Visual Number Beats Abstract Numerical Magnitude: Format-dependent Representation of Arabic Digits and Dot Patterns in Human Parietal Cortex

Jessica Bulthé, Bert De Smedt, and Hans P. Op de Beeck

Abstract

In numerical cognition, there is a well-known but contested hypothesis that proposes an abstract representation of numerical magnitude in human intraparietal sulcus (IPS). On the other hand, researchers of object cognition have suggested another hypothesis for brain activity in IPS during the processing of number, namely that this activity simply correlates with the number of visual objects or units that are perceived. We contrasted these two accounts by analyzing multivoxel activity patterns elicited by dot patterns and Arabic digits of different magnitudes while participants were explicitly processing the represented numerical magnitude. The activity pattern elicited by the digit “8” was more similar to the activity pattern elicited by one dot (with which the digit shares the number of visual units but not the magnitude) compared to the activity pattern elicited by eight dots, with which the digit shares the represented abstract numerical magnitude. A multivoxel pattern classifier trained to differentiate one dot from eight dots classified all Arabic digits in the one-dot pattern category, irrespective of the numerical magnitude symbolized by the digit. These results were consistently obtained for different digits in IPS, its subregions, and many other brain regions. As predicted from object cognition theories, the number of presented visual units forms the link between the parietal activation elicited by symbolic and nonsymbolic numbers. The current study is difficult to reconcile with the hypothesis that parietal activation elicited by numbers would reflect a format-independent representation of number.

INTRODUCTION

Researchers in the field of numerical cognition have proposed that intraparietal sulcus (IPS) contains an abstract module for number processing, which means that IPS comprises neural representations for numerical magnitudes that are independent of format (e.g., Arabic digits or dot patterns). Such an abstract representation account assumes that, for example, the Arabic digit “4” is represented in the same way as a pattern of four dots activating the same neurons, but this representation is different from Arabic digit “8” and a pattern of eight dots. This conclusion is drawn from many studies that observed that IPS is involved in magnitude processing and that this IPS activity is independent of format (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003; Naccache & Dehaene, 2001; Pinel, Dehaene, Rivière, & Le Bihan, 2001; Dehaene & Cohen, 1997).

Many of the relevant brain imaging studies, however, have at least one of the following two major limitations that undermine the observed evidence of IPS as an abstract number module. The first limitation deals with the exclusive use of Arabic digits and/or number words (e.g., “two”) to unravel the abstractness of number processing in parietal cortex, which makes it difficult to make comparisons with the processing of a nonsymbolic numerical magnitude, such as dot patterns (Cohen Kadosh & Walsh, 2009; Shuman & Kanwisher, 2004). When both symbolic and nonsymbolic numbers are used in the task, there is less support for the existence of an abstract representation of numbers (Ansari, Fugelsang, Dhital, & Venkatraman, 2006; Shuman & Kanwisher, 2004).

Second, the notion of abstract number processing predicts null results in neuroimaging studies that compare the mean BOLD signal for Arabic digits and dots in IPS. More specifically, these studies predict no significant differences in the mean activation between different formats in IPS. Such null results are, however, hard to interpret because they might have been due to a lack of statistical power or to the insensitivity of the paradigms that were used (Cohen Kadosh & Walsh, 2009).

Researchers have tried to address this latter limitation by using adaptation fMRI paradigms. In this paradigm, the repetition of the same stimulus (e.g., Arabic digit 4) reduces the BOLD signal. If the adapted BOLD signal changes when the stimulus magnitude changes but not when the stimulus format changes (e.g., four dots), then the inference is made that the underlying neuronal population is sensitive to stimulus magnitude and not to format. This paradigm has allowed researchers in the field of numerical cognition to test whether there are abstract number neuronal populations or not (Holloway, Battista, Vogel, & Ansari, 2013; Roggeman, Santens, Fias, & Verguts, 2011; Santens, Roggeman, Fias, & Verguts, 2010;
The use of adaptation fMRI did not resolve the discussion about the presence of abstract representations of numerical magnitudes in human parietal cortex. For example, the adaptation study of Piazza et al. (2007) with Arabic digits and dots demonstrated that the BOLD signal in right IPS recovered when the numerical magnitude changed and that this recovery was not affected by the stimulus format of the test stimulus. This result fits with the idea of abstract numerical magnitude processing in IPS. However, Piazza et al. (2007) also observed an interaction effect between format and recovery in left IPS, which suggested format-dependent processing of numerical magnitudes (Cohen Kadosh & Walsh, 2009; Ansari, 2007). Together with other fMRI adaptation studies (Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007) and other fMRI evidence (Holloway, Price, & Ansari, 2010), it seems more plausible that both format-independent and format-dependent representations are present in parietal cortex (Ansari, 2007), at least, as far as adaptation provides a reliable measure of neuronal selectivity, which is a point that has been contested (Sawamura, Orban, & Vogels, 2006).

Recently, fMRI studies have included a different methodology to assess neural selectivity, namely multivoxel pattern analysis (MVPA; see Norman, Polyn, Dettre, & Haxby, 2006). MVPA might be a helpful tool to further explore where in the cortex overlapping and/or distributed representations of dots and Arabic digits are present (Dehaene, 2009; Ansari, 2008). Specifically, MVPA has the potential to reveal the presence of neural representations for dots and Arabic digits in the parietal cortex and IPS, and this technique allows one to straightforwardly test whether these neural representations are overlapping or not. For example, a classifier can be trained to differentiate between Arabic digit 2 and Arabic digit 4, and this same classifier can then be used to differentiate between two dots and four dots. If an abstract representation underlies parietal activity, this classifier should be able to generalize from Arabic digits to dots, or vice versa.

Recent studies applying this MVPA technique have been able to extract format-specific magnitude information in parietal cortex and IPS (Bulthé, De Smedt, & Op de Beeck, 2014; Damarla & Just, 2012; Eger et al., 2009); however, the evidence for format-independent representations is weak. Bulthé et al. (2014) and Damarla and Just (2012) found no generalization and thus no overlapping representations between Arabic digits and dots in parietal cortex or IPS, which contradicts the existence of an abstract representation of numerical magnitude in parietal cortex or IPS. On the other hand, Eger et al. (2009) observed weak asymmetrical generalization between dots and digits: the discrimination of Arabic digits generalized to dots was just above chance level, but the generalization from dots to Arabic digits was at chance level. Overall, magnitude representations in IPS seem to a large degree to be format dependent.

However, format dependence does not necessarily mean that there is no relationship between how stimuli from different formats are represented, as the parietal cortex has not only been implicated in magnitude representations. Researchers of object cognition have shown that activity in IPS is associated with the number of visual objects that are presented (Xu, 2008; Xu & Chun, 2007; Song & Jiang, 2006; Todd & Marois, 2005; Vogel & Machizawa, 2004; Wojciulik & Kanwisher, 1999). From this object cognition literature one would predict that Arabic digits will have similar activity to one dot because both contain only one visual object. This prediction is not compatible with the hypothesis of an abstract magnitude representation, which would predict that an Arabic digit and a dot pattern that share the same numerical magnitude (e.g., digit 4 and four dots) would elicit a similar pattern of activity across the neurons in IPS.

Against this background, the associations between the representations of Arabic digits and dots remain unclear. Are the neural patterns in the parietal regions of Arabic digits and dots more alike when they share an underlying magnitude, as expected from the numerical cognition literature (Dehaene, 2009; Piazza et al., 2007; Nieder, Freedman, & Miller, 2002)? Or, are the neural patterns of Arabic digits and dots more related by the number of visual elements they share, as expected from studies in the field of object cognition (Xu, 2008; Xu & Chun, 2007; Song & Jiang, 2006; Todd & Marois, 2005; Vogel & Machizawa, 2004)? This study integrated both research fields (numerical cognition and object cognition) and tested their opposite predictions about the relative similarity of Arabic digits and dot patterns by using MVPA analyses.

**METHODS**

**Participants**

Twelve healthy individuals (three men and nine women, 26.5 ± 2.28 years old, one left-handed) participated in this fMRI study and were paid for their participation. The participants had normal or corrected-to-normal vision, and screening did not indicate a neurological or psychiatric history. The study was approved by the medical ethics committee of KU Leuven. All participants provided written informed consent before scanning.

**Stimuli**

The stimuli (400 × 400 pixels) were presented in a white-centered circle on a black background. Two formats were chosen, Arabic digits and dot patterns, and both comprised 1, 2, 4, and 8 as numerical magnitudes (Figure 1). Using the method and automated program by Dehaene,
Izard, and Piazza (2005), we controlled the dot stimuli for intensive confounding parameters, such as individual item size and interitem spacing, and extensive confounding parameters, for example, total luminance and total area spanned by the dots, by varying them randomly across the dot displays. To avoid adaptation for Arabic digits, the symbols varied in position and size across trials.

Stimuli were presented via PsychoToolbox 3 (Brainard, 1997), and a Barco 6400i LCD projector (resolution: 1024 × 768, refresh rate: 75 Hz) was used to project the stimuli on a vertical screen. The screen was positioned approximately 35 cm from participants’ eyes and was visible via a mirror attached to the head coil.

Design
The experimental procedures were very similar to a previous study (see Bulthé et al., 2014, for more elaborated task details). The critical difference, introduced to be able to test the central hypotheses of the current article, was the inclusion of the numerical magnitude 1 in this study. We used a short block design with variable block duration of either 4, 5, or 6 sec. One run lasted for 280 sec and consisted of 48 experimental blocks (each condition was repeated six times, two times for each block duration) and 7 fixation blocks. In the experimental blocks, one condition (e.g., four dots) was repeated in four, five, or six trials. Each trial comprised in total 1000 msec, including 200-msec stimulus presentation and 800-msec fixation. The first and last fixation blocks were presented for 8 sec. The fixation blocks between experimental blocks lasted for 4, 5, or 6 sec. The experiment comprised 10–12 runs per participant.

Brain imaging data were collected during a number comparison task, which made the participants explicitly access numerical magnitude representations (Zorzi, Di Bono, & Fias, 2011; Piazza et al., 2004; Pinel, Piazza, Le Bihan, & Dehaene, 2004). Participants indicated whether the presented number was smaller or larger than 3 each time format and/or numerical magnitude changed.

The experiment also included a localizer task in which participants had to perform a subtraction task. In this task, participants had to subtract numbers in the number domain 1 to 20 and they indicated whether the solution to a subtraction was odd or even. For each trial, the subtraction problem was presented for 1700 msec followed by a fixation cross for 300 msec. The independent localizer data were used to define the ROIs.

fMRI Data Acquisition
Data were acquired on a 3T Philips Intera Scanner (Department of Radiology, KU Leuven) with a 12-channel head coil. Functional images were obtained with a T2*-weighted EPI sequence (with 48 oblique transverse slices, in-plane resolution = 2.1 mm, slice thickness = 2 mm, interslice gap = 0.1 mm, repetition time = 3000 msec, echo time = 30 msec, flip angle = 90, 104 × 104 matrix). For each participant, a high-resolution T1-weighted anatomical image was obtained (182 slices, resolution 0.98 × 0.98 × 1.2 mm, repetition time = 9.6 msec, echo time = 4.6 msec, 256 × 256 acquisition matrix).

fMRI Preprocessing
The data were processed with Statistical Parametric Mapping software package (SPM8, Welcome Department of Cognitive Neurology, London, United Kingdom). Anatomical images were normalized to the standard brain template defined by the Montreal Neurological 152-brains average. Functional images were corrected for slice timing differences and realigned to the mean image to correct for head movements. Coregistration and spatial normalization were done using the parameters obtained in the normalization of the anatomical images. During normalization functional images were resampled to a voxel size of 2 × 2 × 2 mm. Finally, functional images were spatially smoothed using Gaussian kernels of 4 mm FWHM.

Statistical Analysis
The experimental effects in each voxel were estimated by a multisession design matrix that modeled the data at block level. A general linear model for each run was created with regressors for each participant for each condition. The six motion realignment parameters were additionally included as regressors of no interest to account for signal variations due to head movements. After fitting

![Figure 1. Stimulus examples for all four numerical magnitudes in both formats.](image-url)
the general linear model for each run that was collected, subsequent analyses were performed using t statistics (which resulted from the contrast of each condition vs. baseline), because they take both mean and variance of the activations into account (Misaki, Kim, Bandettini, & Kriegeskorte, 2010).

**ROIs**

For each ROI, we only included voxels that were significantly active in the contrast task minus fixation in the localizer scans. These voxels were restricted to those in the appropriate anatomical mask that was created with the anatomical WFU PickAtlas Toolbox (Wake Forrest University PickAtlas, fmri.wfubmc.edu/cms/software), at least if the ROI was available in the toolbox. The functional contrast was thresholded at $p < .001$ (uncorrected for multiple comparisons).

In view of the literature reviewed above, we mainly focused on parietal cortex, IPS, and its subdivisions (right and left anterior and right and left posterior). In addition, we also included additional ROIs to find out whether similar effects were present in other brain regions (see Bulthé et al., 2014, for a similar rationale). Selection of ROIs was based on ROIs that have been reported to be involved in numerical processes in previous studies (Holloway et al., 2013; Maruyama, Pallier, Jobert, Sigman, & Dehaene, 2012; Zhang, Chen, & Zhou, 2012; Santens et al., 2010; Lyons & Ansari, 2009; Piazza et al., 2007; Dehaene, Piazza, Pinel, & Cohen, 2003; Zago et al., 2001): All regions (all voxels with significant activity vs. baseline in the localizer task in a participant), frontal cortex, parietal cortex, temporal cortex, occipital cortex, left and right superior parietal lobule, inferior occipital cortex, STS, visual word form area, Wernicke’s area, fusiform gyrus, left and right inferior frontal gyri, and left and right superior frontal gyri.

For each participant, the “All Regions” ROI was derived from the subtraction localizer task (i.e., the contrast “task minus fixation”) and comprised all voxels that survived the threshold at $p < .0001$. So, in this ROI, all the voxels that processed and manipulated numerical magnitudes were included. The All Regions ROI was included because it gives a broad overview of the trends in the data (for similar rationale, see Bulthé et al., 2014). For example, when a specific ROI does not show any significant effects, the question remains whether there are just no effects present in that specific ROI or whether there are no measurable effects present in the entire cortex (in the first case, the All Regions ROI will show significant results; in the second case, the All Regions ROI will not yield any significant findings). It is important to point out that including the All Regions is not comparable to a searchlight analysis, because both analyses differ in their spatial scale. Searchlight analysis represents information on a local scale (a very small cluster of neighboring voxels), in contrast to the All Regions ROI, which represents information distributed at global scale (Bulthé, van den Hurk, Daniels, De Smedt, & Op de Beeck, 2014). Figure 2 shows the All Regions ROI across participants derived from a second level analysis.

**Data Analysis**

We implemented decoding and generalization pattern classification with custom code written in Matlab (The MathWorks, Natick, MA). Both pattern recognition analyses were performed with linear support vector machines (SVM) using the OSU SVM toolbox with the following parameters: a radial basis function kernel as decision function with parameter gamma set to 1; a C-SVC classification algorithm was used with parameter C set to 1. Response patterns for every condition in each run were extracted for each ROI and normalized across voxels; the patterns were normalized by subtracting the mean across voxels and then dividing this by the standard deviation across voxels for each condition. We followed a repeated random subsampling cross-validation procedure: The data were randomly divided into 70% training data and 30% test data (the latter were averaged to one response pattern per condition), and this was repeated 100 times. The performance on the test data of all pairwise comparisons between conditions was averaged over different comparisons of interest (e.g., all comparisons of dot conditions or all comparisons of Arabic digits).

**Decoding**

Decoding pattern classification resulted in a $7 \times 8$ decoding matrix for every ROI. Higher decoding accuracies indicate less similar neural representations. The decoding accuracies were then averaged over various comparisons of interest: Arabic digits (mean within-format decoding accuracy for Arabic digits), dots (mean within-format decoding accuracy for dots), same number (mean decoding...
accuracy of every dot condition contrasted with the Arabic digit of the same numerical magnitude), and different number (mean decoding accuracy of every dot condition contrasted with the Arabic digit of a different numerical magnitude). The within-format decoding results for every ROI were tested for significance ($p < .05$) across participants by a two-sided $t$ test with respect to chance level (50%). The same number decoding accuracy was tested for significance ($p < .05$) against the different number decoding by a paired $t$ test.

Furthermore, the decoding accuracy was calculated between every Arabic digit condition (1, 2, 4, 8) with a particular dot condition (e.g., one dot) and averaged across those four decoding accuracies (Figure 5A). This resulted in four averaged decoding accuracies: decoding between digits and one dot, digits and two dots, digits and four dots, and digits and eight dots. A linear regression model was fitted to these four decoding accuracies, resulting in a slope that was tested for significance ($p < .05$) across participants by a $t$ test.

To rule out any bias in our decoding analysis that would lead to chance performance being higher than the theoretically expected proportion of .50, we performed random permutation tests (1000 permutations) for decoding of Arabic digits and dots, within the main ROIs: All Regions, parietal cortex, and IPS. For both dots and Arabic digits, the 95% confidence interval of the null distribution was for all the regions within the range of [0.4929 – 0.5084].

### Generalization

For generalization pattern classification, dot pattern condition pairs (e.g., one dot vs. four dots) were used to train the classifier and the corresponding Arabic number conditions (e.g., Arabic digit 1 and Arabic digit 4) were used to test the performance of the classifier (Figure 6A). All pairs in this generalization analysis included the numerical magnitude “1” as one condition and one of the other numerical magnitudes as the contrast condition. From this analysis, three measures were extracted: the classification accuracy, the correct classification of Arabic digit 1 as one dot, and the confusion classification that indicates how many times another digit is classified as one dot (e.g., Arabic digit 4 classified as one dot). The generalization results for every ROI were tested for significance ($p < .05$) across participants by two-sided $t$ tests with respect to chance level (50%).

### RESULTS

#### Behavioral Results

A two-way repeated-measures (Distance × Format) ANOVA was applied to the accuracy and RTs of the number comparison task for Arabic digits and dots. For accuracy, there was no significant main effect of either Format, $F(1, 11) = 3.88, p = .08$, and Distance, $F(2, 22) = 1.88, p = .18$, and no significant interaction between Distance and Format, $F(2, 22) = 1.04, p = .18$. For the RTs, there was a significant main effect for Distance, $F(2, 22) = 7.38, p = .004$, showing longer RTs for smaller distances than for larger distances. Again, there was no significant main effect of Format, $F(1, 11) = 3.51, p = .09$, and no significant interaction between Format and Distance, $F(2, 22) = 0.81, p = .46$.

#### Classification within Format

Previous MVPA fMRI studies have shown that parietal cortex and IPS contain patterns of activity of dots and Arabic digits, which are informative about which numerical magnitude is represented in a particular format (Bulthé et al., 2014; Damarla & Just, 2012; Eger et al., 2009). We first replicated this finding of previous studies, as a significant decoding within each format is a
prerequisite to find any potential associations between formats. The classification accuracies for dots were significantly above chance ($p < .05$) in all ROIs (Figure 3): All Regions ($86\%, t[11] = 15.19, p < .0001$), parietal cortex ($73\%, t[11] = 9.06, p < .0001$), IPS ($66\%, t[11] = 5.57, p < .0001$), left anterior IPS ($59\%, t[11] = 3.79, p = .003$), right anterior IPS ($56\%, t[11] = 2.73, p = .02$), left posterior IPS ($65\%, t[11] = 4.65, p = .001$), and right posterior IPS ($60\%, t[11] = 3.65, p = .004$). This indicates that in all these ROIs there were distinguishable neural patterns for dots with different numerical magnitudes.

The neural patterns for Arabic digits with different numerical magnitudes were distinct in following ROIs (Figure 3): All Regions ($67\%, t[11] = 5.44, p < .0001$), parietal cortex ($60\%, t[11] = 4.61, p = .001$), IPS ($59\%, t[11] = 4.32, p = .001$), left anterior IPS ($56\%, t[11] = 3.43, p = .006$), and right posterior IPS ($55\%, t[11] = 2.30, p = .04$). However, in the right anterior IPS ($52\%, t[11] = 1.62, p = .13$) and left posterior IPS ($50\%, t[11] = -0.20, p = .85$), there were no distinguishable neural representations present for Arabic digits. These data are in overall agreement with the decoding accuracies obtained in previous research (Bulthé et al., 2014). Also in other parietal and nonparietal ROIs, the results were very similar to the findings of Bulthé et al. (2014).

### Classification between Formats

In this analysis, we compared the activation patterns of Arabic digits and dots to test for their similarity. If an abstract representation underlies the numerical representations, we would expect a lower decoding accuracy (neural patterns are more similar and thus less distinguishable) between Arabic digits and dots sharing the same numerical magnitude compared to Arabic digits and dots that do not have the same numerical magnitude. On the other hand, if the number of visual elements provides the important link between Arabic digits and dots, we expect an increase in decoding accuracies (e.g., neural patterns are less similar) between Arabic digits and a certain dot condition when more dots are visually presented.

The basic output of the decoding analyses were $7 \times 8$ matrices obtained by pairwise classification of the multivoxel patterns of each condition with another condition (Figure 4A–C). These decoding matrices represented the dissimilarity (e.g., higher decoding accuracies) of every condition with another condition. The dissimilarity matrices also allowed us to contrast the object cognition account, that is, the number of visual units is the link between Arabic digits and dots, and numerical cognition account, that is, the numbers are represented in an abstract manner according to their magnitude.

#### Numerical Cognition Account

If number representations in IPS are abstract, we expect a lower decoding accuracy when an Arabic digit and a dot pattern share the same numerical magnitude than when they do not. In Figure 4A–C, this would be visible by lower decoding accuracies in the four cells with the black squares compared to the other cells in the matrix. A visual inspection of the decoding matrix suggested that this was not the case: The four cells with a black square were on average as much or more distinguishable than the other cells in the matrix. This finding was quantified by the lack of any difference in the decoding accuracies between the “same number” data (obtained by averaging the decoding accuracies of the pairwise comparisons of a dot and a digit condition sharing the same magnitude) and the “different number” data (obtained by averaging the decoding accuracies of the pairwise comparisons of a dot and a digit condition with a different magnitude) in all of the ROIs, that is, All Regions ($t[11] = 0.1406, p = .89$), parietal cortex ($t[11] = -0.9503, p = .36$), IPS ($t[11] = -0.1535, p = .89$), left anterior IPS ($t[11] = 0.03, p = .97$), right anterior IPS ($t[11] = 0.74, p = .47$), left posterior IPS ($t[11] = -0.33, p = .75$), and right posterior IPS ($t[11] = 0.17, p = .87$). These results showed that the neural representations of Arabic digits and dots that

---

**Figure 4.** Decoding matrices for All Regions (A), parietal cortex (B), and IPS (C). The cells with a black square represent each an Arabic digit and dot condition sharing the same numerical magnitude (relevant for the “numerical cognition account”). The cells within the black outline are the pairwise comparisons of a dot and a digit condition with a different magnitude (e.g., neural patterns are less similar) between Arabic digits and a certain dot condition when more dots are visually presented. The basic output of the decoding analyses were $7 \times 8$ matrices obtained by pairwise classification of the multivoxel patterns of each condition with another condition (Figure 4A–C). These decoding matrices represented the dissimilarity (e.g., higher decoding accuracies) of every condition with another condition. The dissimilarity matrices also allowed us to contrast the object cognition account, that is, the number of visual units is the link between Arabic digits and dots, and numerical cognition account, that is, the numbers are represented in an abstract manner according to their magnitude.
shared the same numerical magnitude were as distinctive as Arabic digits and dots with different numerical magnitudes in the parietal cortex and in the IPS.

Object Cognition Account

According to the object cognition account, the pattern of activity in the IPS or parietal cortex to numerical stimuli should be related to the number of units contained in a stimulus. In this case, we expected a lower decoding accuracy (more similarity) when any digit was compared with one dot than with two dots, which might in turn result in a lower decoding accuracy to four dots, and so on. In the matrix, this would be visible by a lower decoding value in the black rectangle outline cells than in the columns to the right of this rectangle and an increasing value (higher decoding accuracy) by each shift to the right in the matrix. A visual inspection of the dissimilarity matrices in Figure 4 suggested that this was indeed the case: The similarities of the neural patterns between Arabic digits and one dot were higher than the similarities between a digit condition and dot conditions with more dots.

To test this more formally, a linear regression analysis was applied to the averaged decoding accuracies between Arabic digits and a certain dot condition (e.g., Arabic digits with one dot was calculated by averaging the decoding accuracy between Arabic digit 1 and one dot, the decoding accuracy between Arabic digit 2 and two dots, etc.). Concretely, the four numbers involved in the regression analysis corresponded to the mean of the cells surrounded by the black outline in the matrices shown in Figure 4, followed by the mean of the four cells in the next matrix column to the right, and so on until the last column in the matrices (Figure 5A).

The object cognition account predicts a positive slope as the neural response patterns are expected to reflect the number of objects on the screen. The number account does not predict a particular trend in this regression analysis, because only a low decoding accuracy between Arabic digits and dot patterns that share the same numerical magnitude is expected. The slopes of this linear regression analysis for all ROIs are illustrated in Figure 5B: All Regions ($t(11) = 7.38, p < .0001, R^2 = 0.86$), parietal cortex ($t(11) = 7.69, p < .0001, R^2 = 0.76$), IPS ($t(11) = 4.51, p = .0004, R^2 = 0.56$), left anterior IPS ($t(11) = 3.35, p = .003, R^2 = 0.60$), right anterior IPS ($t(11) = 2.59, p = .01, R^2 = 0.50$), left posterior IPS ($t(11) = 3.94, p = .001, R^2 = 0.48$), and right posterior IPS ($t(11) = 3.45, p = .002, R^2 = 0.68$). In all the other defined ROIs, this slope was significantly (all $p's < .05$) different from zero and positive, except for the visual word form area and left superior frontal gyrus where the same trend towards a positive slope was present but not significant.

Generalization Classification Analyses

We subsequently ran generalization MVPA analyses by training a classifier to differentiate one dot from another dot condition (e.g., eight dots) and by subsequently testing this classifier on Arabic digit “1” and the Arabic digit sharing the same numerical magnitude as the other dot condition (e.g., digit “8”; Figure 6). The numerical cognition account predicts successful generalization because it postulates overlapping neural representations for Arabic digits and dots that share the same numerical magnitude.
magnitude. In contrast, the object cognition account predicts a generalization at chance level and additionally expects the confusion of an Arabic digit of a large size (e.g., digit “8”) with the one-dot condition, as these two conditions share the number of “objects” on the screen.

The generalization accuracy from dots to Arabic digits was not significant (all ps > .26) in any of the 21 ROIs (Figure 6B showing most relevant ROIs). This suggests that there is no abstract coding of numerical magnitude that is independent of format, in contrast to what is expected from the numerical cognition account. This finding is consistent with previous reports that classifiers trained on multivoxel patterns of one format tend to generalize very poorly (Eger et al., 2009) or even not at all towards stimuli with the same numerical magnitudes represented in a different format (Bulthé et al., 2014; Damarla & Just, 2012).

In addition to the absence of generalization, the object cognition account also predicts a confusion of each Arabic digit condition (e.g., digit “8”) with the one-dot condition instead of a classification as the dot condition sharing the numerical magnitude with the Arabic digit (e.g., eight dots). This is what we observed (Figure 6B): All Arabic digits, independent of their numerical magnitude, were more often classified as one dot instead of the dot pattern with the same numerical magnitude (e.g., two dots, four dots and eight dots). This effect was present in All Regions ($t(11) = 11.65, p < .0001$), parietal cortex ($t(11) = 6.48, p < .0001$), IPS ($t(11) = 5.58, p = .0002$), left posterior IPS ($t(11) = 2.64, p = .02$), right posterior IPS ($t(11) = 3.17, p = .01$), right posterior IPS ($t(11) = 3.15, p = .01$), except for left posterior IPS, where the same trend was present but not significant ($t(11) = 1.68, p = .12$). For most of the other defined ROIs, this confusion rate was significant, except for visual word form area, Wernicke, left inferior frontal gyrus, and left and right superior frontal gyrus, where the same trend was again present but not significant.

**DISCUSSION**

The results from this study failed to support the hypothesis that IPS contains abstract numerical magnitude neural representations. Instead, the results confirmed several predictions from the object cognition account. The linear regression on the decoding analyses showed a significant increase in the dissimilarity of neural patterns between Arabic digits and a dot pattern when more dots were being presented. This finding was further bolstered by the generalization MVPA where Arabic digits were significantly classified as one-dot pattern instead of the dot pattern with the same numerical magnitude as the Arabic digit. These findings were observed in IPS, parietal cortex, and other ROIs.

**Absence of Cross-format Generalization**

The abovementioned absence of a cross-format generalization from dots (e.g., four vs. eight dots) to Arabic digits (e.g., Arabic digit 4 vs. Arabic digit 8) suggests that there

---

**Figure 6.** (A) Schematic overview of the generalization classification pattern analysis and the expected results from both hypotheses. (B) The generalization classification results. SVM Generalization: the generalization accuracy of the classifier; “Digit 1 as one dot”: the correct classification performance of digit 1 as one dot. “Confusion of digits as one dot”: the confusion rate. How often is a digit other than “1” confused with one dot instead of the corresponding dot pattern sharing the same numerical magnitude. Error bars represent 95% confidence interval. The filled bullets indicate significant decoding ($p < .05$).

---

**Notes:**

- Bulthé, De Smedt, and Op de Beeck 1383
- Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/jocn_a_00787 by guest on 25 April 2021
were no overlapping neural representations for Arabic digits and dots sharing the same numerical magnitude.

Although this is in line with previous reports by Bulté et al. (2014) and Damarla and Just (2012), it is important to point out that this is still a null result. However, in this study, these null results cannot be explained by a lack of power or to task difficulty differences between both formats. First, the failed cross-format generalization was not due to a lack of power, because it resulted from a significant classification of digit “1” as one dot and a significant and equally large confusion classification of other Arabic digits. Second, the null result of cross-format generalization did not seem to reflect possible task difficulty differences between the Arabic digits and dot comparison tasks, which in theory could modulate IPS and parietal activation in a way that is interfering with generalization across formats. The behavioral data of the number comparison task in our experiment showed no significant differences in accuracy and RTs between both formats. So, neither a lack of power in our data or differences in task difficulty of formats can explain the absence of cross-format generalization.

A similar analysis as our confusion generalization analysis was performed in the study of Eger et al. (2009), which is briefly mentioned in their supplemental data. Their results showed also a high confusion of Arabic digits with one dot, instead of being classified as the corresponding dot condition sharing the same numerical magnitude. These authors concluded that “This could potentially indicate that the classifier has access to a mixture of codes: One of them being a nonsymbolic one (number of objects) and therefore the pattern for a digit (a single object) is most similar to the one for two dots.” So, both our data and Eger et al.’s (2009) data seem to converge with the object cognition account, namely numerical magnitudes processed as number of objects.

The cross-format generalization applied for the confusion analysis in our study was only observed in the direction from dots to Arabic digits. Eger et al. (2009) observed a significant generalization from Arabic digits to dots (but not vice versa). We tested the generalization from Arabic digits to dots in the current data set and found no significant generalization from Arabic digits to dots in any of the ROIs (ps > .30). This failure to replicate the asymmetrical cross-format generalization of Eger et al. (2009) has also been reported in two recent studies (Bulté et al., 2014; Damarla & Just, 2012).

The lack of cross-format generalization in the parietal cortex and IPS does not mean an absence of neural representations of Arabic digits or dots in those regions. The current study and previous studies (Bulté et al., 2014; Damarla & Just, 2012; Eger et al., 2009) have clearly shown that it is possible to distinguish between different neural representations of Arabic digits and dots in parietal regions, which indicates that some numerical aspects of these stimuli are being processed in IPS and parietal cortex. However, the absence of cross-format generalization due to the confusion of Arabic digits as one dot in the current study suggests that these representations are not overlapping.

### Format-specific Processing in Other Brain Regions

Previous neuroimaging studies have pointed to other regions in the cortex that are important when processing Arabic digits and dots, such as TPJ, fusiform gyrus, dorsal pFC, angular gyrus, etc. (Holloway et al., 2013; Roggeman et al., 2011; Santens et al., 2010; Lyons & Ansari, 2009; Ansari, Lyons, van Eimeren, & Xu, 2007; Piazza et al., 2007; Dehaene et al., 2003; Polk, Reed, Keenan, Hogarth, & Anderson, 2001; Menon, Rivera, White, Glover, & Reiss, 2000). To test the presence of neural representations of numerical magnitudes in regions outside IPS and parietal cortex, we included several extraparietal ROIs (see Methods) in our analyses. Many ROIs showed significant within-format decoding for both Arabic digits and dots demonstrating distinct neural representations for both formats.

The significant decoding accuracies in regions outside parietal cortex demonstrate the presence of distinguishable neural representations for Arabic digits and dots in those regions. However, this does not mean these representations reflect an underlying “numerical” magnitude, let alone an abstract numerical magnitude. This can be illustrated by the findings of All Regions ROI that was included in this study. More specifically the All Regions ROI has stronger distinct neural representations of dots and Arabic digits than the parietal cortex or IPS. This does not mean that in all the regions of the human cortex Arabic digits and dots are processed in the same way or that the underlying neural representations are identical across regions. It only reflects the many processes that contribute to the emergence of >symbolic and nonsymbolic numerical magnitude representations. For example, occipital regions are known to represent the visual properties of stimuli and the pFC might process task-related aspects (e.g., process small and large numbers differently).

In this context, it is not surprising that the decoding accuracies in All Regions (compared to parietal cortex and IPS) were much higher for dots than for Arabic digits because nonnumerical features, such as visual characteristics, were much more present in the dots. These visual characteristics probably emerged from the occipital lobe, because when this lobe is excluded from the All Regions ROI the decoding accuracies dropped to the level of the parietal cortex and IPS.

### Other Possible Visual Processes

It is important to point out that the object cognition account is only one of a group of related visually based hypotheses one might evoke to explain our results. The dot pattern conditions, when averaged across all individual trials in each condition, differ from each other on multiple dimensions (other than the number of elements), such as the number of black pixels, clutter, or complexity of the stimulus. All such visually based hypotheses of IPS activity stand in sharp contrast with the idea of an abstract number module and thus serve an equivalent purpose in the context of this study.
On the basis of the results of the current study, we might not be able to pinpoint which visual dimension is the most dominant to the extent that our stimulus set does not fully dissociate them. Although the individual trials vary a lot on these dimensions within conditions (see Methods), more so than the average across all items varies between conditions, there was nevertheless an average difference between conditions on several visual dimensions. For example, there were, first, some differences between the dot conditions in terms of the number of black pixels. The percentages of black pixels relative to the total number of pixels on the screen: 6.74% (one dot), 12.03% (two dots), 12.21% (four dots), and 12.48% (eight dots). The percentages were lower for the symbols conditions (“1”: 0.83%; “2”: 1.49%; “4”: 1.40%; “8”: 1.83%), and as such the number of black pixels could be an explanation for our observation that all the symbol conditions were more similar to one dot than to patterns with more dots. However, the percentages were highly similar for two, four, and eight dots, whereas the decoding accuracy between symbols and two-dot patterns was clearly lower than between symbols and eight-dot patterns (this effect is significant in all parietal ROIs, with all $p < .044$). Thus, at a quantitative level, it is unlikely that this particular visual hypothesis regarding the number of black pixels explains the current findings.

“Clutter” is another visual property that could partially explain our results. This property is very difficult to dissociate from the number of objects. The same applies to the complexity of the total display, although this might depend upon how “complexity” is exactly defined and whether it takes into account the complexity of the individual objects (e.g., an Arabic digit is visually more complex than a dot). Although we cannot precisely pinpoint the exact visual dimension that explains our results, we observed very similar results in all our ROIs, all the way down to primary visual cortex in Brodmann’s area 17, which suggests that at least for some areas the explanation for our findings has to be found in relatively simple visual dimensions. Nevertheless, for parietal areas, the hypothesis in terms of the number of objects comes into the picture as a particularly likely candidate, because studies in the object cognition literature as a whole have controlled for quite a number of visual dimensions and have already revealed the importance of this visually based stimulus property for activation in areas around the IPS.

Reconciling the Object Cognition Account with Recent Studies on Numerical Processing

The current results are consistent with the findings of two very recent fMRI studies performed at high field strength (7T). He, Zuo, Chen, and Humphreys (2014) showed that the IPS activity did not differ between dots and Arabic digits when small numerical magnitudes ($<4$) were presented, but that with increasing numerical magnitude ($>6$), the differences between symbolic and non-symbolic formats became more prominent in the IPS. This result can also be expected by the object cognition account: When numerical magnitude increases, the dot patterns contain a larger number of visual units than the Arabic digits; thus, IPS activity between the formats will become less similar because they do not contain the same number of visual units anymore.

Another 7T fMRI study (Harvey, Klein, Petridou, & Dumoulin, 2013) showed a clear topographic representation of numerical magnitude in the human parietal cortex for dot patterns but not for Arabic digits. On the basis of the object cognition account, one would expect that the Arabic digits ($<10$) are mapped onto the one-dot area in the topographic representation of the number of visual elements and are seen as similar. In light of this account, the finding of Harvey et al. (2013) is not surprising because their study showed no significant differences in activation in parietal regions between the Arabic digits. Because most of them contained the same number of visual elements, namely one visual element, the parietal regions would not handle them differently according to the object cognition account.

Conclusion

By integrating two research domains and applying MVPA analyses, we were able to show that there are no overlapping activity patterns between Arabic digits and dots in IPS and any of its subparts. In line with studies on object cognition, which reported that IPS processes the number of objects presented (Xu, 2008; Xu & Chun, 2007; Song & Jiang, 2006; Todd & Marois, 2005; Vogel & Machizawa, 2004), our data suggest that Arabic digits are more related to one dot than to dot patterns with corresponding numerical magnitude. This significant finding contradicts the hypothesis that numbers would be processed in a format-independent manner in the human parietal cortex.

Acknowledgments

This work was supported by the Fund for Scientific Research-Flanders by a fellowship to J. B., an IDO Project of the KU Leuven (IDO/10/003), an FWO project (G.0946.12), a Federal Research Action (IUAP-P7/11), and an ERC grant (ERC-2011-Stg-284101). We thank the reviewers and Nicky Daniels for their helpful comments and suggestions.

Reprint requests should be sent to Hans P. Op de Beeck, Tiensestraat 102, Leuven, Belgium, 3000, or via e-mail: hans.opdebeeck@ppw.kuleuven.be.

REFERENCES


