word verification task. A picture was followed by a word which participants judged, as quickly as possible, as matching or not matching the image.

**Table 1.** The most frequently listed commonality and the proportion of the participants in the norming study listing that commonality for each category used in the picture-word verification task (used as a proxy for category sparsity in our analyses).

| Category | Most frequently listed commonality (based on mTurk commonality-listing study) | Category sparsity (sparse to dense; proportion of mTurk participants listing most frequent commonality) | Average name agreement (based on mTurk name agreement study) | Average typicality (based on mTurk typicality study; 1–7 scale, 1 = low, 7 = high) | Log word frequency (based on American National Corpus written frequency norms) | Concreteness (based on Nelson et al., 2004; 1–7 scale, 1 = low, 7 = high) | Kind | Feature foil | Semantic foil | Association strength between target category and semantic foil (Nelson et al., 2004) |
|---|---|---|---|---|---|---|---|---|---|---|
| Carrot | Colour | 1.00 | .97 | 3.62 | 2.05 | 6.19 | Natural kind | Pumpkin | Rabbit | .21 |
| Hand | Fingers | .80 | .98 | 4.08 | 3.62 | 5.60 | Natural kind | Glove | Foot | .16 |
| Zebra | Pattern | .77 | 1.00 | 4.42 | 1.32 | 6.52 | Natural kind | Zebra coat | Horse | .19 |
| Ball | Shape | .67 | .99 | 3.02 | 3.14 | 6.18 | Artefact | Orange | Bat | .19 |
| Brush | Bristles | .45 | .92 | 3.44 | 2.36 | 5.66 | Artefact | Porcupine | Hair | .44 |
| Spoon | Handle | .43 | 1.00 | 3.54 | 1.94 | 5.80 | Artefact | Doctor's hammer | Bowl | <.01 |
| Apple | Shape | .38 | 1.00 | 4.45 | 2.72 | 7.00 | Natural Kind | Globe | Banana | .02 |
| Phone | Buttons/numbers | .30 | .86 | 2.72 | 3.40 | 6.02 | Artefact | Calculator | Walkie Talkie | <.01 |
| Desk | Wood | .29 | 1.00 | 2.54 | 2.77 | 5.79 | Artefact | Dresser | Chair | .29 |
| Bird | Feathers | .13 | .97 | 2.58 | 2.81 | 5.77 | Natural kind | Feather duster | Nest | .06 |
| Dog | Four legs | .13 | .97 | 3.80 | 3.27 | 5.75 | Natural kind | Cow | Cat | .67 |
| Purse | Handles | .10 | .82 | 3.54 | 2.06 | 5.64 | Artefact | Bucket | Wallet | .20 |

proportion of participants listing this common feature (see Table 1). For example, if all participants were to list a single common feature for all the items in the category, that category would be maximally sparse with a score of 15/15 or 1.0. Conversely, if each participant listed a different single common feature for all items in the category, that category would be maximally dense and have a category sparsity score of 1/15, .067. Nonsensical answers (e.g. responding "yes" instead of providing an attribute) were excluded (approximately 7% of responses).

### *Name agreement task*

The purpose of the name agreement task was to ensure that the category labels we used were appropriate. Participants were presented with pictures of 10 exemplars from each of the 12 categories, one exemplar at a time, and asked to name each picture. Name agreement was quantified as the proportion of participants listing the same name for each picture. To count as the same, responses had to be identical. Plurals, synonyms, and other variations were not accepted. The five exemplars from each category with the highest name agreement were used in the picture word verification task. Average name agreement was 96% (see Table 1 for full listing).

We collected name agreement primarily to help us select pictures for use in the study, but we additionally sought to ensure that name agreement was not confounded with category sparsity. Average name agreement was positively, though not significantly correlated with the category sparsity of our stimuli, $r(10) = .41$, $p = .24$.

### *Typicality*

Typicality is a well-known predictor of categorisation/naming speed (Rosch & Mervis, 1975; Rosch, Simpson, & Scott, 1976), and therefore it was important to rule out the possibility that effects of category sparsity were confounded by typicality. We collected typicality ratings for all pictures by presenting each item one at a time, and asking participants to rate its typicality by responding to the following prompt: "On a scale of 1 to 5, with 1 being the least typical and 5 being the most, how typical is this [dog] of [dogs] in general?" Mean category typicality correlated positively, $r(10) = .40$, $p = .26$, though not-reliably, with category sparsity; see Table 1.

The goal of Experiment 1 was to obtain a measure of category sparsity for our picture verification stimuli. In Experiment 2 we validate our measure of category sparsity using a picture-word verification task.

## Experiment 2: validating category sparsity

The goal of Experiment 2 was to validate our measure of category sparsity. If categories for which participants were most likely to list only a few relevant features (e.g. stripes for the category of ZEBRAS) are represented in a way that accentuates that feature, then participants should have some trouble rejecting non-category members that share that property (e.g. rejecting an image of a woman wearing a zebra-striped coat when prompted by "zebra"). In contrast, because no one feature is critical to membership in a dense category such as PURSE, participants should have little trouble rejecting feature-based foils such as a bucket (even though purses and buckets both have handles), though they may be slowed to reject more general similarity-based foils (e.g. a wallet). To test this prediction we conducted a picture-word verification task designed to assess difficulty in rejecting different types of foils for categories varying in sparsity.

### *Participants*

Eighteen monolingual English-speaking undergraduates participated for course credit.

### *Methods*

On each trial participants heard one of the words listed in Table 1 (e.g. "zebra") followed by a picture from one of the matching or non-matching categories. The task was to press a "match" button if the picture matched the word, which it did on 45% of the trials, and a mismatch button otherwise. On match trials, the picture was one of the five exemplars per category selected based on the judgments in the name agreement task. On the remaining trials, participants saw either a feature-based or semantic-based foil that they had to reject. Feature foils were pictures of categories that shared a common feature with the target category based on the most commonly listed feature in the commonality-listing task. For example, for ZEBRA, the *feature foil* was a woman wearing a zebra-striped coat (critical feature: stripes), for purse it was a bucket (critical feature: handle). Semantic foils were pictures of categories that shared a general semantic relationship to the target category, without depicting the critical feature. We selected the most strongly associated imageable category based on word association norms (Nelson, McEvoy, & Schreiber, 2004). For example, the *semantic foils* for "zebra" and "purse" were pictures of a horse and a wallet, respectively (see Table 1 for complete list). Because the different categories varied in the extent to which the semantic foil and feature foil related to the target, we included semantic association strength (associative strength between the target and semantic foil from based on Nelson et al., 2004 norms) as a covariate in our analyses. In our stimuli, semantic association strength was positively, though not significantly correlated with category sparsity, $r(10) = .25$, $p = .44$. We selected three exemplars of each foil category from online image collections. Each of the 12 category labels was presented 11 times for a total of 132 trials.

## Analytic approach

All analyses in this experiment were conducted using mixed effects regression models. Reaction time analyses were conducted using linear mixed models. To determine the best-fit model, we used chi-square tests comparing models with and without the factor of interest. For interactions, we report coefficients and confidence intervals from the full model, and the chi-square test of model fit from the comparison to a model with the interaction removed. For main effects, we report coefficients and confidence intervals from the full model, and the chi-square test of model fit from the comparison to a model with the predictor main effect removed. To determine appropriate random effects, we began with completely specified random effects structures including random slopes for all variables in a given model. Using model comparison, we systematically removed uninformative random effects (Jaeger, 2009). Unless otherwise specified, final models included random intercepts for subjects and items. Factors were centered (e.g. ±0.5).

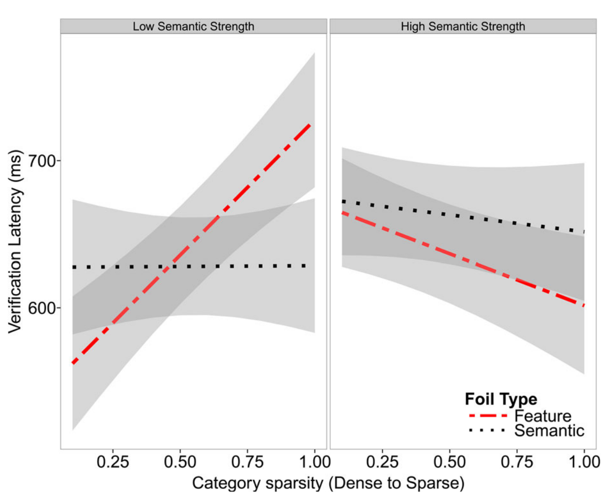


**Figure 3.** Time (ms) to correctly reject a foil in the sparsity validation study. Category sparsity is represented from left to right as dense to sparse based on the highest proportion of participants in the norming study listing the same commonality. Note: Error bands depict standard error of predicted means.

All final models are listed in the appendix and referenced by number in the main manuscript (e.g. M1a).

## Results and discussion

Our first analysis examined the interaction between category sparsity, foil type, and semantic association strength on RTs. The full model included the interaction between semantic association strength (a continuous measure rather than a median split was used in all analyses), category sparsity, and foil type (feature vs. semantic) and main effects of semantic association strength, category sparsity, and foil type for mismatch trials (Appendix M1a). There was a marginally significant two-way interaction between foil type and category sparsity $b = -256$, $CI_{95\%} = [-423, -91]$, $\chi^2(1) = 2.74$, $p = .098$ such that participants were generally slower to reject feature-based foils for sparse than dense categories but were not affected by category sparsity for semantic-based foils. There was no interaction between foil type and semantic association strength, $b = -239$, $CI_{95\%} = [-440, -38]$; $\chi^2(1) = .10$ $p = .76$ or between category sparsity and semantic association strength, $b = -145$, $CI_{95\%} = [-1252, 963]$; $\chi^2(1) = .67$, $p = .41$. There were no main effects of category sparsity, $b = 64$, $CI_{95\%} = [-179, 307]$; $\chi^2(1) = .03$, $p = .86$; foil type, $b = 80$, $CI_{95\%} = [19, 142]$; $\chi^2(1) = .03$, $p = .87$; or semantic association strength, $b = 98$, $CI_{95\%} = [-196, 392]$; $\chi^2(1) = .73$, $p = .39$ (see Appendix M1a)[5]. However, as can be seen in Figure 3, there was also a three-way interaction such: speed to reject foils on feature vs. semantic foil trials varied as a function of semantic association strength and category sparsity, $b = 997$, $CI_{95\%} = [241, 1752]$, $\chi^2(1) = 6.69$, $p = .01$.

To examine this three-way interaction, we looked at the interaction between category sparsity and foil-type separately for the categories that had relatively low association strength with their semantic foils (see rightmost column of Table 1) and those with higher semantic strength. For categories having low semantic strength between targets and semantic foils, there was no main effect of foil type, $\chi^2(1) = 1.81$, $p = .18$, a marginal effect of category sparsity, $b = 189$, $CI_{95\%} = [65, 315]$, $\chi^2(1) = 2.79$, $p = .09$ (Appendix M1b) and a significant interaction between category sparsity and foil type, $b = -187$, $CI_{95\%} = [-290, -85]$, $\chi^2(1) = 12.66$, $p = .0004$ (Figure 3 – left). Feature foils were harder to reject from sparse categories compared to dense categories, $b = 187$, $CI_{95\%} = [99, 276]$, $\chi^2(1) = 8.16$, $p = .004$ (Appendix M1c). For example, participants were slower to reject a striped coat from the category zebra (sparser category) compared to a feather duster from the category bird (denser category) Participants rejected semantic-similarity foils with equal speeds regardless of category sparsity, $\chi^2(1) = .003$, $p = .95$ (Appendix M1d).

For categories having relatively high semantic association strength between targets and semantic foils there were no main effects of category sparsity, $\chi^2(1) = 1.04$, $p = .31$, or foil type, $\chi^2(1) = 2.09$, $p = .15$, and no significant interaction between category sparsity and foil type, $\chi^2(1) = 1.26$, $p = .26$ (Appendix M1e).

Together these findings demonstrate that rejecting a feature-based foil (e.g. rejecting a striped coat as not being a zebra) is more difficult for sparse categories (for which the feature is proposed to be highly relevant) than dense (where the feature is one of many relevant features). This effect of sparsity was only observed for categories lacking strong semantic associates (based on association norms: Nelson et al., 2004) (Figure 3 – left). Nonetheless, because none of the categories we use are strongly associated with one another, we take these results to validate our measure of sparsity for use in Experiments 3 and 4.

## Experiment 3

The goal of Experiment 3 was to examine the extent to which individual differences in recognising familiar objects from sparse and dense categories relate to individual differences in selective representation. To this end, we correlated people's performance on a picture-word verification task and a flanker task.

The picture-word verification task measures speed of implicit naming. Participants who are best at representing information in a selective way should show an advantage for verifying the names of sparse category members. The flanker task required participants to indicate the direction of a central target while ignoring irrelevant distractors on either sides (flankers) (see Figure 1). Participants generally show an advantage (faster RTs and/or higher accuracy) on trials in which the target and flankers face the same way (congruent trials) and a cost (slower RTs and/or lower accuracy) on incongruent trials, suggesting a difficulty in ignoring irrelevant information. Differences in the size of participants' cost/advantage can be interpreted as reflecting an ability to selectively represent the target stimulus/response (see Cagigas, Filoteo, Stricker, Rilling, & Friedrich, 2007; Eriksen, 1995; Roberts, Anderson, & Husain, 2010). Participants who are best at representing information in a selective way should show the least facilitation of congruent flankers and the least interference of incongruent flankers. We reasoned that people who show the most selective processing in the flanker task (as measured by the size of their congruent advantage and/or incongruent cost) will be faster to recognise pictures belonging to sparse categories than participants with less selective (more "integrative") processing.

### Methods

#### Participants

Thirty-six monolingual English-speaking undergraduates participated for course credit.

#### Picture-word verification task

On each trial, participants saw a picture of a familiar object or animal presented on a computer screen, heard a familiar label, and responded via button press whether the label matched the picture (see Figure 2). Participants received auditory feedback: a bell for correct responses and a buzz for incorrect responses. Note that unlike Experiment 2, the word now followed rather than preceded the image. This allowed us to measure the extent to which viewing the image activates its name. Across the experiment, each picture was presented six times in a random order, for a total of 360 trials. The auditory label matched the picture on half of the trials. On the remaining mismatch trials, the label was randomly selected from one of the other categories. All labels were recorded by a female speaker, and all sound files were length-normalised to 750 ms. Presentation of labels followed presentation of pictures after a delay of 200, 300, or 600 ms. Each delay interval was used twice for each exemplar – once when the label and picture matched, and once when they did not. We anticipated that if participants are implicitly naming the pictures (or to put it more passively, if the pictures automatically activate their names), then the more time participants have to generate a name, the faster their responses will be once the name is played making verification RTs a good proxy for speed of picture-naming. This kind of task has certain advantages over overt naming tasks because unlike a naming response, there is no ambiguity as to when a response has occurred. The primary measure of interest was the difference in speed with which people implicitly named pictures from categories of varying sparsity. Verification response times were measured relative to label onset.

#### Flanker congruity task

Participants next completed a version of a flanker task (e.g. Eriksen, 1995). Each trial began with a fixation cross (700–900 ms.) followed by a target and flanker display containing seven small triangles, as shown in Figure 1. The participants' task was to report the direction of the central target. Trials were evenly split among three types: on *congruent* trials, the target and flankers all faced the same direction. On *incongruent* trials, the target faced opposite the flankers. We also included *neutral* trials on which the flankers faced upward. Because "up" is not a possible response, upward facing flankers are predicted to interfere minimally with responding to the direction of the target. The neutral trials served as a baseline allowing us to distinguish between a congruent advantage and an incongruent cost. The trials were additionally split between three delay conditions: simultaneous presentation (standard flanker display), and flanker-first presentation in which the flankers preceded the target by 150 or 500 ms. A longer flanker-to-target delay provides additional time during which people can attempt to selectively attend to the central location in which the target will appear and perhaps inhibit the processing of the irrelevant flankers. As the delay between the onset of the flankers and the onset of the target is increased, the detrimental influence of the incongruent flankers should decrease and the effect of seeing response-*congruent* flankers may increase (for review see Botella, Barriopedro, & Joula, 2002). Examining performance on delay trials allowed us to examine individual differences in selectivity with respect to difficulties in ignoring incongruent information, ignoring congruent (though still irrelevant) information, and the change to this selectivity as a function of time. Participants completed a total of 270 trials with all trial types intermixed. Incorrect responses were signalled by a buzzing sound.

The primary measures of interest in the flanker task were the difference in speed and accuracy of identifying the target in the presence of various flankers (distractors). Less selective representation of the target is expected to lead to increased difficulty in responding when the flankers do not match the target and faster or more accurate responding when they do match the target.

### Analytic approach

We first report the results for each task separately and then examine the relationship between them. All analyses were conducted using mixed effects regression models. Reaction time analyses were conducted using linear mixed models and accuracy analyses were conducted using logistic mixed models. Best fit models and best random effects structure were determined in the same manner as in Experiment 2. All models of picture-word verification include random effects of subject and picture category. All models of flanker performance include random effects of subject. Factors were centered (e.g. ±0.5). All final models are listed in the appendix and referenced by number in the main manuscript (e.g. M1a).

### Results and discussion

#### Picture-word verification RT performance

We first examined how category sparsity and the delay between picture and label affected verification RTs. We excluded incorrect trials and trials on which participants responded faster than 150 ms or slower than 1500 ms (approximately 4% of trials). The full model (Appendix M2a) included the interaction between category sparsity and delay length and main effects of category sparsity and delay length. Participants responded more quickly with increasing delays (200 ms delay: $M = 678$ ms; 300 ms delay: $M = 648$ ms; 600 ms delay: $M = 617$ ms), $b = -136$, $CI_{95\%} = [-171, -101]$, $\chi^2 = 224$, $p < .00001$. A parsimonious way to think about name-verification RTs decreasing as the delay increased is that on seeing a picture, participants automatically generate its name and then compare it with the word they hear. There was no main effect of category sparsity on RTs, $\chi^2 = .90$, $p = .34$, and no interaction between delay and category sparsity $\chi^2 = .13$, $p = .72$.

#### Picture-word verification accuracy

We next examined how category sparsity and delay affected verification accuracy. The full model (Appendix M2b) included the interaction between category sparsity and delay length and main effects of category sparsity and delay length. We again found an effect of delay, $b = 1$, $CI_{95\%} = [-.3, 2]$, $\chi^2 = 9.16$, $p = .002$. Participants were faster to verify pictures with increasing delays. There was no interaction between delay and category sparsity, $\chi^2 = .01$, $p = .92$, nor a main effect of category sparsity, $\chi^2 = .01$, $p = .93$.

Overall, people's recognition/naming of pictures was independent of those pictures' category structure, as measured by our sparsity measure. The critical question is whether performance on categories of varying sparsity related to differences in performance on the flanker task.

#### Flanker congruency task performance

Our primary analysis examined how trial type and delay affected responses to the target direction. We excluded data from incorrect trials and trials on which participants responded faster than 150 ms or slower than 1100 ms (approximately 4% of trials). The full model (Appendix M3a) included the interaction between trial type and delay and main effects of trial type and delay. Our basic flanker results were typical of this task: there was an effect of trial type, $\chi^2(2) = 456.84$, $p < .0001$ such that people were slower to respond on incongruent trials, $M = 504$ ms, than on neutral, $M = 460$ ms, or congruent trials, $M = 450$ ms, and an effect of delay length, $\chi^2(2) = 1873.1$, $p < .0001$, such that participants were overall faster to respond with increasing delays. However, as can be seen in Figure 4(a), there was a significant trial-type by delay interaction, $\chi^2(2) = 25.29$, $p < .0001$. Follow-up models comparing performance on neutral trials to performance on each of the other trial types revealed that this interaction was driven by the incongruent cost decreasing with the delay, $b = 63$, $CI_{95\%} = [39, 87]$, $\chi^2(1) = 26.89$, $p < .001$ (Appendix M3b) and the congruent advantage increasing with the delay, $b = -40$, 95% CI = [−64, −17], $\chi^2(1) = 11.18$, $p = .0008$ (Appendix M3c). Participants were much slower on incongruent than neutral trials on trials when the target and flanker appeared simultaneously (simultaneous trials) and were only slightly slower on incongruent than neutral trials when the flankers appeared first (150 ms and 500 ms delay trials). On the other hand, participants were only faster on congruent than neutral trials when there was a delay between target and flanker onset.

Turning now to an analysis of accuracy, our full model (Appendix M3d) included the interaction between trial type and delay and main effects of trial type and delay. We found a significant effect of trial type, $\chi^2(1) = 119.42$, $p < .0001$, such that although participants were overall equally accurate on congruent trials, $M = .99$, and neutral trials, $M = .99$, they were slightly less accurate on incongruent trials, $M = .95$. There was no significant interaction between trial type and delay, $\chi^2(2) = 3.23$, $p = .20$ nor a significant effect of delay, $\chi^2(1) = .20$, $p = .66$. In other words, unlike the RT cost, the accuracy cost on incongruent trials did not decrease with greater delays.

### *Individual differences in flanker effectiveness*

To relate people's flanker performance to their picture-word verification performance we needed a way to quantify the effectiveness of the flankers – the extent to which incongruent flankers slowed people down and/or the extent to which congruent flankers sped people up. Each person's congruent-trial advantage was calculated by dividing their average RT on neutral trials by their average RT on congruent trials during sequential presentations. Each person's incongruent-trial RT cost was calculated by dividing their average RT on incongruent trials by their average RT on neutral trials during sequential presentations. We focused on performance on sequential trials because we were interested in examining both the incongruent cost and the congruent advantage, and although participants demonstrated an incongruent cost at each delay length, they only demonstrated an congruent advantage existed only trials (a result typical of other flanker studies, e.g. Botella et al., 2002).

The incongruent accuracy cost was calculated in the same manner:

$$\text{Congruent - trial RT advantage} = \text{Neutral} - \text{trial RT}/\text{Congruent} - \text{trial RT}, \ \text{Incongruent} - \mathit{trialRTcost} = \text{Incongruent} - \text{trialRT}/\text{Neutral} - \text{trialRT}, \ \text{Incongruent - trial accuracy cost} = \text{Incongruent} - \text{trialaccuracy}/\text{Neutral} - \text{trialaccuracy}.$$

The size of a participant's incongruent-trial cost and congruent-trial advantage indicates the extent to which they utilised more integrative or selective representations in the flanker task. If categorising objects belonging to sparse categories (e.g. ZEBRA) requires a more selective representation than categorising objects belonging to dense (e.g. DOG), then participants who were the most selective in the flanker task (i.e. those who have a small incongruent-trial cost and/or small congruent-trial advantage) should be faster to categorise sparse categories than the participants who were the least selective.

## *Relationship between flanker effectiveness and picture-word verification*

### *Congruent-trial advantage*

We first examined how category sparsity and participants' selectivity in the flanker task, as measured by congruent-trial advantage size, affected verification RTs. The full model included the interaction between category sparsity and congruent-trial RT advantage, the interaction between category sparsity and congruent-trial accuracy advantage and main effects of category sparsity, congruent-trial RT advantage, and congruent-trial accuracy advantage (Appendix M4a). As can be seen in Figure 4(b), there was a significant interaction between category sparsity and congruent-trial RT advantage, $b = 198$, $CI_{95\%} = [1, 394]$, $\chi^2(1) = 3.87$, $p = .049$. People showing a smaller congruent-trial RT advantage were more likely to show an effect of category sparsity (being faster to verify pictures from sparse than from dense) than people showing a larger congruent-trial RT advantage. There was no interaction between category sparsity and congruent-trial accuracy advantage, $\chi^2(1) = .02$, $p = .87$. There was no main effect of category sparsity, $\chi^2(1) = .89$, $p = .34$. Overall performance was likewise not predicted by either the congruent-trial RT advantage, $\chi^2(1) = .68$, $p = .41$ or the congruent-trial accuracy advantage, $\chi^2(1) = 2.38$, $p = .12$.

The interaction between category sparsity and the size of participants' congruent-trial RT advantage in the flanker task indicated that those with the best ability to selectively represent relevant information in the flanker

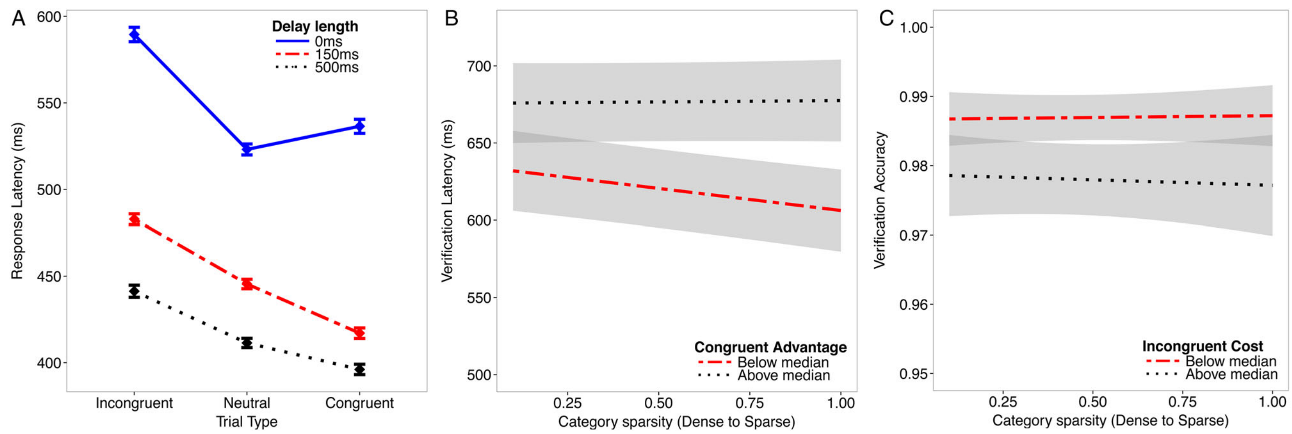


**Figure 4.** (a) Average reaction time (ms) for each trial type and delay periods in the flanker task. Note: Error bars depict standard error of the means with between-subject variance removed (Morey, 2008). (b) Picture-word verification reaction time (ms) based on category sparsity and congruent advantage (in RTs) during sequential flanker presentations (delay 150 and 500 ms). (c) Picture-word verification accuracy based on category sparsity and incongruent cost during sequential flanker presentations. "Below-median" and "above-median" groups show performance of participants who have a below and above-median incongruent cost on sequential flanker trials, respectively. Category sparsity is represented from left to right as dense to sparse based on the highest proportion of participants in the norming study listing the same commonality. Note: Error bands depict standard error of predicted means.

task were also best at selectively representing relevant information in sparse categorisation (as measured by picture-word verification).

### *Incongruent-trial cost*

We next examined how category sparsity and participants' selectivity in the flanker task, as measured by incongruent-trial cost size, affected verification accuracy. The full model included the interaction between category sparsity and incongruent-trial RT cost, the interaction between category sparsity and incongruent-trial accuracy cost and main effects of category sparsity, incongruent-trial RT cost, and incongruent-trial accuracy cost (Appendix M4b). Figure 4(c) shows that there was a marginal interaction between category sparsity and incongruent-trial RT cost on verification accuracy, $b = 7$, $CI_{95\%} = [-.6, 15]$, $\chi^2(1) = 3.51$, $p = .06$ and a marginal interaction between category sparsity and incongruent-trial accuracy cost, $b = 6$, $CI_{95\%} = [-.6, 13]$, $\chi^2(1) = 3.18$, $p = .07$. Participants with the largest costs (in RTs and in accuracy) were slightly less accurate in verifying the names of pictures from sparse categories. We also found a significant main effect of incongruent-trial accuracy cost, $b = 6$, $CI_{95\%} = [-.6, 13]$, $\chi^2(1) = 10$, $p = .002$, such that participants with the largest costs were slightly less accurate overall. There were no main effects of incongruent-trial RT cost, $\chi^2(1) = 2.06$, $p = .15$ or of category sparsity, $\chi^2(1) = .01$, $p = .93$. Although participants were overall quite accurate in both the flanker and the picture-word verification tasks (Figure 4(c)), the small, but systematic relationship between verification accuracy and the effect of incongruent flankers add to the evidence from the congruent-trial advantage effect on verification RTs (see Figure 4(b)). To summarise, those with the best ability to selectively represent relevant information in the flanker task were somewhat faster and more accurate to categorise pictures from sparse categories.

### *Conclusions*

The results of Experiment 3 suggest that selective representation demands are slightly higher for sparse than dense categorisation – even for the categorisation of highly familiar objects like zebras and carrots. Left unclear, however, is what role, if any, language may have in supporting this selective representation. In Experiment 4, we examined whether perturbing neural activity associated with labelling affects picture-word verification with respect to the sparsity of the categories to which pictures belong.

## Experiment 4

If verbal labelling helps to selectively represent category-relevant information, then manipulating the ease with people can use labels should perturb this process. We manipulated labelling through direct current stimulation of Wernicke's area (see Perry & Lupyan, 2013, 2014 for rationale).

tDCS is a painless, noninvasive technique used to temporarily alter cortical excitability through weak electrical current to the scalp which is thought to affect cortical excitability through changes in transmembrane potential causing changes in spontaneous firing (Iyer et al., 2005;

Nitsche & Paulus, 2000; Wagner et al., 2007). Placing the cathode over the site of interest is generally thought to decrease excitability; placing the anode is thought to increase it (Nitsche & Paulus, 2000). tDCS over Wernicke's area has been previously used to in studies examining the role of labelling in, for example, learning of novel categories (Perry & Lupyan, 2014). To explore whether labelling supports selective representation in categorisation of even highly familiar objects, we asked whether tDCS-induced differences in picture-word verification mimic the individual differences found in Experiment 3.

## Method

### Participants

Twenty monolingual English-speaking undergraduates participated for course credit. We randomly assigned participants to one of two between-subjects conditions: anodal-stimulation ($n = 10$), and cathodal-stimulation ($n = 10$). Exclusion criteria included being left handed, having a history of neurologic or psychiatric disease, or use of anti-convulsants, anti-psychotic, or sedative medications.

### tDCS procedure

tDCS was delivered by a battery-driven constant direct-current stimulator (Soterix 1 × 1 low-intensity stimulator). Rubber electrodes were inserted into saline-soaked 5 × 7 $cm^2$ sponges. Placement of the stimulation electrode was made by reference to the 10–20 system: intersection of T5–C3 and T3–P3 for Wernicke's area, posterior region of BA22 (Homan, Herman, & Purdy, 1987). The reference electrode (cathodal or anodal depending on condition) was attached to the right cheek. At the start, current was increased over 30 s. to 1.75 mA, and the task than began. Current lasted 20 minutes, the approximate task length.

### Picture-word verification task

We administered the same picture-word verification task as used in Experiment 3.

### Analytic approach

As in Experiment 3, we excluded data from incorrect trials and trials on which participants responded faster than 150 ms or slower than 1500 ms (<3% of trials). All analyses were conducted using mixed effects regression models. Reaction time analyses were conducted using linear mixed models and accuracy analyses were conducted using logistic mixed models. Best fit models and best random effects structure were determined in the same manner as in Experiments 2–3. All models included random effects of subject and picture category. Factors were centered (e.g. ±0.5). All final models are listed in the appendix and referenced by number in the main manuscript (e.g. M1a).

## Results and discussion

Our main prediction was that up-regulating activity over Wernicke's area – insofar as it enhances labelling – should selectively enhance performance on categorising items from sparse categories, and/or that down-regulating activity over Wernicke's area – insofar as it impairs the labelling process – should selectively impair performance on categorising items from sparse category trials.

### Effects of tDCS on picture-word verification RTs

We first examined whether stimulation condition, category sparsity, and delay between picture and label had an effect on verification speed (see Figure 5(a)). The full model included the three-way interaction between stimulation condition, category sparsity, and delay, and interactions between stimulation condition and category sparsity, between stimulation condition and delay, and between category sparsity and delay, and main effects of stimulation condition, category sparsity, and delay (Appendix M5a). As can be seen in Figure 5(a), we found that participants in the cathodal-stimulation condition had marginally slower RTs ($M = 689$ ms), than those in the anodal-stimulation condition ($M = 635$ ms), $b = 38$ ms, $CI_{95\%} = [-30, 107]$ $\chi^2(1) = 2.71$, $p = .10$. We also found a significant interaction between stimulation condition and category sparsity, $b = 36$ ms, $CI_{95\%} = [6, 65]$ $\chi^2(1) = 5.67$, $p = .02$ such that those in the anodal-stimulation condition were faster to verify pictures from sparse than dense categories, but those in the cathodal-stimulation condition were not affected by category sparsity. Finally, we found a significant main effect of delay such that participants were faster to respond with increasing delays, $b = -100$ ms, $CI_{95\%} = [-134, -66]$, $\chi^2(1) = 65.47$, $p < .0001$. There was no significant three-way interaction, $\chi^2(1) = .01$, $p = .91$ no interaction between stimulation condition and delay, $\chi^2(1) = .01$, $p = .94$, no interaction between category sparsity and delay, $\chi^2(1) = .98$, $p = .32$; and no main effect of category sparsity, $\chi^2(1) = 1.65$, $p = .20$.

Next, we conducted follow-up comparisons to examine the significant two-way interaction between stimulation condition and category sparsity. We first looked at the effect of category sparsity for only those in the anodal-stimulation condition (Appendix M5b). There was a marginally significant effect of category dimensionality, $b = 38$ ms, $CI_{95\%} = [-30, 107]$, $\chi^2(1) = 3.42$, $p = .06$ such that these participants were faster to verify the names of pictures from sparse than dense

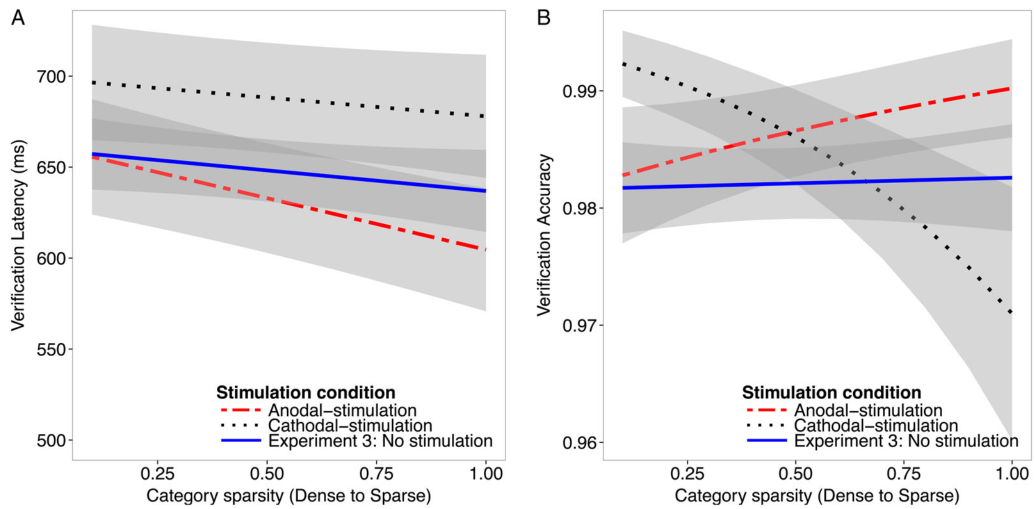


**Figure 5.** (a) Picture-word verification reaction time (ms) based on category sparsity and stimulation condition. (b) Picture-word verification accuracy based on category sparsity and stimulation condition. Note: Error bands depict standard error of predicted means.



categories. On the other hand, we did not find an effect of category sparsity for participants in the cathodal-stimulation condition, $\chi^2(1) = .39$, $p = .53$ (Appendix M5c).

Next, we compared the effect of category sparsity for participants in each stimulation group to the no-stimulation participants from Experiment 3 treating the latter as a baseline control. We first looked only at participants in the anodal-stimulation and no-stimulation conditions (Appendix M5d). There was a significant interaction between stimulation condition and category sparsity, $b = 34$ ms, $CI_{95\%} = [-30, 107]$, $\chi^2(1) = 8.11$, $p = .004$ such that those in the anodal-stimulation condition were faster to verify pictures from sparse than dense categories, but those in the no-stimulation condition were not affected by category sparsity. There were no main effects of stimulation condition, $\chi^2(1) = .16$, $p = .269$ or category sparsity, $\chi^2(1) = 1.46$, $p = .23$.

We next looked at participants in the cathodal-stimulation and no-stimulation conditions (Appendix M5e). For these participants, there was no interaction between stimulation condition and category sparsity, $\chi^2(1) = .02$, $p = .88$ and no significant effects of stimulation condition, $\chi^2(1) = 1.44$, $p = .23$ or category sparsity, $\chi^2(1) = .75$, $p = .39$. Thus, although up-regulating activity over Wernicke's area (via anodal tDCS) appears to have facilitated the categorisation speed of objects belonging to sparse categories, down-regulating activity (via cathodal tDCS) did not affect verification speed when contrasted with performance of no-stimulation controls from Experiment 3.

### *Effects of tDCS on picture-word verification accuracy*

We next examined whether stimulation condition, category sparsity, and delay between picture and label had an effect on verification accuracy (see Figure 5(b)). The full model included the three-way interaction between stimulation condition (anodal, cathodal), category sparsity, and delay, interactions between stimulation condition and category sparsity, between stimulation condition and delay, and between category sparsity and delay, main effects of stimulation condition, category sparsity, and delay (Appendix M6a). We found a significant two-way interaction between stimulation condition and category sparsity, $b = -2$, $CI_{95\%} = [-4, -1]$, $\chi^2(1) = 12.38$, $p = .0004$ such that those in the cathodal-stimulation condition were slightly less accurate at verifying pictures from sparse than dense categories and those in the anodal-stimulation condition were not affect by category sparsity (see Figure 5(b)). There was no interaction between stimulation condition and delay, $\chi^2(1) = .96$, $p = .33$ or between category sparsity and delay, $\chi^2(1) = 10$, $p = .75$ and no main effects of stimulation, $\chi^2(1) = .00001$, $p = .99$, category sparsity, $\chi^2(1) = .69$, $p = .41$, or delay, $\chi^2(1) = .89$, $p = .35$. However, we did find a marginal three-way interaction between stimulation category sparsity and delay condition, $b = -8$, $CI_{95\%} = [-15, .03]$, $\chi^2(1) = 3.67$, $p = .06$.

Follow-up comparisons revealed that the three-way interaction between stimulation, category sparsity and delay was driven by performance on the longer delay trials. A model of the interaction between stimulation condition and category sparsity on only the longer delay trials (Appendix M6b) revealed a significant interaction between stimulation condition and category sparsity, $b = -3$, $CI_{95\%} = [-4, -1]$, $\chi^2(1) = 12.82$, $p = .0003$ such that those in the cathodal-stimulation condition are less accurate to verify pictures from sparse than dense categories, $b = -2$, $CI_{95\%} = [-3, -1]$, $\chi^2(1) = 8.64$, $p = .003$

(Appendix M6c), but those in the anodal-stimulation condition were not affected by category sparsity, $\chi^2(1) = 1.07$, $p = .30$ (Appendix M6d). There were no main effects of stimulation condition, $\chi^2(1) = .05$, $p = .83$, or category sparsity, $\chi^2(1) = 1.03$, $p = .31$.

Additional models of the effect of category sparsity across delays for only those in each stimulation condition revealed a significant effect of category sparsity for those in the cathodal-stimulation condition (Appendix M6e), $b = -1$, $CI_{95\%} = [-2, -1]$, $\chi^2(1) = 6.52$, $p = .01$ such that participants were less accurate at verifying the names of sparse than dense categories, regardless of delay length, but those in the anodal-stimulation condition did not reveal an effect of category sparsity (Appendix M6f), $\chi^2(1) = 1.33$, $p = .25$. Together these results – although admittedly very small in magnitude – complement the pattern of RTs reported above by showing stimulation differentially affects verification of sparse vs. dense categories.

Finally, we compared the effect of category sparsity on accuracy for participants in each stimulation condition to the no-stimulation participants from Experiment 3. We first compared those in the cathodal-stimulation and no-stimulation conditions using a model of the interaction between stimulation condition and category sparsity and main effects of stimulation condition and category sparsity (Appendix M6g). There was a significant interaction between stimulation condition and category sparsity, $b = 2$, $CI_{95\%} = [1, 2]$, $\chi^2(1) = 10.39$, $p = .001$ such that participants were less accurate at verifying pictures from sparse than dense categories. There were no main effects of stimulation condition, $\chi^2(1) = .56$, $p = .46$, or category sparsity, $\chi^2(1) = .33$, $p = .57$. On the other hand, when we compared those in the anodal-stimulation and no-stimulation conditions using a model of the interaction between stimulation condition and category sparsity and main effects of stimulation condition and category sparsity (Appendix M6h), did not find an interaction between stimulation condition and category sparsity, $\chi^2(1) = 1.36$, $p = .24$, nor main effects of stimulation condition, $\chi^2(1) = .57$, $p = .45$, or category sparsity, $\chi^2(1) = .17$, $p = .68$. These results suggest that cathodal, but not anodal stimulation affected accuracy relative to baseline performance (Experiment 3). Down-regulating activity over Wernicke's area (via cathodal tDCS) appears to have disrupted the categorisation of objects belonging to sparse categories (although by a very small degree).

### *Conclusions*

The results of Experiment 4 suggest that anodal stimulation selectively decreased participants' picture-word verification RTs (i.e. improved performance), specifically for more sparse categories. Cathodal stimulation did not selectively affect RTs as a function of category sparsity, but did lead to small decreases in accuracy for more sparse categories. Insofar as our stimulation regime affects a process related to labelling, these findings, although small in size, provide some evidence that labelling may selectively affect recognition and/or naming of objects from sparse categories.

## General discussion

Our main goal was to explore the processes required for familiar object recognition, especially recognition of sparse category members. We asked (1) whether selective representation demands are greater for recognising members of sparser than denser categories and (2) whether verbal labelling is involved in identifying members of sparse categories.

In Experiment 1, we quantified the category sparsity of 12 familiar categories that serve as stimuli in Experiments 3 and 4. In Experiment 2, we validated our measure of category sparsity, demonstrating that when participants distinguish category members from non-members, specific features (e.g. zebra stripes) are more central to recognising members of sparse vs. dense categories.

In Experiment 3, we found that participants demonstrating more selective processing in the flanker task – as measured by their smaller congruent-trial flanker advantage – were faster to verify the names of pictures belonging to sparse categories than the participants exhibiting more integrative processing. Additionally, we found that participants' incongruent-trial cost in the flanker task was marginally related to their verification performance – though only in accuracy. We discuss implications of these findings below.

In Experiment 4, we found that up-regulating neural activity over Wernicke's area led to an increase in the speed with which participants verified the names of sparse category members. Down-regulation led a slight decrease in the accuracy with which participants verified the names of sparse. Finally, down-regulating neural activity over this same area led to slightly slower overall verification RTs. Together, these findings offer preliminary evidence that verbal labelling increases the selectivity of representations, important for recognising even highly familiar objects, especially those belonging to sparser categories.

Earlier studies have demonstrated a link between language and category structure: sparse categories have previously been argued to have easily verbalisable rules (see Ashby & Maddox, 2011) and interfering with language appears to disrupt learning these categories.

However, it has remained unclear *why* such a link might exist. The present studies move beyond earlier work, demonstrating that selective representation demands are greater for recognising members of sparse compared to dense categories and that verbal labelling may aid this process of selective representation.

### Feature names or category names

In Experiment 4, we sought to up- and down-regulate the labelling process via tDCS over Wernicke's area, finding that up-regulation led to a systematic change in picture-word verification speed, particularly for pictures of objects from sparse categories. There are two (not necessarily mutually exclusive) ways in which labelling may have affected recognition/naming of such items: (1) labelling of the category-relevant feature (e.g. STRIPES) and/or (2) labelling of the category itself (e.g. ZEBRA). Evidence from developmental psychology suggests that participants do not have to actively label the feature itself for there to be an effect of labelling on selective attention to the feature. For example, children prioritise shape when recognising members of familiar categories (e.g. Yee, Jones, & Smith, 2012) and when labelling novel categories (e.g. Perry & Samuelson, 2011; Perry, Samuelson, Malloy, & Schiffer, 2010) even at an age when they do not know any names for specific shapes, such as "square" or "triangle" (Dale & Fenson, 1996). Additionally, they do not prioritise shape in non-labelling contexts, such as grouping novel nameless objects by similarity (Landau, Smith, & Jones, 1988). Thus, although children can prioritise an object's shape without labelling that *shape*, labelling the *whole object* as a member of a category is what draws their attention to that critical feature.

An account for how labelling may influence selective representation is Lupyan's label-feedback hypothesis (Lupyan, 2007, 2012a). In learning a name for a sparse category like ZEBRA, the category label becomes strongly associated with a few visual features (those most critical to category membership). The activation of the label feeds back to lower-level visual representations and helps to "clean up" the representation, down-weighting irrelevant features and highlighting the relevant ones (Lupyan, 2012a, 2012b). Nevertheless, it remains possible that feature labels also play a role in the recognition process.

### Behavioural consequences of tDCS

Cathodal-stimulation led to a general increase in RTs – regardless of category sparsity and to a small, but statistically significant decrease in verification accuracy for sparse categories. Anodal-stimulation specifically led to a decrease in verification RTs for sparse categories. Although anodal- and cathodal-stimulation are often described as having opposing effects via increases and decreases of cortical excitability, respectively, this may not always be the case (Batsikadze, Moliadze, Paulus, Kuo, & Nitsche, 2013; Lupyan et al., 2012; Nozari, Woodard, & Thompson-Schill, 2014), and opposing effects of cortical excitability do not necessarily lead to exactly opposite behavioural effects (e.g. Nitsche et al., 2003).

The size of effects of tDCS on picture-word verification performance as a function of category sparsity (Experiment 4) was quite small, particularly on accuracy, but this not detract from the importance of the present findings. After all, tDCS in no way entirely disrupts verbal labelling. These effects are nonetheless important in demonstrating causality: we are creating slight perturbations in a cognitive process to explore how that process relates to some aspect of behaviour. The small but systematic effects thus provide preliminary support for the idea that labelling is an important process involved in the selective representation needed for sparse categorisation.

### Within-category commonalities vs. between-category differences

If one knows that something is a carrot, one can, with high probability, expect it to be orange. Knowing that something is orange, however, will only slightly increase the probability that it is a carrot. Orange-ness is thus a highly useful cue in helping to group carrots together, but not very useful in distinguishing between carrots and non-carrots if the non-carrots also happen to be orange. Because the particular categories we used in Experiments 3 and 4 did not share critical features and were not semantically associated with the other categories we used, within-category feature commonalities were also useful for making between-category discriminations.

Nevertheless, more work is needed to understand how category sparsity bears on distinguishing category members from nonmembers when the set of categories is not as limited, for example, when distinguishing between related categories (e.g. cat and dog) or multiple categories sharing a critical feature (e.g. carrot and pumpkin). An important consideration for future work is that distinguishing category members from non-members is highly context dependent – the information needed to decide that a pumpkin is not a carrot (shape) is different from the information needed to decide an orange habanero pepper is not a carrot (texture, taste), and different still from the information needed to decide a goldfish is not a carrot (shape, texture, animacy). Such context-dependency will likely mean that the sparsity of a category may be quite flexible, depending on the contextually relevant set of contrasting categories.

## *Conclusion*

When recognising an object as a member of a category, people must selectively represent category-relevant information. This process is more central when recognising members of sparse categories such as a zebra, and less important for recognising members of dense categories such as dog. Modulating processes involved in labelling via tDCS led to small, but systematic changes in the speed and accuracy with which people could recognise and/or name members of sparse categories. We provide preliminary evidence that the process of verbal labelling supports selective representation even in the context of recognising and naming members of highly familiar categories.

## Notes

1. This characterisation of categorisation leaves unspecified whether it is the perceptual inputs that are thus transformed, or whether the transformation happens as part of post-perceptual processing.
2. Although certain features like spatial and acoustic frequency, orientation, size, motion direction, and a number of geometric properties like concavity may be basic in the sense of being ready-made objects of attentional selection (Wolfe & Horowitz, 2004), attempts to derive a vocabulary of basic features in perception, and much less in higher-level cognition and language (Evans & Levinson, 2009) has not been successful. It seems to us more likely that features are emergent higher-level units derived from learning environmental co-occurrences (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Schyns, Goldstone, & Thibaut, 1998) and thus will vary depending on the experiences of an organism as well as current task demands. A familiar example of such emergent units are the "chunks" first described by Miller (1956).
3. We include in the definition of dense categories ad-hoc categories such as THINGS COMMONLY FOUND IN A KITCHEN (Lupyan et al., 2012)
4. There are many components to the labelling/naming process, and we cannot claim that our direct current stimulation procedure is manipulating a specific process. For simplicity we will use the general term "labelling" throughout this paper, with the understanding that that labelling is a complex and multifaceted process.
5. It was somewhat unexpected to not find a main effect of semantic association strength, or especially an effect of semantic association strength just on semantic foil trials, $\chi^2(1) = .67$, $p = .41$. However, we had selected the particular categories we did to be evenly spread across the sparsity continuum rather than to be evenly distributed in semantic association strength, and this could have been an artefact of the particular categories chosen.


## Acknowledgments

We thank Jesse Sherman, Ishaan Guptasarma, Oliver Roe, and Callie Porter-Borden for their help in data collection. We would also like to thank Bob McMurray and two anonymous reviewers for their thoughtful feedback on an earlier version of this manuscript.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The study was partially supported by NSF#1331293 to G. L.


## References


Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*(1), 147–161. doi:10.1111/j.1749-6632.2010.05874.x

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*(1), 3–26.

Banich, M. T. (2009). Executive function the search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89–94. doi:10.1111/j.1467-8721.2009.01615.x

Batsikadze, G., Moliadze, V., Paulus, W., Kuo, M. F., & Nitsche, M. A. (2013). Partially non-linear stimulation intensity-dependent effects of direct current stimulation on motor cortex excitability in humans. *The Journal of Physiology*, *591*(Pt 7), 1987–2000. doi:10.1113/jphysiol.2012.249730

Botella, J., Barriopedro, M. I., & Joula, J. F. (2002). Temporal interactions between target and distractor processing: Positive and negative priming effects. *Psicológica*, *23*(2), 371–400.

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *The Journal of Neuroscience*, *35*(25), 9329–9335. doi:10.1523/JNEUROSCI.5111-14.2015

Cagigas, X. E., Filoteo, J. V., Stricker, J. L., Rilling, L. M., & Friedrich, F. J. (2007). Flanker compatibility effects in patients with Parkinson's disease: Impact of target onset delay and trial-by-trial stimulus variation. *Brain and Cognition*, *63*(3), 247–259. doi:10.1016/j.bandc.2006.09.002

Couchman, J. J., Coutinho, M. V. C., & Smith, J. D. (2010). Rules and resemblance: Their changing balance in the category learning of humans (Homo sapiens) and Monkeys (Macaca mulatta). *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(2), 172–183. doi:10.1037/a0016748

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125–127. doi:10.3758/BF03203646

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*(2–3), 101–118. doi:10.1080/13506289508401726

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(05), 429–448. doi:10.1017/S0140525X0999094X

Gainotti, G. (2014). Old and recent approaches to the problem of non-verbal conceptual disorders in aphasic patients. *Cortex*, *53*, 78–89. doi:10.1016/j.cortex.2014.01.009

Homan, R. W., Herman, J., & Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, *66*(4), 376–382. doi:10.1016/0013-4694(87)90206-9

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *The Behavioral and Brain Sciences*, *24*(5), 849–878; discussion 878–937.

Iyer, M. B., Mattu, U., Grafman, J., Lomarev, M., Sato, S., & Wassermann, E. M. (2005). Safety and cognitive effect of frontal DC brain polarization in healthy individuals. *Neurology*, *64*(5), 872–875. doi:10.1212/01.WNL.0000152986.07469.E9

Jaeger, T. F. (2009). Random effect: Should I stay or should I go? Retrieved February 8, 2013, from http://hlplab.wordpress.com/2009/05/14/random-effect-structure/

Kan, I. P., & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(4), 466–482. doi:10.3758/CABN.4.4.466

Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General*, *137*(1), 52–72. doi:10.1037/0096-3445.137.1.52

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321. doi:10.1016/0885-2014(88)90014-7

Lupyan, G. (2007). *The label feedback hypothesis: Linguistic influences on visual processing*. PhD. Thesis. Pittsburgh, PA: Carnegie Mellon University.

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review*, *16*(4), 711–718. doi:10.3758/PBR.16.4.711

Lupyan, G. (2012a). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Cognition*, *3*, 54. doi:10.3389/fpsyg.2012.00054

Lupyan, G. (2012b). What do words do? Toward a theory of language-augmented thought. *Psychology of Learning and Motivation–Advances in Research and Theory*, *57*, 255.

Lupyan, G., & Mirman, D. (2013). Linking language and categorization: Evidence from aphasia. *Cortex*, *49*(5), 1187–1194. doi:10.1016/j.cortex.2012.06.006

Lupyan, G., Mirman, D., Hamilton, R., & Thompson-Schill, S. L. (2012). Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex. *Cognition*, *124*(1), 36–49. doi:10.1016/j.cognition.2012.04.002

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking redundant labels facilitate learning of novel categories. *Psychological Science*, *18*(12), 1077–1083. doi:10.1111/j.1467-9280.2007.02028.x

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, *141*(1), 170–186. doi:10.1037/a0024904

Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research*, *74*(2), 219–236. doi:10.1007/s00426-009-0245-z

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. doi:10.1037/h0043158

Minda, J. P., Desroches, A. S., & Church, B. A. (2008a). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(6), 1518–1533. doi:10.1037/a0013355

Minda, J. P., Desroches, A. S., & Church, B. A. (2008b). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1518–1533. doi:10.1037/a0013355

Minda, J. P., & Miles, S. J. (2010). The influence of verbal and nonverbal processing on category learning. In Brian H. Ross (Ed.), *Psychology of learning and motivation* (Vol. *52*, pp. 117–162). Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079742110520036

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition*, *80*(3), B11–B20.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. doi:10.3758/BF03195588

Nitsche, M. A., Liebetanz, D., Lang, N., Antal, A., Tergau, F., & Paulus, W. (2003). Safety criteria for transcranial direct current stimulation (tDCS) in humans. *Clinical Neurophysiology*, *114*(11), 2220–2222. doi:10.1016/S1388-2457(03)00235-9

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of Physiology*, *527*(Pt 3), 633–639.

Nozari, N., Woodard, K., & Thompson-Schill, S. L. (2014). Consequences of cathodal stimulation for behavior: When does it help and when does it hurt performance? *PLoS ONE*, *9*(1), e84338. doi:10.1371/journal.pone.0084338

Perry, L. K., & Lupyan, G. (2013). What the online manipulation of linguistic activity can tell us about language and thought. *Frontiers in Behavioral Neuroscience*, *7*, 122. doi:10.3389/fnbeh.2013.00122

Perry, L. K., & Lupyan, G. (2014). The role of language in multi-dimensional categorization: Evidence from transcranial direct current stimulation and exposure to verbal labels. *Brain and Language*, *135*, 66–72. doi:10.1016/j.bandl.2014.05.005

Perry, L. K., & Samuelson, L. K. (2011). The shape of the vocabulary predicts the shape of the bias. *Frontiers in Psychology*, *2*. doi:10.3389/fpsyg.2011.00345

Perry, L. K., & Samuelson, L. K. (2013). The role of verbal labels in attention to dimensional similarity. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuch (Eds.), *Proceedings of the thirty-fifth annual conference of the cognitive science society*. (pp. 3217–3222). Austin, TX: Cognitive Science Society.

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894–1902. doi:10.1177/0956797610389189

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–81. doi:S0010–0277(07)00108-4

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences*, *28*(01), 1–14. doi:10.1017/S0140525X05000014

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, *197*(Pt 3), 335–359. doi:10.1046/j.1469-7580.2000.19730335.x

Reppen, R., Ide, N., & Suderman, K. (2005). *American National Corpus (ANC) second release [DVD]*. Philadelphia.

Roberts, R. E., Anderson, E. J., & Husain, M. (2010). Expert cognitive control and individual differences associated with frontal and parietal white matter microstructure. *The Journal of Neuroscience*, *30*(50), 17063–17067. doi:10.1523/JNEUROSCI.4879-10.2010

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. doi:10.1016/0010-0285(75)90024-9

Rosch, E., Simpson, C., & Scott, R. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 491–502. doi:10.1037/0096-1523.2.4.491

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences*, *21*(01), 1–17.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*(7), 1244–1286. doi:10.1111/j.1551-6709.2010.01129.x

Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology*, *24*(2), 279–298. doi:10.1016/0022-0965(77)90007-8

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. doi:10.1038/nature04171

Wagner, T., Fregni, F., Fecteau, S., Grodzinsky, A., Zahn, M., & Pascual-Leone, A. (2007). Transcranial direct current stimulation: A computer-based human model study. *NeuroImage*, *35*(3), 1113–1124. doi:10.1016/j.neuroimage.2007.01.027

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–11. doi:10.3758/BF03202594

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. doi:10.1038/nrn1411

Yee, M., Jones, S. S., & Smith, L. B. (2012). Changes in visual object recognition precede the shape bias in early noun learning. *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00533

## Appendix

### *Experiment 2*

sparsity = proportion of people listing the same common feature for the target category (coded low to high as dense to sparse)

distractorType = whether or not the picture was a feature foil, semantic foil, or matched the target word. Coded as ±0.5. (Match trials were excluded from this analysis).

semanticStrength = strength of association between target and semantic foil

strength = median split of semanticStrength

```
pv$strength<-factor(pv$semanticStrength>=median(pv$semanticStrength),
        levels=c("FALSE","TRUE"),
        labels=c("Low Semantic Strength","High
            Semantic Strength"))
```

```
M1a<-lmer (latency ~ sparsity * distractorType * semanticStrength + (1|subjCode) + (1|picCategory), data=subset(pv, distractorType!="match"))
```

```
M1b<-lmer (latency ~ sparsity * distractorType + (1|subjCode) + (1|picCategory), data=subset(pv, distractorType!="match" & strength== "Low Semantic Strength"))
```

```
M1c<-lmer (latency ~ sparsity + (1|subjCode) + (1|picCategory), data=subset(pv, distractorType="feature" & strength== "Low Semantic Strength"))
```

```
M1d<-lmer (latency ~ sparsity + (1|subjCode) + (1|picCategory), data=subset(pv, distractorType="semantic" & strength== "Low Semantic Strength"))
M1e<-lmer (latency ~ sparsity * distractorType + (1|subjCode) + (1|picCategory), data=subset(pv, distractorType!="match" & strength== "High Semantic Strength"))
```

### Experiment 3

#### *Picture-word verification*

sparsity = proportion of people listing the same common feature for the target category (coded low to high as dense to sparse)

delay = length of time between presentation of picture and word (200, 300, or 600ms)

pictureCategory = category to which each picture belongs

M2a<-lmer (latency ~ sparsity * delay + (1|subjCode) + (1|pictureCategory), data=p)

M2b<-glmer (isRight ~ sparsity * delay + (1|subjCode) + (1|pictureCategory), data=p, family="binomial")

### *Flanker task*

trialType = how the flanker stimuli corresponded to the target stimulus (contrast coded: incongruent: −.5, neutral: 0, congruent: .5)

delay = length of time between onset of flankers and onset of target stimulus (0, 150, 500ms)

M3a <-lmer (latency ~ trialType * delay + (1|subjCode), data=f)

M3b <-lmer (latency ~ trialType *delay + (1|subjCode), data = subset (f, trialType != "congruent")

M3c <-lmer (latency ~ trialType *delay + (1|subjCode), data = subset (f, trialType != "incongruent")

M3d <-glmer (isRight ~ trialType * delay + (1|subjCode), data=f, family= "binomial")

### *Between-task comparison*

congruentAd = size of congruent advantage in RT on sequential trials
congruentAdACC= size of congruent advantage in accuracy on sequential trials

incongruentCost = size of incongruent cost in RT on sequential trials

incongruentCostACC = size of incongruent cost in accuracy on sequential trials

M4a<-lmer (latency ~ congruentAd * sparsity + congruentAdACC * sparsity + (1|subjCode) + (1|pictureCategory), data=p)

M4b<-glmer (isRight ~ incongruentCost * sparsity + incongruentCostACC * sparsity + (1|subjCode) + (1|pictureCategory), data=p, family=binomial)

## Experiment 4

electrode = stimulation condition (anodal: .5, cathodal: −.5)

M5a<-lmer (latency ~ electrode * sparsity * delay+ (1|subjCode) + (1|pictureCategory), data=t)

M5b<-lmer (latency ~ sparsity + (1|subjCode) + (1|pictureCategory), data= subset(t, electrode== "anodal"))

M5c<-lmer (latency ~ sparsity + (1|subjCode) + (1|pictureCategory), data=subset(t, electrode== "cathodal"))

M5d<-lmer (latency ~ electrode * sparsity + (1|subjCode) + (1|pictureCategory), data=subset(all, electrode!= "cathodal"))

M5e<-lmer (latency ~ electrode * sparsity + (1|subjCode) + (1|pictureCategory), data=subset(all, electrode!= "anodal"))

M6a<-glmer (isRight ~ electrode * sparsity * delay + (1|subjCode)+(1|pictureCategory), data=t, family= "binomial")

t$Mdelay<-factor(t$delay>=median(t$delay),

levels=c("FALSE","TRUE"),

labels=c("Short Delay","Long Delay"))

l<-subset(t, Mdelay=="Long Delay")
M6b<- glmer (isRight ~ electrode * catDim + (1|subjCode) + (1|pictureCategory), data=l, family="binomial")

M6c<- glmer (isRight ~ electrode * catDim + (1|subjCode) + (1|pictureCategory), data=subset(l, electrode== "cathodal"), family="binomial")

M6d<- glmer (isRight ~ electrode * catDim + (1|subjCode) + (1|pictureCategory), data=subset(l, electrode== "anodal"), family="binomial")

M6e<-glmer (isRight ~ sparsity * delay + (1|subjCode) +(1|pictureCategory), data=subset(t, electrode== "cathodal"), family= "binomial")

M6f<-glmer (isRight ~ sparsity * delay + (1|subjCode)+(1|pictureCategory), data=subset(t, electrode== "anodal"), family= "binomial")

M6g<-glmer (isRight ~ electrode * sparsity + (1|subjCode)+(1|pictureCategory), data=subset(all, electrode!= "anodal"), family= "binomial")

M6h<-glmer (isRight ~ electrode * sparsity + (1|subjCode)+(1|pictureCategory), data=subset(all, electrode!= "cathodal"), family= "binomial")