CHAPTER 1

Neuronal foundations of human numerical representations

E. Eger

INSERM Cognitive Neuroimaging Unit, NeuroSpin Center, CEA DSV/I2BM, Université Paris-Sud, Université Paris-Saclay, Gif/Yvette, France

Corresponding author: Tel.: +33-1-69 08 19 06; Fax: +33-1-69 08 79 73; e-mail address: evelyn.eger@gmail.com

Abstract

The human species has developed complex mathematical skills which likely emerge from a combination of multiple foundational abilities. One of them seems to be a preverbal capacity to extract and manipulate the numerosity of sets of objects which is shared with other species and in humans is thought to be integrated with symbolic knowledge to result in a more abstract representation of numerical concepts. For what concerns the functional neuroanatomy of this capacity, neuropsychology and functional imaging have localized key substrates of numerical processing in parietal and frontal cortex. However, traditional fMRI mapping relying on a simple subtraction approach to compare numerical and nonnumerical conditions is limited to tackle with sufficient precision and detail the issue of the underlying code for number, a question which more easily lends itself to investigation by methods with higher spatial resolution, such as neurophysiology. In recent years, progress has been made through the introduction of approaches sensitive to within-category discrimination in combination with fMRI (adaptation and multivariate pattern recognition), and the present review summarizes what these have revealed so far about the neural coding of individual numbers in the human brain, the format of these representations and parallels between human and monkey neurophysiology findings.

Keywords

Number representation, fMRI, Parietal cortex, Adaptation, Multivariate decoding

1 INTRODUCTION

High-level numerical abilities appear at the heart of many inventions of technologically advanced human societies. It is, therefore, not surprising that a substantial amount of neuroscientific effort is dedicated to understanding what a “number” is for the human brain. Answering this question is made complex in the first place...
by the multiple meanings in which we use the term number: in its most basic sense, “number” refers to a property characterizing any set of concrete objects, such as its cardinality ( numerosity). Humans, nonhuman primates, and many other animals do share the ability to rapidly extract and compare the numerosity of sets of objects in an approximate fashion, and the behavior of both human and nonhuman primates in such tasks is characterized by Weber’s law: the accuracy with which the numerosity of two sets of items can be discriminated depends linearly on their ratio, at least over an intermediate range of (not too small and not too large) numerosities (eg, Cantlon and Brannon, 2006; Piazza et al., 2004). It has been suggested that numerosity is not a mere abstract concept but a perceptual property, since it is subject to adaptation after-effects in a similar way as other visual features, for example, orientation, color, motion (Burr and Ross, 2008).a Numerosity, however, is a more complicated property in the sense that it is not bound to any single input modality or presentation mode, and the way it is extracted by sensory systems is far less understood than it is for the other features mentioned. Interestingly, perceptual adaptation to numerosity can occur across changes in sensory modality (visual, auditory) and presentation mode (simultaneous vs sequential) (Arrighi et al., 2014), suggesting that the neuronal populations coding for it within each modality may be at least intricately connected, if not feeding into a common representation.

The second meaning of the term “number” is an abstract mathematical object referred to by symbols and used to count, measure, or rank virtually everything. Although this might appear quite removed from the perceptual property of numerosity, a lot of evidence has accumulated to show that across the whole lifespan, in humans there exists a profound link between the capacity to enumerate/compare concrete sets and more abstract numerical/mathematical abilities: behavioral performance for distinguishing two symbolic numerals, although usually more precise overall than the one to distinguish two nonsymbolic numerical stimuli, is less precise and more slow for numerical quantities separated by a smaller ratio, suggesting that the system for comparing the numerical magnitude of symbols is inheriting parts of its metric from the processing of nonsymbolic numerical input (Buckley and Gillman, 1974; Dehaene et al., 1990). Interindividual differences in the precision with which numerosity is discriminated can be correlated with, and even longitudinally predictive of children’s success in symbolic skills such as numerical comparison and calculation (eg, Gilmore et al., 2007; Halberda et al., 2008), even though sensitivity to numerosity is not necessarily the only significant predictor and also other visuospatial abilities (eg, sensitivity to orientation) have been found to correlate with mathematical performance (Tibber et al., 2013). In some children suffering from dyscalculia, the capacity to discriminate visual numerosity can be strongly impaired with respect to age and intelligence matched controls (eg, Mazzocco et al., 2011; Mussolin et al., 2010; Piazza et al., 2010), and interestingly, the

aAfter prolonged exposure (adaptation) to a given numerosity, a set of items of smaller numerosity than the one adapted to is perceived as smaller than its actual value and the opposite for a larger one.
impairment seems to be mainly related to situations where other properties of the stimuli such as, for example, size or area covered provide incongruent magnitude information and have to be discarded to extract a rather abstract representation of cardinality (Bugden and Ansari, 2015; Szucs et al., 2013). Training on approximate additions and subtractions of dot numerosities appeared to have positive transfer effects onto performance in symbolic numerical tasks (Park and Brannon, 2013), while reciprocally, learning symbols for number and/or learning to count has been suggested to enhance the precision of visual numerosity discrimination (Piazza et al., 2013). However, other studies did not find a relation between nonsymbolic and symbolic numerical skills (see, eg, De Smedt et al., 2013, for a review), it has been observed that the relation between numerosity discrimination capacities and mathematical skills is weaker than other relations, for example, the one between symbolic comparison and calculation (Schneider et al., 2016), and some developmental studies did not find a relation between nonsymbolic processing capacities and acquisition of numerical symbols (Sasanguie et al., 2014).

Taken together, even though no definitive consensus has been achieved, there is some evidence to suggest that the cognitive systems for processing nonsymbolic numerical input and more abstract (symbolic) numerical concepts may share some common resources. This raises the questions of whether and how in the human brain the representations of nonsymbolic and symbolic numerical information may be linked, and what is the nature of the neuronal code of numerical magnitude. The present review will give an overview of neuroscientific findings related to the underpinnings of numerical representations in humans, with a particular focus on functional imaging methods. Starting by outlining the regions that have emerged as important substrates of numerical processing and placing them into the context of the more general functional neuroanatomy, the review will then focus on what techniques providing enhanced sensitivity to finer-scale brain representations in combination with fMRI have so far revealed about some crucial stages of the representation of individual numerical magnitudes within these key regions.

2 A CORE NUMERICAL REPRESENTATION IN PARIETAL CORTEX

2.1 NUMERICAL PROCESSING AND THE COARSE SCALE FUNCTIONAL NEUROANATOMY OF PARIETAL CORTEX

Long before the introduction of functional brain imaging methods, neuropsychology had already demonstrated that damage to preferentially left-sided parts of the parietal lobe can result in profound deficits in calculation and other tasks requiring to represent and manipulate numerical information (eg, Cipolotti et al., 1991; Dehaene et al., 1998). Since then, the implication of parts of the parietal (and frontal) lobes in different numerical tasks has been studied extensively with fMRI. Synthesizing findings from neuropsychology and early fMRI studies, it has been hypothesized that
central parts of the human intraparietal sulcus (IPS) constitute a key node for the abstract representation of numerical magnitude (Dehaene et al., 2003). Intraparietal cortex is recruited during a wide range of symbolic and nonsymbolic numerical tasks and is one of the most consistently activated regions in a recent metaanalysis of fMRI studies of numerical processing, both for nonarithmetic and arithmetic tasks (Arsalidou and Taylor, 2011), see Fig. 1A.

As part of high-level association cortex, the IPS is endowed with a rather complex functionality beyond the domain of numerical cognition. This includes, for example, spatial and action-related aspects of perception (Culham and Valyear, 2006), multisensory, and sensory-motor integration. Sensory-motor integration is achieved within a series of spatial field maps which are characterized by coding for space by a progression of reference frames (see, eg, Hubbard et al., 2005; Sereno and Huang, 2014). Superior parts of the intraparietal cortex further play a crucial role

![Image](image_url)

**FIG. 1**

Cortical regions important for numerical processing in the human and macaque monkey brain. (A) Overview of regions revealed by a recent metaanalysis of human fMRI studies of numerical processing, separately for nonarithmetic tasks (top) and arithmetic tasks (bottom), in that case color coding separately different types of arithmetic operations. (B) Overview of regions of the macaque monkey brain where different percentages of numerically selective neurons have been found during delayed match-to-sample tasks with visual numerosities. While the similar regions found across the two species suggests a close homology, it is important to bear in mind that rather different kinds of comparisons provided the basis for the different figures: discrimination within dimension (between individual numerosities) in the case of the neurophysiological findings, and in most cases subtractions between numerical and nonnumerical control conditions in the fMRI findings, where controls differed not only in the type of stimulus but also different instrumental processes recruited.

in cognitive functions such as attention, working memory, episodic retrieval, and mental imagery, which are traditionally conceived of and studied as separate entities, but have also been conceptualized in terms of top-down modulation of externally (or internally) evoked representations as a common substrate, these regions are also referred to as part of the “dorsal attention system” (see Lueckmann et al., 2014 for a review). Slightly more lateral parts of the IPS have been implicated in cognitive control functions as part of a so-called multiple-demand system (Duncan, 2010) which has been suggested to be important for controlling subtask assembly in complex goal-directed behavior. These regions have been shown to be modulated by task difficulty across a variety of tasks, for example, spatial and verbal working memory, STROOP and multisource interference tasks in addition to mental arithmetics (Fedorenko et al., 2013).

In the face of this multifaceted functionality of intraparietal cortex, it can be difficult to disentangle whether activations during often complex numerical tasks reflect preferential responses to numerical stimuli as opposed to other processes that might be instrumental to and differ between the numerical and control tasks employed. Some earlier studies have attempted to isolate more precisely responses to numerical stimuli by using tightly matched control conditions: when subjects were presented with numerals, letters, or colors either visually (Arabic digits) or auditorily (spoken words) while instructed to detect prespecified target items for each category (one letter, one numeral, and one color), the IPS was activated for (nontarget) numerals over (nontarget) letters and colors in both input modalities, albeit more weakly than during other tasks requiring more explicit numerical processing (Eger et al., 2003). But preferential activation of the IPS for numerical over tight control conditions in an orthogonal task is not commonly observed: the IPS was found to be activated similarly during “pseudo-calculation” or substitution tasks involving digits and letters where tasks shared equivalent resources in terms of visuospatial processing, exchange, and manipulation of items in working memory (Gruber et al., 2001), or during decisions about which Arabic digit was numerically larger or which letter came earlier in the alphabet, as opposed to a dimming detection task on the same stimuli (Fias et al., 2007). Using nonsymbolic numerical stimuli, the IPS was activated similarly by numerosity comparison and color comparison tasks, and an increase in task difficulty led to an equivalent increase in activation during the numerosity and color tasks (Shuman and Kanwisher, 2004). Also, comparison tasks on nonsymbolic numerical stimuli do share a lot of large-scale activation overlap with comparison tasks on other quantitative dimensions (Dormal and Pesenti, 2009; Dormal et al., 2012; Fias et al., 2003; Pinel et al., 2004), and it is not entirely clear which components of the specific tasks account for the overlap as opposed to the differences.

To summarize, the human IPS is an area of high-level association cortex participating in a wide range of functions. While this region is found activated during a wide range of numerical tasks, fMRI studies relying on simple subtraction methods (testing for coarse scale preferential activations) could not unambiguously assign such activations to the mere stimulus category (numerical material/magnitude)
as opposed to other types of processing differing between numerical and control tasks. Therefore, at least at the macroscopic level, there is no strong evidence for an entirely category-specific substrate of numerical processing. On the other hand, the studies focusing on macroscopic preferential activations did not directly address the coding of numerical information as this term is understood for the purpose of the current review: referring to the discriminability and/or degree of similarity of neuronal signals within dimension, thus between individual numerical stimuli/magnitudes (an approach orthogonal to the one inherent to the studies reviewed so far). This kind of investigation is more easily afforded by methods with higher inherent spatial resolution (for example, neurophysiology).

2.2 FINE-SCALE REPRESENTATION OF NUMERICAL INFORMATION: FINDINGS FROM MACAQUE NEUROPHYSIOLOGY

Neurophysiological recordings have identified single neurons with numerical responses in different subregions of the parietal and prefrontal cortex of the macaque monkey, with different stimuli and paradigms: in posterior parietal cortex (PPC) and lateral prefrontal cortex (PFC), see Fig. 1B, during an active numerical matching task on visual sets of items (Nieder and Miller, 2004), and in area 5 and 2 of superior parietal cortex during the execution of sequences of actions (Sawamura et al., 2002). Small visual numerosities (1–5 items) have been most extensively studied, but see Nieder and Merten (2007) for a study of responses to up to 30 items in PFC. The tuning curves of these neurons resemble bell-shaped functions peaking at a given numerosity and showing reduced firing to other numerosities as a function of numerical distance, indicating an approximate code where discriminability increases with numerical ratio, compatible with Weber’s law which also underlies the animals’ behavioral performance in numerical tasks.

For visual sets of dots, PPC neurons respond with shorter latencies than PFC neurons, suggesting that visual numerosity is initially extracted in the dorsal visual stream, and only later amplified for task purposes in PFC (Nieder and Miller, 2004). Within parietal cortex, such numerical responses during delayed match-to-sample tasks were most concentrated (up to ~20% of the neurons tested responded to the numerosity of dot displays without a significant effect of changes in low-level parameters as overall number of pixels, item size and shape, different levels of spacing) in the ventral intraparietal area (VIP) (see Nieder and Dehaene, 2009). Since in the earlier studies monkeys had received extensive training with numerosity matching, a relevant question was in how far these responses could merely be the result of learning. Recently, numerical responses were confirmed in this region (13% of the neurons, 10% “pure” numerosity selective, without an effect of low-level stimulus factors) and PFC (14% of all neurons, 10% “pure” numerosity selective) in not numerically trained monkeys during a delayed match-to-sample task on colored sets of dots where color instead of numerosity was the task-relevant parameter, and after training, numerical representation became further enhanced in prefrontal, but not parietal cortex (Viswanathan and Nieder, 2015). In addition to VIP, numerical
responses without prior training of the animals were also observed in the lateral intraparietal area (LIP) (Roitman et al., 2007). In that case, ~60% of the neurons responded to numerosity, without a significant effect of low-level properties of the dot sets (matching either the overall number of pixels or item size, and either overall extent or spacing), during a saccade task where numerosity was not explicitly relevant, but indicative of the reward the monkey was going to receive (nevertheless, the neurons response reflected numerosity and not reward status per se). In contrast to the findings from VIP and PFC, where bell-shaped tuning curves had been reported, almost all of the neurons in region LIP either monotonically increased or decreased their response with numerosity.

Very few neurophysiological experiments have been conducted (all of them in trained animals) to test for selectivity to numerosity in VIP and/or prefrontal neurons across stimulus modality and mode. For responses to small numerosities of visual items with either simultaneous or sequential presentation recorded in VIP (Nieder et al., 2006), neurons with numerical preferences across presentation modes were found during the delay period of the match-to-sample task, when the numerosity was held in working memory (19% of all neurons tested) and these were largely distinct from the ones coding for numerosity in a mode-specific way during the initial stimulation. With sequential presentation of dots in either the visual or auditory modality, numerical responses specific for each modality were found in both VIP and PFC (Nieder, 2012). Identical numerical preferences in both modalities during the sample phase occurred in 11% of prefrontal neurons, but only 3% of VIP neurons, while bimodal preferences during the delay period were found in both regions (13% in PFC, 10% in VIP). Monkey PFC thus was the region with most pronounced cross-modal responses. One study has performed electrophysiological recordings in macaques trained to associate small numerosities of dots (1–4) with the corresponding number symbols (Arabic numerals) (Diester and Nieder, 2007). While again neurons with numerical selectivities specific to either format were common in both PPC and PFC, “association neurons” which had similar tuning functions for a given nonsymbolic numerical stimulus and the corresponding symbol were largely restricted to PFC (23% of the neurons in that area), and very rare in PPC (2% of the neurons, just above chance level). Thus, it appears that in the macaque, PFC is the area where associations between numerosities and the corresponding symbols seem to be formed. Since the monkeys in this case were trained for only 2 months, the question remains in how far with longer training and/or higher behavioral proficiency this kind of association could also develop in monkey parietal cortex. On the other hand, qualitative differences are likely to exist between the acquisition and manipulation of symbols in humans and nonhuman primates, and studies in nonhuman primates can only be partly informative on the neural mechanisms by which humans assign numerical meaning to symbols.

To summarize, neurophysiological work has implicated macaque PPC and PFC in the coding of nonsymbolic numerical information. Findings suggest that visual numerosity is first extracted in PPC and later amplified in PFC for task purposes. A few studies that directly compared numerical responses in the same neurons across presentation
modes and input modalities indicate some degree of convergence onto a common mo-
dality and mode invariant representation of numerosity in frontoparietal areas when
these numerosities are the explicitly task-relevant categories. Studying numerical pro-
cessing in humans at a level of spatial precision (also allowing for within-category
discrimination) closer to the one of neurophysiology is highly desirable for a better
understanding of what human numerical representations share, or in which way they
are different from the ones of other animals. However, such studies are made difficult
by the fact that numerically responsive neurons appear highly intermingled within
small parts of intraparietal cortex as shown by neurophysiology, and that due to their
invasive nature the same techniques cannot be applied in humans.

2.3 FINE-SCALE REPRESENTATION OF NUMERICAL INFORMATION:
fMRI IN HUMANS

Many perceptual features are represented in the brain in a distributed and overlap-
ping fashion at a fine scale, and two different techniques have been applied to dis-
entangle fMRI responses to such features in different perceptual and cognitive
domains: fMRI adaptation which is tracking the cortical response to changes along
a given perceptual dimensions, and multivariate pattern recognition which is testing
for differences in fine-scale evoked activity across voxels, as explained in more
detail in Box 1. These two approaches are orthogonal and complementary in the
sense that each one theoretically has advantages in different situations.

Beyond mere discrimination performance, both approaches can be used to study
representational invariance, by changing a selected property of the stimulus between
repeated presentations, or by training and testing a classifier on evoked activity pat-
terns that differ in that selected property (eg, same numerical magnitude, but differ-
ent low-level properties or format).

Applying fMRI adaptation to visual numerosity, after habituating subjects by pre-
senting a constant number of items (16 or 32) while varying associated low-level-
features such as dot size, cumulative area, and spacing (Fig. 2A) in a way that none
of those individual features was predictive of numerical change in a given numerical
deviant trial, a change in numerosity led to a release from adaptation in bilateral intra-
parietal cortex (Fig. 2B), the size of which was related to the ratio between adaptation
and deviant numerosity, in agreement with Weber’s law, and closely reflecting the
behavioral sensitivity to numerical change as measured outside the scanner (Piazza
et al., 2004). Adaptation to numerosity in parietal cortex has been confirmed with
fMRI (and other methods such as event-related potential measurements and near-
infrared spectroscopy) even in children/infants (Cantlon et al., 2006; Hyde et al.,
2010; Izard et al., 2008), suggesting that it is reflective of a very primordial processing
capacity preceding language and explicit learning. Beyond these studies which
attempted to dissociate numerosity from low-level factors by varying the dot size,
cumulative area, etc., a recent demonstration is lending further support to the notion
that adaptation to visual numerosities is reflecting the number of objects instead of
simple lower level properties: when presenting subjects with dot sets that did either
have or not have some of the dots connected by lines (where connecting led to under-
estimation of the total number of dots), fMRI adaptation curves in the IPS were shifted
into the direction reflecting the number of resulting perceived units rather than the orig-
inal dot numerosity (He et al., 2015a). On the other hand, fMRI adaptation has been
observed in frontal and parietal areas even for quantitative proportion stimuli (varying
the ratio between two intermingled sets of dots displayed in different colors, or the ratio
between two lines of different length) (Jacob and Nieder, 2009). This suggests that
adaptation effects in these regions are not restricted to simple numerosities but can ex-
tend to higher order, relative aspects of quantity processing.

**BOX 1 APPROACHES TO DISENTANGLE EFFECTS FROM FINE-SCALE
NEURONAL REPRESENTATIONS WITH fMRI**

The fMRI adaptation technique is making use of the reduced signal which accompanies repeated
presentations of the same or related stimuli, with a signal rebound observed when an unrelated
stimulus is subsequently presented (Grill-Spector and Malach, 2001; Naccache and Dehaene, 2001).
This method is thus relying on a repetition-related memory phenomenon to enable inferences about
neuronal populations with different selectivities coexisting at the subvoxel level. While this is an
elegant and highly successfully used approach, its neural underpinnings still remain not fully
understood, and both neuronal habituation (reduced firing of neurons selective to the repeated
stimulus), and experience-related sharpening (reduced firing of neurons unselective to the repeated
stimulus) have been evoked as explanatory mechanisms (Grill-Spector et al., 2006). Adaptation
effects in fMRI can also depend on factors such as attention, familiarity, and in some cases reflect
perceptual expectations instead of mere stimulus repetition (Summerfield et al., 2008). Finally, using
this technique to make an inference about representational characteristics is only possible to the extent
to which a region shows any repetition-related decrease, and this appears to be less the case in early
sensory than in higher level regions.

Multivariate pattern recognition compares direct evoked activity across multiple voxels between
conditions. This technique cannot resolve intermixed selectivities within a single voxel, but by
considering simultaneously the activity of multiple voxels these methods can accumulate any
potential small biases that different individual voxels may have for one or the other condition to
enable discrimination when individual voxels tested in isolation would not yield reliable results (see,
eg, Haynes and Rees, 2006; Norman et al., 2006). Pattern-based analysis methods come in different
flavors: on the one hand, a machine learning algorithm can be used within a cross-validation
procedure to learn an association between stimulus condition and data and subsequently predict the
condition of left-out data (multivariate decoding), after which prediction accuracies between different
conditions can be compared to the chance level to evaluate significance of the discrimination, or
between different pairs of conditions to infer characteristics of the representational space. On the other
hand, and most useful when facing a large number of experimental conditions, a simple dissimilarity
measure (eg, Euclidean or correlation distance) can be computed between the average activation
patterns evoked by different conditions to obtain a measure of the representational space
(representational similarity analysis—RSA), while not necessarily being able to evaluate significance
for discrimination between individual conditions unless the similarity measure is computed in a cross-
validated fashion from independent parts of data (Kriegeskorte et al., 2008). Not relying on a memory
phenomenon as adaptation, pattern recognition methods provide a more direct means to reveal
characteristics of fine-scale distributed representations. However, this approach is most sensitive as
long as the underlying representation is relatively distributed and at the same time sampled
heterogeneously across individual voxels. This seems to be relatively common for features in early to
mid-level sensory areas, but is less clear for higher levels.
FIG. 2
Findings from fMRI adaptation methods to support the coding of individual numerical stimuli in human intraparietal cortex: (A) when presenting subjects with a stream of habituation stimuli of constant number of dots (16 or 32) but varying associated low-level properties (dot size, density, cumulative area) so that occasionally occurring numerical deviants were novel in numerosity only, such deviant numerosities induced a release from adaptation in bilateral intraparietal cortex (B). The release from adaptation followed the ratio of difference between adaptation and deviant numerosity (consistent with Weber’s law), in a way very similar to the profile of behavioral discrimination obtained for the same stimuli in a same–different judgment task outside the scanner. (C) In a similar paradigm using numerical stimuli in both symbolic and nonsymbolic formats, release from adaptation occurred in human intraparietal cortex as a function of the numerical ratio between deviant and habituation stimulus, irrespective of format, suggesting an abstract-semantic level of numerical representation.

Regarding the cortical representation of symbolic numerical stimuli, using both nonsymbolic (visual sets of dots) and symbolic (Arabic digits) stimuli, Piazza et al. (2007) found that within both formats, intraparietal and frontal cortices responded more to numerically far deviant stimuli than to numerically close ones, indicating a similar quantitative metric which characterizes the response to change in both formats (Fig. 2C). Later studies found distance-dependent recovery from adaptation for Arabic digits only in the left intraparietal cortex (Holloway et al., 2012; Notebaert et al., 2011). This ratio-dependent adaptation increases with age (6–12 years), presumably reflecting the extent to which the representation of the meaning of these numerals became sharpened with experience (Vogel et al., 2014). Numerical adaptation effects have also been observed to generalize across different symbolic notations: using Arabic digits or written number words in a masked priming paradigm, reduced activation in bilateral parietal cortex was found when prime and target were of the same as opposed to different numerical magnitude, across changes in symbolic notation (Naccache and Dehaene, 2001). Again, in some cases such adaptation effects across different symbolic notations were only detected in left parietal cortex (Cohen Kadosh et al., 2007). Although mere comparisons of same vs different numerical magnitude in adaptation paradigms support some common representation across different symbolic notations, further evidence for a magnitude basis of these transfer effects was later obtained by studies that manipulated the numerical distance (Notebaert et al., 2010): release from adaptation in bilateral parietal cortex followed numerical distance across changes in symbolic notation (Arabic digits and number words). Adaptation across changes in format can also be found when using both symbolic (digits) and nonsymbolic (sets of dots) stimuli (Piazza et al., 2007): even when numerosity deviants were presented after adaptation to digits, or digit deviants after adaptation to numerosities, the release from adaptation in both frontal and intraparietal areas was related to numerical distance (Fig. 2C), suggesting that a representation of numerical magnitude is commonly accessed by numerosities and symbolic numerical stimuli. However, another study found that a format change (e.g., from dots to digits) without accompanying change in numerical magnitude also did lead to a release from adaptation in the IPS, and even to a larger degree than a change in numerical magnitude within format (Cohen Kadosh et al., 2011), compatible with multiple mechanisms contributing to numerical representation in these regions.

Multivariate pattern recognition was introduced slightly later to test for numerical information in direct evoked activity patterns (instead of repetition-related changes). Support vector machine classification was applied during a delayed visual numerosity comparison task to discriminate patterns evoked by different sample numerosities that the subjects were seeing and holding in mind (Eger et al., 2009). A multivariate searchlight analysis scanning the whole volume for differences in local activation patterns showed that individual numerosities could be most significantly discriminated in bilateral intraparietal cortex. Classification of individual numerosities based on activation patterns in intraparietal cortex generalized across changes in low-level stimulus properties (overall number of pixels or dot size equated between numerosities in different stimulus sets), see Fig. 3A. When
FIG. 3
Findings from fMRI pattern recognition methods revealing distinct multivoxel response patterns for individual numerical stimuli in human intraparietal cortex: during a delayed comparison paradigm, subjects were seeing and holding in mind a given numerical sample stimulus (Eger et al., 2009). In experiment 1 (A), dot numerosities (4–32) were presented with either equated dot size or cumulative area. Based on an intraparietal ROI comprising the most activated voxels (across all stimuli vs baseline) in each subject, numerosities could be successfully discriminated within and across the different stimulus sets, indicating invariance to these low-level factors. In experiment 2 (B), numerical magnitudes 2–8 were either presented in symbolic or nonsymbolic format. A classifier trained on data from numbers of dots which was highly accurate when tested on dots themselves, yielded chance performance when tested on digit evoked patterns. Still, the digit-trained classifier, which had overall much more modest prediction accuracy, completely generalized its performance to numbers of dots, suggesting that format-specific and format-invariant components coexist in the complete activation pattern, but could not yet be further (Continued)
focusing, using specific neurophysiologically motivated localizer paradigms, on the intraparietal subregions functionally equivalent to those (areas LIP and VIP) were numerosity-selective neurons have been observed by neurophysiology (Fig. 3C, left), information discriminative between individual numerosities was present in both regions in humans, and generalized across the spatial location of the stimuli (Eger et al., 2015). Multivariate decoding also provided evidence for a graded nature (quantitative metric) of the numerical representation in intraparietal cortex: numerical distance effects on the classification accuracies for sample numerosities were observed for both small (Eger et al., 2009) and larger numerosities (Eger et al., 2015), see Fig. 3C, right. These findings related to numerical distance confirm a quantitative metric of the code under orthogonal task conditions during mere viewing and holding in mind of a given numerosity, where activation differences cannot be explained by decision difficulty as during comparison, or be secondary to the degree of perceived change between consecutive stimuli as possible during adaptation. Discriminable multivoxel activation patterns and/or distance effects for visual numerosities have been recently confirmed multiple times by other studies using slightly different task contexts: comparison (Bulthé et al., 2014, 2015), matching (Lyons et al., 2015), or simple viewing (Damarla and Just, 2013).

In how far the differences between response patterns found for the numerosity of simultaneously visually presented sets dots would also generalize to other presentation modes (sequential vs simultaneous) or input modalities (auditory vs visual) is a remaining question. One study investigated responses to serially presented numerosity (between 5 and 16 dots) in either the visual or the auditory modality (Cavdaroglu et al., 2015), and failed to find discrimination between sample numerosities that were sequentially presented. This could imply a real difference in the degree to which individual sequential (nonspatial) vs simultaneous (spatial) numerosities are represented in the areas in question or reflect a limitation in sensitivity due to the use of numerosities separated by a smaller ratio than in previous studies. Interestingly, another very recent study investigating cross-modal numerosity responses was able to find generalization between activation patterns for small numerosities presented in either the visual (1, 3, or 5 dots) or the auditory (1, 3, or 5 tones) domain, under conditions of presentation which allowed for counting (Damarla et al., 2016).

**FIG. 3—CONT’D**

Disentangled in this case (right). (C) When focusing with neurophysiologically motivated localizer scans specifically on the subregions functionally equivalent to those where numerical responses have been observed in macaques (LIP and VIP), both regions were found to encode information on individual nonsymbolic numerosities (8–34 dots) in humans (Eger et al., 2015). The functional equivalent of area LIP showed a more pronounced effect of numerical distance, compatible with a coarser representation of numerosity, and speculatively, a summation code. A hypothesis which has not yet received explicit experimental confirmation is that a format invariant representation of numerical magnitude would arise in area VIP or a later stage.
Intraparietal activation patterns for Arabic digits could also be discriminated above chance in the study of Eger et al. (2009), see Fig. 3B. However, performance of the classifier for Arabic numerals was considerably lower than the one for nonsymbolic numerical stimuli and did not show a significant numerical distance effect. The finding of a weaker numerical distance effect for symbolic than nonsymbolic stimuli was replicated by another study (Lyons et al., 2015) using correlation-based representational similarity analysis and hypothesized to reflect a sharper representation of symbolic numbers. However, some caution needs to be exercised with this interpretation as long as digit-related patterns cannot be discriminated with equal or better accuracy than patterns evoked by numerosities, since the absence of the numerical distance effect could simply reflect weaker/more noisy activation patterns in the case of digits. Interestingly, when training and testing a multivariate classifier across symbolic and nonsymbolic formats, a classifier trained on dot numerosities (which had been highly accurate for discrimination of dot numerosities themselves) yielded chance performance when tested on Arabic digit-related activity patterns (Eger et al., 2009). Nevertheless, the classifier trained to discriminate between digits completely generalized its performance to the corresponding dot numerosities (Fig. 3B). The fact that generalization was unidirectional, and that furthermore for the same given numerical magnitude, the two formats could be clearly discriminated, suggests that the complete pattern within intraparietal cortex does not reflect a single, or entirely abstract representation. Nevertheless, the generalization from symbolic to nonsymbolic stimuli suggests that the existence of a format-invariant component which might coexist with a format-specific representation of numerosity in a way that the methods did not allow to distinguish yet, either in closely neighboring subregions, or even within the same area. Others studies have more recently failed to replicate generalization of the evoked activation patterns from Arabic digits to dot numerosities (Bulthé et al., 2014, 2015), and therefore, concluded that the parietal representation is format specific and reflecting the number of objects rather than abstract numerical magnitude. In these studies the subjects carried out a comparison task at the appearance of each stimulus, rather than separating the sample stimuli from the comparison process as in the earlier study (Eger et al., 2009). Neurophysiological studies have described parietal neuronal responses generalizing across presentation modes (although not format) mainly during the delay period of a working memory task and not during the sample phase where responses were specific to the mode of presentation (Nieder et al., 2006). One could, therefore, hypothesize that fMRI activity in human parietal cortex might also reflect a combination of format-specific and invariant components, and that the latter one (which may reflect the final extracted magnitude) could be hard to detect especially in situations of a direct comparison/response.

To summarize, fMRI adaptation and pattern recognition methods in humans have established a close parallel between human and monkey intraparietal cortex for what concerns the coding (or within-category discrimination) of visual numerosities. Although a critical role for human intraparietal cortex in abstract representation of numerical magnitude had already been hypothesized early on the basis of neuropsychological and pioneering imaging findings, the degree of format
invariance of the representations in this region has remained an issue of controversies with the introduction of new techniques sensitive to within-category discrimination. Both fMRI adaptation and multivariate decoding studies found that intraparietal cortex is sensitive to both numerical magnitude and input format (symbolic vs nonsymbolic). While with adaptation generalization across formats (symbolic vs nonsymbolic, or different symbolic formats) and a numerical distance-dependent metric of the effect for numerical symbols have been observed multiple times, distance effects in the symbolic format could not yet be detected with pattern recognition, and some decoding studies have failed to find generalization of numerical information across formats (symbolic vs nonsymbolic). Beyond the issue already mentioned earlier of fMRI adaptation likely being more sensitive to neuronal representations intermingled at the finest spatial scale (ie, the subvoxel level) in contrast to multivariate decoding which can be predicted to have best performance when the evoked activity patterns are relatively distributed across many voxels (also see Drucker and Aguirre, 2009), it is possible that adaptation studies, relying on a memory phenomenon between sequentially presented stimuli, are more influenced by semantic representations (the abstract quantitative meaning extracted from either the dot set or the symbol) instead of mere perceptual/stimulus-evoked activity. Any potential format-invariant neuronal populations, which can only arise as a result of associative learning at the endpoint of two separate processing pathways for symbolic and for nonsymbolic numerical stimuli, might be sparse, and therefore, not necessarily sampled in a way which is easily detectable by fMRI pattern recognition. Nevertheless, at least some converging evidence from both methods is compatible with a hierarchical model where after initial format-specific stages, the processing culminates within parts of human intraparietal cortex in some neuronal populations coding for different numerical magnitudes in a way accessible across formats, which may correspond to the neuronal mechanisms by which number symbols acquire their meaning. It remains to be confirmed which are the precise intraparietal subregions implementing such a format-invariant stage. One hypothesis is that a format-invariant code for numerical magnitude arises at the level where numerical values are represented by a place coding scheme (Verguts and Fias, 2004), thus in the equivalent of macaque area VIP or a later region. It remains to be understood in more detail under which task conditions, and via which earlier (format-specific) neuronal computations this format-invariant (abstract-semantic) level of numerical representation can be reached.

3 THE EXTRACTION OF NUMERICAL INFORMATION: FORMAT-SPECIFIC CONTRIBUTIONS WITHIN AND BEYOND PARietAL CORTEX

3.1 THE EXTRACTION OF NUMEROSITY FROM CONCRETE SETS OF OBJECTS

The way in which the cardinality of concrete sets of object is extracted from sensory signals in the brain is not entirely understood and different theoretical/computational models have been proposed to account for this capacity. Roughly, there are two types
of models, according to which numerosity is either quantified directly on the basis of segmented perceptual units (Dehaene and Changeux, 1993; Sengupta et al., 2014; Stoianov and Zorzi, 2012; Verguts and Fias, 2004), or indirectly on the basis of a summary statistics of low-level visual features (combination of spatial filters, potentially similar to the estimation of texture density) (Dakin et al., 2011). It has been suggested that both object-based and texture-based processing mechanisms do contribute when performing numerosity tasks, as a function of which one is more performant with the particular task or stimuli at hand (Anobile et al., 2014). It is becoming more and more evident that also other low-level quantitative properties exert an influence on numerosity discrimination performance, which can thus depend on the precise way the stimuli are defined in an individual study (eg, DeWind et al., 2015; Gebuis and Reynvoet, 2012; Hurewitz et al., 2006). The question of whether this indeed speaks against dedicated extraction mechanism for numerosity per se, or rather suggests that subjects are combining information from different “channels” at the level of comparative decisions, warrants further research.

Some of the computational models proposed for the extraction of numerosity include as an important component monotonically responsive units through which segmented objects are accumulated. A potential correlate of this mechanism has been observed in the already earlier mentioned numerical responses of area LIP of the macaque monkey (Roitman et al., 2007), where approximately equal proportions of neurons either monotonically increase or decrease their firing rate with numerosity. Parametric increases of overall activation level with increasing numerosity have also been observed for small sets of items in human superior parietal cortex (He et al., 2015b; Santens et al., 2010), however, it remains to be explained how such large-scale increases in the BOLD signal, which tend to level off for larger numerosities (beyond ~8 items) (Eger et al., 2015) could result from similar mechanisms as the monkey findings, where firing rates of intermingled individual neurons either increase or decrease over a rather wide range of numerosities tested (2–32 dots). In this context it is of interest that in the human equivalent of area LIP, the discrimination of individual numerosities showed a more pronounced numerical distance effect than in area VIP (Eger et al., 2015), see Fig. 3C. The more pronounced distance effect in decoding could be compatible with an underlying summation code (for which the distinctiveness of activation patterns can be hypothesized to increase with numerical distance without leveling off as expected for neurons with bell-shaped tuning when their tuning curves do no longer overlap). However, it is currently not possible to disambiguate this possibility from the one of a broader tuning or a different spatial layout of numerical preferences on the cortex.

Area LIP has also been proposed as the neuronal correlate of a saliency or priority map (Koch and Ullman, 1985). Computational models of such architectures in the form of artificial neural networks are composed of multiple nodes exhibiting both self-excitation and mutual inhibition (eg, Itti and Koch, 2001; Roggeman et al., 2010). One critical parameter is the amount of mutual inhibition: with higher inhibition, individual items/locations can be represented with a high precision at the cost of being restricted to very few of them (lower capacity limit), but as the inhibition
decreases, more items/locations can be represented, albeit with less precise representation of their associated features. The idea that a saliency map architecture in area LIP could represent multiple visual items and thus underlie both the extraction of numerosity and multiple objects’ features tracking (as in visual working memory tasks) was tested recently (Knops et al., 2014): when performing alternatively an enumeration, or a visual short-term memory task for orientation, on between 1 and 6 presented Gabor stimuli (Fig. 4A), an identical set of voxels in the human equivalent of area LIP increased and then reached a plateau of its overall activation profile for lower numerosities during the working memory than during the enumeration task (Fig. 4B), which reflected the differential behavioral set size limits in the two tasks. These overall activation profiles could be explained by a salience map model, using a high inhibition setting for the short-term memory task, and a lower inhibition setting for the enumeration task. Finally, multivariate pattern recognition was applied to test for different capacity limits in information encoding during the different tasks: during enumeration, the number of items could be discriminated across the complete range (1–6 items), and discrimination performance showed the typical effect of numerical distance. During the working memory task, however, only the lowest numbers of items could be precisely decoded, compatible with a lower capacity limit. Saliency maps thus provide a biologically plausible mechanism for the extraction of numerosity from at least small sets of items, as tested in that study. Interestingly, it has been recently described that within the range of up to ~4 items where enumeration is typically precise and perceived as effortless, a phenomenon traditionally referred to as subitizing (see, eg, Trick and Pylyshyn, 1994), accurate performance does actually depend on attention (Burr et al., 2010), suggesting that additional mechanism to those operating across the whole numerical range may be at play for small numerosities. These additional resources could be provided by the saliency map representation, or on the other hand, the saliency map could constitute a general processing step underlying also the extraction of larger numerosities with even more reduced levels of lateral inhibition (Roggeman et al., 2010; Sengupta et al., 2014).

Beyond modulations of overall response profiles by small numbers of items and discrimination of individual numerosities on the basis of multivoxel response patterns, when presenting subjects with small numerosities in a gradually increasing and decreasing fashion and applying advanced encoding models to estimate the selectivity of individual voxels, an orderly spatial layout of responses to small numbers of visual items could recently be revealed in a superior parietal lobule area, more pronounced in the right hemisphere (Harvey et al., 2013), see Fig. 4C and D. This layout was very similar across multiple stimulus sets (with constant dot size, area, or circumference, across high- and low-density conditions, and with circular items only or variable shape). The area in question, although not predefined by a functional localizer in this case, corresponds well in terms of its average coordinates to the human equivalent of area LIP targeted in the other studies, it did respond much less to larger numerosities (20 dots), and did not show any layout for symbolic numerical stimuli. Beyond these parallels in terms of the cortical location, and the
FIG. 4
See legend on opposite page.
responsiveness to nonsymbolic numerical stimuli only, it remains to be understood in how far such findings of spatial layouts can be reconciled with summation coding schemes and/or saliency map architectures, or in how far they arise from different neuronal/computational mechanisms in the same or a nearby area.

### 3.2 THE EXTRACTION OF NUMBER FROM SYMBOLS

The triple-code model of numerical representation (Dehaene and Cohen, 1995) had first proposed an important node for the processing of Arabic numerals in ventral visual cortex, corresponding to a stage specialized in the (presemantic) processing of Arabic digits, where their shapes are identified, but not yet associated to their meaning. The ventral stream of primate visual cortex plays a central role in object recognition, with some subregions responding preferentially to certain object categories, not only natural categories such as faces, bodies, and scenes but also categories without a long evolutionary history such as written words (see Op de Beeck et al., 2008). For Arabic digits, such macroscopic functional specialization (detectable by fMRI) is not consistently observed. However, early electrophysiological recordings

**FIG. 4**

Specific findings concerning the extraction of numerosity from concrete sets of objects. (A) In an fMRI experiment where subjects were processing multiple stimuli (1–6 Gabors) but were either asked to merely enumerate them or to keep in short-term memory the orientation of the Gabors. (B) A common set of voxels in functionally defined area LIP was differentially modulated by numerosity across tasks (left), and these activation profiles could be explained by a saliency map model, using different amounts of lateral inhibition. Confusion matrices from multivariate decoding of activation patterns evoked by the different numerosities (right) indicate that while in the enumeration task the number of items could be discriminated across the entire range, during the visual working memory task only the lowest numbers of items could be accurately discriminated, indicating a lower capacity limit. These results are compatible with the notion of a saliency map architecture (with different amounts of lateral inhibition, leading to different capacity limits) underlying both visual object working memory and enumeration in area LIP. (C) When scanning subjects using high-field fMRI during passive viewing of dot patterns that increase and decrease in numerosity (1–7 dots) over time and applying advanced encoding models to estimate the selectivity of individual voxels, a small region in the posterior superior parietal cortex was found to show an orderly layout of responses to small numerosities, forming a continuum from medial to lateral. (D) Exemplar time courses of two voxels: top, a voxel preferring a single dot, bottom, a voxel preferring seven dots. Both voxels were deactivated for larger numerosities (20 dots presented as baseline), and no spatial layout was observed in that region for responses to symbolic numbers.

from the surface of inferotemporal cortex in epilepsy patients had reported some scarce and variably located recording sites with preferential responses to Arabic digits over a few other object categories such as faces, words, and false fonts (Allison et al., 1994). In addition, a more recent intracranial recording study comparing the processing of Arabic digits to control conditions well matched in terms of visual (letters, false fonts), semantic (number words), or phonological (phonologically similar nonnumber words) factors, reported a region in the inferotemporal gyrus which responded most strongly to digits (Shum et al., 2013). Beyond some responsiveness to simple viewing of digits, this region, originally termed visual numeral area, considerably increases its response to symbolic numerical stimuli when presented in the context of a mental arithmetic task, and that equally for Arabic digits and written number words (Hermes et al., 2015). Thus, it appears that the function of the ventral numeral responsive region goes beyond perceptual identification of the shape of Arabic digits, being most strongly recruited when some kind of explicit numerical processing of a given set of stimuli is occurring. The weaker responsiveness of the inferotemporal area during mere viewing of numerals (likely in combination with other factors such as signal loss near the areas in question) might explain why fMRI studies comparing ventral stream activation for Arabic digits to well-matched controls (letters or letter strings) during mere perceptual tasks, in many cases found no preferential activation for digits (eg, Eger et al., 2003; Polk et al., 2002; Price and Ansari, 2011). However, Park et al. (2012) described a positive finding in a slightly different location (lateral occipital cortex), and Pinel et al. (1999) as well as Abboud et al. (2015) described preferential fMRI activations for digits in ventral or lateral inferotemporal regions during tasks requiring some explicit processing of quantitative meaning. Interestingly, another recent study was able to detect preferential activations for digits over letters and meaningless characters in the bilateral inferotemporal gyri during a one-back task, while explicitly minimizing signal drop-out by using high spatial resolution for the acquisition (Grotheer et al., 2015).

While as a consequence of the influential triple-code model the ventral visual cortex has received most attention regarding its role in the processing of numerical symbols, another region which has been suggested to be important for processing of Arabic digits is the angular gyrus or temporoparietal junction. This region was found to be activated more strongly when performing number comparisons on Arabic digits than on concrete sets of dots, especially for small quantities (He et al., 2015b; Holloway et al., 2010). Since for equal numerical distances, comparison tasks on digits are likely easier than on numerosities of dots, the role of this factor needs to be considered, since these effects arise in the vicinity of areas being part of the default mode network (Buckner et al., 2008). However, a preferential activation of the angular gyrus for digits was also observed when subjects were merely viewing digits, letters, or scrambled symbols during an orthogonal color change detection task (Price and Ansari, 2011). The suggested interpretation of this finding (reflecting conceptual identification) seems in line with the fact that the angular gyrus is one of the regions that have been shown to be more strongly activated for conceptual
(magnitude, time of year, or day) as opposed to perceptual (color) judgments on Arabic numerals (Cappelletti et al., 2010).

To summarize, the exact cortical processing stages via which numerical information can be extracted from format-specific sensory input are only beginning to be understood: for visual numerosity (at least with simultaneous presentation), a likely important step involves a salience map architecture in the equivalent of area LIP, which is not exclusively dedicated to numerical processing but can also underlie, among other things, more specific feature tracking of objects as in visuo-spatial working memory. This component and/or a summation coding mechanisms, which has been demonstrated by neurophysiology in the same area, could explain the only partial or absent generalization across numerical formats found in fMRI studies investigating the coding of individual numerical values when focusing on rather large parts of intraparietal cortex. Two different regions, one in the ventral object-processing pathway and one in the angular gyrus part of the parietal lobe, have been proposed as possible substrates of format-specific processing of symbolic numerical stimuli (Arabic digits). We still miss studies relating their location to the more general functional neuroanatomy of the areas in question (eg, preferential responses for other object categories in the ventral stream, language areas, or default mode system in the angular gyrus). Neither of the two regions has so far been explicitly targeted by fMRI adaptation or multivariate pattern recognition methods in a way similar to what has been done for the IPS to reveal what, if any, is their role in distinguishing between individual numerical symbols. Sufficient sensitivity permitting, the representational metrics revealed by such investigations could give further hints toward the precise role of these regions in the processing of number symbols.

4 CONCLUDING REMARKS

In recent years, techniques which allow to disentangle features represented in fine-scale cortical activity have made it possible to study with fMRI not only the overall contribution of different macroscopic areas of the human brain to a given task but also to reveal more directly the information coded within individual areas. Within the domain of numerical processing, research along these lines has allowed to establish a more direct parallel to monkey neurophysiology regarding the coding of visual numerosities in subregions of the IPS, and has shown that in humans both format-specific (visual numerosity) and more abstract-semantic (common to concrete sets of dots and numerical symbols) components contribute to numerical representation in these areas. The critical role of the representations of individual numerosities identified in parietal cortex for interindividual differences in the behavioral precision of numerical discrimination is also beginning to be revealed (Lasne et al., 2015). For the future, one would hope for a gain in sensitivity which could potentially be provided by ultra-high-field, high-resolution imaging (see, eg, Olman and Yacoub, 2011), to provide more detailed answers to outstanding questions as for example:
What are the cortical layouts underlying the coding of not only individual non-symbolic but also symbolic numerical stimuli? What are the processing steps and corresponding cortical stages by which numerical magnitude can be, in different modalities and presentation modes, abstracted from other perceptual features? What, if any, is the role of the numerical representations identified so far in explicit numerical computations? Previous studies have mainly focused on stimulus-evoked representations and it remains to be shown whether (some of) the same precise representations are recruited when internally generating and manipulating a given numerical magnitude, instead of merely extracting numerical information from the sensory input. Finally, it is still largely unknown how, on the basis of neuronal populations encoding numerical information, several quantities could be combined in the context of a calculation, and models making clear predictions testable by neuroscientific methods are needed to make further progress in this direction.

REFERENCES


Piazza, M., Pica, P., Izard, V., Spelke, E.S., Dehaene, S., 2013. Education enhances the acuity of the nonverbal approximate number system. Psychol. Sci. 24, 1037–1043.
Sasanguie, D., Defever, E., Maertens, B., Reynvoet, B., 2014. The approximate number system is not predictive for symbolic number processing in kindergarteners. Q. J. Exp. Psychol. 67, 271–280.
Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., De Smedt, B., 2016. Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. Dev. Sci. Epub ahead of print.