Can responses to basic non-numerical visual features explain neural numerosity responses?

Ben M. Harvey[a,b,*], Serge O. Dumoulin[b,c]

[a] Faculty of Psychology and Education Sciences, University of Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal
[b] Experimental Psychology, Helmholtz Institute, Utrecht University, Heidelberglaan 1, Utrecht 3584 CS, Netherlands
[c] Spinoza Centre for Neuroimaging, Amsterdam, Netherlands

ABSTRACT

Humans and many animals can distinguish between stimuli that differ in numerosity, the number of objects in a set. Human and macaque parietal lobes contain neurons that respond to changes in stimulus numerosity. However, basic non-numerical visual features can affect neural responses to and perception of numerosity, and visual features often co-vary with numerosity. Therefore, it is debated whether numerosity or co-varying low-level visual features underlie neural and behavioral responses to numerosity. To test the hypothesis that non-numerical visual features underlie neural numerosity responses in a human parietal numerosity map, we analyze responses to a group of numerosity stimulus configurations that have the same numerosity progression but vary considerably in their non-numerical visual features. Using ultra-high-field (7T) fMRI, we measure responses to these stimulus configurations in an area of posterior parietal cortex whose responses are believed to reflect numerosity-selective activity. We describe an fMRI analysis method to distinguish between alternative models of neural response functions, following a population receptive field (pRF) modeling approach. For each stimulus configuration, we first quantify the relationships between numerosity and several non-numerical visual features that have been proposed to underlie performance in numerosity discrimination tasks. We then determine how well responses to these non-numerical visual features predict the observed fMRI responses, and compare this to the predictions of responses to numerosity. We demonstrate that a numerosity response model predicts observed responses more accurately than models of responses to simple non-numerical visual features. As such, neural responses in cognitive processing need not reflect simpler properties of early sensory inputs.

Introduction

Humans and many animals can distinguish between stimuli that differ in numerosity, the number of objects in a set. Human and macaque parietal lobes contain neural populations that respond to changes in stimulus numerosity, and neural responses here predict numerosity discrimination behavior (Piazza et al., 2004; Tudusciuc and Nieder, 2007). However, it is passionately debated whether these behavioral discriminations and neural responses follow numerosity per se or rather rely on low-level visual features that often co-vary with numerosity. One side of this debate proposes that numerosity is sensed directly using numerosity-selective neurons (Anobile et al., 2014; DeWind et al., 2015; Nieder et al., 2002; Park et al., 2016; Ross and Burr, 2010; Tudusciuc and Nieder, 2007), while the other proposes that neural and behavioral responses to numerosity reflect non-numerical visual features that change with numerosity (Clearfield and Mix, 2001; Feigenson et al., 2002; Gebuis and Gevers, 2011; Durgin 2008; Dakin et al., 2011; Morgan et al., 2014; Gebuis et al., 2014).

Convincing arguments and experiments support both viewpoints. Data from our previous fMRI study (Harvey et al., 2013) has been proposed to support both sides of the debate (Anobile et al., 2016; Gebuis et al., 2014), but in each case without quantitative comparisons against the predictions made by the alternative viewpoint. Here we quantitatively and formally test the predictions of alternative models of apparently numerosity-selective responses against this data.

Behavioral studies have proposed that numerosity discrimination relies on: total item perimeter (Clearfield and Mix, 2001), total item luminance, total item surface area (Feigenson et al., 2002), area covered by the set outline (mathematically, the convex hull) (Gebuis and Gevers, 2011), density (Durgin, 2008), or the contrast energy at high spatial frequencies (Dakin et al., 2011; Morgan et al., 2014). Others have argued that patterns of behavioral responses are inconsistent with such non-numerical cues (Anobile et al., 2014; DeWind et al., 2015; Ross and Burr, 2010). It is not possible to design a stimulus configuration where none of these visual features co-varies with numerosity: eliminating covariance of one feature with numer-

osity introduces covariance with another features, which observers could use to make judgments. Similarly, using multiple stimulus configurations with different relationships between visual features and numerosity cannot eliminate potential feature-based strategies because the observer can change strategies between configurations. As such, although observers are typically instructed to distinguish between stimuli based on numerosity, in remains unclear to what extent they can distinguish based on the co-varying low-level visual features listed above.

At a neural level, populations of neurons responding to presentation of specific numerosities are found in macaques and humans (Eger et al., 2009; Harvey et al., 2013; Nieder et al., 2002; Nieder and Miller, 2004; Piazza et al., 2004). Experiments to characterize such responses typically use multiple stimulus configurations to demonstrate responses are robust to changes in visual features. Indeed, a recent event-related potential study of early visual responses that increase monotonically with numerosity (distinct from the tuned numerosity-selective responses we examine here) shows they are effected very little by non-numerical cues that can co-vary with numerosity (Park et al., 2015). Unlike human observers, neurons are unlikely to change which visual feature they respond to when stimulus configurations change. However, preferences of later, tuned numerosity-selective neurons can change between stimulus configurations, so at least some of these neurons may also respond to other features (Harvey et al., 2013; Nieder and Miller, 2004; Tudusciuc and Nieder, 2007). Alternatively, apparent responses to numerosity might reflect responses to a co-varying non-numerical visual feature, with no numerosity responses present (Gebuis et al., 2014). The critical test here is whether models of responses to co-varying visual features predict the measured neural responses better than responses to numerosity.

Here we describe a method to test predictions of alternative models of cognitive neural responses functions from fMRI data. We use this to test the hypothesis that neural responses to numerosity reflect responses to co-varying non-numerical visual features in two ways. First, we quantify these features across the multiple stimulus configurations used to investigate numerosity responses in a recent experiment that shows effects of visual features on neural numerosity responses particularly clearly (Harvey et al., 2013). This quantifies relationships between several visual features and numerosity in these different stimulus configurations, revealing where these features can be used to distinguish between stimuli of differing numerosity. While none of these features is perfectly correlated with numerosity in all stimulus configurations, the effects of stimulus configuration on numerosity tuning raise the possibility that non-numerical features underlie these responses (Gebuis et al., 2014). Second, we test the possibility by comparing the ability of visual feature response model predictions and numerosity response model predictions to explain fMRI responses measured in an area of posterior parietal cortex whose responses to numerosity changes are believed to reflect numerosity-selective activity. We demonstrate that numerosity response models predict observed fMRI responses more accurately than non-numerical visual features response models. As such, human fMRI responses to numerosity do not reflect responses to the proposed low-level, non-numerical visual features.

Methods

All data are taken from (Harvey et al., 2013), where experimental methods for their acquisition are described in full. Here we outline methodological details that aid interpretation of results, and describe new analyses.

Subjects

We present data from eight subjects (two female, age range 19–38 years). Two were left-handed. All were well educated, with good mathematical abilities. All had normal or corrected to normal visual acuity. All were trained with tasks requiring numerosity judgments before scanning. Experiments were undertaken with the informed written consent of each subject. All experimental procedures were cleared by the ethics committee of University Medical Center Utrecht.

Stimuli

Visual stimuli were presented by back-projection onto a 15.0×7.9 cm screen inside the MRI bore. The subject viewed the display through prisms and mirrors, and the total distance from the subject's eyes (in the scanner) to the display screen was 41 cm. Visible display resolution was 1024×538 pixels.

The stimuli were generated in Matlab using the PsychToolbox (Brainard, 1997; Pelli, 1997). A large diagonal cross, composed of thin red lines, crossed the entire display, a design that allows very accurate fixation (Schira et al., 2009). Subjects (n=8) fixated the intersection of the cross. Stimuli consisted of groups of items randomly positioned at each presentation so that each item fell entirely within 0.75° of fixation. As such, contrast energy was distributed equally across the stimulus area for all numerosities, avoiding any links between numerosity and the visual field position of the contrast energy. To prevent perceptual grouping, individual items were distributed roughly homogenously across the stimulus area (except for the high density configuration described below).

We used various stimulus configurations (Fig. 1) (Harvey et al., 2013; Nieder et al., 2002) that ensure that responses to other visual features did not follow the same time course in different configurations. The first stimulus configuration ('constant area') kept the total surface area of all of items combined constant across numerosities, ensuring equal luminance across numerosities. The second ('constant item size') kept individual item size constant. The third ('constant perimeter') kept the total perimeter constant, ensuring equal edge density. We have previously referred to this as the 'constant circumference' configuration: 'perimeter' is a more general term, applying to all shapes, not only circles. The fourth stimulus configuration ('high density') contained the same items as the constant area configuration, but at higher density, with all items falling within a 0.375° radius circle that was randomly placed inside the stimulus area. The fifth stimulus configuration ('variable features') contained the same item sizes as the constant item size (and did not vary item size), but used various shapes instead of the circles used in other configurations.

All patterns were presented as black items on a gray background. Patterns were presented briefly (300 ms) to ensure subjects did not have time to count (Saltzman and Garner, 1948). This was repeated every 750 ms, each time with a new random pattern presented, with 450 ms presentation of a uniform gray background between pattern presentations. This was repeated six times, over 4500 ms, corresponding to three fMRI volume acquisitions (TRs), before the numerosity changed. On 10% of pattern presentations, patterns were shown in white instead of black. Subjects were instructed to press a button when this happened to ensure they were paying attention to the patterns during fMRI acquisition. No numerosity judgments were required. Subjects responded on 90–100% of white presentations within each scanning run.

The numerosities one through seven were shown as the main stimulus, first presented in ascending order, followed by a longer period (13.5 s) where the stimulus contained twenty items, followed by the numerosities in descending order, followed by another long period of twenty items. This sequence was repeated four times in each scanning run. Despite fMRI's lack of a specific baseline, the long period of twenty items allowed us to distinguish between very small tuning width (which never respond to the 1−7 range) and very large tuning widths (which always respond to the 1−7 range) (Dumoulin and Wandell, 2008). While both tuning widths capture responses to the 1−7 range, responses to the twenty item presentation will decrease for a large tuning width but not a small tuning width.
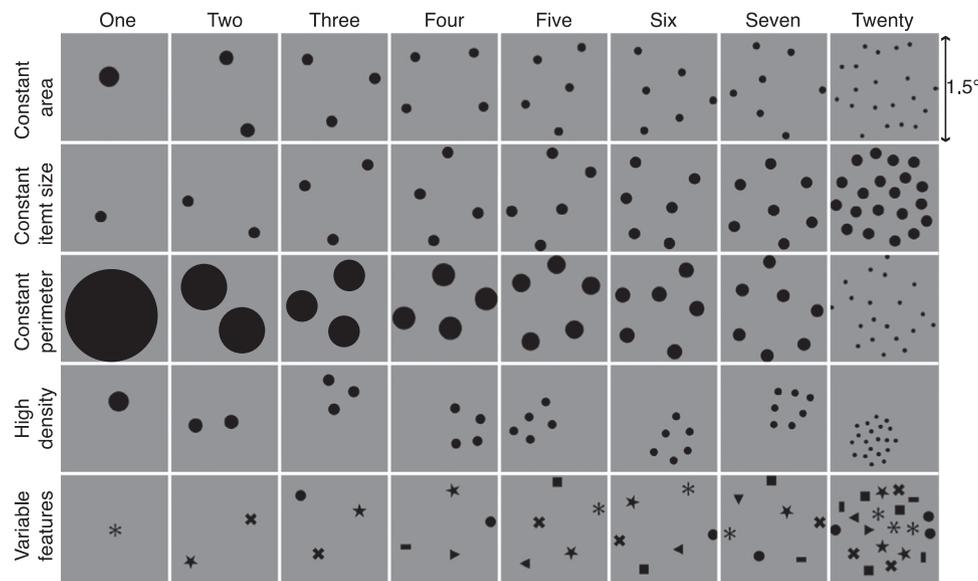
Fig. 1. Example stimuli in for all presented numerosities in all stimulus conditions. For the same numerosity, low-level visual features varied considerably between conditions.

Some non-numerical features do not change much between seven and twenty items. Including or excluding a large change in these features will not affect the variance explained by the model, the criterion we use to compare alternative response models. However, features that are similar between seven and twenty item presentations are unlikely to capture our responses well, because the recorded responses to seven and twenty items are quite different.

Furthermore, during this twenty item presentation, little neural response was expected from numerosity-selective neurons with small preferred numerosities, as such a large numerosity should have been well outside of the numerosity range that elicits strong responses. However, neural populations responding to the contrast energy of the stimulus should respond most strongly during the twenty item presentation, avoiding confusion with populations preferring a specific large numerosity. Finally, the twenty item presentation increases the range of presented numerosities and non-numerical features, and the range of responses. Therefore, this provides extra information to distinguish between response models.

Quantification of visual features

For every stimulus configuration and numerosity, we quantified several visual features of the presented stimuli. Some of these features were linearly proportional to each other. Neural response models were general linear models, so gave identical predictions for linearly proportional visual features.

Many of these visual features (individual item radius, perimeter and area, total item perimeter and area, and display luminance) relied only on item size and numerosity. As item size had a fixed relationship to numerosity in all stimulus configurations, and most stimulus configurations used circular items, we quantified these features geometrically. For the variable features configuration, where various randomly-chosen shapes form the stimulus configuration, we measured the areas and perimeters of these shapes over 1200 presentations of each numerosity and used the means of these distributions to build visual feature response models.

As all items were the same color (black), display luminance was linearly proportional to total item area. Radius was also linearly proportional to perimeter where items are circular, and items that are not circular had no radius. Therefore, to test predictions of responses to the visual features listed above, we only needed to test models of responses to individual item perimeter, individual item area, total item perimeter and total item area.

Other visual features (convex hull area, convex hull perimeter, numerical density, luminance density and edge density) depended on item placement, which was randomized for each presentation. For these features, we measured the distribution across 1200 presentations of each stimulus configuration, and used the mean of this distribution to make neural response models. We defined the convex hull area as the area under the entire pattern, and convex hull perimeter as the length of the line surrounding the entire pattern.

Numerical density, luminance density and edge density described the number of items, total item area, and total perimeter per unit stimulus area respectively. We quantified numerical density, luminance density and edge density within the convex hull: numerical density, luminance density and edge density within the total possible stimulus area were proportional to numerosity, total item area and total perimeter respectively because the total possible stimulus area was always the same. Density within the convex hull was a more parsimonious measure, as this density measure increased when items were grouped more closely. However, we density is not a meaningful measure where a display contains a single object.

We also quantified stimulus contrast. As all stimuli had the same maximum and minimum luminance, they contained no variance in Michelson contrast, so Michelson contrast could not explain any variance in our measured responses. Root mean square (RMS) contrast, however, reflects the standard deviation of luminance values within a particular area. We quantified RMS contrast within both the total stimulus area (display RMS contrast) and within the convex hull (convex hull RMS contrast).

Finally, we examined contrast energy at high spatial frequencies (Dakin et al., 2011; Morgan et al., 2014). Many different metrics could be used to characterize the spatial frequency distribution, and we did not aim to examine all possible visual features. High spatial frequency contrast energy has been used to explain performance in numerosity discrimination tasks (Dakin et al., 2011; Morgan et al., 2014). Which spatial frequencies are involved depends on item size and display size. We quantified contrast energy at frequencies above 4 cycles/deg from Fourier transformations of our stimulus images. This frequency band approximated that used in the cited studies, considering differences in item sizes and display size. We also observed that this frequency band co-varied closely with numerosity in some of our stimulus configurations.

Having quantified these visual features, we examined their relationships to numerosity in each configuration, and compared predictions of models of responses to these features to predictions of models of responses to numerosity.

FMRI acquisition

Functional T2*-weighted 2D echo planar images were acquired on a 7 T scanner using a 32 channel head coil at a resolution of 1.98×1.98×2.00 mm, with a field of view of 190x190×50 mm. TR was 1500 ms, TE was 25 ms, and flip angle was 80°. Functional runs were each 248 time frames (372 s) in duration, of which the first eight time frames (12 s) were discarded to ensure the signal was at steady state. Four repeated runs were acquired within the same session for each stimulus configuration. Responses to different stimulus configurations were recorded on different days.

MRI data analysis

T1-weighted anatomical scans were automatically segmented using Freesurfer (Dale et al., 1999) and then hand-edited to minimize segmentation errors (Teo et al., 1997) using ITK-SNAP (Yushkevich et al., 2006). Functional MRI analysis was performed in the mrVista software package, which is freely available at (http://white.stanford.edu/software/). FMRI data were corrected for subject motion and aligned to anatomical scans (Nestares and Heeger, 2000) in which gray and white matter had been labeled to allow reconstruction of the cortical surface (Wandell et al., 2000). Data from several sessions, resulting from all stimulus configurations, was interpolated into the same anatomical segmentation space, allowing us to compare responses of the same anatomical locations to different stimulus configurations. Data collected on different scanning runs and stimulus cycles in which the same stimulus configuration was presented were averaged together before model fitting, following typical pRF modeling approaches. Areas with overlying draining veins, which spatially and temporally distort the fMRI signal, were identified by their low mean blood oxygenation level dependent (BOLD) signal intensity (Harvey and Dumoulin, 2011; Yacoub and Hu, 2001) and excluded from further analysis.

We evaluated predictions of responses to each candidate visual feature using population receptive field (pRF) modeling, fitting an independent response for each recording site within gray matter. Population receptive field models described the aggregate tuning of the neural population within each fMRI recording site (Fig. 2). A forward model predicted neuronal responses at each stimulus time point depending on the quantity of the feature shown. The model described a tuned response to the candidate feature using a Gaussian characterized by a preferred feature quantity (mean of the Gaussian distribution) and tuning width (standard deviation of the Gaussian). By examining the overlap of the stimulus at each time point with this tuning model, a prediction of the neuronal response time course was generated. By convolving this with a hemodynamic response function (HRF), a predicted fMRI time course was generated. The predicted fMRI time courses were generated for all combinations of a large range of preferred feature quantity and tuning width parameters. For each feature tested (including numerosity), we used 101 candidate values for preferred feature quantity, evenly spaced between the minimum and maximum value of that feature found in all stimulus configurations (the feature range). At each of these candidate preferred feature values, we tested 401 candidate tuning width values, evenly spaced between 0.01 times the feature range and the maximum feature value. This yielded 40,501 candidate combinations of pRF parameters, scaled to the range of feature values presented over all stimulus configurations. For each recording site, the parameters were chosen from the prediction that fit the data most closely by minimizing the sum of squared errors between the predicted and observed fMRI time series (and so maximizing R^2).

Region of interest (ROI) definitions

We used the same regions of interest here as in our initial description of the posterior parietal numerosity map (Harvey et al.,
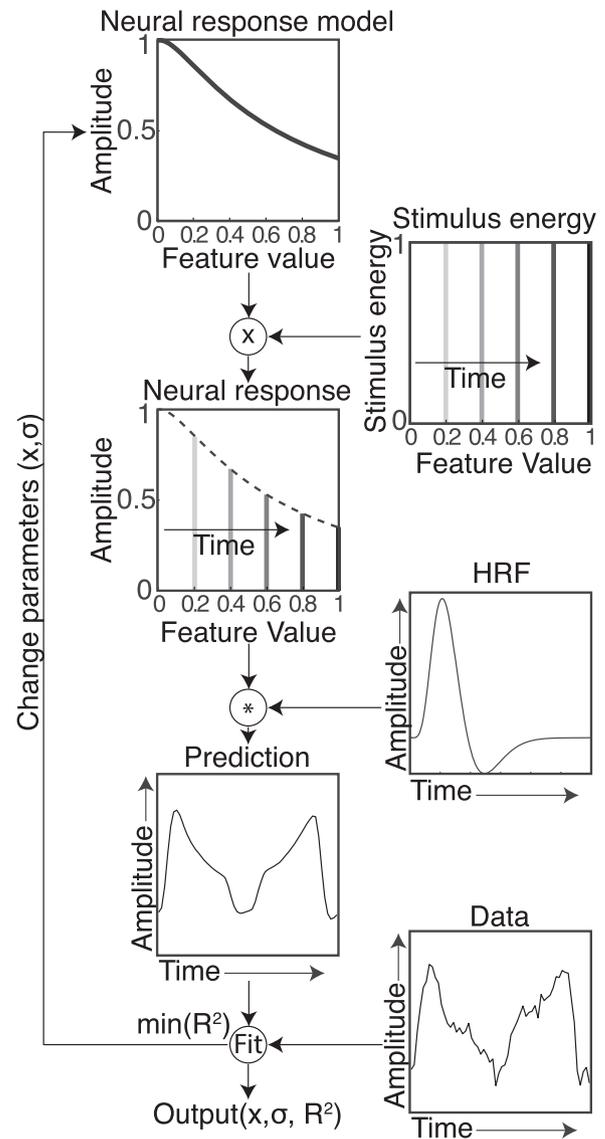
Fig. 2. A flow chart describing the pRF model fitting principle for a single recording point (fMRI voxel). A neural response model describing one possible neural response configuration, with a particular set of parameters describing the preferred feature value and tuning width of the neural population within a voxel. The predicted neural response over time is calculated by multiplying the presented feature value at each time point by the normalized response to that feature in the neural response model (X). This predicted neural response time course is convolved with a hemodynamic response function (*) to produce a predicted fMRI time course, given this set of neural response model parameters. Best fitting model parameters are found by minimizing the difference between the predicted and the recorded data. This process was repeated for every feature response model tested.

2013). We determined which parts of the cortex appeared to respond in a numerosity-selective way in all stimulus configurations using a conjunction analysis by taking the minimum goodness of fit (R^2) of the numerosity tuning model in any stimulus configuration. The resulting minimum R^2 values were rendered onto an inflated cortical surface. This analysis highlighted a consistent region in the posterior parietal lobe in all subjects. This region formed the basis our ROI.

We then rendered the preferred numerosities of each recording site on the cortical surface. The model from the responses averaged across stimulus configurations consistently gave the clearest and largest region of topographic representation, covering the topographically organized regions in all other configurations, so we used this data to define our ROI. Medial and lateral borders of the ROI each followed lines of equal preferred numerosity at the low and high ends of the
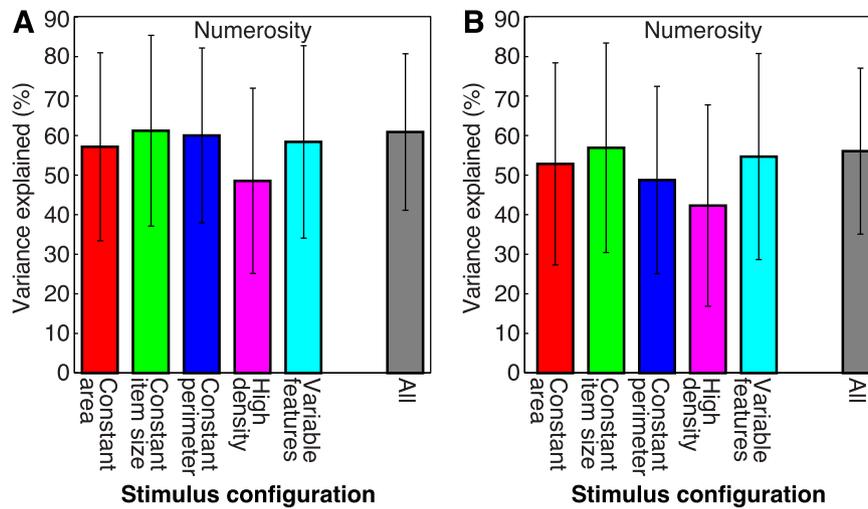
Fig. 3. Variance explained by neural models responding to numerosity. (A) When numerosity model parameters were allowed to differ between models of responses to different stimulus configurations, models explained response variance well for all stimulus configurations. (B) When numerosity model parameters were constrained to use identical parameters to predict responses to all stimulus configurations, these models predicted less variance than models where parameters could vary between stimulus configurations (shown in A). The decrease in variance explained was largest in the constant perimeter configuration where numerosity preferences differed significantly from other configurations (Harvey et al., 2013). Bars show the average variance explained in the numerosity map, and error bars show the standard deviation.

preferred numerosity range seen in each subject. Anterior and posterior borders describe the edges of the topographic organization, which typically also coincided with decreases in the goodness of model fits. All recording sites within these borders were included in our ROI.

We repeated our analyses to capture responses of a larger area that is not selected by its response to numerosity. We used freesurfer to label the right superior parietal lobule from anatomical data alone, and used this as our region of interest. This produced very similar results, although all feature response models fit responses less well averaged over this whole cortical lobule than they did in the numerosity map ROI.

Comparing visual feature response models

This goodness of fit measure (R^2) quantified the amount of variance in each recording site's response that was explained by each model ('variance explained'). We determined the distribution of variance explained by every visual feature model in every stimulus configuration. To quantify the variance explained across all stimulus configurations, we quantified the amount of variance in each recording site's response to all stimulus configurations was explained by each model. To determine whether responses to a particular feature explained measured data better than responses to numerosity, we performed Wilcoxon signed rank tests (non-parametric paired difference test), comparing the variance explained by numerosity models in each recording site to that explained by visual features models at the same sites. This test makes a single comparison over the whole population of recording sites and across all subjects (n=906). We also repeat this test in each individual subject (number of recording sites differs between subjects: S1=107, S2=134, S3=103, S4=160, S5=168, S6=87, S7=46, S8=101). To account for the fact that neighboring fMRI voxels do not represent entirely independent measures of neural activity, we reduce the number of measures in our test to assume an effective independent voxel size of 6 mm. This is highly conservative, as the full-width at half-maximal-height (FWHM) BOLD point spread function in gray matter is estimated to be around 2mm at 7 T (Shmuel et al., 2007).

As neurons were unlikely to change tuning between stimulus configurations, we constrained neural models to use identical preferred feature quantity and tuning width parameters to predict responses to all stimulus configurations (constrained models). As responses to different stimulus configurations were recorded on different days, response amplitudes (GLM betas) could differ between configurations.

In previously presented analyses (Harvey et al., 2013), preferred numerosity and tuning width could differ between stimulus configurations. We also fit models where preferred feature quantity, tuning width and response amplitude could all change between stimulus configurations (configuration-specific models).

Numerosity models described tuned responses as Gaussian functions in logarithmic numerosity space, although Gaussians in linear space predicted responses almost as well. Responses to any feature might follow either its logarithmic or linear progression. Therefore, we always fit visual feature models as Gaussian functions in both logarithmic and linear spaces. When comparing model prediction accuracy, we first determined for each feature whether log or linear models gave the highest average variance explained across all recording sites and stimulus configurations (i.e. which best predicted measured responses). We then used this model for all recording sites.

Results

Common numerosity tuning for all stimulus configurations predicted responses well

Numerosity models predicted measured responses to every stimulus configuration well, explaining 60.9% of measured response variance over all stimulus configurations (Fig. 3A). Using the a numerosity model with identical parameters across all stimulus configurations still explained 56.1% of measured response variance over all stimulus configurations (Fig. 3B). In the constant perimeter configuration, numerosity preferences differed significantly from other stimulus configurations (Harvey et al., 2013), so explained variance fell more here. As such, while numerosity preferences were significantly affected by the stimulus configuration used (Harvey et al., 2013), this had only modest effects on the responses to each stimulus configuration: using the same numerosity model parameters (and predicted response) for all stimulus configurations had only a modest impact on the predictive accuracy of these models.

Feature co-variance and model fits

Quantifying how each feature varied with numerosity revealed relationships that distinguished predictions of neural models responding to numerosity or low-level visual features. Here, we describe these possible relationships and their effects on visual feature model predictions. We use the root mean square (RMS) contrast of the display area
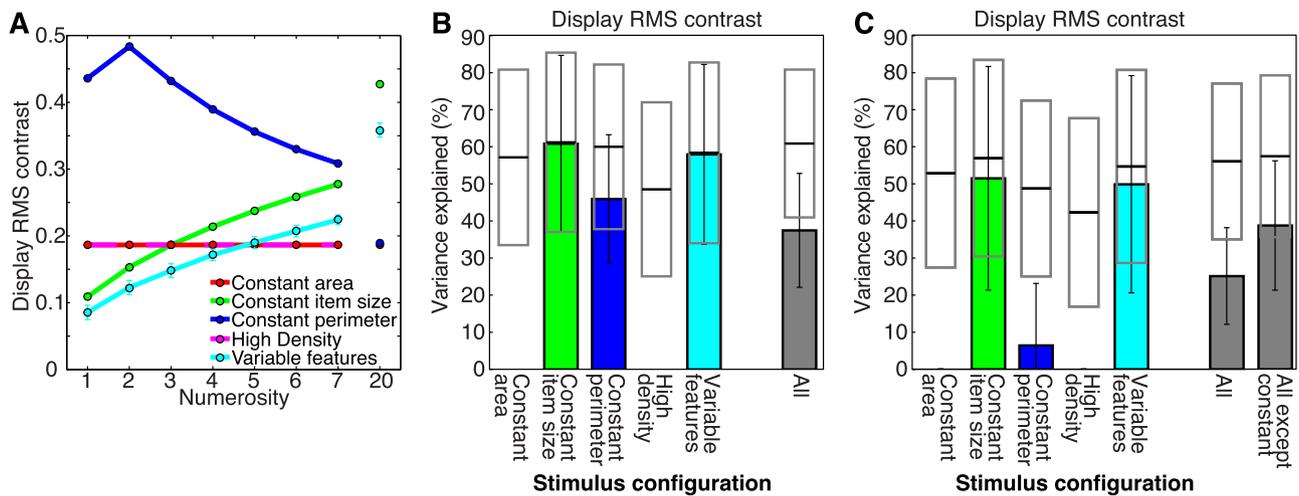
Fig. 4. Patterns of visual feature co-variance with numerosity and the predictive accuracy of visual feature models, for the example of display RMS contrast. (A) Display RMS contrast co-varied approximately linearly with numerosity in some stimulus configurations, varied nonlinearly with numerosity in some stimulus configurations, and did not vary in some stimulus configurations. Colored points, joined by colored lines, show the visual feature quantity for each numerosity in each stimulus configuration. (B) In stimulus configurations where display RMS contrast co-varied approximately linearly with numerosity, the display RMS contrast model predicted responses approximately as well as the numerosity model did (constant item size and variable features stimulus configurations). Where it varied nonlinearly with numerosity, the display RMS contrast model predicted responses with different accuracy than the numerosity model did (constant perimeter stimulus configuration). Where it was constant for all numerosities within a stimulus configuration, the display RMS contrast model could not predict a changing response and so could not explain any response variance (constant area and high density stimulus configurations). However, constant stimulus features may affect the response baseline, and general linear modeling approaches to fMRI cannot capture such effects. Bars show the mean variance explained in responses at many recording sites, and error bars show the standard deviation. Black lines represent the mean response variance explained by numerosity models in each stimulus configuration, and gray boxes represent the standard deviation, taken from Fig. 3. (C) When numerosity models and display RMS contrast models were constrained to use identical parameters to predict responses to all stimulus configurations, these predicted less response variance than separate models fit to individual stimulus configurations (shown in B), particularly when features in different stimulus configurations co-varied with numerosity over different ranges or in different directions, as seen in (A). These models therefore explain less response variance calculated over all stimulus configurations, or over only those stimulus configurations where the display RMS contrast varied (gray bars).
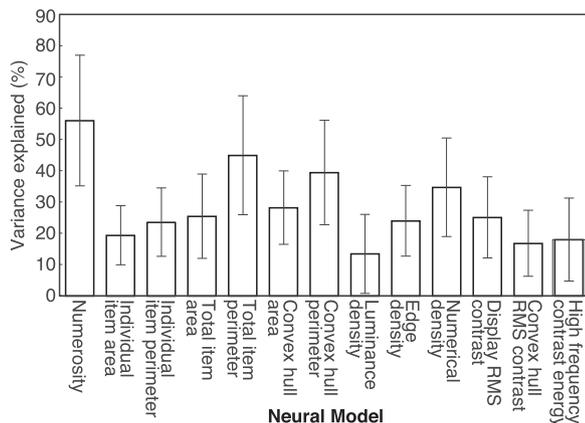
Fig. 5. The numerosity model captured significantly more variance in the data compared to visual feature models. These values the taken from the variance explained by the constrained models across all stimulus configurations (gray bars in Figs. 3B, 4C, and right side panels of Data in Brief article figures).

as an illustrative example (Fig. 4). Similar step-by-step analyses underlie the comparisons made for all stimulus features in Fig. 5. These are presented in full in the accompanying Data in Brief article.

First, where a visual feature did not change with numerosity in a particular stimulus configuration, the feature would not change during that scanning run, so responses to the feature could explain no response variance. Specifically, fMRI responses lack an explicit baseline, so a stationary predictor is added to general linear models of fMRI responses to capture a constant stationary baseline. Adding a further stationary predictor for a constant feature can explain no variance that is not already captured by the first stationary predictor. Therefore, our approach provides no evidence for or against the hypothesis that a constant feature contributes to the baseline, but a constant feature cannot predict a changing response. For example, display RMS contrast did not vary in the constant area or high density stimulus configurations (Fig. 4A) so predicted no response variance here (Fig. 4B).

Second, because we used general linear models here, responses to a visual feature that co-varied linearly with numerosity in a particular stimulus configuration would predict responses to that stimulus configuration exactly as well as numerosity does. This is true regardless of the slope of the covariance, whether negative or positive. Here, it was necessary to use responses to other stimulus configurations to distinguish between visual feature and numerosity models. For example, display RMS contrast varied approximately linearly with numerosity in the constant item size stimulus configuration (Fig. 4A), and responses to RMS contrast therefore predicted response variance approximately as accurately as responses to numerosity did (Fig. 4B).

Third, where a visual feature varied nonlinearly with numerosity in a particular stimulus configuration, the visual feature model would make different predictions from the numerosity model, allowing us to distinguish these models. For example, display RMS contrast decreased nonlinearly with increasing numerosity in the constant perimeter stimulus configuration (Fig. 4A), so the RMS contrast model predicted response variance less accurately than the numerosity model did (Fig. 4B).

We also tested models constrained to use identical parameters to predict responses to all stimulus configurations. In these constrained models, different relationships between that feature and numerosity in different stimulus configurations could distinguish numerosity models from visual feature models, even if the feature co-varied closely with numerosity within each stimulus configuration. Predictions of constrained visual feature models and numerosity models would differ if the feature either: 1) had a range of magnitudes in different stimulus configurations or 2) increased with numerosity in some configurations and decreased in others. For example, display RMS contrast increased with numerosity in the constant item size and variable features stimulus configurations, but decreased with numerosity in the constant perimeter stimulus configuration. Models constrained to use identical parameters to predict these responses explained responses to the constant perimeter stimulus configuration poorly, suggesting responses followed stimulus numerosity rather than display RMS contrast

(Fig. 4C). The range of display RMS contrasts present in the stimulus configuration was also quite different (Fig. 4A). Again, models constrained to use identical parameters predicted responses to the constant perimeter configuration poorly. As this constraint only modestly affected the predictive accuracy of numerosity models (Fig. 3), this again suggested responses followed stimulus numerosity rather than display RMS contrast. Features needed to co-vary similarly with numerosity in all stimulus configurations for visual feature models to consistently predict measured responses as well as numerosity models did. Any visual feature that co-varies differently with numerosity in different stimulus configurations predicted different responses that could be distinguished from numerosity responses.

This detailed analysis is presented in the accompanying Data in Brief article for all visual features.

Numerosity models predicted responses better than models of responses to proposed non-numerical visual features

The critical test of model performance is how well a single constrained model predicted responses to all stimulus configurations. Fig. 5 summarizes this measure of model performance for all tested visual features. When grouping responses of recording sites across all subjects, the numerosity model predicted responses better than any visual feature model, in all cases at $p < 10^{-16}$. The numerosity model also predicted responses better than any visual feature model in every individual subject, in all cases at $p < 0.038$. Visual feature models predicted responses well only in stimulus configurations where visual features co-varied approximately linearly with numerosity, and never predicted responses to any single stimulus configuration better than the numerosity model did (Data in Brief Article Figs. 1–5, Tables 1–5).

The best performing visual feature was total item perimeter, which co-varied approximately linearly with numerosity in all except the constant perimeter stimulus configuration (Data in Brief article Figs. 1J-1L). Responses to this stimulus configuration still showed responses to changing numerosity, though numerosity preferences were significantly different to those from other stimulus configurations (Harvey et al., 2013). Even if we excluded the constant perimeter stimulus configuration from this analysis, the numerosity model still predicted responses better than the total item perimeter model did ($p < 10^{-6}$). As such, the numerosity model predicted responses to numerosity-varying stimuli within the posterior parietal 'numerosity map' (Harvey et al., 2013) better than models of responses to any proposed non-numerical visual feature did.

Discussion

Using fMRI analyses that can distinguish between alternative neural models, we demonstrate that numerosity models predict measured fMRI responses better than models responding to non-numerical visual features that co-vary with numerosity, in a posterior parietal area with apparently numerosity-selective responses. Non-numerical visual features only predict fMRI responses to specific stimulus conditions well when they co-vary with numerosity, supporting the hypothesis that neural responses primarily reflect stimulus numerosity.

Numerosity tuning differences between stimulus conditions

Our stimulus configurations were designed to contain very different relationships between numerosity and low-level, non-numerical visual features (Harvey et al., 2013; Nieder et al., 2002). If responses were identical regardless of visual features, we could conclude the underlying neural populations responded to numerosity only. While responses are similar across stimulus configurations, they differs significantly between stimulus configurations in both the populations within our fMRI recording sites (which each contain approximately 80,000 numerosity selective neurons) (Harvey et al., 2013) and in at

least some single neurons in macaque parietal cortex (Nieder and Miller, 2004).

Two interpretations of variation in numerosity responses with visual features have been proposed. First, they may reflect responses for another visual feature, together with numerosity, in the same neural population (Harvey et al., 2013) or the same single neurons (Tudusciuc and Nieder, 2007). Specifically, tuned responses to object size together with numerosity would explain effects of stimulus condition well, particularly if neural populations responding preferentially to small numerosities responded preferentially to small object sizes. We have recently demonstrated such a pattern of associated numerosity and object size response preferences (Harvey et al., 2015).

Second, effects of visual features on numerosity responses may reflect responses to a low-level, non-numerical visual feature *instead* of numerosity, with no responses to numerosity present in the brain (Gebuis et al., 2014). No experimental evidence has been proposed to support this proposal, beyond behavioral inference effects (de Hevia, 2011; de Hevia and Spelke, 2009; Gebuis and Gevers, 2011; Gebuis et al., 2014; Hurewitz et al., 2006) and the variability of numerosity responses between stimulus configurations (Gebuis et al., 2014; Harvey et al., 2013; Nieder and Miller, 2004). As no specific feature has been proposed, we tested predictions of responses to several non-numerical visual features that have been proposed to underlie numerosity discrimination performance in behavioral experiments.

Our statistical approach groups recording sites throughout a region of interest before comparing the variance explained by different response models. This demonstrates that responses within this region are generally better predicted by stimulus numerosity than non-numerical features. Nevertheless, it is possible that smaller groups of recording sites or neurons within this region respond to non-numerical features, either with or without responding to numerosity (Harvey et al., 2015; Tudusciuc and Nieder, 2007).

Previous evidence for direct numerosity estimation

It has been passionately debated whether numerosity is directly estimated by the visual system or whether behavioral and neural responses to numerosity instead reflect responses to a low-level, non-numerical visual feature that covaries with numerosity. Several recent studies using both behavioral and neural measures are in line with our conclusion that responses are incompatible with responses to low-level visual features alone.

Studies describing numerosity-selective neural responses have consistently used several visual stimulus configurations that vary considerably in their visual appearance (Eger et al., 2009; Harvey et al., 2013; Nieder et al., 2002; Nieder and Miller, 2004; Piazza et al., 2004). Despite two reports of effects of stimulus configuration (Harvey et al., 2013; Nieder and Miller, 2004), these effects are remarkably small for the changes in low-level features and seem likely to reflect responses to other stimulus quantities in the same neural populations (Harvey et al., 2015; Tudusciuc and Nieder, 2007).

More recently, event-related potentials have been found at occipital and occipito-parietal sites that increase in amplitude monotonically with numerosity (Ester et al., 2012; Park et al., 2015). These do not seem likely to reflect tuned, numerosity-selective responses, but instead seem to reflect monotonic summation responses from which numerosity-selective responses are thought to be derived (Dehaene and Changeux, 1993; Verguts and Fias, 2004). Even in these early responses, non-numerical visual features predict little or no response variance (Park et al., 2015).

These neuroimaging and neurophysiological studies (including the current study) do not demonstrate that the neural responses they measure are relevant for behavior. However, there is also extensive recent behavioral evidence for direct estimation of numerosity. Repeated presentation of numerosity patterns leads to systematic adaptation effects on perception of subsequently presented numeros-

ities (Burr and Ross, 2008). If numerosity perception is adaptable, perceived numerosity must be represented in the brain. Some studies have claimed co-varying visual features (primarily texture density) underlie this adaptation effect (Dakin et al., 2011; Durgin, 2008; Morgan et al., 2014). However, this seems unlikely because adaptation to numerosity in one sensory modality transfers to other modalities (Arrighi et al., 2014), and numerosity comparisons are unaffected by density (Ross and Burr, 2010).

Further behavioral evidence for direct estimation of numerosity comes from the demonstration that numerosity can be estimated accurately from second-order stimuli that do not change luminance with numerosity (Kramer et al., 2011). Finally, numerosity discriminations in stimuli that vary randomly in number, object size and spacing are better explained by numerosity than any other feature (DeWind et al., 2015). Together, these findings support the view that numerosity is directly sensed in the brain and perception, although our perception and neural representation of numerosity is linked to that of other stimulus quantities.

Visual feature estimation and combination

A biologically plausible mechanism whose output follows numerosity closely would allow the animal to use this output to make decisions based on numerosity, for example during foraging, which would be selectively advantageous (Harper, 1982; Vallentin and Nieder, 2008). The proposal that numerosity perception or neural responses to numerosity instead reflect responses to a co-varying non-numerical visual feature suggests that the brain cannot accurately and straightforwardly estimate numerosity from analysis of visual inputs. If so, it seems likely that other, more easily quantified features would be used instead.

We quantify many of these candidate non-numerical visual features using simple mathematical operations. Similarly simple mathematical operations can also yield numerosity. While it is uncertain how numerosity is computed in the visual system, it is similarly uncertain how several other proposed non-numerical visual features could be computed. It seems premature to assume neural responses to visual features whose biological estimation is just as unclear, particularly when neurons responding to these features have not been described.

Mathematically, the total area of all items in the set divided by the area of an individual item gives the number of items in the set. If the visual system can quantify a visual feature of the whole set, quantify the same feature in an individual item, and divide one by the other, this would provide a perfect estimate of numerosity. Therefore, if such combinations of visual features can be computed, straightforward combinations do not necessarily support the viewpoint that numerosity responses actually reflect responses to non-numerical visual features. Rather, they may suggest mechanisms that could determine numerosity perfectly. Despite this, recent analyses of early visual event-related potential responses to dot patterns demonstrate that these responses are well predicted by numerosity, and not preceded by responses to individual item area, total item area, individual item perimeter or total item perimeter (Park et al., 2015).

Models of numerosity estimation

However, such mathematical mechanisms are unlikely to describe the brain's numerosity estimation mechanism. Given that many animals can make decisions based on numerosity (Brannon and Terrace, 1998; Brannon et al., 2001; Cantlon and Brannon, 2007; Nieder et al., 2002; Rugani et al., 2009), and that the brain contains neurons whose output follows numerosity closely with little effect of other visual features (Eger et al., 2009; Harvey et al., 2013; Nieder et al., 2002; Nieder and Miller, 2004; Piazza et al., 2004), such a mechanism does seem to exist. As such, a biologically-plausible mechanism whose output follows numerosity closely is required.

Recently, such a mechanism has been proposed (Dakin et al., 2011; Morgan et al., 2014). Following established computational properties of the early visual system, this mechanism decomposes the visual input by spatial frequency. It then uses the outputs of high spatial frequency filters to estimate numerosity. This predicts systematic errors in numerosity estimation that follow errors made in psychophysics experiments. However, we test this mechanism and find it does poorly in explaining responses to some of our stimulus configurations. Specifically, Dakin, Morgan and colleagues use stimulus configurations where item size does not vary with numerosity. We confirm high spatial frequency contrast energy increases in proportion to numerosity in this case. However, where item size decreases with increasing numerosity to keep total item area constant, high spatial frequency contrast energy varies little with numerosity. Where item size decreases more strongly to keep total item perimeter constant, high spatial frequency contrast energy decreases with increasing numerosity. Responses do not follow these changes in high spatial frequency contrast energy between stimulus configurations.

Despite problems with this exceptionally simple model in explaining numerosity responses in some stimulus configurations, it is a biologically plausible mechanism that makes quantitative predictions against which data can be tested. Further development of such models may predict response differences between these stimulus configurations. Testing predictions of such models using the methods we describe here can validate them against neural response data, just as careful psychophysics can test their predictions against perception.

While such models describe how responses that scale with numerosity might arise, emergence neurons with *tuned* responses to the output of such mechanisms rely on further models. One plausible candidate mechanism (Dehaene and Changeux, 1993), which pre-dates the discovery of numerosity-tuned neurons, shows how such tuned responses could arise through the interaction of neurons whose response increases monotonically with numerosity and neurons whose response decreases monotonically with numerosity.

A numerosity estimation mechanism need not perform perfectly to produce behaviorally useful information. It would only need to estimate numerosity fairly accurately in most real world visual input to provide a selective advantage. For example, visual motion is detected by a mechanism that responds to consecutive stimulation of two spatially separated points, known as a Reichardt detector. While experimenters can create stimuli that do contain motion but do not activate the Reichardt detector (and vice versa) the Reichardt detector accurately detects motion in most real world stimuli. As such, it is generally considered a motion detector, with deviations from accurate motion detection revealing its underlying mechanism. Similarly, a numerosity estimation mechanism need only estimate numerosity accurately in most real-world stimuli, but inaccurate estimation may reveal its underlying mechanisms.

Even if numerosity is accurately detected, subsequent modulation of numerosity responses by other visual features certainly can yield inaccurate perception of numerosity. For example, larger objects are systematically perceived to be more numerous (Hurewitz et al., 2006). This finding might suggest that numerosity perception relies on a single visual feature that is affected by both object size and numerosity. Alternatively, neurons responding to object size and numerosity might arise through independent mechanisms and subsequently interact (Harvey et al., 2015), causing interference between neural representations of object size and numerosity.

Numerosity and visual feature range

Here we examine response to stimuli with small numerosities (one through seven), where numerosity can be determined accurately at a glance. For larger numerosities, we cannot judge numerosity accurately without counting, and instead perceive numerosity only approximately. It has been proposed that different mechanisms process different

numerosity ranges (Anobile et al., 2014). Indeed, it has been proposed that numerosities below four and numerosities above four are represented by different systems, the object tracking system and approximate number system respectively (Hyde, 2011; Piazza, 2010). Our stimuli cross this boundary, but we find very few recording sites with numerosity preferences above five, suggesting that the neural numerosity-selective responses we (and others) describe are primarily in the object tracking system. Conversely, experiments investigating approximate enumeration of larger numerosities typically avoid the subitizing range where number is perceived accurately. As accurate numerosity estimation becomes much harder for higher numerosities, performance in approximate numerosity judgments may rely on very different computations, textural cues (Dakin et al., 2011; Morgan et al., 2014) or co-varying visual features, rather than a numerosity estimation mechanism that yields numerosity-tuned responses. This is a matter of some debate (Anobile et al., 2014).

Furthermore, if other visual features vary over a far larger range than numerosity does, tuning for these features may come to dominate response variance (Park et al., 2015). We see some evidence of such effects in the different tuning parameters found in the constant perimeter condition, where individual item area varies by a factor of 49, while numerosity varies by only a factor of 7. We present no evidence that the results described here generalize to larger numerosities, or where other visual features vary over a far larger range than numerosity does. We quantify differences in the range of the various stimulus features in Tables 1–5 of the Data in Brief article.

Similarly, our stimulus was designed to examine responses to numerosity, rather than the other visual features we investigate here. Some other features did not vary systematically and progressively throughout the stimulus sequence in every stimulus configuration (see accompanying Data in Brief article). A more ideal stimulus design might randomly vary all stimulus features and determine which features affected responses by regression (DeWind et al., 2015; Park et al., 2015). However, such designs are less well suited to fMRI: BOLD responses are slow, requiring longer intervals between stimulus presentations to distinguish different responses (Buracas and Boynton, 2002). Thoroughly testing all of these visual features with stimulus sequences that progress linearly through each feature would have required impractical amounts of data collection.

Furthermore, we believe this difference is unlikely to underlie the differences in predictive accuracy of visual features that we report. First, numerosity-selective responses follow the logarithm of numerosity (Harvey et al., 2013; Nieder and Miller, 2003), which changes non-linearly in our stimulus sequences. In most cases other visual features similarly progress systematically but non-linearly through any stimulus configuration. Second, in most other cases the feature does not vary at all. A feature that does not change throughout the stimulus sequence cannot predict a changing response, and the recorded responses clearly change in response to our stimulus sequence. Third, there are a few cases where a visual feature changes non-systematically through the stimulus sequence. Even here, the widely accepted general linear modeling approach to fMRI analysis shows that responses to non-systematic variations should be just as predictable by a pRF model as responses to systematic variations (Boynton et al., 1996; Friston et al., 1998). Researchers typically use systematic stimulus sequences to map pRFs because the responses to these sequences are easier for the researcher and reader to interpret than responses to randomly ordered stimulus sequences, and because pRF parameter estimates are less sensitive to differences in HRF timing parameters when nearby stimulus time points show similar stimuli. As such, most other stimulus features progress similarly to numerosity through our stimulus sequence, and established general linear modeling principles demonstrate that fMRI responses are independent of the sequence used.

Finally, due to random placement of the objects, some stimulus features have different magnitudes on different presentations of the

same stimulus configuration, while numerosity always has the same values. We may miss some trial-by-trial variation in responses that could result from trial-by-trial variations in visual feature magnitude. The vast majority of fMRI experiments similarly average responses to several trials. We predict this average response using the average visual feature magnitude, again following general linear modeling principles.

We examine responses from one area only. Other areas are likely to respond to quite different features. For example, luminance contrast strongly predicts responses in primary visual cortex (Boynton et al., 1996), so the luminance contrast of our stimuli is likely to predict responses more closely than numerosity here.

Conclusions

We describe a statistically powerful method to distinguish between different models that are hypothesized to describe fMRI responses to complex stimulus properties. We demonstrate that fMRI responses to stimuli that vary in numerosity are better predicted by models of responses to numerosity than responses to any non-numerical visual feature that co-varies with numerosity. This approach can be used to compare models of numerosity estimation, or emergence of responses to other complex stimulus features, from quantifiable low-level visual features.

Acknowledgements

References

Anobile, G., Cicchini, G.M., Burr, D.C., 2014. Separate mechanisms for perception of numerosity and density. Psychol. Sci. 25, 265–270.

Anobile, G., Cicchini, G.M., Burr, D.C., 2016. Number as a primary perceptual attribute: a review. Perception 45, 5–31.

Arrighi, R., Togoli, I., Burr, D.C., 2014. A generalized sense of number. Proc. Biol. Sci., 281.

Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. J. Neurosci. 16, 4207–4221.

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vision. 10, 433–436.

Brannon, E.M., Terrace, H.S., 1998. Ordering of the numerosities 1–9 by monkeys. Science 282, 746–749.

Brannon, E.M., Wusthoff, C.J., Gallistel, C.R., Gibbon, J., 2001. Numerical subtraction in the pigeon: evidence for a linear subjective number scale. Psychol. Sci. 12, 238–243.

Buracas, G.T., Boynton, G.M., 2002. Efficient design of event-related fMRI experiments using M-sequences. Neuroimage 16, 801–813.

Burr, D., Ross, J., 2008. A visual sense of number. Curr. Biol. 18, 425–428.

Cantlon, J.F., Brannon, E.M., 2007. Basic math in monkeys and college students. PLoS Biol. 5, e328.

Clearfield, M.W., Mix, K.S., 2001. Amount versus number: infants' use of area and contour length to discriminate small sets. J. Cogn. Dev. 2, 243–260.

Dakin, S.C., Tibber, M.S., Greenwood, J.A., Kingdom, F.A., Morgan, M.J., 2011. A common visual metric for approximate number and density. Proc. Natl. Acad. Sci. USA 108, 19552–19557.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

de Hevia, M.D., 2011. Sensitivity to number: reply to Gebuis and Gevers. Cognition 121, 253–255.

de Hevia, M.D., Spelke, E.S., 2009. Spontaneous mapping of number and space in adults and young children. Cognition 110, 198–207.

Dehaene, S., Changeux, J.P., 1993. Development of elementary numerical abilities - a neuronal model. J. Cogn. Neurosci. 5, 390–407.

DeWind, N.K., Adams, G.K., Platt, M.L., Brannon, E.M., 2015. Modeling the approximate number system to quantify the contribution of visual stimulus features. Cognition 142, 247–265.

Dumoulin, S.O., Wandell, B.A., 2008. Population receptive field estimates in human visual cortex. Neuroimage 39, 647–660.

Durgin, F.H., 2008. Texture density adaptation and visual number revisited. Curr. Biol.

18, (R855-856; author reply R857-R858).

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., Kleinschmidt, A., 2009. Deciphering cortical number coding from human brain activity patterns. Curr. Biol. 19, 1608–1615.

Ester, E.F., Drew, T., Klee, D., Vogel, E.K., Awh, E., 2012. Neural measures reveal a fixed item limit in subitizing. J. Neurosci. 32, 7169–7177.

Feigenson, L., Carey, S., Spelke, E., 2002. Infants' discrimination of number vs. continuous extent. Cogn. Psychol. 44, 33–66.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. Neuroimage 7, 30–40.

Gebuis, T., Gevers, W., 2011. Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). Cognition 121, 248–252.

Gebuis, T., Gevers, W., Cohen Kadosh, R., 2014. Topographic representation of high-level cognition: numerosity or sensory processing? Trends Cogn. Sci. 18, 1–3.

Harper, D.G.C., 1982. Competitive foraging in mallards:"ideal free'ducks. Anim. Behav. 30, 575–584.

Harvey, B.M., Dumoulin, S.O., 2011. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. J. Neurosci. 31, 13604–13612.

Harvey, B.M., Fracasso, A., Petridou, N., Dumoulin, S.O., 2015. Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. Proc. Natl. Acad. Sci. USA 112, 13525–13530.

Harvey, B.M., Klein, B.P., Petridou, N., Dumoulin, S.O., 2013. Topographic representation of numerosity in the human parietal cortex. Science 341, 1123–1126.

Hurewitz, F., Gelman, R., Schnitzer, B., 2006. Sometimes area counts more than number. Proc. Natl. Acad. Sci. USA 103, 19599–19604.

Hyde, D.C., 2011. Two systems of non-symbolic numerical cognition. Front. Hum. Neurosci. 5, 150.

Kramer, P., Di Bono, M.G., Zorzi, M., 2011. Numerosity estimation in visual stimuli in the absence of luminance-based cues. PLoS One 6, e17378.

Morgan, M.J., Raphael, S., Tibber, M.S., Dakin, S.C., 2014. A texture-processing model of the 'visual sense of number'. Proc. R. Soc. B: Biol. Sci. 281, 20141137.

Nestares, O., Heeger, D.J., 2000. Robust multiresolution alignment of MRI brain volumes. Magn. Reson. Med. 43, 705–715.

Nieder, A., Freedman, D.J., Miller, E.K., 2002. Representation of the quantity of visual items in the primate prefrontal cortex. Science 297, 1708–1711.

Nieder, A., Miller, E.K., 2003. Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. Neuron 37, 149–157.

Nieder, A., Miller, E.K., 2004. A parieto-frontal network for visual numerical information in the monkey. Proc. Natl. Acad. Sci. USA 101, 7457–7462.

Park, J., DeWind, N.K., Woldorff, M.G., Brannon, E.M., 2015. Rapid and direct encoding of numerosity in the visual stream. Cereb. Cortex..

Park, J., DeWind, N.K., Woldorff, M.G., Brannon, E.M., 2016. Rapid and Direct Encoding of Numerosity in the Visual Stream. Cereb. Cortex 26, 748–763.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–442.

Piazza, M., 2010. Neurocognitive start-up tools for symbolic number representations. Trends Cogn. Sci. 14, 542–551.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., Dehaene, S., 2004. Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron 44, 547–555.

Ross, J., Burr, D.C., 2010. Vision senses number directly. J. Vis. 10 (10), 11–18.

Rugani, R., Fontanari, L., Simoni, E., Regolin, L., Vallortigara, G., 2009. Arithmetic in newborn chicks. Proc. Biol. Sci 276, 2451–2460.

Saltzman, I.J., Garner, W.R., 1948. Reaction time as a measure of span of attention. J. Psychol. 25, 227–241.

Schira, M.M., Tyler, C.W., Breakspear, M., Spehar, B., 2009. The foveal confluence in human visual cortex. J. Neurosci. 29, 9050–9058.

Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K., Ugurbil, K., 2007. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 T. Neuroimage 35, 539–552.

Teo, P.C., Sapiro, G., Wandell, B.A., 1997. Creating connected representations of cortical gray matter for functional MRI visualization. IEEE Trans. Med. Imaging 16, 852–863.

Tudusciuc, O., Nieder, A., 2007. Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. Proc. Natl. Acad. Sci. USA 104, 14513–14518.

Vallentin, D., Nieder, A., 2008. Behavioral and prefrontal representation of spatial proportions in the monkey. Curr. Biol. 18, 1420–1425.

Verguts, T., Fias, W., 2004. Representation of number in animals and humans: a neural model. J. Cogn. Neurosci. 16, 1493–1504.

Wandell, B.A., Chial, S., Backus, B.T., 2000. Visualization and measurement of the cortical surface. J. Cogn. Neurosci. 12, 739–752.

Yacoub, E., Hu, X., 2001. Detection of the early decrease in fMRI signal in the motor area. Magn. Reson. Med. 45, 184–190.

Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31, 1116–1128.