Task-context-dependent Linear Representation of Multiple Visual Objects in Human Parietal Cortex

Su Keun Jeong[1] and Yaoda Xu[2]

Abstract

■ A host of recent studies have reported robust representations of visual object information in the human parietal cortex, similar to those found in ventral visual cortex. In ventral visual cortex, both monkey neurophysiology and human fMRI studies showed that the neural representation of a pair of unrelated objects can be approximated by the averaged neural representation of the constituent objects shown in isolation. In this study, we examined whether such a linear relationship between objects exists for object representations in the human parietal cortex. Using fMRI and multivoxel pattern analysis, we examined object representations in human inferior and superior intraparietal sulcus, two parietal regions previously implicated in visual object selection and encoding, respectively. We also examined responses from the lateral occipital region, a ventral object processing area. We obtained fMRI response patterns to object pairs and their constituent objects shown in isolation while participants viewed these objects and performed a 1-back repetition detection task. By measuring fMRI response pattern correlations, we found that all three brain regions contained representations for both single object and object pairs. In the lateral occipital region, the representation for a pair of objects could be reliably approximated by the average representation of its constituent objects shown in isolation, replicating previous findings in ventral visual cortex. Such a simple linear relationship, however, was not observed in either parietal region examined. Nevertheless, when we equated the amount of task information present by examining responses from two pairs of objects, we found that representations for the average of two object pairs were indistinguishable in both parietal regions from the average of another two object pairs containing the same four component objects but with a different pairing of the objects (i.e., the average of AB and CD vs. that of AD and CB). Thus, when task information was held consistent, the same linear relationship may govern how multiple independent objects are represented in the human parietal cortex as it does in ventral visual cortex. These findings show that object and task representations coexist in the human parietal cortex and characterize one significant difference of how visual information may be represented in ventral visual and parietal regions. ■

INTRODUCTION

Decades of neuroscience research have associated visual object representation primarily with occipital and temporal regions of the primate brain (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982). Despite the primate parietal cortex's major involvement in visuospatial-related processing (Kravitz, Saleem, Baker, & Mishkin, 2011; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982), a growing body of evidence suggests that parietal cortex also participates in nonspatial aspects of visual information processing. For example, macaque monkey lateral intraparietal neurons show selectivity to color, shape, and task-defined category membership (Freedman & Assad, 2006; Toth & Assad, 2002; Sereno & Maunsell, 1998). Likewise, fMRI responses in human intraparietal sulcus (IPS) have been associated with the processing of color, orientation, shape, and object category (Bettencourt & Xu, 2016; Jeong & Xu, 2013, 2016; Xu & Jeong, 2015; Christophel, Hebart, & Haynes, 2012; Xu & Chun, 2006, 2009; Konen & Kastner, 2008; Todd & Marois, 2004).

In monkey neurophysiology studies, neural response to a pair of unrelated objects could be reliably predicted by the average response of each object shown alone in inferotemporal (IT) cortex (Zoccolan, Cox, & DiCarlo, 2005; see also Reynolds, Chelazzi, & Desimone, 1999). It is argued that IT neuronal outputs are either normalized by summed synaptic drive into IT or normalized by spiking activity originated within IT (Zoccolan et al., 2005). Such a linear averaging relationship has also been found in the human occipital–temporal cortex whereby fMRI response pattern to a pair of unrelated objects could be well approximated by the average response pattern of each constituent object shown in isolation (MacEvoy & Epstein, 2009, 2011; Reddy, Kanwisher, & VanRullen, 2009). This is similar to divisive normalization previously proposed to explain attentional effects in early visual areas (Carandini & Heeger, 2012; Reynolds & Heeger, 2009; Heeger, 1992). MacEvoy and Epstein (2009) provided one good reason this particular mechanism may be implemented by our brain to represent multiple objects. Specifically, because of the finite firing rates

[1]Korea Brain Research Institute, [2]Harvard University

Journal of Cognitive Neuroscience 29:10, pp. 1778–1789
doi:10.1162/jocn_a_01156

of individual neurons, a summation of responses to multiple objects would quickly saturate neuronal responses, especially those neurons responding vigorously to both objects. Such neuronal response saturation would result in the lost of object identity representation at the population level. By normalizing neuronal responses according to the number of objects shown, this problem may be avoided (see also Heeger, 1992).

Although fMRI responses in human IPS have been shown to track the encoding of multiple unrelated objects in visual short-term memory (VSTM; Xu & Chun, 2006; Todd & Marois, 2004), whether a similar response averaging relationship exists in the human parietal cortex during the representation of these objects is unknown. The answer to this question not only would be informative regarding the nature of visual object representation in parietal cortex but also could reveal how visual representation may differ in occipito-temporal and parietal regions.

To address this question, this study examined the representation of a pair of unrelated objects in the human parietal cortex. We focused on two functionally defined parietal regions, with one located at the inferior and one at the superior part of IPS (hence forward referred to as inferior and superior IPS, respectively, for simplicity). These two parietal regions have been shown to be involved in object selection via locations and the encoding of object features, respectively (Xu, 2007, 2009; Xu & Chun, 2006, 2009; see also Todd & Marois, 2004). Recent fMRI multivoxel pattern analysis (MVPA) studies have reported the representation of single object in these parietal regions (Bettencourt & Xu, 2016; Jeong & Xu, 2016; Xu & Jeong, 2015). Here, we tested whether a pair of unrelated objects could be represented in these brain regions and whether a response averaging relationship exists between the representations for an object pair and its constituent objects shown in isolation. To compare and contrast, we also examined responses from the object-selective lateral occipital (LO) region (Grill-Spector, Kushnir, Hendler, & Malach, 2000; Kourtzi & Kanwisher, 2000; Malach et al., 1995).

METHODS

Participants

Ten paid participants (eight women) took part in the study with informed consent. They all had normal or corrected-to-normal visual acuity, were right-handed, and between 18 and 35 years old (M_{age} = 29.33 years, SD = 3.08 years). One additional participant was tested but excluded from data analysis because of excessive head motion (>3 mm). The study was approved by the institutional review board of Harvard University.

Experimental Design

Main Experiment

Participants viewed blocks of images in the main experiment (Figure 1). Each single-object block contained a sequential presentation of 10 images from the same object category either all above or all below the central fixation. Each object pair block contained a sequential presentation of two streams of 10 exemplars from two different categories with one always above and one always below the central fixation. Participants performed a 1-back repetition detection task and pressed a response key whenever they detected an immediate repetition of the exemplar. A repetition occurred twice in each block. In object pair blocks, the repetition occurred randomly in either the upper or lower location. The presentation order of the stimulus blocks and the presentation order of the exemplars within each block were randomly chosen. Four object categories (shoe, bike, guitar, and couch) with each containing 10 different exemplars were shown. All the exemplars from a given object category were

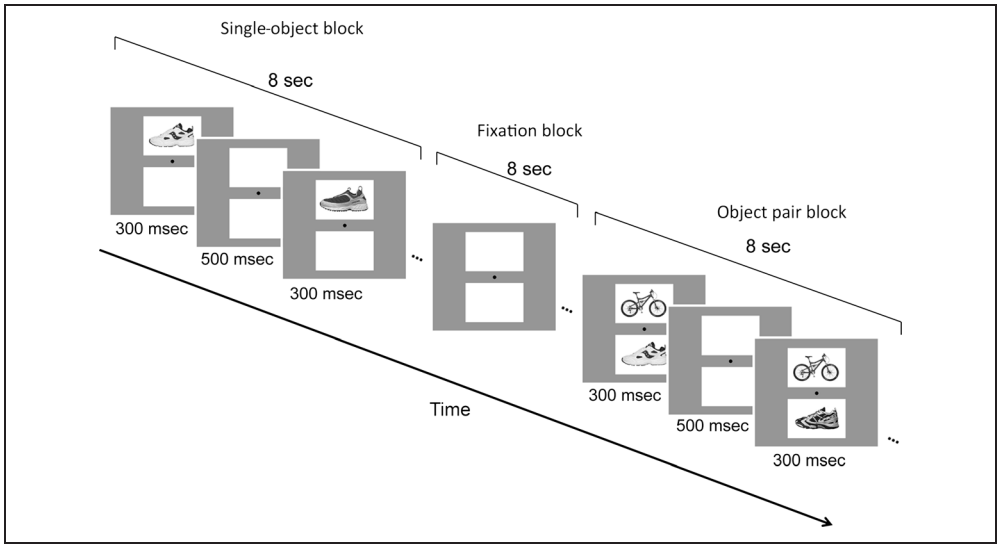

Figure 1. Example trials of the main experiment. In single-object blocks, participants viewed a sequential presentation of single objects (different exemplars) from the same category at a fixed spatial location. In object pair blocks, they viewed a sequential presentation of object pairs from two different categories. The location of the objects from each of the two categories was fixed within a block. The task was to detect a 1-back repetition of the same object. In object pair blocks, the repetition could occur in either object category. The repetition occurred twice in each block.

shown in the same view and thus shared a similar outline. This allowed us to increase the difficulty of the 1-back task, thereby increasing participants' attentional engagement on the task. This also resulted in objects from different categories to be more distinctive from each other, allowing them to be more easily decoded in our MVPA. Category decoding in this study was thus equivalent to shape decoding. There were eight unique single-object blocks (4 object categories × 2 locations) and 12 unique object pair blocks (with all possible combinations of object categories and locations included).

Each exemplar image subtended approximately 5.5° × 2.8°. Two white square placeholders (7° × 4.7°), marking the two exemplar locations, were shown above and below the central fixation throughout a block of trials. The distance between the central fixation and the center of each placeholder was 3.2°. Each stimulus block lasted 8 sec and contained 10 images, with each appearing 300 msec followed by a 500-msec blank interval. Fixation blocks, which lasted 8 sec, were inserted at the beginning and end of the run and between each stimulus block. Each run contained 20 stimulus blocks, with each unique stimulus block appearing once, and 21 fixation blocks. Each participant was tested with 10 runs, each lasting 5 min 36 sec. Participants' eye movements during the main experiment were monitored with an EyeLink 1000 eye tracker to ensure proper central fixation.

Inferior IPS/LO Localizer

Following Xu and Chun (2006), in this localizer, participants viewed blocks of sequentially presented object and noise images (both subtended approximately 12° × 12°). Each object image contained four unique objects shown above, below, and to the left and right of the central fixation (the distance between the fixation and the center of each object was 4°). Gray-scaled photographs of everyday objects (including those that appeared in the main experiment) were used as the object stimuli. To prevent grouping between the objects, objects appeared on white placeholders (4.5° × 3.6°) that were visible throughout an object image block. Noise images were generated by phase-scrambling the object images used. Each block lasted 16 sec and contained 20 images, with each image appearing for 500 msec followed by a 300-msec blank display. Participants were asked to detect the direction of a slight spatial jitter (either horizontal or vertical), which occurred randomly once in every 10 images. Eight object blocks and eight noise blocks were included in each run. Each participant was tested with two or three runs, each lasting 4 min 40 sec.

Superior IPS Localizer

Following Xu and Chun (2006) and Todd and Marois (2004), to localize superior IPS, we conducted an object VSTM experiment. Participants viewed a simultaneous presentation of either one, two, three, or four objects around the central fixation at the four possible locations as outlined in the inferior IPS localizer run. After a brief delay, they judged whether the probe object appearing at one of the four locations shared the same category or was from a different category compared with the object shown at the same location in the initial display. Objects from the same four categories in the main experiment (shoe, bike, guitar, and couch) were shown. Each trial lasted 6 sec and contained a fixation period (1000 msec), a sample display (200 msec), a delay period (1000 msec), a test display/response period (2500 msec), and a feedback (1300 msec). The size of the individual object and the whole display were identical to those used in the inferior IPS/LO localizer. With a counterbalanced trial history design, there were 15 stimulus trials for each display set size as well as 15 fixation trials in which only a fixation dot appeared for 6 sec. Three filler trials were also added for practice and trial history balancing purposes but were excluded from further analysis. Each participant was tested with two runs, each lasting 8 min.

fMRI Methods

fMRI data were acquired from a Siemens (Erlangen, Germany) Tim Trio 3-T scanner at the Harvard Center for Brain Science (Cambridge, MA). Participants viewed images back-projected onto a screen at the rear of the scanner bore through an angled mirror mounted on the head coil. All experiments were controlled by an Apple MacBook Pro laptop running MATLAB (The MathWorks Natick, MA) with Psychtoolbox extensions (Brainard, 1997). For the anatomical images, high-resolution T1-weighted images were acquired (repetition time = 2200 msec, echo time = 1.54 msec, flip angle = 7°, 144 slices, matrix size = 256 × 256, and voxel size = 1 × 1 × 1 mm). Functional data in the main experiment and in the LO/inferior IPS localizer were acquired using the same gradient-echo echo-planar T2*-weighted sequence (repetition time = 2000 msec, echo time = 30 msec, flip angle = 90°, 31 slices, matrix size = 72 × 72, voxel size = 3 × 3 × 3 mm, 168 volumes for the main experiment and 140 volumes for the LO/inferior IPS localizer). Functional data in the superior IPS localizer were acquired using a slightly different gradient-echo echo-planar T2*-weighted sequence (repetition time = 1500 msec, echo time = 29 msec, flip angle = 90°, 24 slices, matrix size = 72 × 72, voxel size = 3 × 3 × 5 mm, 320 volumes).

Data Analysis

fMRI data were analyzed in each participant's native space with BrainVoyager QX (www.brainvoyager.com). 3-D motion correction, slice acquisition time correction, and linear trend removal were performed during data preprocessing. The first two volumes of all functional runs were also discarded.

LO and inferior IPS ROIs were defined separately in each participant as voxels showing higher activations to the object images than to the noise images (false discovery rate $q < .05$, corrected for serial correlation) in lateral occipital cortex and inferior part of IPS, respectively. Superior IPS was defined separately in each participant as voxels tracking that participant's behavioral VSTM capacity estimated using Cowan's K formula (Cowan, 2001). Specifically, we performed a multiple regression analysis with the regression coefficients for each VSTM set size weighted by the participant's corresponding behavioral VSTM capacity for that set size. Superior IPS was defined as voxels showing significant activations in the regression analysis (false discovery rate $q < .05$, corrected for serial correlation) and whose Talairach coordinates matched those reported in Todd and Marois (2004; see Xu & Chun, 2006, for more details).

For each participant, a general linear model was applied to the fMRI data from the main experiment, and beta values were extracted from each stimulus block in each run and in each voxel of each ROI. Following Haxby et al. (2001), we divided the data into odd and even runs and correlated voxel response patterns between odd and even runs in a number of different comparisons. The resulting correlation coefficients were transformed into Fisher's z scores and pooled across participants to allow us to perform group level statistical analyses.

To decode the representation for a single object, we calculated four correlation coefficients using data from the single-object blocks: both same (a correlation of voxel response patterns of the same object category at the same location between odd and even runs), category different (a correlation of two different object categories at the same location between odd and even runs), location different (a correlation of the same object category at different locations between odd and even runs), and both different (a correlation of two different object categories at different locations between odd and even runs; see Figure 2A). Main effects from repeated-measures ANOVA with Object category and Location as factors were then obtained at the group level to evaluate whether Object category and Location were represented in a given brain region. Greenhouse–Geisser correction was used when sphericity assumption was violated, and Bonferroni correction was applied to control for multiple pairwise comparisons.

To decode the representation for object pairs, we calculated two relevant correlation coefficients using data from the object pair blocks: both same (a correlation of the same object pair between odd and even runs) and both different (a correlation of two different object pairs between odd and even runs; Figure 3A). The representation for the different object pairs was then evaluated by comparing these correlation coefficients at the group level.

The main question of this study was whether the representation of an object pair could be approximated by the average representation of its constituent objects shown in isolation. To address this question, we compared the voxel response pattern correlation of the same object pair between odd and even runs (actual pair) with

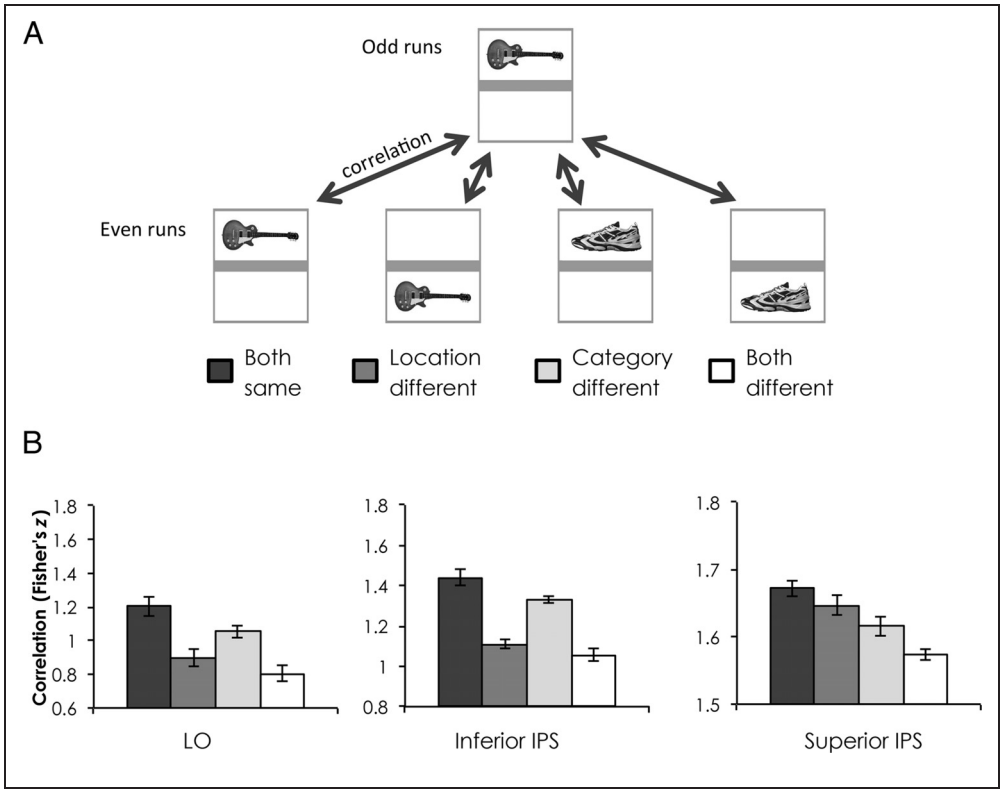

Figure 2. The decoding of single-object representations. (A) A schematic illustration of the correlations computed among the different single-object blocks across odd and even runs. (B) Fisher-transformed correlation coefficients (z) of the correlations computed in LO, inferior IPS, and superior IPS. Dark gray, gray, light gray, and white bars indicate both same, location-different, category-different, and both-different conditions, respectively. Main effects of category and location were significant in all three ROIs, suggesting the existence of object and location representations in these brain regions. Error bars indicate within-participant *SEM*.

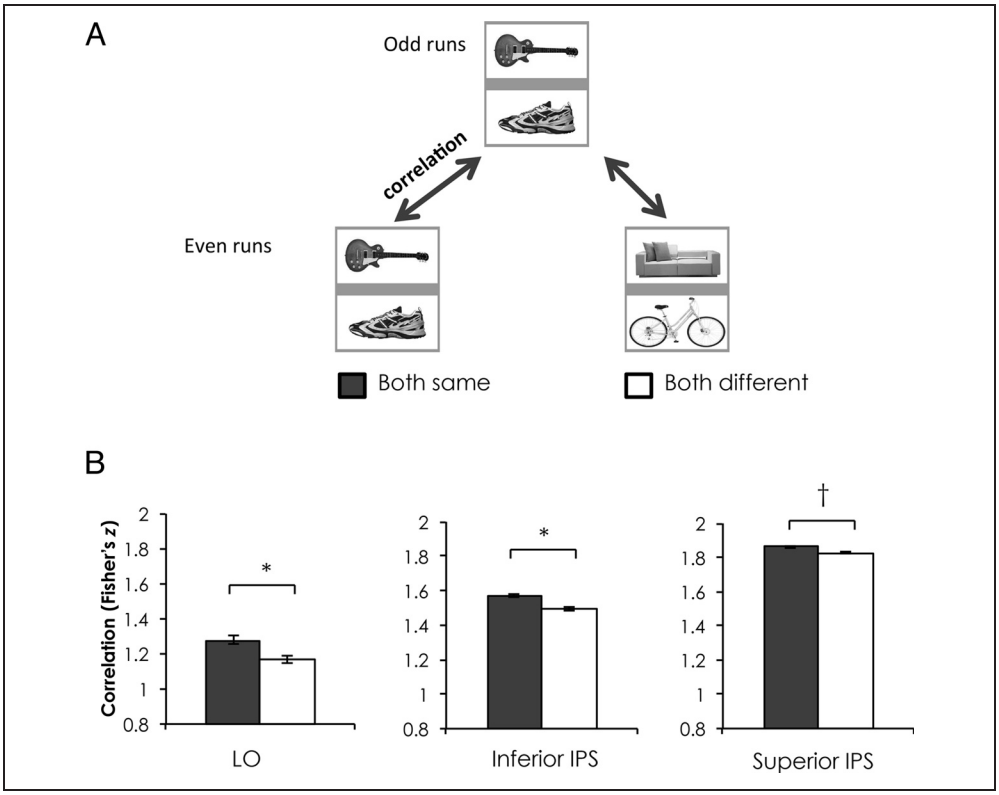

Figure 3. The decoding of object pair representations. (A) A schematic illustration of the correlations computed among the different object pair blocks across odd and even runs. (B) Fisher-transformed correlation coefficients (z) of the correlations computed in LO, inferior IPS, and superior IPS. Dark gray and white bars indicate both-same and both-different conditions, respectively. Both-same condition was greater than both-different condition in all three ROIs, suggesting the existence of object pair representations in these brain regions. Error bars indicate within-participant *SEM*. $*p < .05$; $\dagger p < .1$. *ns* = not significant.

the voxel response correlation of an object pair and the average of its constituent objects between odd and even runs (averaged pair; Figure 4A). If the strength of the two correlations does not differ significantly in a brain region, then the representation of an object pair can be approximated by the average representation of its constituent objects shown in isolation. It is worth noting that, because we used a correlation measurement that normalized response amplitudes into z scores when a correlation was computed, averaging and summing the patterns from the individual objects to predict the actual object pair would produce the same result.

To examine whether mechanisms other than a simple average or sum can predict the representation of an object pair, we additionally tested how patterns generated by a MAX model (Riesenhuber & Poggio, 1999) and a weighted average model would predict the actual object pair patterns. To generate a MAX object pair pattern, in each voxel, we compared responses of the two constituent objects shown in isolation. We then selected the higher response of the two and assigned that as the response of that voxel to predict the object pair pattern. To generate a weighted average object pair pattern, we systematically varied the contribution of the two isolated objects' patterns from 0% versus 100% and 10% versus 90% to 90% versus 10% and 100% versus 0%, with 10% incremental changes in between. The resulting MAX and weighted average object pair patterns were correlated with the actual object pair pattern across odd and even runs. This correlation was then compared with the correlation of the actual object pair patterns between odd and even runs.

Although participants performed the same 1-back repetition detection task in both the single-object and object pair conditions, attending to one versus two locations imposed different attentional/task demand and put participants under different task contexts. To examine whether the representation of task information needed to be taken into account when evaluating the response averaging relationship between object pairs and single object, we performed an additional analysis and compared three additional correlation coefficients. We first averaged two different object pairs (e.g., bike/couch and guitar/shoe) and correlated this pair averaged response pattern between odd and even runs (actual two-pair). We then correlated the pair average and the average of the four constituent objects shown in isolation between odd and even runs (averaged four-single). Finally, we correlated the pair average from two sets of pairs containing the same four constituent objects at the same locations but combined differently (e.g., the average of bike/couch and guitar/shoe and the average of bike/shoe and guitar/couch) (averaged two-pair) between odd and even runs (see Figure 5A).

RESULTS

Behavioral Results

Behavioral performance for single-object and object pair blocks were 92.22% (*SD* = 8.56%) and 51.46% (*SD* = 4.46%),

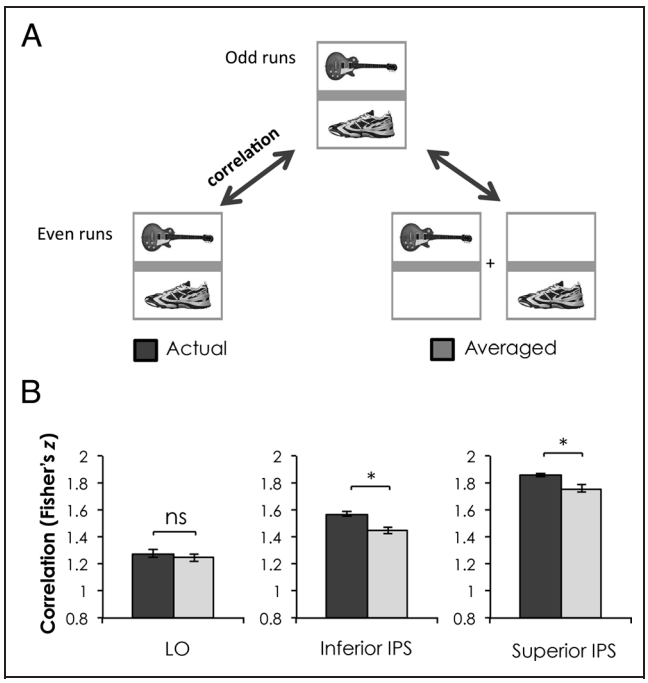

Figure 4. Reconstructing object pair representation with single-object representations. (A) A schematic illustration of the correlations computed. An object pair was correlated with itself across odd and even runs (actual) and the average of its constituent objects shown in isolation across odd and even runs (average). (B) Fisher-transformed correlation coefficients (z) of the correlations computed in LO, inferior IPS, and superior IPS. Dark gray and white bars indicate actual pair and average pair, respectively. Although representation of an object pair could be approximated by its constituent objects shown in isolation in LO, such an averaging relationship was not observed in inferior or superior IPS. Error bars indicate within-participant *SEM*. *$p < .05$. *ns* = not significant.

respectively, showing that object pair blocks were significantly more difficult to monitor than single-object blocks ($t(9) = 18.06, p < .001$).

Univariate Analysis

Response amplitudes of the single object did not differ among the four object categories (i.e., shoe, bike, couch, and guitar) in any of the three ROIs ($ts < 2.33, ps > .264$, Bonferroni corrected). As expected, response amplitudes for object pairs were higher than those for single objects in all three ROIs because of the greater amount of visual stimulation evoked by object pairs ($ts > 2.363, ps < .042$). Consistent with the lower field bias previously reported (Strother, Aldcroft, Lavell, & Vilis, 2010; Sayres & Grill-Spector, 2008), response amplitudes for single objects shown in the lower visual field were greater than those shown in the upper visual field in all three ROIs ($ts > 2.37, ps < .042$).

Despite these differences in response amplitudes in the different conditions, by using a pattern correlation measure in our MVPA, any response amplitude differences were normalized, thereby removing their impact on our decoding results.

The Representation of Single Objects and Object Pairs

To examine whether distinctive representations for different single objects and their locations exist in LO, inferior IPS, and superior IPS, we compared fMRI response pattern correlation coefficients when the object categories across the odd and even runs of comparisons were the same and when they were different and when the object categories were shown at the same and different locations. We performed a repeated-measures ANOVA with Object category (same, different) and Location (same, different) as the main factors (Figure 2A). Because all the exemplars in an object category shared the same overall outline, object category decoding here was equivalent to object shape decoding. In all three ROIs, significant main effects of object category ($Fs > 10.93, ps < .009$) and

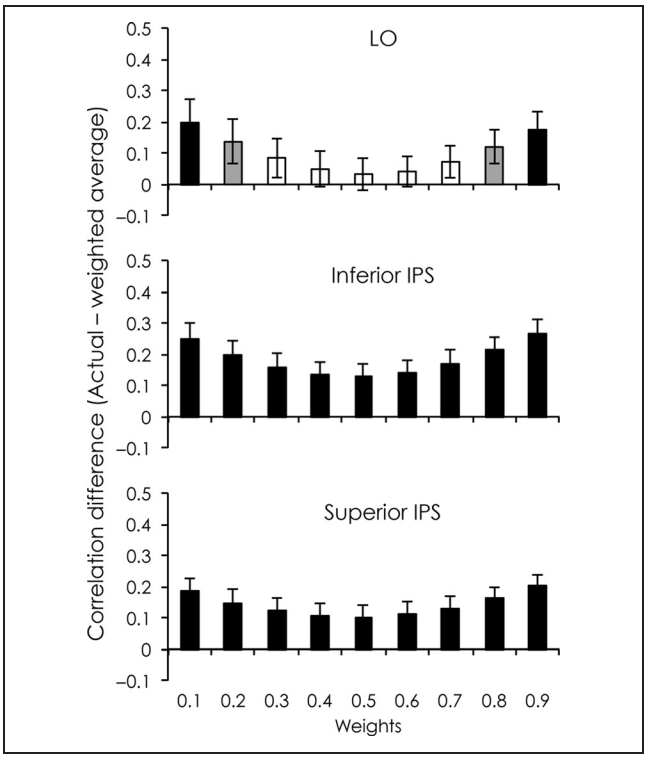

Figure 5. Comparisons between the actual pair pattern and the weighted average of the constituent object patterns. y Axis indicates Fisher-transformed correlation difference between an actual pair to itself and an actual pair to the average of the constituent objects with different weights across odd and even runs. x Axis indicates the different weight combinations used with respect to the top object (e.g., 0.1 indicates 10% top object and 90% bottom object). Black, gray, and white bars indicate that the correlation difference is significantly greater than zero ($p < .05$), marginally greater than zero ($p < .1$), and not different from zero ($p \geq .1$), respectively. In both parietal regions, the weighted average pattern was significantly different from that of an actual pair across all the weight combinations.

location (Fs > 9.16, ps < .014) were found (Figure 2B). The interaction between category and location was not significant in inferior IPS ($F(1, 9) = 2.74$, $p = .13$) but was marginally significant in LO and superior IPS (Fs > 3.45, ps < .096). Overall, these results show that all three ROIs carried representations for object category (i.e., shape) and location.

We then examined whether distinctive representations for different object pairs exist in the three ROIs by comparing fMRI response pattern correlation coefficients when the object pairs across the odd and even runs were the same and when they were completely different (see Methods and Figure 3A). Changing both objects in a pair reduced correlation in all three ROIs ($t(9) = 2.22$, $p = .053$, for superior IPS; ts > 2.88, ps < .018, for LO and inferior IPS). These results indicate that representations for distinctive object pairs exist in LO and both parietal regions examined.

Reconstructing Object Pair Representation with Single Object Representations

The main question of this study was whether the representation of an object pair could be approximated by the average representation of its constituent objects shown in isolation. To address this question, we compared the voxel response pattern correlation of the same object pair between odd and even runs (actual pair) with the voxel response correlation of an object pair and the average of its constituent objects shown in isolation between odd and even runs (averaged pair; Figure 4A). If the strength of the two correlations did not differ significantly in a brain region, then the representation of an object pair can be approximated by the average representation of its constituent objects shown in isolation. Because we used a correlation measurement that normalized response amplitudes into z scores when a correlation was computed, averaging and summing the patterns from the constituent objects to predict the actual object pair would produce the same result.

In LO, actual pair and averaged pair were indistinguishable from each other ($t < 1$, $p = .541$), showing that a pair of objects could be approximated by the linear average of its constituent objects. However, such a linear averaging relationship was not found in either inferior or superior IPS, as actual pair showed significantly greater correlations than did averaged pair (ts > 2.49, ps < .034). An interaction between region and condition was observed between LO and inferior IPS ($F(1, 9) = 6.96$, $p = .027$), but not between LO and superior IPS ($F(1, 9) = 1.43$, $p = .262$) or between inferior and superior IPS ($F < 1$, $p = .529$).

We also generated a synthetic object pair pattern by using the MAX model following Riesenhuber and Poggio (1999) and, in each voxel, selected the higher of the two responses from the constituent objects shown alone as the predicted voxel response for the object pair. Using the MAX model did not improve pattern prediction in either parietal regions and still resulted in a lower correlation with actual pair than when actual pair was correlated with itself across odd and even runs (ts > 3.41, ps < .008).

Because of the existence of possible location encoding bias within the object processing regions (Strother et al., 2010; Sayres & Grill-Spector, 2008; see also Silson, Chan, Reynolds, Kravitz, & Baker, 2015), two objects in a pair may not contribute equally to the final representation. To test this possibility, we generated weighted average patterns from the constituent object patterns and examined whether a particularly weighted average pattern would better predict the actual object pair pattern. We systematically varied the contribution of the two constituent objects from 0% to 100% with 10% increments. In both inferior and superior IPS, none of the weighted patterns predicted that of the actual object pair, and all were significantly different from that of the actual object pair (ts > 2.55, ps < .031; see Figure 5). Interestingly, in LO, the weighted average patterns and an actual object pair pattern were indistinguishable (ts < 1.45, ps > .19), unless the weight was heavily biased toward one of the constituent objects (ts > 2.54, ps < .031 for [10% weight on the top object, 90% weight on the bottom object in a pair] and [90% top, 10% bottom] and ts > 1.93, ps < .085 for [20% top, 80% bottom] and [80% top, 20% bottom]).

LO contained more voxels than either inferior or superior IPS did (LO: mean = 291.9, $SD = 39.96$; inferior IPS: mean = 256.1, $SD = 59.91$; superior IPS: mean = 259.3, $SD = 69.15$). However, this difference could not explain why an averaging relationship existed in LO but not in either parietal regions. This is because, when we selected up to 100 voxels from each ROI based on their averaged response amplitude across all the conditions (in some participants, an ROI contained less than 100 voxels), we found that the average pair still showed a significantly lower correlation with an actual pair than the actual pair was to itself in both parietal regions (ts > 3.2, ps < .011). In LO, this difference was not found ($t(9) = 1.02$, $p = .33$), showing a good prediction of the average pair to the actual pair as was found before.

Overall, these results showed that the representation of an object pair in LO could be well approximated by the average representation of its constituent objects shown in isolation, replicating previous findings (MacEvoy & Epstein, 2009). Such a relationship, however, was absent in the two parietal regions examined. A MAX and a weighted average pattern of the constituent objects also failed to perform any better (if not worse) than the averaged pattern in predicting the pattern of the actual object pair in parietal cortex. Thus, the fMRI response pattern evoked by an object pair in parietal cortex could not be approximated by a simple average (or sum, as they are equivalent in our correlation analysis), a weighted average, or a MAX model of the fMRI response patterns of the constituent objects shown in isolation.

Reconstructing Object Pair Representation with the Addition of Task Context

In the current experiment, participants performed the same 1-back repetition detection task in both the single-object and object pair blocks. Nevertheless, attending to one versus two locations imposed different attentional/task demand and put participants under different task contexts. Indeed, as reported earlier, behavioral performance was lower for the object pair than the single- object condition. Because parietal regions have been shown to encode task-related information (Woolgar, Hampshire, Thompson, & Duncan, 2011; Woolgar, Thompson, Bor, & Duncan, 2011; Balan & Gottlieb, 2006; Stoet & Snyder, 2004), differences in task information representation in the single-object and object pair conditions could have obscured an otherwise averaging relationship between the two stimulus conditions in the two parietal regions examined.

To test this possibility, we compared three correlation coefficients (see Figure 6A). We first averaged two different object pairs (e.g., bike/couch and guitar/shoe) and correlated this pair averaged response pattern between odd and even runs (actual two-pair). We then correlated the pair average and the average of the four constituent objects shown in isolation between odd and even runs (averaged four-single). Finally, we correlated the pair average from two sets of pairs containing the same four constituent objects at the same locations but combined differently (e.g., the average of bike/couch and guitar/shoe and the average of bike/shoe and guitar/couch; averaged two-pair; see Figure 6A). Among these three conditions, task context still differed between actual two-pair and average four-single, but not between actual two-pair and average two-pair. If a linear averaging relationship exists between an object pair and its constituent objects in parietal cortex and if task information is also represented

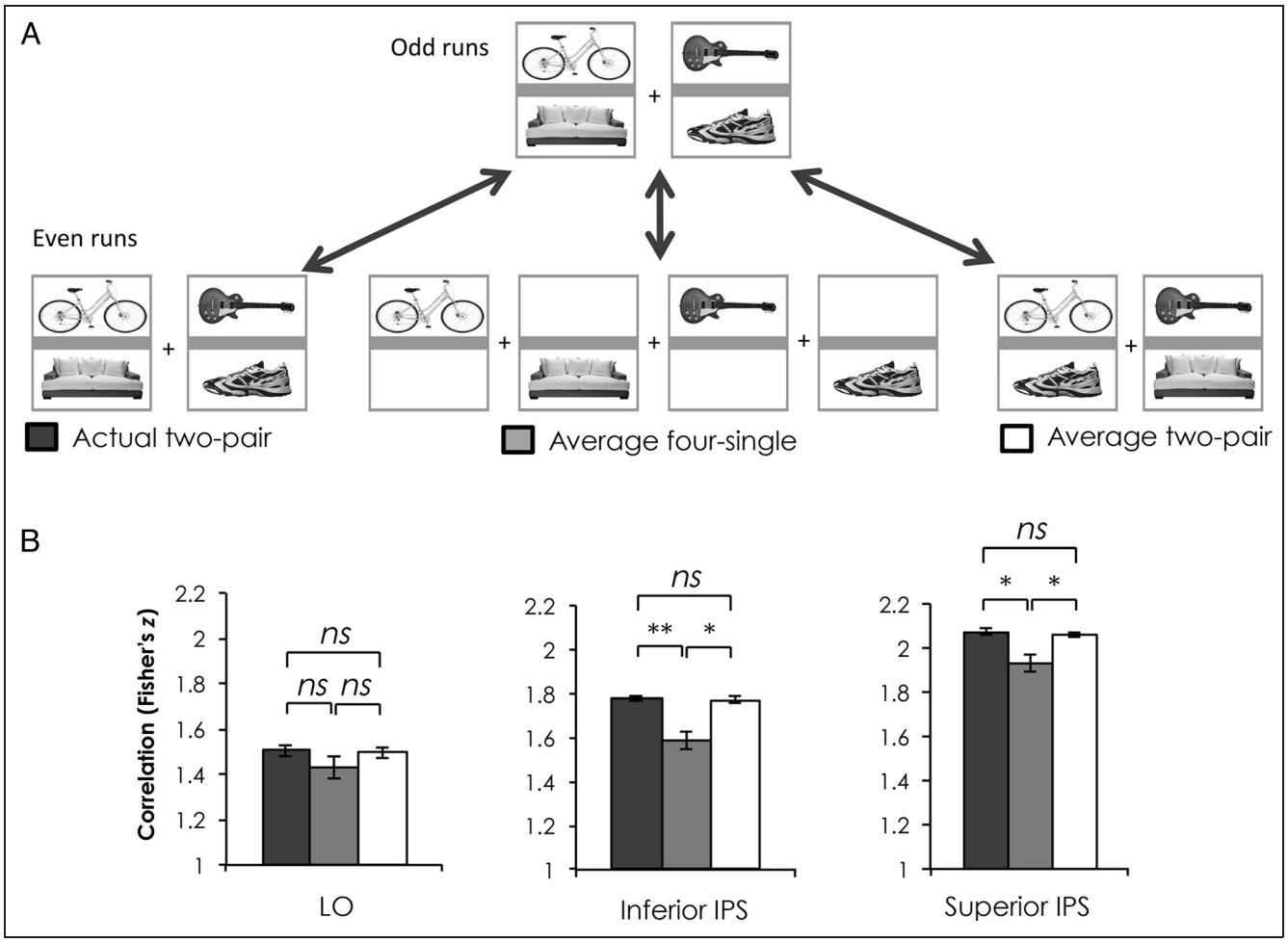

Figure 6. Reconstructing object pair representation with the addition of task context. (A) A schematic illustration of the correlations computed. Two object pairs were averaged and were correlated across odd and even runs with either themselves, with the four constituent objects shown in isolation, or with the average of two object pairs containing the same four objects but paired differently. (B) Fisher-transformed correlation coefficients (z) of the correlations computed in LO, inferior IPS, and superior IPS. In LO, the representation of an object pair could be approximated by those of its constituent objects shown in isolation. Such an averaging relationship could be seen in inferior and superior IPS only when the task context was also matched. Error bars indicate within-participant *SEM*. $*p < .05$; $**p < .01$. *ns* = not significant.

in this brain region, then by taking task context into account, an actual and an averaged two-pair should evoke very similar representations in parietal cortex.

In LO, a repeated-measures ANOVA with three conditions (Actual two-pair, Averaged four-single, and Averaged two-pair) failed to reveal a significant main effect ($F(2, 18) = 1.044$, $p = .372$). There was no difference among actual two-pair, averaged four-single, and average two-pair (ts < 1.07, ps $> .938$, corrected; Figure 6B). In contrast, in both parietal regions, we found a significant main effect of conditions (Fs > 9.44, ps $< .002$). In inferior and superior IPS, average four-single was significantly lower than actual two-pair (ts > 3.2, ps $< .032$, corrected), similar to the difference we observed between actual pair and averaged pair in Figure 4B. Critically, in both parietal regions, average two-pair was no different from actual two-pair (ts < 1.11, ps $> .879$, corrected) but significantly higher than average four-single (ts > 3.006, ps $< .044$, corrected). The interaction between region and condition was significant between inferior IPS and LO ($F(2, 18) = 4.76$, $p = .022$), but not between inferior and superior IPS ($F < 1$, $p = .46$) or between LO and superior IPS ($F < 1$, $p = .41$).

Overall, these results indicated that, by taking task information representation into account, actual and averaged two-pair evoked very similar representations in parietal cortex, thereby showing the existence of a linear averaging relationship between an object pair and its constituent objects in parietal cortex under the same task context.

DISCUSSION

This study examined the neural representation of a pair of unrelated objects in the human parietal cortex. We focused on inferior and superior IPS, two parietal regions previously implicated in object selection and encoding, respectively (Xu & Chun, 2006, 2009). We also examined responses from LO, an object processing region in ventral visual cortex. By measuring fMRI response pattern correlations, we found that these regions contained representations for single objects and object pairs. In LO, the representation for a pair of objects could be reliably approximated by the average representation of its constituent objects shown in isolation. This result is consistent with previous monkey neurophysiology and human fMRI studies in ventral visual cortex (MacEvoy & Epstein, 2009; Zoccolan et al., 2005) and is similar to divisive normalization previously proposed to explain attentional effects in early visual areas (Carandini & Heeger, 2012; Reynolds & Heeger, 2009; Heeger, 1992). It has been argued that such a linear averaging relationship could prevent neuronal response saturation when multiple objects are shown simultaneously and allow our visual system to maintain representation fidelity (MacEvoy & Epstein, 2009; see also Heeger, 1992).

Although, in the real world, visual objects are rarely shown in isolation but are surrounded by contextually meaningful objects, it is useful to document the linear averaging relationship between a pair of unrelated objects and their constituent objects. We can then see when this linear relationship is violated as a way to measure the interaction between objects when contextual effects may be at play. With this approach, MacEvoy and Epstein (2011) showed participants four categories of scenes and eight categories of "signature" objects. By measuring fMRI response patterns, they found that in LO scenes were still well approximated by the average of the patterns elicited by their signature objects; this relationship, however, was not observed in parahippocampal place area, a scene processing region. Similarly, Baeck, Wagemans, and Op de Beeck (2013) found a linear relationship between object pairs and single objects regardless of whether the two objects in a pair interacted in a meaningful way. Thus, although object processing may be influenced by context in behavioral studies, such an effect could occur at a semantic level, likely corresponding to object processing in anterior temporal lobe (Peelen & Caramazza, 2012); the available evidence has not shown that object processing in LO is influenced by context.

Because we measured response pattern correlation that normalized response amplitudes into z scores when a correlation was computed, averaging or summing the response patterns from the constituent objects to predict the actual object pair would produce the same result. Thus, our result in LO was consistent with both an average and a sum model and indicated a simple linear relationship between an object pair and its constituent objects. Such a simple linear relationship, however, was not readily observed in either parietal region examined. Neither did a MAX model nor a weighted average representation from the constituent objects predicted that of the actual object pair in parietal cortex. Nevertheless, when we equated the task information present by examining responses from two pairs of objects, we did observe a simple linear average/sum relationship between an object pair and its constituent objects in both parietal regions. Thus, when the task context was held consistent, the same linear mechanism seems to operate in parietal regions as it does in ventral visual cortex.

The representation of task information in inferior and superior IPS is consistent with previous research reporting neural activities related to task information such as task rule and task context in posterior parietal cortex (PPC; Woolgar, Thompson, et al., 2011; Gottlieb & Snyder, 2010; Balan & Gottlieb, 2006; Stoet & Snyder, 2004). Inferior and superior IPS regions have been shown to overlap substantially with parietal topographic areas (Bettencourt & Xu, 2016). Although characteristics of the parietal topographic areas may not fully capture all the response properties of these two functionally defined parietal regions (Bettencourt & Xu, 2016; Jeong & Xu, 2016), given that both of these two regions exhibited similar response profiles in this study, the effect of task context on visual object representation is likely a

common characteristic of PPC. Consistent with this view, in a recent study, by asking participants to perform two different tasks on the same stimuli, direct decoding of task has revealed task representation in V3A, V3B, and IPS0 to IPS2 as well as inferior and superior IPS (Vaziri-Pashkam & Xu, under review).

In this study, in the single-object condition, participants only had to attend to one spatial location, and the 1-back task was less demanding. In the object pair condition, participants had to attend to two spatial locations simultaneously and monitor two streams of objects, resulting in a more demanding 1-back task. These differences plus the difference in task instruction (monitoring one vs. two streams of objects) likely contributed to distinctive task representations in the two conditions. Because we used a correlation measure that normalized response amplitudes, difference between the single-object and object pair conditions could not have been driven entirely by a difference in response amplitudes, but rather, a significant change in response pattern had to occur to account for our results. Thus, although both ventral visual cortex and parietal cortex can represent a variety of visual information, parietal cortex additionally represents the task context from which visual information is extracted that is beyond a simple modulation of response amplitude. This likely characterizes one significant difference between these two brain regions in their representation of visual information.

Indeed, although brain regions in both the ventral and dorsal pathways have been shown to be involved in visual object processing (Konen & Kastner, 2008; Lehky & Sereno, 2007; Sereno & Maunsell, 1998), a growing body of evidence suggests that they play different roles. Lehky and Sereno (2007) found that PPC showed weaker shape selectivity than IT cortex when monkeys were performing a passive fixation task. However, other studies showed that PPC representations were task dependent in both monkeys and humans (Bracci, Daniels, & Op de Beeck, 2017; Xu & Jeong, 2015; Toth & Assad, 2002). We additionally showed in a recent study that superior IPS representation better tracked perceptual similarity (as measured in a behavioral task) than that in LO (Jeong & Xu, 2016). Using representation similarity measures, it was further shown that, although the representation structure in ventral regions is driven predominantly by variations in object shapes, both object shape and task influence the representation structure in dorsal regions (Vaziri-Pashkam & Xu, under review; see also Bracci et al., 2017).

Some have argued that dorsal object representations support object-directed action (Kastner, Chen, Jeong, & Mruczek, 2017; Freud, Plaut, & Behrmann, 2016; Monaco et al., 2014). However, we do not fully agree with this view. On the basis of results from others as well as our own, we would like to argue that dorsal pathway plays a greater role of flexibly supporting visual object representation depending on the current task demands, whether an action is required. That is, dorsal visual representation contains not only object information but also additional information related to the task being performed on the object. Thus, whereas visual representations in the ventral pathway are more invariant and reflect "what an object is," those in the dorsal pathway are more adaptive and reflect "what we do with it" (Vaziri-Pashkam & Xu, under review).

Xu and Chun (2006, 2009) argued that the role of superior IPS is to encode detailed object feature information whereas that of inferior IPS is to index object locations via coarse object representation. This was based on the observation that fMRI response amplitudes in inferior IPS followed the number of object locations encoded (up to four) regardless of object complexity, whereas responses in superior IPS correlated with the amount of object information encoded (Xu, 2009; Xu & Chun, 2006). Although these studies indicated a functional dissociation between the two IPS regions, the current study showed that the representations of single objects and pairs of objects existed in both regions (Figures 2B and 3B). Because highly distinctive object shapes were used in this study to facilitate object category decoding, it seemed that distinctive object representations could still be formed even in a coarse object representation region such as inferior IPS. Future studies are needed to systematically vary the similarity of the objects to directly compare the representation resolution of inferior and superior IPS.

Although the representation of single isolated objects has been the focus of most neuroscience research, given that objects rarely occur in isolation in the real world, understanding the neural representation of multiple objects is essential if we want to thoroughly understand the neural mechanisms supporting visual object representation. This study shows that a simple linear relationship previously characterized for object representations in the ventral visual cortex also applies to those in the human parietal cortex when the task context was held consistent. Thus, the same linear mechanism may govern how multiple independent objects may be represented in the human brain across multiple different brain regions.

Acknowledgments

This research was supported by NIH grant 1R01EY022355 to Y. X. and the KBRI basic research program through the Korea Brain Research Institute funded by the Ministry of Science, ICT & Future Planning (No. 2231-415) to S. K. J.

Reprint requests should be sent to Su Keun Jeong, Korea Brain Research Institute, Cheomdan-ro 61, Dong-gu, Daegu, South Korea 41068, or via e-mail: skjeong@kbri.re.kr.

REFERENCES

Baeck, A., Wagemans, J., & Op de Beeck, H. P. (2013). The distributed representation of random and meaningful object pairs in human occipitotemporal cortex: The weighted average as a general rule. *Neuroimage, 70,* 37–47.

Balan, P. F., & Gottlieb, J. P. (2006). Integration of exogenous input into a dynamic salience map revealed by perturbing attention. *Journal of Neuroscience, 26,* 9239–9249.
Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience, 19,* 150–157.
Bracci, S., Daniels, N., & Op de Beeck, H. (2017). Task context overrules object- and category-related representational content in the human parietal cortex. *Cerebral Cortex, 27,* 310–321.
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.
Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13,* 51–62.
Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *Journal of Neuroscience, 32,* 12983–12989.
Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences, 24,* 87–114.
Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature, 443,* 85–88.
Freud, E., Plaut, D. C., & Behrmann, M. (2016). "What" is happening in the dorsal visual pathway. *Trends in Cognitive Sciences, 20,* 773–784.
Gottlieb, J. P., & Snyder, L. H. (2010). Spatial and non-spatial functions of the parietal cortex. *Current Opinion in Neurobiology, 20,* 731–740.
Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3,* 837–843.
Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425–2430.
Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197.
Jeong, S. K., & Xu, Y. (2013). Neural representation of targets and distractors during object individuation and identification. *Journal of Cognitive Neuroscience, 25,* 117–126.
Jeong, S. K., & Xu, Y. (2016). Behaviorally relevant abstract object identity representation in the human parietal cortex. *Journal of Neuroscience, 36,* 1607–1619.
Kastner, S., Chen, Q., Jeong, S. K., & Mruczek, R. E. B. (2017). Neuropsychologia. *Neuropsychologia,* 1–0.
Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience, 11,* 224–231.
Kourtzi, Z., & Kanwisher, N. G. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20,* 3310–3318.
Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience, 12,* 217–230.
Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences, 17,* 26–49.
Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology, 97,* 307–319.
MacEvoy, S. P., & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology, 19,* 943–947.
MacEvoy, S. P., & Epstein, R. A. (2011). Constructing scenes from objects in human occipitotemporal cortex. *Nature Neuroscience, 14,* 1323–1329.
Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, U.S.A., 92,* 8135–8139.
Monaco, S., Chen, Y., Medendorp, W. P., Crawford, J. D., Fiehler, K., & Henriques, D. Y. P. (2014). Functional magnetic resonance imaging adaptation reveals the cortical networks for processing grasp-relevant object properties. *Cerebral Cortex, 24,* 1540–1554.
Peelen, M. V., & Caramazza, A. (2012). Conceptual object representations in human anterior temporal cortex. *Journal of Neuroscience, 32,* 15728–15736.
Reddy, L., Kanwisher, N. G., & VanRullen, R. (2009). Attention and biased competition in multi-voxel object representations. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 21447–21452.
Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19,* 1736–1753.
Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron, 61,* 168–185.
Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2,* 1019–1025.
Sayres, R., & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology, 100,* 249–267.
Sereno, A. B., & Maunsell, J. H. (1998). Shape selectivity in primate lateral intraparietal cortex. *Nature, 395,* 500–503.
Silson, E. H., Chan, A. W.-Y., Reynolds, R. C., Kravitz, D. J., & Baker, C. I. (2015). A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *Journal of Neuroscience, 35,* 11921–11935.
Stoet, G., & Snyder, L. H. (2004). Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron, 42,* 1003–1012.
Strother, L., Aldcroft, A., Lavell, C., & Vilis, T. (2010). Equal degrees of object selectivity for upper and lower visual field stimuli. *Journal of Neurophysiology, 104,* 2075–2081.
Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751–754.
Toth, L. J., & Assad, J. A. (2002). Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature, 415,* 165–168.
Ungerleider, L. G., & Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinion in Neurobiology, 4,* 157–165.
Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
Vaziri-Pashkam, M., & Xu, Y. (under review). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations.
Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience, 31,* 14592–14599.
Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *Neuroimage, 56,* 744–752.
Xu, Y. (2007). The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *Journal of Neuroscience, 27,* 11676–11686.

Xu, Y. (2009). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience, 21,* 511–518.
Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440,* 91–95.
Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences, 13,* 167–174.
Xu, Y., & Jeong, S. K. (2015). The contribution of human superior intra-parietal sulcus to visual short-term memory and perception. In P. Jolicoeur & J. Martinez-Trujillo (Eds.), *Mechanisms of sensory working memory: Attention and performance XXV*. London: Elsevier.
Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience, 25,* 8150–8164.