Puddles, Parties, and Professors: Linking Word Categorization to Neural Patterns of Visuospatial Coding

Susanne Quadflieg[1,2,3], Joset A. Etzel[2], Valeria Gazzola[2,4],
Christian Keysers[2,4], Thomas W. Schubert[5],
Gordon D. Waiter[1], and C. Neil Macrae[1]

Abstract

■ Behavioral evidence suggests that during word processing people spontaneously map object, valence, and power information to locations in vertical space. Specifically, whereas "overhead" (e.g., *attic*), positive (e.g., *party*), and powerful nouns (e.g., *professor*) are associated with "up," "underfoot" (e.g., *carpet*), negative (e.g., *accident*), and powerless nouns (e.g., *assistant*) are associated with "down." What has yet to be elucidated, however, is the precise nature of these effects. To explore this issue, an fMRI experiment was undertaken, during which participants were required to categorize the position in which geometrical shapes appeared on a computer screen (i.e., upper or lower part of the display). In addition, they also judged a series of words with regard to location (i.e., up vs. down), valence (i.e., good vs. bad), and power (i.e., powerful vs. powerless). Using multivoxel pattern analysis, it was found that classifiers that successfully distinguished between the positions of shapes in subregions of the inferior parietal lobe also provided discriminatory information to separate location and valence, but not power word judgments. Correlational analyses further revealed that, for location words, pattern transfer was more successful the stronger was participants' propensity to use visual imagery. These findings indicate that visual coding and conceptual processing can elicit common representations of verticality but that divergent mechanisms may support the reported effects. ■

INTRODUCTION

Although many people would find it a rather simple task to describe the defining characteristics of a *fork* or a *hat*, the question of how the mind represents conceptual knowledge is far from trivial. A person's repository of such knowledge provides not only the basis for interacting successfully with encountered objects (including other people) but also for retrieving the meaning of literally thousands of words. In addition, concepts form the elementary units of many higher-order cognitive operations, including reasoning and decision making. Given these observations, it is unsurprising that psychologists, neuroscientists, linguists, and philosophers have expended considerable effort in attempts to understand how the mind represents and organizes conceptual knowledge (e.g., Barsalou, 2008; Boroditsky & Prinz, 2008; Lakoff, 2008; Jackendoff, 2002; Jeannerod, 2001).

To elucidate how concepts are represented in the brain, a large corpus of neuroimaging and patient data has been collected (see Mahon & Caramazza, 2009; Martin, 2007). Of interest herein, much of this work has revealed that concep-

tual knowledge resides in the neural systems involved in action and perception. For instance, thoughts about tools have been found to elicit activity in brain regions related to object use, whereas pondering about animals has yielded activation in areas associated with the encoding of shape and color (e.g., Hauk, Davis, Kherif, & Pulvermüller, 2008; Marques, Canessa, Siri, Catricalà, & Cappa, 2008; Chao, Weisberg, & Martin, 2002; Chao, Haxby, & Martin, 1999). What this suggests is that thinking about a specific concept seems to activate—at least partially—a similar neural state to that generated when one interacts with the object in question. In this way, the representation of conceptual knowledge has been hypothesized to rely on a diverse collection of sensorimotor simulation mechanisms (Barsalou, 2008; Boroditsky & Prinz, 2008; Lakoff, 2008; Gallese & Lakoff, 2005; Jeannerod, 2001; Glenberg & Robertson, 2000).

Importantly, it is not only neuroscientific methods that have been used to examine whether concepts are enriched by sensorimotor traces. Further support has been garnered from studies exploring language processing (see Zwaan, 2009; Meteyard & Vigliocco, 2008). Besides providing additional evidence for the activation of motor reenactments (e.g., Zwaan & Taylor, 2006; Glenberg & Kaschak, 2002), these studies have revealed a tight relationship between conceptual thought and spatial representations. Numerous

experiments have found that the arrangement of words on a computer screen can influence how rapidly semantic knowledge is retrieved. For instance, when asked to judge the relatedness of words, which are presented one above the other (e.g., *branch* above *root* vs. *root* above *branch*), participants can affirm word relatedness faster when the arrangement follows that observed in the world (Zwaan & Yaxley, 2003). Similarly, participants required to judge the animacy of stimuli respond faster when the spatial appearance of the words on a computer screen reflects the veridical position of the objects they denote (e.g., when the word *attic* appears at the top rather than at the bottom of the screen; Šetić & Domijan, 2007).

Intriguingly, similar effects have been reported for words that have no obvious spatial component. In an influential study, Meier and Robinson (2004) revealed that participants asked to sort words according to valence (i.e., whether they referred to something positive or negative) responded faster when positive items (e.g., *love*) were presented in the upper half and negative items (e.g., *danger*) were in the lower half of a computer screen (compared with the opposite arrangement). In addition, it has been suggested that words referring to people in powerful or powerless positions (e.g., *master* vs. *slave*) may also elicit associations of verticality (Casasanto, 2009; Zwaan, 2009). When, however, Schubert (2005) used an experimental paradigm, in which single words appeared in upper or lower positions on a computer screen—a setup that successfully revealed spatial congruency effects for both object and valence words (Šetić & Domijan, 2007; Meier & Robinson, 2004)—only partial evidence for power space mapping was observed. Although words denoting influential people were judged more quickly as powerful when presented in the upper compared with the lower part of the screen, a speed difference for words referring to powerless individuals on the basis of screen position was not observed. Notwithstanding this finding, it has repeatedly been claimed that people map object, valence, and power-related words to locations in vertical space (e.g., Casasanto, 2009; Zwaan, 2009; Meier, Hauser, Robinson, Kelland Friesen, & Schjeldahl, 2007).

Of relevance to the current inquiry, the precise mechanism supporting word verticality mappings remains a matter of debate (see Pecher, van Dantzig, Boot, Zanolie, & Huber, 2010). Most commonly, it has been suggested that retrieving the meaning of object, valence, and power words may activate spatial representations that either conflict or converge with those elicited by different screen positions (Schubert, 2005; Meier & Robinson, 2004; Zwaan & Yaxley, 2003). In the case of object words (e.g., *attic*), the activation of such spatial information has been argued to reflect the everyday occurrence of encountering concrete objects at specific locations in vertical space (Šetić & Domijan, 2007; Zwaan & Yaxley, 2003). For valence and power words, however, such an experience-based explanation seems less applicable, given that words such as *freedom* and *manager* have no direct spatial connotations (but see

Schubert, Waldzus, & Seibt, in press, for a discussion of this issue).

Instead, it has been suggested that, in the English language, positive and negative concepts, as well as powerful and powerless concepts, may be linked to locations in vertical space via metaphors. For example, we are taught that the righteous go *up* to heaven, whereas sinners go *down* to hell (Meier & Robinson, 2004). Similarly, we learn to *look up* to those at the height of their power, whereas we *look down* on subordinates (Schubert, 2005). Thus, via metaphorical mapping, space-unrelated target domains (i.e., valence and power) can be linked to a concrete source domain (i.e., verticality; Lakoff & Johnson, 1999). Importantly, it is thought that "to the extent that mental representations in perceptuomotor source domains constitute abstract concepts, these concepts can be instantiated by the same neural […] structures that simulate perception and action in the physical world" (Casasanto, 2009, p. 352).

But what are the neural structures that allow humans to experience vertical space in the first place? It is widely agreed that a person's sense of orientation in space relies heavily on a combination of vestibular and visual cues. Although the existence of primary vestibular cortex remains a matter of debate, accumulating evidence indicates that processing vestibular information at the cortical level is highly localized (for a review, see Brandt & Dietrich, 1999). The artificial induction of vestibular sensations by means of caloric irrigation, galvanic stimulation, or specific tone bursts, for instance, has been found to elicit enhanced activity in the inferior parietal lobe (IPL) (Schlindwein et al., 2008; Eickhoff, Weiss, Amunts, Fink, & Zilles, 2006; Stephan et al., 2005). Along similar lines, lesions in the IPL have been found to cause vestibular syndromes such as vertigo (Naganuma et al., 2006; Urasaki & Yokota, 2004). Finally, the coding of verticality based on visual information has also been associated with IPL activity regardless of whether participants are required (i) to judge the vertical position of a target dot in relation to a reference line (up vs. down, Baciu et al., 1999), (ii) to report the position of both hands of an analogue clock at a certain time on the clock face (e.g., upper vs. lower half; Trojano et al., 2002), or (iii) to determine the type of spatial relation (e.g., above/below) between object pairs (Amorapanth, Widick, & Chatterjee, 2010; Corradi-Dell'Aqua, Hesse, Rumiati, & Fink, 2008).

Given these observations, the claim that space-unrelated target domains such as valence and power may be linked to representations of verticality via metaphorical mapping raises an intriguing question. Does the processing of verticality-related words share a common neural signature in the IPL with verticality coding based on visual input? Addressing this possibility, the current investigation used multivoxel pattern classification analysis (MVPA; O'Toole, Abdi, Pénard, Dunlop, & Parent, 2007) to explore if the neural patterns that differentiate between the spatial coding of "up" or "down" based on percepts can also be used to separate conceptual judgments of words pertaining to objects (e.g., *carpet*, *ceiling*), valence (e.g., *delight*, *disaster*),

and power (e.g., *manager*, *assistant*). Notably, by utilizing MVPA, we refrained from merely establishing the involvement of a specific brain region that has previously been associated with spatial processing—a pitfall that has been termed "reverse inference." Such inferences are known to be problematic, particularly when regions are found to be active during many different mental operations (Poldrack, 2006). In contrast, neural pattern classifiers offer the possibility to identify mental states from distributed neural activity, thereby enabling researchers to test the similarity structure of mental processes across tasks (Poldrack, Halchenko, & Hanson, 2009).

METHODS

Participants

Twenty-one undergraduate students from the University of Aberdeen (11 men), with ages between 18 and 24 years (mean = 21 years) participated in the study. All volunteers were native English speakers, right-handed as determined by the Edinburgh handedness inventory (Oldfield, 1971), and reported to have normal or corrected-to-normal vision. None of the participants had a history of neurological or neuropsychiatric disorders or were currently taking psychoactive medications. Informed consent was obtained from all individuals, and the study protocol was approved by the School of Psychology Ethics Committee of the University of Aberdeen.

Experimental Task

Both the pilot task as well as the main experiment entailed participants making several types of judgment. During position judgment trials (based on Baciu et al., 1999), a geometric shape (e.g., a *circle*) was presented either above or below a horizontal line, and the participants' task was to judge whether the shape was "up" or "down" in relation to this reference line. During location judgment trials, participants saw words that referred to a concrete entity (e.g., *attic*) and were asked to indicate its typical everyday location ("up" in space vs. "down" in space). To introduce a standard of comparison, participants were told to consider everything as "up" in space if the object was typically encountered above the height of their eyes and everything as "down" in space if it was typically found below eye level. During valence judgment trials, participants were asked to categorize nouns (e.g., *anger*) according to whether they referred to something "positive" or "negative." Finally, during power judgment trials, participants judged nouns (e.g., *captain*) according to whether they referred to people in "powerful" or "powerless" positions.

Stimulus Materials

To select target nouns for the main experiment, a pilot study with 16 native English speakers (students of the University of Aberdeen, five men, average age = 25 years) was conducted. Volunteers were seated in front of a MacBook Pro laptop computer (15-in. monitor, resolution of 1440 × 900 pixels) and informed that they would be taking part in a word categorization task that required them to make location, valence, and power judgments (as described above). During location judgment trials, stimuli comprised 20 nouns associated with "upper" vertical space and 20 nouns associated with "lower" vertical space (based on Zwaan & Yaxley, 2003). During valence judgment trials, stimuli comprised 20 positive and 20 negative nouns (taken from Bradley & Lang, 1999). During power judgment trials, 20 nouns referring to people in powerful positions and 20 nouns referring to people in powerless roles were presented (based on Schubert, 2005).

All stimuli were displayed on a uniform black background using Psyscope presentation software (Version 1.2.5; Cohen, MacWhinney, Flatt, & Provost, 1993). Each trial began with a centrally presented white fixation cross shown for 1000 msec followed by an instruction screen.

This screen displayed one of three words ("*Valence*?," "*Power*?," "*Location*?") in capitalized red letters (bold 32-point Arial font) for 1000 msec. Underneath this prompt, two response options were listed [i.e., up/down, pos/neg (for positive/negative), pf/pl (for powerful/powerless)] in green letters (italicized 32-point Arial font). The presentation order of the response options from left to right determined the required button press, such that the first word (e.g., *up*) always mapped onto a participant's right index finger and the second (e.g., *down*) onto his or her right middle finger. Response buttons (the letters "o" and "p" on the computer keyboard) were counterbalanced across participants, such that half of the participants used their index finger to indicate "powerful," "positive," and "up" and their middle finger to indicate "powerless," "negative," and "down," whereas the other half of the participants used the reverse mapping. After the instruction screen, the target stimulus was displayed in the middle of the screen, in white 36-point Verdana font, until either the participant made a response or 4000 msec elapsed. A 1000-msec intertrial interval separated each trial from the next. Following an initial practice session of 16 trials that familiarized participants with the task (using stimuli not displayed during the subsequent pilot experiment), the critical block of 120 randomized trials was completed. Participants were asked to maximize the accuracy and speed of their responses.

On the basis of the recorded error rates during this task, a subset of 10 stimuli per experimental condition was selected. Selection was limited to words that had elicited no more than one misclassification across all participants to ensure that the observed error was likely to reflect a mistaken button press rather than a stimulus-related conceptual ambiguity. Given that most errors occurred during powerless word trials, the first 10 most accurately classified stimuli were chosen from this category. Afterwards, 10 words from each of the remaining conditions were

selected so that error rates were minimized and a matching of words with regard to the average number of syllables ($M = 1.9$) and the average number of letters ($M = 6.3$) was achieved across all experimental conditions (see Table 1). Subsequent to stimulus selection, accuracy rates and median response times were examined for the selected items. These analyses showed that—as intended—errors occurred rarely during the relevant trials ($M = 2\%$). Submitting the accuracy scores of each experimental condition to a 3 (Word Type: location, valence, power) × 2 (Spatial Mapping: up vs. down) repeated measures ANOVA revealed no significant interaction or main effects [word type: $F(2, 30) = 2.05$; spatial mapping: $F(1, 15) = .09$; interaction: $F(2, 30) = 3.30$, all ps > .05], showing that errors were distributed equally across experimental conditions. Similarly, submitting the median response times of the accurate trials to the same ANOVA also failed to return any significant effects [word type: $F(2, 30) = 2.87$; spatial mapping: $F(1, 15) = .27$; interaction: $F(2, 30) = .45$; all ps > .05], indicating that respondents needed a similar amount of time to make all types of judgment (overall $M = 809$ msec).

Table 1. Set of Nouns as Presented in the Imaging Experiment

Location	Valence	Power
Up/Positive/Powerful		
airplane	beauty	boss
attic	birthday	captain
ceiling	delight	chief
chimney	freedom	emperor
cloud	friend	judge
giraffe	kiss	leader
moon	laughter	manager
roof	miracle	master
satellite	music	officer
treetop	party	professor
Down/Negative/Powerless		
basement	accident	assistant
carpet	anger	butler
doormat	danger	cleaner
floor	death	intern
mushroom	disaster	maid
pavement	fear	novice
puddle	hardship	pupil
river	horror	servant
root	stench	slave
subway	tragedy	trainee

In addition to the pilot study, for the position judgment trials, 15 geometric shapes were selected (e.g., *rectangle*, *square*, *triangle*, *circle*) from the Powerpoint "basic shapes" menu (Microsoft Office Professional Edition, Version 2003) and stretched until they were of similar height (15–20 pixels). They were then paired with a white horizontal line of fixed length (108 pixels) so that the center of each shape divided the line horizontally in two equal halves, whereas in the vertical dimension, the distance between the center of the shape and the center of the line was kept constant (30 pixels). Using Adobe Photoshop (Version 8.0), each shape–line pairing was then inserted in white font on a standardized black canvas of a common height and width (50 × 120 pixels). To create a set of stimuli in which each shape appeared in the exact same position once above and once below the white line, the resulting pictures were subsequently mirrored along their vertical axis. Finally, 10 of the 15 shapes (in their up and down version, that is, 20 images) were chosen at random to be included in the final experiment, whereas the remaining images were used during practice or catch trials (see below).

fMRI Paradigm

For the scanner task, we eliminated button presses that would force participants to map their answers either in horizontal or vertical space (depending on how the response buttons would be arranged on a button box). Thus, an event-related fMRI paradigm was developed, which comprised trials of interest that did not require an overt response. During these trials, participants first saw a centrally presented target item (i.e., shape or noun). Simultaneously, a word (30-point green Arial Narrow font) instructing them what type of judgment to perform ("*position*," "*location*," "*valence*," "*power*") was displayed above the target. Participants were instructed to silently voice the appropriate answer (options: *up*, *down*, *positive*, *negative*, *powerful*, *powerless*) in their head as quickly and accurately as possible upon the appearance of this screen. Once 1000 msec elapsed, both items disappeared and were replaced by a centralized white fixation cross. The duration of the cross varied randomly between 11 and 17 sec, allowing the stimulus-related hemodynamic response function to return to baseline between trials. Such a slow event-related design was chosen to ensure that training as well as test examples used for the pattern classifier could independently be drawn from the obtained source distribution of events (see Pereira, Mitchell, & Botvinick, 2009) (Figure 1).

To be able to check whether participants paid attention to the task, trials of interest were intermixed with a series of catch trials (see Figure 2). During catch trials, the same first screen as during trials of interest was shown, however, once 1000 msec elapsed an additional screen displaying a specific combination of button labels appeared. Depending on the type of task participants would
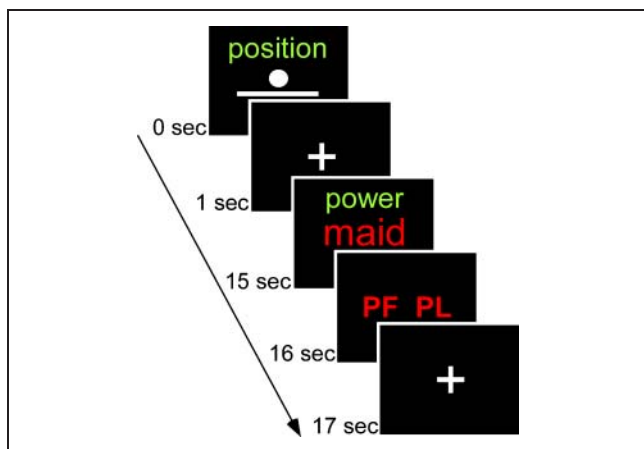
Figure 1. Trial of interest separated from a catch trial by a white fixation cross with a duration of 14 sec.

either see the labels "UP" and "DOWN," "PF and "PL," or "POS" and "NEG" in the middle of the screen (30-point red Arial Narrow font). Participants were instructed to translate their covert reply into an overt button press whenever they saw a response screen. Responses during catch trials were given by pressing one of two horizontally aligned buttons on a button box with the index or middle finger of the right hand. The presentation order of the response labels from left to right informed participants about the required button press such that the first label (e.g., "UP") always mapped onto their right index finger and the second (e.g., "DOWN") onto their right middle finger. Finger–label combinations were randomized across catch trials, and the response screen disappeared automatically after 1000 msec. Afterwards, a white fixation cross was shown for a random duration of 10–16 sec so that the average duration of catch trials matched those of trials of interest.

In total, 192 trials were presented in the course of the experiment. During trials of interest, each of the selected target stimuli was presented twice, resulting in a total of 160 trials, comprising 40 position judgments (20 "up" shapes, 20 "down" shapes), 40 location judgments (20 nouns implying "up," 20 nouns implying "down"), 40 valence judgments (20 "positive" nouns, 20 "negative" nouns), and 40 power judgments (20 "powerful" nouns, 20 "powerless" nouns). In addition, participants encountered 32 catch trials throughout the experiment (eight position judgments, eight location judgments, eight valence judgments, eight power judgments). Because catch trials were of no further interest, target stimuli comprised a random selection of items excluded during pilot testing. Each of these items was presented only once, allowing participants in theory to figure out that repeated items would never require an overt response. Given that such an insight was by definition not possible during the initial presentation of the stimuli, however, we considered the influence of such an effect on the attention of our participants over the course of the experiment to be negligible.

For each participant, a uniquely randomized sequence of trials was presented with the total number of trials being equally distributed across four runs. Within each run, five trials of interest and one catch trial were presented for each of the eight experimental conditions (position up, position down, location up, location down, valence positive, valence negative, powerful, powerless) resulting in 48 trials per run. After each run, the word "Rest" appeared on the screen, and scanning was resumed after a minute unless the participant indicated the need for a longer break. While in the scanner, all stimuli were back projected onto a screen visible via a mirror mounted on the MRI head coil (visual angle = ~6.5° × 6.5°). Stimulus presentation and recording of participants' responses and associated latencies were accomplished using Presentation software (version 9.13, Neurobehavioral Systems, Inc., Albany, CA).
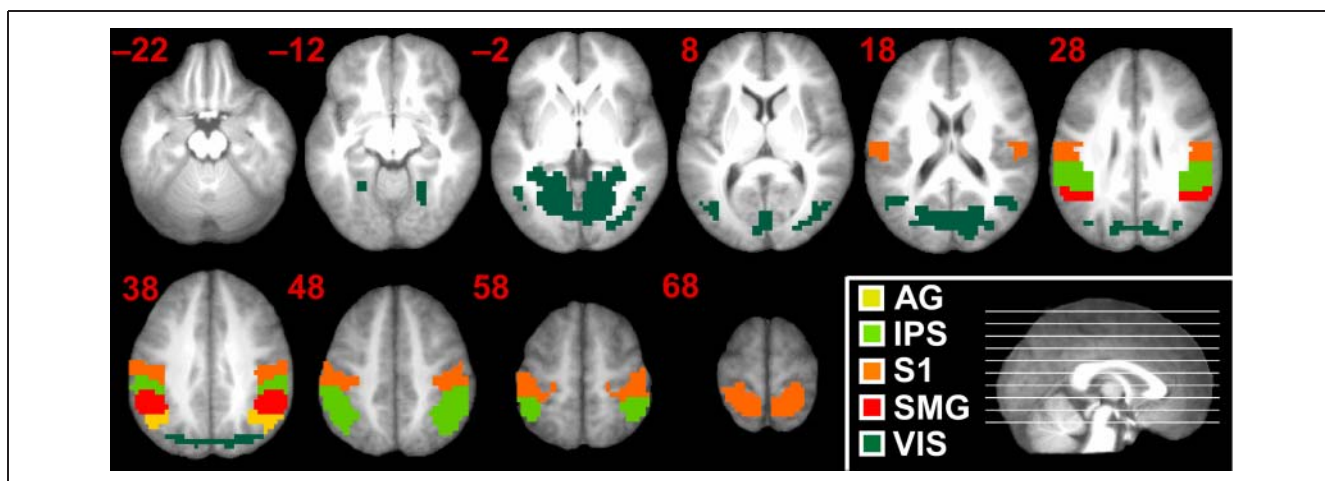

Figure 2. ROI masks overlaid on the structural anatomy averaged across all 21 participants. Red numbers in the left upper corner of each brain denote the z value of each slice according to MNI coordinates. AG = angular gyrus; IPS = banks of the intraparietal sulcus within the inferior parietal lobe; S1 = primary somatosensory cortex; SMG = supramarginal gyrus; VIS = visual association cortex.

All stimuli appeared on a uniform black background. Target stimuli (words and images) were presented in white with words displayed in 38-point Arial Narrow font.

To ensure that shape judgments would not require any up- and down-related eye movements, the height of all shape–line pairs resembled the height of the presented words (visual angle = ~1.2°). Also, the center of mass of each pair always appeared in the middle of the screen. As such, the position judgment did not require participants to move their eyes up and down in relation to the reference line to be able to detect the shape. Rather, solving the task always required the integration of two objects that covered the same visual area on the screen regardless whether an up or a down trial was displayed. To familiarize participants with the shape as well as the word judgments, they completed 32 practice trials on a Toshiba Laptop computer outside the scanner. None of the target items used during practice was included in the experiment proper.

Postscanning Questionnaire

Previous investigations have rarely studied individual differences in the spontaneous activation of spatial representations during word comprehension. It could be argued, however, that at least for the processing of words referring to clearly localized objects, the observed effects may be driven by an individual's tendency to use visual imagery. Given this possibility, participants' disposition to visualize information was measured following the imaging experiment (i.e., outside the scanner) by administering the Style-of-Processing Scale (SOPS, Childers, Houston, & Heckler, 1985) as a paper-and-pencil questionnaire. The SOPS measures individual differences in the disposition to engage in visual or verbal processing. It contains 22 items, 11 of which assess the propensity to process visually (e.g., "*My thinking often consists of mental 'pictures' or images.*") and the other 11 of which assess the propensity to process verbally (e.g., "*I enjoy doing work that requires the use of words.*"). Participants were required to rate the extent to which each of the 22 items was characteristic of them on a 4-point scale ranging from 1 (*always true*) to 4 (*never true*). The relative disposition to process information visually is captured by the difference between the mean response to the visual items and the mean response to the verbal items (i.e., mean visual − mean verbal). Hence, a low difference score is indicative of a preference to engage in visual processing.

Image Acquisition

Image acquisition was performed on a 3-T whole body scanner (Philips Medical Systems, Best, the Netherlands) with an eight-channel phased array head coil. For registration purposes, anatomical images were acquired using a high-resolution 3-D fast field echo sequence (170 sagittal slices, TE = 3.8 msec, TR = 8.2 msec, flip angle = 8°, voxel size = 0.94 × 0.94 × 1 mm). Functional images were collected using a gradient-echo, echo-planar sequence sensitive to BOLD contrast (TR = 1300 msec, T2* evolution time = 30 msec, flip angle = 90°, in-plane resolution = 3.5 × 3.5, matrix = 80 × 80, field of view = 28 cm^2). For each volume, 20 axial slices (5-mm slice thickness, 0-mm skip between slices) were acquired. Because of randomizing the duration of the fixation cross between events, the exact number of slices collected within each run varied slightly across runs and participants ranging from 545 to 590 scans. For each run, the first six volumes were discarded to account for T1 saturation effects.

ROI

Classification algorithms are known to perform poorly when faced with many irrelevant features (i.e., voxels; see Formisano, De Martino, & Valente, 2008)—especially when the number of training samples is rather limited as is typically the case in fMRI studies. Hence, analysis of the current study was limited to an anatomical region of theoretical interest (see Etzel, Gazzola, & Keysers, 2009; Poldrack et al., 2009). According to previous imaging studies, coding of verticality on the basis of visual input has predominantly been associated with activation in the IPL (Amorapanth et al., 2010; Corradi-Dell'Aqua et al., 2008; Baciu et al., 1999). Neuroimaging as well as lesion data also suggest that categorical (e.g., *up* vs. *down*) compared with continuous spatial coding relies predominantly on inferior parietal resources located in the left rather than the right hemisphere (Amorapanth et al., 2010; Kemmerer, 2006; Trojano et al., 2002; Baciu et al., 1999). Given these empirical findings, the IPL was selected as the ROI in the current study and the WFU pickatlas was used to create the required ROI mask separately for the left and the right hemispheres (see Table 2). To enhance the specificity of our analysis, the IPL was further subdivided into prominent anatomical subregions as defined by the pickatlas: the angular gyrus (AG), the supramarginal gyrus (SMG), and the remaining IPL including the banks of the intraparietal sulcus (IPS). Additionally, visual association cortex (VIS) and primary somatosensory cortex (S1) were included as ROIs to examine the discriminatory validity of the classification procedures. It was expected that activity in VIS could be used to reliably distinguish between "up" and "down" shapes given the systematic visual differences between these items, but not between any of the word items given their visual similarity. In addition, it was hypothesized that none of the judgments should lead to a systematic activation difference within or across tasks in S1.

Data Analysis

Behavioral data were analyzed using SPSS for Windows. Preprocessing of the neuroimaging data and the creation of parameter estimate images (PEIs) was performed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). Subsequent handling of the PEIs as well as

Table 2. Number of Voxels (4 × 4 × 4 mm) in Each ROI and the WFU Pickatlas Areas Used to Create Each Mask

ROI (including Abbreviations Used)	WFU Pickatlas Label	Number of Voxels Left	Number of Voxels Right
Primary Somatosensory Cortex (S1)	Postcentral Gryrus	451	456
Visual Association Cortex (VIS)	BA 18 and BA 19	1459	1489
Angular Gyrus (AG)	Angular Gyrus	44	49
Supramarginal Gyrus (SMG)	Supramarginal Gyrus	130	131
Banks of the Intraparietal Sulcus (IPS)	Inferior Parietal Lobe	380	361
Inferior Parietal Lobe (IPL)	Angular Gyrus, Supramarginal Gyrus, Inferior Parietal Lobe	554	541

Whereas the software's default dilation was kept when regions could be specified as lobes or gyri, a dilation of 1 in 3-D was applied for areas defined based on Brodmann's areas, that is, for the VIS.

conducting the multivariate pattern classifications was undertaken in R (version 2.8.0, R Foundation for Statistical Computing, Vienna, Austria). All classifications were performed using the support vector machine command in the e1071 R package with a linear kernel and the cost parameter fixed at 1.

Data preprocessing began by correcting differences in acquisition time between slices for each whole-brain volume. Functional data were then realigned to the first volume acquired for each participant using a least squares approach and a six-parameter (rigid body) spatial transformation to minimize the effects of head movements on data analysis. The direction and magnitude of motion for each participant over the course of each run were examined. Fifteen participants moved less than 1.5 mm in any direction within each of the four runs and their complete data were considered during analyses. One participant showed sudden movement in the first run, causing us to replace two scans with a weighted average consisting of the preceding and following scans. Importantly, the two replaced scans fell between two events and were not collected within the first 10 sec subsequent to the first event; thus, the chance of introducing bias by this replacement was considered to be negligible. One further participant showed significant movement because of coughing at the beginning of the fourth run. As a result, all scans affected by the coughing (less than 1/5 of the scans in the run) were excluded, and only the remaining scans were used during analyses. Finally, four participants displayed significant movement (>1.5 mm) from the third run onward; thus, only data collected during their first two runs of the experiment were considered. Despite these exclusions, the data set for each participant comprised at least 10 instances within each of the eight experimental conditions (i.e., up shapes, down shapes, up nouns, down nouns, positive nouns, negative nouns, powerful nouns, powerless nouns).

Following realignment, each participant's mean EPI image (based on the valid scans only) was registered to the individual's high-resolution gray matter segment, applying SPM8's rigid body transformation. Individual gray matter segments were extracted using SPM8's "segment" function. To ensure that the procedure was successful, all extracted segments were inspected visually and if necessary cleaned from remaining bits of dura matar using MRIcroN (www.cabiatl.com/mricro/mricron/index.html). Subsequent to their successful coregistration, functional data were transformed into standard anatomical space by determining the normalization parameters required to warp each individual's coregistered gray matter segment onto the gray matter MNI template. These parameters were then applied to a person's functional and structural volumes using an isotropic voxel size of 4 × 4 × 4 mm for the former and of 1 × 1 × 1 mm for the later. Finally, the normalized functional images were spatially smoothed applying an 8-mm FWHM Gaussian kernel (Kriegeskorte, Cusack, & Bandettini, 2010; Op de Beeck, 2010). Both preprocessing steps (i.e., normalization as well as smoothing of the data) were undertaken to alleviate the interparticipant variation of functional activations before classifying the fMRI images of multiple participants (Fan, Shen, & Davatzikos, 2006).

The goal of the current investigation was to perform three types of multivariate classification analyses. First, to verify that reliable stimulus-distinguishing information was present during shape judgments, a within-task between-subject analysis was performed to detect activation patterns, which consistently distinguish between up and down shape judgments across all participants. Second, cross-task between-subject analyses were conducted to examine whether patterns that distinguished between these two types of shapes could also be used to separate the two word types presented within each of the other three judgment tasks. Third, cross-task within-subject analyses were computed to examine whether the overlap in neural patterns across tasks within participants was related to their inclination to visualize.

For all three analysis types, PEIs for each relevant event in the trial sequence were created. Thus, for each target

stimulus, a boxcar function convolved with the canonical hemodynamic response function was fitted to the functional data using a general linear model approach (Formisano et al., 2008). The PEIs were further processed to remove voxels that had zero variance in any individual participant or run. Removing zero-variance voxels from all participants allows analyses to be performed using the same voxels in all participants, a necessary requirement for between-subjects analyses. Subsequently, the voxels within each ROI mask (as discussed in ROI) were extracted from the PEIs. For illustrative purposes, Figure 2 displays the remaining voxels arranged by ROIs in the left and right hemisphere. All further classification analyses were conducted considering only data within these ROIs (individually). The rows of the relevant data matrices (e.g., the voxels of each PEI, separately for each ROI) were then scaled to have zero mean and unit variance—a standard procedure for multivariate fMRI data analyses (Pereira et al., 2009).

For the between-subject analyses, data were averaged across all events within an experimental condition for each participant to reduce intrasubject variability and improve the signal to noise ratio of the data (Pereira et al., 2009; Fan et al., 2006). The averaged PEIs were again scaled to have unit variance to minimize differences in the absolute range of data across experimental conditions and individuals. To create training and testing data for the within-task between-subject analyses, a leave-one-out cross-validation procedure was chosen. Thus, the classification algorithms were trained by using data of all participants except for one, whose averaged data were used for testing. This procedure was then repeated in turn for each individual allowing the computation of an overall accuracy rate based on the average of accurate predictions made for all the participants (see Pereira et al., 2009; Poldrack et al., 2009).

For the cross-task between-subject analyses, shape judgments were considered either as training or testing data. When considered as training data, the classification algorithm was trained on the up and down shape judgments of all 21 participants and applied to predict the up and down location judgments of all 21 participants. The same procedure was then repeated to predict valence judgments (positive vs. negative) and power judgments (powerful vs. powerless). Paralleling this approach, when the shapes were considered as testing data, the classification algorithm was trained on the location, valence, or power judgments and applied to predict shape judgments across all participants.

For both within-task and cross-task between-subject classification analyses, the statistical significance of the relevant outcome measure (i.e., of the average classification accuracy) was determined by permutation testing (Etzel et al., 2009; Pereira et al., 2009; Golland & Fischl, 2003). The permutation test was performed by repeating each analysis 1000 times, randomly permuting the data labels (stimulus type) each time. The labels were permuted within runs per person (before averaging) to preserve the variance structure of the data. The significance was calculated as the proportion of permuted data sets returning a higher accuracy rate than the true data, for a maximum significance level of .001. A p value of less than .05 was considered statistically significant.

Finally, within-subject cross-task classification analyses were computed to examine whether individual differences in spontaneous visualization would correlate with differences in the extent to which a shape-trained classifier could be applied to separate word judgments for each participant. Hence, the neural responses elicited by the exemplars within each experimental condition were not averaged per person. Rather, for each person, the full set of shape examples was used as training data and the computed algorithm was then applied to predict the same person's full set of word examples for each type of judgment. When the number of examples within the training or testing data were unequal (because of the exclusion of scans related to movement as described above), examples were removed at random from the larger class to achieve a balanced set of training and testing data. In this case, each analysis was repeated 10 times, and the results were averaged to counteract the effect of removing samples at random. The returned average classification accuracies for each participant and for each of the three relevant across-task multivariate analyses (i.e., shapes–object words, shapes–valence words, shapes–power words) were then correlated with participants' SOPS scores. A p value of less than .05 was considered statistically significant.

RESULTS

Behavioral Analysis of Catch Trials

The first set of analyses determined how accurately participants responded during the 32 catch trials. Participants' accuracy scores on these trials ranged from 88% to 100% with a mean score of 94% ($SD = 4\%$). A one-sample t test showed that these scores were significantly better than chance [$t(20) = 45.82, p < .001$]. In addition, a repeated measure ANOVA comparing the four judgment types failed to yield a significant effect [position: $M = 95\%$; location: $M = 96\%$; valence: $M = 93\%$, power: $M = 91\%$; $F(3, 60) = 1.21, p = .313$] revealing that participants performed equally well on all four tasks. As an indicator of task involvement, catch trials were discarded from any further analyses.

Within-task Between-subject Classification Data

Processing of "up" and "down" shapes could be predicted with greater than chance accuracy on the basis of activation in the left VIS (accuracy = 67%, $p = .013$), the right VIS (accuracy = 64%, $p = .037$), the left AG (accuracy = 71%, $p = .009$), and marginally so in the right AG (accuracy = 62%, $p = .091$). As predicted, such a discrimination of shape judgments was not possible on the basis of activation

Figure 3. Within-task between-subject classification results for shape judgments (up vs. down). The green bars indicate the acceptance range of the permutation test. The bar's upper limit is the maximum accuracy observed in any of the 1000 permutations, and the minimum is the accuracy corresponding to 95% of the range (i.e., the $p = .05$ cutoff). The dots indicate the measured accuracy for each ROI, and a dot falling within a green bar is considered significant. The red plus signs represent the average proportion correct for randomly labeled data as determined by the permutation test.
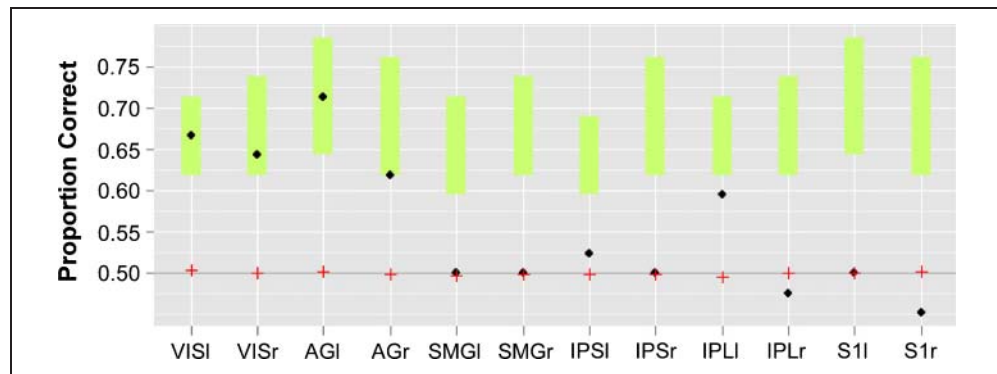
in the control region S1 (left, accuracy = 50%; right, accuracy = 45%; both $ps > .54$; see Figure 3).

Cross-task Between-subject Classification Data

Cross-task analyses revealed that classifiers, which successfully distinguished between "up" and "down" shapes in subregions of the inferior parietal lobe, also provided discriminatory information to separate location and valence judgments (see Figure 4). More specifically, activation patterns in the left AG (accuracy = 60%, $p = .083$) and the entire left IPL (accuracy = 60%, $p = .078$) could be used to sort "overhead" and "underfoot" words, although these effects were only marginally significant. In addition, significant transference effects were found from shapes to valence words in the left IPS (accuracy = 67%, $p = .006$) and the right SMG (accuracy = 62%, $p = .044$). In contrast, for power words, no reliable prediction accuracy was achieved in any of the ROIs (all $ps > .181$).

Finally, the reverse classification (i.e., training classifiers on valence, power, and location words and applying them to shape judgments) did not yield any statistically significant results.

Relation of the Cross-task Within-subject Classification Data with SOPS Scores

Examining responses on the postscanning questionnaire showed that participants' SOPS scores ranged from -1.00 to 1.18 with an average of .19 ($SD = .59$). Correlational analyses revealed that, as participants' tendency to visualize increased, the more successfully the neural activity observed during the perception of shapes could be used to classify the processing of up and down words in the left IPS [$r(19) = -.47, p < .032$] and the left IPL [$r(19) = -.51, p < .019$; see Figure 5]. No other correlations reached significance. To determine whether the observed correlations between the extent of pattern transferability and

Figure 4. Across-task between-subject classification results for shape–location transfer and shape–valence transfer. The green bars indicate the acceptance range of the permutation test. The bar's upper limit is the maximum accuracy observed in any of the 1000 permutations, and the minimum is the accuracy corresponding to 95% of the range (i.e., the $p = .05$ cutoff). The dots indicate the measured accuracy for each ROI, and a dot falling within a green bar is considered significant. The red plus signs represent the average proportion correct for randomly labeled data as determined by the permutation test.
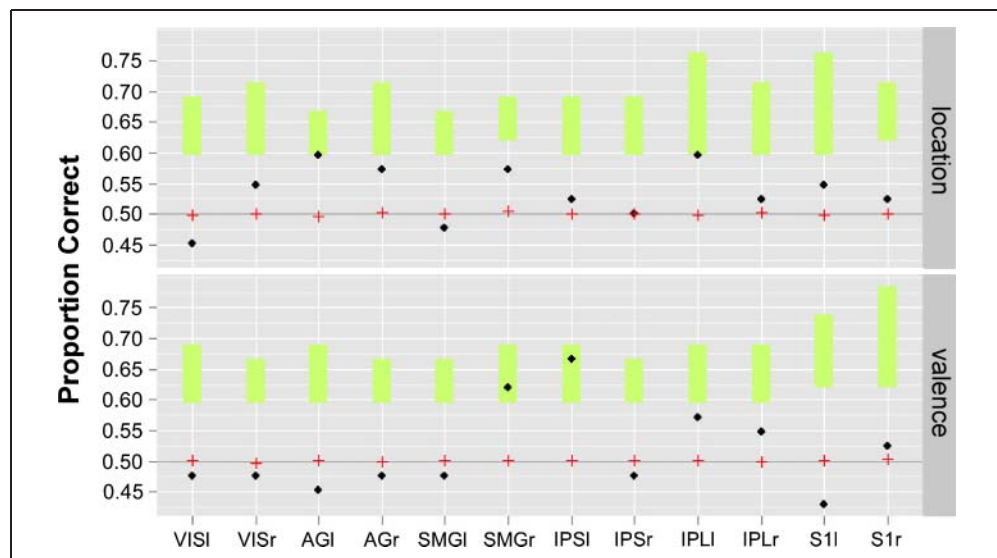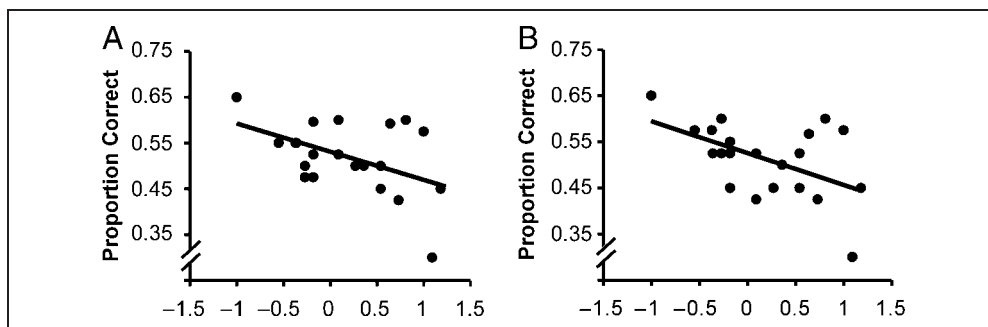
Figure 5. Relationship between SOPS scores and within-person cross-task classification accuracy for shape judgments to object word judgments in the left IPS (A) and the left IPL (B).

SOPS scores differed significantly across tasks, we computed William's *t* for the comparison of dependent correlation coefficients where appropriate (see Steiger, 1980). For the left IPL, it was found that the correlation of the SOPS scores with the pattern transferability score for up and down words was significantly different from those obtained for both, power [$r(19) = .03, ns; t(18) = -2.01, p < .03$] and valence words [$r(19) = -.02, ns; t(18) = -1.73, p = .05$]. For the left IPS, the correlation of the SOPS scores with the pattern transferability score for up and down words was different from that observed for power words [$r(19) = .17, ns; t(18) = -2.10, p < .025$] but not significantly so from that obtained for valence words [$r(19) = -.29, ns; t(18) = -0.70, ns$].

GENERAL DISCUSSION

In recent years, the observation that the spatial arrangement of words on a computer screen can influence how rapidly their meaning is retrieved has raised an intriguing question. Does the processing of certain concepts evoke spatial representations that are similar to those elicited by interacting with or perceiving the physical world (Casasanto, 2009)? To address this question, the current investigation applied MVPA to examine whether neural activity corresponding with the spatial coding of "up" and "down" on the basis of visual information could also be used to separate nouns pertaining to objects, valence, and power (Šetić & Domijan, 2007; Schubert, 2005; Meier & Robinson, 2004; Zwaan & Yaxley, 2003). Given that previous studies have repeatedly reported recruitment of inferior parietal resources during orienting in space on the basis of vestibular and visual information (Amorapanth et al., 2010; Corradi-Dell'Aqua et al., 2008; Schlindwein et al., 2008; Eickhoff et al., 2006; Naganuma et al., 2006; Stephan et al., 2005; Urasaki & Yokota, 2004; Trojano et al., 2002; Brandt & Dietrich, 1999), the IPL and its prominent subregions were chosen as brain sites of major interest in this inquiry.

In an initial within-task classification analysis, it was established that neural activity during shape categorization contained sufficient information to successfully separate "up" and "down" judgments. As expected, the data revealed better than chance classification of "up" and "down" shapes on the basis of the patterns of activity located in the

left AG, a region that has previously been associated with verticality judgments involving shape–line images (Baciu et al., 1999). In a next step, we examined whether neural patterns associated with the spatial coding of verticality on the basis of visual information could be transferred to separate word judgments. It is important to note that because of the diverging statistical methods underlying the computation of the within-task and the cross-task between-subject classifiers, the returned algorithms differed with regard to their sensitivity. For within-task analyses, a leave-one-subject-out approach was adopted to examine whether the two types of shapes could be separated by neural activity within the chosen ROIs. In contrast, cross-task classifiers investigated the transferability of discriminatory signal and profited from enhanced statistical sensitivity, because they were trained and tested on data from all participants. Therefore, these algorithms could detect patterns of neural activity that remained hidden from the within-task classifier.

As predicted, cross-task analyses revealed that the neural pattern separating "up" from "down" shapes could successfully be transferred to distinguish "overhead" and "underfoot" as well as positive and negative word judgments. In particular, the neural pattern underlying the visuospatial coding of "up" allowed the categorization of "overhead" and positive words, whereas the pattern underlying the visuospatial coding of "down" was similar to the pattern observed during the processing of "underfoot" and negative words. More specifically, in the case of object words, successful pattern transfer was localized in the left AG and the entire left IPL (albeit only marginally significantly so), whereas in the case of valence brain sites with better-than-chance classifications, successful pattern transfer was localized in the left IPS and the right SMG. By establishing better-than-chance cross-task classifications for object and valence words, the current findings support the idea that conceptual processing can elicit representations of verticality as previously suggested by behavioral work (Casasanto & Dijkstra, 2010; Estes, Verges, & Barsalou, 2008; Meier, Sellbom, & Wygant, 2007; Šetić & Domijan, 2007; Crawford, Margolies, Drake, & Murphy, 2006; Meier & Robinson, 2004; Zwaan & Yaxley, 2003).

Depending on word type, however, different mechanisms may have evoked spatial representations during conceptual processing. In the case of object words, it has previously been argued that the activation of verticality

may be related to everyday experiences of encountering the denoted items at specific locations in vertical space. In support of this view, the current study revealed that for object words the extent of pattern transferability was associated with participants' inclination to use visual imagery. Put differently, the more likely it was that people visualized words such as *treetops* and *puddle* (i.e., the more they re-envoked everyday experiences with these items), the more successfully IPL activity during the perception of shapes could be used to classify the processing of up and down words. This result fits perfectly within the wider literature on visual imagery which indicates that imagery activates the same neural representations that are activated by corresponding real world stimulation (Stokes, Thompson, Cusack, & Duncan, 2009; Mechelli, Price, Friston, & Ishai, 2004; O'Craven & Kanwisher, 2000). In this respect, spatial traces during object word processing seem at least partially constituted by processes of visual imagery. At the same time, however, it needs to be kept in mind that, in the current inquiry, a participants' task was to explicitly decide for object words whether they referred to items typically encountered up or down in space. In this respect, it could be argued that the task encouraged the use of visual imagery, which might not necessarily contribute to the comprehension of these words in other contexts. This question merits future empirical attention.

Importantly, though, in the context of the current investigation, successful cross-task classification was also observed for valenced words for which a visual-imagery-based explanation seems less applicable. First, words such as *beauty* and *hardship* are harder to visualize than object words because of their inherent abstractness. Second, these words have no direct spatial connotation on the basis of everyday experience. Indeed, the current study failed to observe any modulation of the extent of pattern transferability for positive and negative words on the basis of participants' inclination to visualize. These data indicate that spatial representations during valence processing are likely to be triggered by other mechanisms than visualization. One possibility through which an abstract target domain (i.e., valenced concepts) can be linked to a concrete source domain (i.e., vertical space) is that of metaphorical mapping. In several languages, including English, it has been noted that metaphorical expressions link positive and negative valence with the top and bottom of a vertical spatial continuum (Lakoff & Johnson, 1999). For instance, depending on their mood, people may be described as *upbeat* or *down*, depending on their character, they might be told that they will go *up to heaven* or *down to hell*, and depending on their outfit, they might get the *thumbs up* or *down* (see Casasanto, 2009; Meier & Robinson, 2004). More specifically, the observation that people often recruit metaphors from concrete and/or experientially rich domains to talk about abstract things (such as valenced constructs) has led to the hypothesis that the human mind may recruit phylogenetically older structures of the brain for new uses so that sensory and motor representations from physical interactions with the world may be adopted to support abstract thought (Casasanto, 2010).

In line with this argument, the current inquiry revealed that representations of verticality as elicited by the categorization of valence words shared a partially overlapping neural signature with representations of verticality based on visual information (i.e., based on a "physical interaction with the world"). In so doing, the data emphasize that, even in the representation of abstract, "nonlocalized" concepts such as valenced words can comprise spatial traces. This observation converges with results from previous neuroimaging studies, which have revealed parietal involvement during other types of "nonlocalized" conceptual processing, such as during number and social distance judgments that are hypothesized to trigger spatial representations in the horizontal dimension (Knops, Thirion, Hubbard, Michel, & Dehaene, 2009; Yamakawa, Kanai, Matsumura, & Naito, 2009). Despite this overlap, however, it is important to note that it remains uncertain whether the observed common neural signature is of modal or amodal nature or reflects a gradual combination of these two possibilities (Chatterjee, 2010).

According to amodal accounts of cognition (e.g., Jackendoff, 2002; Fodor, 1975; Dennett, 1969), spatial representations can be elicited by the spreading of activation in an associative network from target words to semantic nodes symbolizing "up" and "down." In contrast, embodied theories of cognition argue that the linkage between words and space occurs because of the activation of sensorimotor states during word processing that are identical to those underlying the perception and experience of vertical arrangements in the real world (Barsalou, 2008; Gallese & Lakoff, 2005; Jeannerod, 2001). In the domain of spatial representations, however, these two accounts are particularly hard to tease apart. In contrast to purely visual or auditory sensations, spatial experiences are inherently multimodal and often comprise the integration of vestibular, visual, and even somatosensory or auditory cues. In line with this observation, the IPL is a multisensory region coding spatial experiences across various modalities (Renier et al., 2009; Indovina et al., 2005; Brandt & Dietrich, 1999). How exactly the integration of different types of sensory information is achieved (via amodal vs. multimodal coding), however, remains a matter of debate (Macaluso & Driver, 2005; Calvert, Spence, & Stein, 2004).

Nevertheless, some initial data speak against mere amodal representation in IPL. Neuroimaging work has shown that, whereas spatial coding based on visual information recruits inferior parietal resources (Amorapanth et al., 2010; Corradi-Dell'Aqua et al., 2008; Baciu et al., 1999), the processing of spatial relations in more abstract terms recruits neural resources beyond the IPL. In particular, the comprehension of so-called locative prepositions such as *above*, *below*, and *between* has been associated with additional activity in the left inferior frontal gyrus (IFG) (Wu, Waller, & Chatterjee, 2007; Tranel & Kemmerer, 2004; Damasio et al., 2001). Because locative prepositions are considered special

linguistic modifiers that form a natural link between verbal (i.e., symbolic) representations and spatial experiences (Noordzij, Neggers, Ramsey, & Postma, 2008), a view has begun to emerge according to which the IFG may represent spatial relations on a more symbolic (i.e., amodal) level than the IPL by abstracting away from specific sensory input. Further support for this view comes from patient studies. Whereas some brain-damaged patients successfully solve problems that require visuospatial coding but fail when probed for the adequate use of locative prepositions, others show the reverse pattern of impairments (Kemmerer & Tranel, 2000). Such data indicate that the representation of spatial relations on diverging levels of abstraction [amodal vs. (multi)modal] may occur independently. The current data, however, do not provide any evidence in favor of one of the two formats and this issue requires further investigation.

Future research is also needed to elucidate why the current study failed to reveal brain pattern transferability in the IPL from shape to power judgments. There are several possible explanations as to why this may have been the case. First, behavioral evidence for the activation of spatial associations during the comprehension of single power words is modest. Previous data indicate that powerful is related to "up" in space, but powerlessness may not be related to "down" (Schubert, 2005). As the current methodology relied on contrasting up and down, it cannot investigate the two poles separately. Second, it should be noted that participants were asked to lie down in the scanner during task completion. As a result, they were looking through a mirror at stimuli that were projected onto a screen perpendicular to the floor of the scanner room and thus perpendicular to their own momentary position. Hence, the definition of "up" and "down" in the shape judgment task was aligned with an environment-based (i.e., a gravity-based) reference frame, but not with a viewer's frame of reference (i.e., "up" was not where the head was; see Carlson, 1999). The resulting misalignment between these two frames, which is unusual in everyday life (because people are mostly in an upright position), may also have decreased the likelihood of finding neural patterns of verticality that transferred from the visual task to conceptual judgments that rely on a viewer's frame of reference. Again, this possibility requires additional experimentation.

It is also noteworthy that attempts at reverse classification (i.e., training classifiers on valence, power, and location words and applying them to shape judgments) did not yield any statistically significant results. What this suggests is that algorithms that were free to optimize the separation of "overhead" and "underfoot" words (positive and negative words, powerful and powerless words) did not comprise a significant amount of spatial information in the current inquiry. This lack of reverse classification could have been caused by several factors. On the one hand, given the richness of associations triggered by concepts such as *danger* or *basement*, it would be expected that, for a classifier to be able to detect the commonality of spatial representations across such a wide set of words, a lot of unique concept-specific information (i.e., noise) needs to be controlled. In this case, training the classifier on a larger pool of examples than the limited set of words included in the current study should increase prediction accuracy for spatial probes. On the other hand, however, it is also possible that successful word classification could be achieved based on other pivotal verticality-unrelated dimensions and that spatial representations make only a minor contribution to determining the meaning of these concepts. In line with this argument, it has previously been suggested that spatial representations during conceptual processing may only play an epiphenomenal rather than causal role during actual word comprehension (Barsalou, 2008). In context of the current study, however, the lack of reverse transference success represents a classical null finding that warrants further investigation.

In summary, the current inquiry examined whether classifiers that successfully distinguished between neural patterns underlying the perception of "up" and "down" shapes in several subregions of the IPL also provided discriminatory information to separate object, valence, and power words. In support of previous behavioral work (Meier, Hauser, et al., 2007; Šetić & Domijan, 2007; Meier & Robinson, 2004; Zwaan & Yaxley, 2003), it was found that discriminatory neural patterns underlying the spatial processing of visual input generalized without further training to object and valence word judgments. In so doing, the data suggest that spatial traces can be used to enrich conceptual representations of various classes of knowledge. Correlational analyses further revealed, however, that depending on the type of concepts probed (concrete vs. abstract spatial connotation) traces of verticality may be elicited by diverging mechanisms. To our knowledge, these are the first findings elucidating how word verticality mappings fit within the neuroscience of concept representation (Mahon & Caramazza, 2009; Martin, 2007).

REFERENCES

Amorapanth, P. X., Widick, P., & Chatterjee, A. (2010). The neural basis for spatial relations. *Journal of Cognitive Neuroscience, 22,* 1739–1753.

Baciu, M., Koenig, O., Vernier, M.-P., Bedoin, N., Rubin, C., & Segebarth, C. (1999). Categorical and coordinate spatial

relations: fMRI evidence for hemispheric specialization. *NeuroReport, 10,* 1373–1378.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Boroditsky, L., & Prinz, J. J. (2008). What thoughts are made of. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 98–115). Cambridge: Cambridge University Press.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings.* Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brandt, T., & Dietrich, M. (1999). The vestibular cortex. Its locations, functions, and disorders. *Annals of the New York Academy of Sciences, 871,* 293–312.

Calvert, G. A., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes.* Cambridge, MA: The MIT Press.

Carlson, L. A. (1999). Selecting a reference frame. *Spatial Cognition and Computation, 1,* 365–379.

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138,* 351–367.

Casasanto, D. (2010). Space for thinking. In V. Evans & P. Chilton (Eds.), *Language, cognition, and space* (pp. 453–478). London: Equinox Publishing.

Casasanto, D., & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition, 115,* 179–185.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2,* 913–919.

Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience dependent modulation of category-related cortical activity. *Cerebral Cortex, 12,* 545–551.

Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition, 2,* 79–116.

Childers, T. L., Houston, M. J., & Heckler, S. E. (1985). Measurement of individual differences in visual versus verbal information processing. *Journal of Consumer Research, 12,* 125–134.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

Corradi-Dell'Aqua, C., Hesse, M. D., Rumiati, R. I., & Fink, G. R. (2008). Where is a nose with respect to a foot? The left posterior parietal cortex processes spatial relationships among body parts. *Cerebral Cortex, 18,* 2879–2890.

Crawford, L. E., Margolies, S. M., Drake, J. T., & Murphy, M. E. (2006). Affect biases memory of location: Evidence for the spatial representation of affect. *Cognition and Emotion, 20,* 1153–1169.

Damasio, H., Grabowski, T. J., Tranel, D., Ponto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). Neural correlates of naming actions and of naming spatial relations. *Neuroimage, 13,* 1053–1064.

Dennett, D. C. (1969). *Content and consciousness.* London: Routledge & Kegan Paul.

Eickhoff, S. B., Weiss, P. H., Amunts, K., Fink, G. R., & Zilles, K. (2006). Identifying human parieto-insular vestibular cortex using fMRI and cytoarchitectonic mapping. *Human Brain Mapping, 27,* 611–621.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science, 19,* 93–97.

Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification analysis. *Brain Research, 1282,* 114–125.

Fan, Y., Shen, D., & Davatzikos, C. (2006). Detecting cognitive states from fMRI images by machine learning and multivariate classification. *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop* ISBN: 0-7695-2646-2.

Fodor, J. A. (1975). *The language of thought.* New York: Crowell.

Formisano, E., De Martino, F., & Valente, G. (2008). Multivariate analysis of fMRI time series: Classification and regression of brain responses using machine learning. *Magnetic Resonance Imaging, 26,* 921–934.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensorymotor system in reason and language. *Cognitive Neuropsychology, 22,* 455–479.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9,* 558–565.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43,* 379–401.

Golland, P., & Fischl, B. (2003). Permutation tests for classification: Towards statistical significance in image-based studies. In C. J. Taylor & J. A. Noble (Eds.), *Information processing in medical imaging* (pp. 330–341). Berlin/Heidelberg: Springer.

Hauk, O., Davis, M. H., Kherif, F., & Pulvermüller, F. (2008). Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *European Journal of Neuroscience, 27,* 1856–1866.

Indovina, I., Maffei, V., Bosco, G., Zago, M., Macaluso, E., & Lacquaniti, F. (2005). Representation of visual gravitational motion in the human vestibular cortex. *Science, 308,* 416–419.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford University Press.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage, 14,* 103–109.

Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia, 44,* 1607–1621.

Kemmerer, D., & Tranel, D. (2000). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology, 17,* 393–414.

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324,* 1583–1585.

Kriegeskorte, N., Cusack, R., & Bandettini, P. (2010). How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex spatiotemporal filter? *Neuroimage, 49,* 1965–1976.

Lakoff, G. (2008). The neural theory of metaphor. In R. W. Gibbs (Ed.), *Cambridge handbook of metaphor and thought* (pp. 17–38). Cambridge, MA: Cambridge University Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* Chicago: University of Chicago Press.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences, 28,* 264–271.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology, 60,* 27–51.

Marques, J. F., Canessa, N., Siri, S., Catricalà, E., & Cappa, S. (2008). Conceptual knowledge in the brain: fMRI evidence for featural organization. *Brain Research, 1194,* 90–99.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25–45.

Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom–up meets top–down: Neuronal interactions during perception and imagery. *Cerebral Cortex, 14,* 1256–1265.

Meier, B. P., Hauser, D. J., Robinson, M. D., Kelland Friesen, C., & Schjeldahl, K. (2007). What's "up" with god? Vertical space as a representation of the divine. *Journal of Personality and Social Psychology, 93,* 699–710.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up. Associations between affect and vertical position. *Psychological Science, 15,* 243–247.

Meier, B. P., Sellbom, M., & Wygant, D. B. (2007). Failing to take the moral high ground: Psychopathy and the vertical representation of morality. *Personality and Individual Differences, 43,* 757–767.

Meteyard, L., & Vigliocco, G. (2008). The role of sensory and motor information in semantic representation: A review. In P. Calvo & A. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 293–312). London: Academic Press.

Naganuma, M., Inatomi, Y., Yonehara, T., Fujioka, S., Hashimoto, Y., Hirano, T., et al. (2006). Rotational vertigo associated with parietal cortical infarction. *Journal of the Neurological Sciences, 246,* 159–161.

Noordzij, M. L., Neggers, S. F. W., Ramsey, N. F., & Postma, A. (2008). Neural correlates of locative prepositions. *Neuropsychologia, 46,* 1576–1580.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013–1023.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate analyses? *Neuroimage, 49,* 1943–1948.

O'Toole, A. J., Abdi, F. J. H., Pénard, H., Dunlop, J. P., & Parent, M. A. (2007). Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. *Journal of Cognitive Neuroscience, 19,* 1735–1752.

Pecher, D., van Dantzig, S., Boot, I., Zanolie, K., & Huber, D. E. (2010). Congruency between word position and meaning is caused by task-induced spatial attention. *Frontiers in Psychology, 1,* 1–8.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45,* S199–S209.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.

Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science, 20,* 1364–1372.

Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., Van Meter, J., & Rauschecker, J. P. (2009). Multisensory integration of sounds and virotactile stimuli in processing streams for "what" and "where." *Journal of Neuroscience, 29,* 10950–10960.

Schlindwein, P., Mueller, M., Bauermann, T., Brandt, T., Stoeter, P., & Dieterich, M. (2008). Cortical representation of saccular vestibular stimulation: VEMPs in fMRI. *Neuroimage, 39,* 19–31.

Schubert, T. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality & Social Psychology, 89,* 1–21.

Schubert, T. W., Waldzus, S., & Seibt, B. (in press). More than a metaphor: How the understanding of power is grounded in experience. In T. W. Schubert & A. Maass (Eds.), *Spatial dimensions of social thought.* Berlin: Mouton de Gruyter.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language & Cognitive Processes, 22,* 297–312.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.

Stephan, T., Deutschländer, A., Nolte, A., Schneider, E., Wiesmann, M., Brandt, T., et al. (2005). Functional MRI of galvanic vestibular stimulation with alternating currents at different frequencies. *Neuroimage, 26,* 721–732.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top–down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience, 29,* 1565–1572.

Tranel, E., & Kemmerer, D. (2004). Neuroanatomical correlates of locative prepositions. *Cognitive Neuropsychology, 21,* 719–749.

Trojano, L., Grossi, D., Linden, D. E., Formisano, E., Goebel, R., Cirillo, S., et al. (2002). Coordinate and categorical judgments in spatial imagery. *Neuropsychologia, 40,* 1666–1674.

Urasaki, E., & Yokota, A. (2004). Rotational vertigo caused by cerebral lesions: Vertigo and areas 3av, 2v and 7. *Journal of Clinical Neuroscience, 13,* 114–116.

Wu, D. H., Waller, S., & Chatterjee, A. (2007). The functional neuroanatomy of thematic role and locative relational knowledge. *Journal of Cognitive Neuroscience, 19,* 1542–1555.

Yamakawa, Y., Kanai, R., Matsumura, M., & Naito, E. (2009). Social distance evaluation in human parietal cortex. *PLoS ONE, 4,* e4360. doi:10.1371/journal.pone.0004360.

Zwaan, R. A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology, 39,* 1142–1150.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135,* 1–11.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10,* 954–958.