

Task-Related Dynamic Division of Labor Between Anterior Temporal and Lateral Occipital Cortices in Representing Object Size

Rocco Chiou and  Matthew A. Lambon Ralph

Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom

Object size is represented by functionally distinct sectors along the ventral visual pathway. The early visual cortex encodes objects' sensory-retinal size. Subsequently, the occipitotemporal cortex computes objects' canonical size based on statistical regularities of visual features. Although the neurocomputation of size has been studied in a "bottom-up" sensory-driven framework, little is known about how perceptual size information is transformed into conceptual knowledge and how this computation is modulated by "top-down" goal-driven signals. Using continuous theta burst stimulation, we demonstrated that behavioral goal shapes the neurocognitive network underpinning object size. We manipulated the congruency of perceptual versus conceptual object size, which provides a robust behavioral probe reflecting implicit size knowledge. Neurostimulation was targeted at the lateral occipital cortex (LOC), a key region for object perception, or the anterior temporal lobe (ATL), a "hub" of supramodal conceptual processing. We observed striking contextual modulation of the neurocognitive architecture: when human participants judged perceptual size, the congruency effect was significantly attenuated by LOC stimulation but stayed resilient to ATL stimulation. By contrast, when they judged conceptual size, both LOC and ATL stimulation eradicated the otherwise robust effect. Our findings demonstrate disparate functional profiles of the LOC and ATL, providing the first evidence of a malleable network adaptively altering its division of labor with top-down states. The LOC, regardless of task demand, automatically represents "bottom-up" statistical regularities of visual conformation (reflecting typical object size), whereas the ATL contributes to this computation when the context requires semantically based linkage of visual attributes to object recognition.

Key words: anterior temporal lobe; continuous theta burst stimulation; lateral occipital cortex; object size knowledge; semantic memory; transcranial magnetic stimulation

Significance Statement

In the present study, we provide compelling evidence that the "top-down" cognitive state of an observer changes the dynamic interaction between different subregions of the ventral temporal cortex. Using inhibitory neurostimulation combined with a novel paradigm, we demonstrate a flexible division of labor in the neurocognitive architecture that underpins size knowledge: the lateral occipital cortex codes perceptually based aspects (statistical visual configuration of small/large objects), whereas the anterior temporal lobe represents semantically based aspects (object identity), with their involvement interactively weighted by task demand. The interactive nature of the ventral temporal cortex highlights how top-down modulation constrains and shapes neural representations in the visual system.

Introduction

The physical volume of an object is a key dimension that determines how we interact with it: we manipulate smaller objects

manually, whereas we lean against big objects or use them to orient direction. Apart from its functional significance, object size also constrains the topography of object representation in

Received July 28, 2015; revised Jan. 6, 2016; accepted Jan. 25, 2016.

Author contributions: M.A.L.R. and R.C. designed research; R.C. performed research; M.A.L.R. and R.C. analyzed data; M.A.L.R. and R.C. wrote the paper.

This work was supported by Medical Research Council Programme Grant MR/J004146/1 to M.A.L.R. and a Stepping Fellowship to R.C., funded through Wellcome Trust Institutional Strategic Support Fund Award 097820 to the University of Manchester.

The authors declare no competing financial interests.

This article is freely available online through the *JNeurosci* Author Open Choice option.

Correspondence should be addressed to Dr. Rocco Chiou or Dr. Matthew A. Lambon Ralph, Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of Manchester, England, United Kingdom, E-mail: rocco.chiou@manchester.ac.uk or matt.lambon-ralph@manchester.ac.uk.

DOI:10.1523/JNEUROSCI.2829-15.2016

Copyright © 2016 Chiou and Lambon Ralph

This is an Open Access article distributed under the terms of the Creative Commons Attribution License Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

visual cortices. Konkle and colleagues (Konkle and Oliva, 2012a; Konkle and Caramazza, 2013) have demonstrated a “size preference map” of the occipitotemporal cortex, with the inferior–temporal and lateral–occipital cortices comparatively more sensitive to small objects and the bilateral parahippocampal regions more responsive to large objects. Furthermore, the brain uses a network of cortical areas to represent object size at different levels of a “perceptual-to-conceptual” continuum, mapped onto different stages of the visual processing cascade. An object’s sensory-retinal size is coded in the early visual cortex (EVC), which changes its response with the spatial extent of visual stimuli regardless of object category (Konkle and Oliva, 2012a). Following initial sensory registry, selectivity to different sizes begins to emerge when visual processing proceeds to the occipitotemporal cortex. Unlike the EVC that veridically represents size to reflect changes of retinal contents and ratios, the occipitotemporal cortex is tolerant to variation of retinal coverage, maintaining preference to objects’ real-world canonical size; the occipitotemporal size-selective activity is not reliant on sensory input as it persists even when objects are imagined (Konkle and Oliva, 2012a). This led Konkle and Oliva (2012a) to conclude that the neurocomputation performed in the occipitotemporal cortex is “object-centered,” above and beyond retinotopic biases from the EVC.

The occipitotemporal cortex, however, is not the endpoint of visual processing hierarchy. Rather, there is accruing evidence that the occipitotemporal cortex is an intermediate stage of an expansive network, which relays feedforward signals to the anterior temporal lobe (ATL) and receives its feedback (Chan et al., 2011; Kravitz et al., 2013). Regions along the ventral visual pathway, from EVC through occipitotemporal areas to ATL, are orchestrated to represent visual stimuli at different depths of cognitive specificity. Face perception epitomizes this neural orchestration: the occipital face area, a patch located in the lateral surface of the inferior occipital gyrus, is the first node of a face-processing network tuned to sensory alterations of facial constituent features but insensitive to facial identity (Rotshtein et al., 2005). By contrast, the fusiform face area is sensitive to integrated facial configuration and identity but indifferent to physical changes of features (Grill-Spector et al., 2004). Finally, at the apex of the processing stream (the ATL), there is accumulating evidence of a region underpinning the linkage between facial representation and person-specific semantic knowledge (Collins and Olson, 2014). In a similar vein to face perception, the visual system might analogously deploy a constellation of cortical regions to represent different aspects of size representation. The EVC and lateral occipital cortex (LOC) seem to underpin sensory-retinal and real-world sizes, respectively. However, it remains unknown whether the most rostral parts of the temporal cortex, the ATL structure, contribute to representing object size and how they share the division of labor with posterior regions. Given the established role of the ATL as a “hub” where modality-specific channels converge to form modality-invariant conceptual knowledge (Patterson et al., 2007; Lambon Ralph, 2014), this warrants speculation of its participation in size processing at a more conceptual level.

In the present study, we explored two principal issues regarding the neurocognitive architecture of size knowledge. First, with a “virtual lesion” approach using continuous theta burst stimulation (cTBS) that suppresses a targeted area, we explored whether and how the ATL and LOC contribute to different facets of object size. Second, rather than adopting a passive “bottom-up” framework focusing on how the visual system passively reacts to size, we manipulated the “top-down” task requirement to in-

vestigate how the brain orchestrates the involvement of different areas contingent upon different cognitive states, enabling inference about the dynamics between behavioral goal and the neural substrate of size representation.

Materials and Methods

Participants. Twenty-four volunteers (Experiment 1: $n = 12$, 9 females, mean age: 25 ± 4 years; Experiment 2: $n = 12$, 7 females, age: 24 ± 5 years) gave informed consent. All reported right-handedness, had normal vision, completed safety screening for TMS and MRI before the experiment, and reported no history of neurological disease/injury. This study was reviewed and approved by the local research ethics committee.

Apparatus. In the initial session, we acquired a high-resolution T1-weighted structural image for each participant using a 3T Philips Achieva scanner and an 8-element head-array coil, with in-plane resolution of 0.94 mm and slice thickness of 0.9 mm. In the subsequent sessions, we conducted the transcranial magnetic stimulation (TMS) experiments. Visual stimuli were presented using MATLAB with Psychophysics (Brainard, 1997; Pelli, 1997) on a computer monitor (29×39.5 cm; 75 Hz refresh rate; 1024×768 resolution). Participants’ head position was stabilized with a chin-rest, keeping a viewing distance of 57 cm from the screen. Brain stimulation was applied via a Magstim Super Rapid² system and with a figure-of-eight coil (70 mm) and guided using a frameless stereotaxic neuronavigation system (Brainsight 2, Rogue Research) to ensure precise localization (for details of TMS protocol, see below).

Design and stimuli. For the psychophysical experiments, we modified a paradigm that has been used to probe the interplay between sensory-retinal versus real-world size in object recognition (Konkle and Oliva, 2012b). In each trial, we presented the images of two real-world objects (one canonically large in volume, the other small) at two different sizes on the screen. The real-world size of the objects could be congruent (e.g., a small stapler and a big coach) or incongruent (e.g., a big stapler and small coach) with their visual size. In separate experiments, we required participants to make judgments about on-screen size of the visual stimuli (Experiment 1: “regardless of object identity, which image looks smaller/bigger on the screen?”) or their real-world size (Experiment 2: “in the real world, which object has a canonically smaller/bigger volume?”). The merit of this paradigm is that it elicits a robust Stroop-type behavioral effect that serves as a proxy allowing us to gauge the impact of size knowledge on visual processing. This characteristic was particularly crucial for assessing automatic access to canonical size representation, which we sought to address its relationship with the LOC in Experiment 1 in which there was little incentive for participants to explicitly retrieve information of real-world size.

In Experiment 1, we manipulated the target of TMS (ATL vs LOC vs the vertex as a control site), visual stimuli (objects vs numbers as control stimuli; for illustration, see Fig. 1a), task (smaller vs bigger), and congruency of sizes (congruent vs incongruent) as the repeated-measure factors. To ensure the specificity of anterior–temporal and lateral–occipital TMS to the processing of object stimuli, we included a control condition based on number stimuli. Numbers have long been exploited in neuropsychological and TMS investigations as control stimuli to pit against semantic/object processing while controlling for overall task difficulty (e.g., Halpern et al., 2004; Chiou et al., 2014). In this control condition, we presented a pair of Arabic digits in each trial and manipulated the congruency between numerical and font sizes (e.g., congruent: “₂ 6”, incongruent: “₂ 6”). This Stroop-type display induces a comparable behavioral effect to that observed in the object size task but operates through distinct mechanisms supported by the inferior parietal lobule that underpin numerical quantity (Kaufmann et al., 2005). In each separate session, we stimulated one of the three sites, and participants judged on-screen size (responding to the visually smaller/bigger item of the pair, performed in separate blocks of trials) of the two types of stimuli (objects or numbers, also in separate blocks). The orders of the three TMS targets, the “smaller versus bigger” tasks, and the “objects versus numbers” stimuli were counterbalanced across participants.

Each trial began with a black fixation dot on a white background (250 ms), followed by the target image (either a pair of objects or numbers)

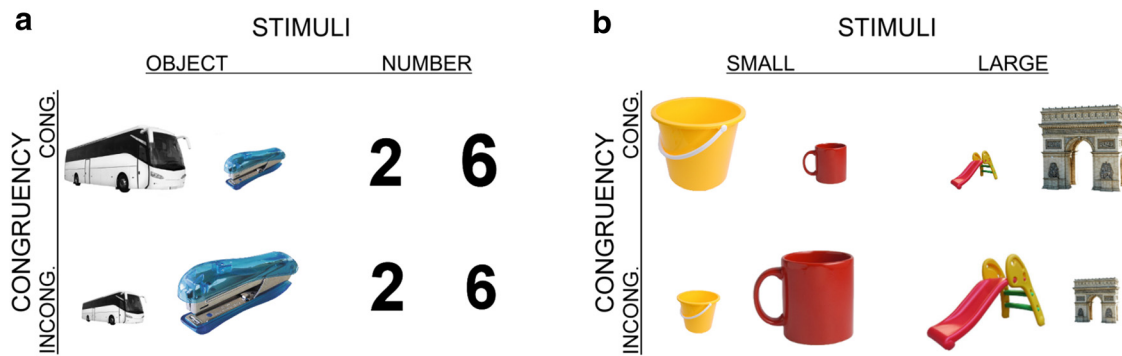


Figure 1. *a*, Example stimuli from Experiment 1, which varied stimulus type (objects vs. numbers) and congruency. *b*, Example stimuli from Experiment 2, which manipulated stimulus type (small vs. large objects) and congruency.

presented for 4 s or until a response was detected. In different blocks, participants compared the visual size of the stimuli and responded to the one that appeared smaller/larger by pressing a designated button using their left/right index finger. There was a 750 ms interval between trials in which feedback of accuracy of the preceding trial was presented. We stressed that object identity and numerical magnitude were irrelevant to the task and should be ignored. Instead, we emphasized focusing on visual size and to respond as quickly as possible while maintaining accuracy.

For the object condition, we selected 20 images of canonically small objects (e.g., stapler, coin, etc.) and 20 images of large objects (e.g., coach, piano, etc.). Like Konkle and Oliva (2012b), we avoided using objects with internal holes or elongated shapes. Care was taken to control for the area of image pixels, ensuring that size ratio was comparable between conditions. Specifically, we ascertained that, when the two object images of each trial were presented on a 1024×768 screen, the ratio of their pixel area was equated between the congruent and incongruent conditions. Independent *t* test showed that the visually smaller stimuli differed highly significantly from their accompanying visually larger stimuli in both the congruent and incongruent conditions (both *p* values $< 1 \times 10^{-8}$), indicating readily discernible image sizes in both conditions. Importantly, the ratios of visually “smaller versus larger” stimuli were matched between the two conditions: on average, the pixel counts of visually larger stimuli were 1.85 ± 0.35 and 1.82 ± 0.57 times greater than those of the visually smaller items in the congruent and incongruent condition, respectively, with no difference between ratios ($t < 1$, $p = 0.83$). For the number condition, the digits were shown in Arial bold format, with a font size of 300 for the visually smaller and 330 for larger stimuli, on a 1024×768 display.

We counterbalanced all experimental parameters for the stimuli so that each individual stimulus, be it an object or a number, was equally likely to be shown in a visually small/large size, situated on the left/right side of the screen, responded to by the left/right hand, and presented in the congruent/incongruent condition. There were 8 blocks of 40 trials in each session (4 blocks of each task), yielding totally 320 trials in the experiment (in the object/number condition, each congruency condition contained 80 trials). Each block consisted of an equal number of congruent and incongruent trials, randomly intermingled.

In Experiment 2, the design was similar to Experiment 1 with a small number of modifications. We manipulated the target of TMS (the ATL, the LOC, and the control site vertex), real-world size (small graspable objects vs. large nongraspable objects, in separate blocks; see Fig. 1*b*), task (smaller vs. bigger), and congruency (congruent vs. incongruent) as the repeated-measure factors. Rather than pairing a small graspable and a large nongraspable object together in each trial's display as in Experiment 1, here we presented them in separate blocks to explore whether stimulation to the LOC would selectively affect decisions to small objects, as there is evidence that this region exhibits preference to small objects over large ones (Konkle and Oliva, 2012a). The control number condition was not included. The trial procedure was identical to Experiment 1, but the task was to indicate which one of the pair is canonically smaller/larger in

real world. We again controlled for the relative pixel area between the visually smaller and larger stimuli in each display, ensuring comparable visual ratios and equal difficulty between conditions. In all conditions, the visually smaller stimuli differed highly significantly from their accompanying larger stimuli in pixel area (all *p* values $< 1 \times 10^{-6}$). Comparisons of the ratio of visually smaller versus larger stimuli showed that, for both small and large objects, the ratios of pixel area between visually small and large items did not differ between congruent and incongruent conditions (both *t* values < 1 , both *p* values > 0.51). There were 8 blocks of 40 trials in each session, giving 80 trials in each congruency condition. The order of conditions and settings for stimuli were fully counterbalanced across participants.

TMS procedure. For brain stimulation, we adopted offline cTBS using a Magstim Rapid² system and a 70 mm figure-of-eight induction coil. cTBS was delivered onto the targeted site in repeated trains of 200 bursts (3 magnetic pulses per burst; 50 Hz) with an intertrain interval of 200 ms (5 Hz); the stimulation lasted for 40 s, with a total number of 600 magnetic pulses (Huang et al., 2005). Participants received cTBS before the cognitive tasks, and their performance was probed immediately following stimulation. This offline cTBS approach avoids nonspecific interferences, such as discomfort, noise, muscle twitches, and so on, that online TMS (i.e., concomitant stimulation during task execution) usually produces and is suggested to be effective for probing high-level cognitive functions (Sandrini et al., 2011). The stimulation was set at 80% of resting motor threshold (the minimum stimulation intensity on the motor cortex that caused a visible finger twitch; for testing individual resting motor threshold, we applied single-pulse stimulation to the left primary motor cortex; the value was defined as the minimum strength sufficing to trigger visible twitches in the right abductor pollicis muscle on six out of ten contiguous trials). In Experiment 1, the averaged intensity of stimulation was $42 \pm 5\%$ of the stimulator maximum output (range: 37%–52%); in Experiment 2, it was $43 \pm 3\%$ of the stimulator output (range: 36%–46%).

Target sites for cTBS were localized individually based on T1-weighted MR structural scan and cerebral–scalp coregistration. Neuroanatomical definitions for each site were based on relevant neuroimaging studies exploring the neural correlates of object size and conceptual knowledge: For lateral occipital stimulation, we selected the peak activation of a lateral occipital cluster that exhibited preferential responses to smaller real-world size over larger size (Talairach coordinates: $-42, -61, -2$) (Konkle and Oliva, 2012a). For anterior temporal stimulation, we selected the peak activation of a ventral ATL cluster that showed modality-independent responses when participants were engaged in semantic processing on visual and auditory stimuli (MNI coordinates: $-36, -9, -36$) (Visser and Lambon Ralph, 2011). For each individual, we normalized their structural image into the standardized space of MNI system using SPM8 (Wellcome Department of Imaging Neuroscience, London) and then converted the coordinates of our literature-defined LOC and ATL sites to derive the corresponding coordinates in each participant's anatomical native space. As the location of the ATL site is slightly too ventral and medial to be accessed by stimulation on the scalp, we adjusted

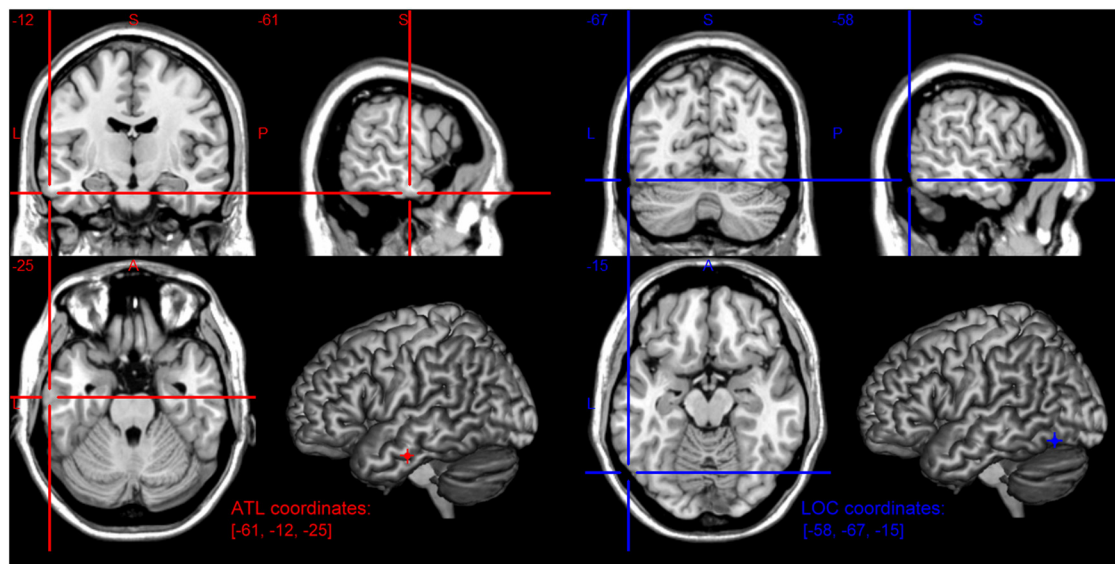


Figure 2. Target stimulation sites pinpointed on the MNI cortical template. Red represents ATL. Blue represents LOC.

the coordinates of this stimulation site based on individual anatomy, making it more lateral and dorsal to the original site and hence accessible to TMS. Meticulous care was taken to strike a good balance between adjacency to the original converted site and accessibility to the more lateral scalp stimulation point. In Experiment 1, the averaged MNI coordinates of the ATL and LOC sites across participants were as follows: -61 ± 3 , -12 ± 3 , -25 ± 4 and -58 ± 5 , -67 ± 5 , -15 ± 5 , respectively (see Fig. 2). In Experiment 2, the averaged MNI coordinates of the ATL and LOC sites across participants were as follows: -61 ± 3 , -14 ± 2 , -25 ± 4 and -57 ± 4 , -69 ± 5 , -13 ± 4 , respectively. The control site vertex was defined as the midpoint between each individual's nasion and inion, along the sagittal midline of the scalp.

Before the behavioral experiments, we performed a coregistration procedure mapping the cerebral site of TMS target of each session onto the corresponding point on the scalp surface using theBrainsight neuro-navigation system, which tracked the position of the coil during stimulation and allowed online adjustment to achieve precise positioning. For all three stimulation sites, the coil was placed tangentially to the scalp with the handle pointing posteriorly (parallel to the rostrocaudal axis). For each individual, the TMS sessions were separated by at least 48 h.

Results

Errors (1.8% in Experiment 1, 1.9% in Experiment 2) and outliers (reaction times [RTs] faster than 100 ms or slower than 3 SDs above the condition mean; 4.8% in Experiment 1, 4.6% in Experiment 2) were excluded before analysis for RTs. Consistent with previous demonstrations (Konkle and Oliva, 2012b), the behavioral index that taps into automatic activation of object size (i.e., the size congruency effect) was only evident in RT, presumably because of swift processing of size in these simple binary-choice tasks (especially for Experiment 1 where decision was required at a sheer perceptual level). The nature of such tasks led to ceiling-level accuracy in all conditions. Therefore, we focused on the RT data as this was the dependent measure in which the congruency effect manifested itself.

For Experiment 1 (on-screen visual size judgment), we undertook separate three-way repeated-measure ANOVAs for the object and control number stimuli, including within-participant factors of stimulation site (ATL, LOC, and vertex), task (small, big), and congruency (congruent, incongruent). Analysis for responses to objects revealed a significant main effect of congruency ($F_{(1,11)} = 34.42$, $p < 0.001$, $\eta_p^2 = 0.75$) and a significant

task \times congruency interaction ($F_{(1,11)} = 21.34$, $p = 0.001$, $\eta_p^2 = 0.66$). Crucially, we found a significant stimulation site \times congruency interaction ($F_{(2,22)} = 9.42$, $p = 0.001$, $\eta_p^2 = 0.46$; see Fig. 3a). This significant interaction suggests that the pattern of congruency effect might differ between cTBS conditions. To test this speculation, we performed a posteriori comparisons (paired-sample *t* test) by stimulation site, inspecting the congruency effect separately for each cTBS location. Results revealed that perceptual decision on the incongruent trials was significantly slower than on the congruent trials in all three conditions of stimulation sites (all *p* values < 0.02). However, as evident in the inset box of Figure 3a, the interaction originated from the fact that the magnitude of congruency effect, indexed by the difference of incongruent RTs minus congruent ones, was significantly reduced when cTBS was applied to the LOC (20 ms), compared with the same effect following stimulation to the control vertex site (36 ms; “LOC vs vertex”: $p = 0.009$, Cohen's $d = 0.7$) and to the ATL (48 ms; “LOC vs ATL”: $p = 0.003$, Cohen's $d = 1.0$). The strength of congruency effect did not differ between vertex and ATL stimulation ($p = 0.1$, Cohen's $d = 0.5$). As an exploratory test, we analyzed the interaction by performing a posteriori comparisons by congruency, comparing between cTBS sites for each congruency condition; results revealed no reliable difference (all *p* values > 0.07). Together, the pattern of statistics indicates the source of the interaction to be the significant attenuation of congruency effect driven by LOC stimulation relative to the two other sites.

The pattern of data was markedly different in the responses to numbers. The ANOVA only revealed significant main effects of task ($F_{(1,11)} = 5.41$, $p = 0.04$, $\eta_p^2 = 0.33$) and, pertinent to our interest, congruency ($F_{(1,11)} = 39.27$, $p < 0.001$, $\eta_p^2 = 0.78$). The congruency effect indicated that reacting to incongruent numeral pairs (508 ms) took a significantly longer time than for congruent stimuli (474 ms). No other statistics of this analysis reached significance; particularly, none of the effects involving stimulation site was significant (all other *p* values > 0.09).

For Experiment 2 (real-world actual size judgment), we performed a four-way repeated-measure ANOVA on the RTs, including within-participant factors of stimulation site (ATL, LOC, and vertex), stimuli (small graspable objects, large nongrasbable

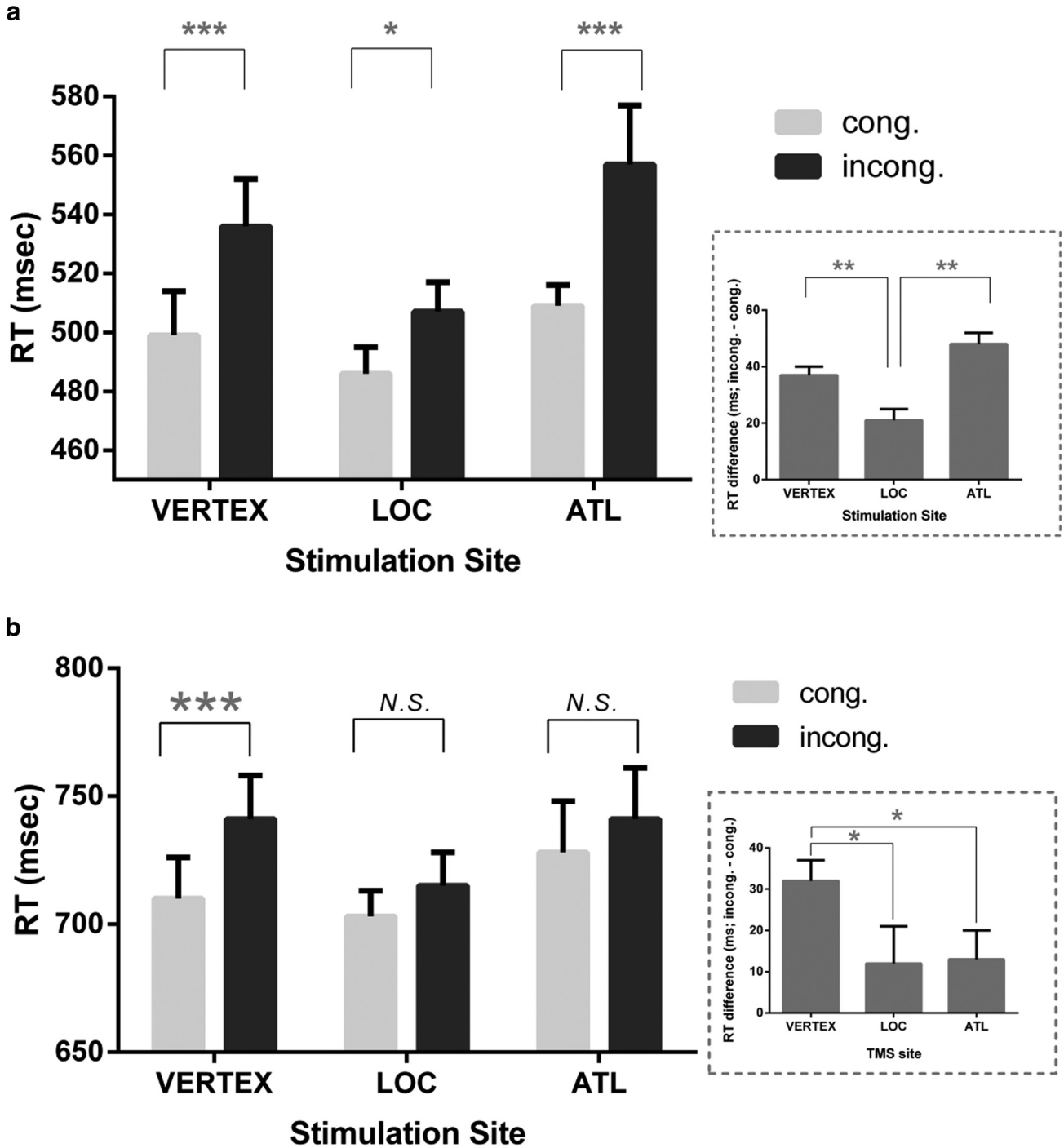


Figure 3. *a*, Experiment 1: the significant stimulation site × congruency interaction, plotted as a function of RTs. Inset box, The bar represents the TMS effect magnitude for each stimulation site. *b*, Experiment 2: the significant stimulation site × congruency interaction, plotted as a function of RTs. Inset box, The bar represents the TMS effect magnitude for each stimulation site. Error bars indicate SEM. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$. *N.S.*, Not significant.

objects), task (small, big), and congruency (congruent, incongruent). We found significant main effects of task ($F_{(1,11)} = 26.73, p < 0.001, \eta_p^2 = 0.70$) and congruency ($F_{(1,11)} = 11.93, p = 0.005, \eta_p^2 = 0.52$). Neither the main effect nor any interaction involving stimuli was significant (all p values > 0.07), indicating parallel patterns in small graspable and big nongrasbable objects. More importantly, there was a significant stimulation site × congruency interaction ($F_{(2,22)} = 3.65, p = 0.04, \eta_p^2 = 0.25$; see Fig. 3*b*). To identify the source of this interaction, we performed a posteriori comparisons by stimulation site. In the control condition in

which participants received stimulation to the vertex, we found a congruency effect, indexed by significantly slower incongruent RTs than congruent ones ($p < 0.001$). Critically, this robust effect was eliminated following stimulation to the LOC ($p = 0.23$) and to the ATL ($p = 0.11$). Illustrated in the inset box of Figure 3*b*, further analysis revealed that the magnitude of congruency effect in the control vertex condition (31 ms), serving as a baseline, was significantly greater than the effect after LOC stimulation (11 ms; “LOC vs vertex”: $p = 0.01$, Cohen’s $d = 0.8$) and ATL stimulation (12 ms; “ATL vs vertex”: $p = 0.01$, Cohen’s $d = 0.9$). We also

performed a posteriori tests by congruency to check whether RTs differed between sites; results revealed no reliable difference for any contrast (all p values >0.23). This suggests the driving force of the interaction to be the congruency effect persevering under vertex stimulation but selectively getting eradicated by LOC and ATL stimulation.

For completeness, we also conducted the same analysis on the accuracy data. There were no significant effects of TMS, consistent with previous suggestions (e.g., Pobric et al., 2007) that, unlike neuropsychological data, the impact of stimulating the ATL is more likely to manifest in RT than accuracy changes.

Discussion

By gauging the modulatory impact of theta burst stimulation on a robust behavioral indicator of size processing (the size congruency effect), we established that knowledge about object size is underpinned by a synergistic neural network entailing the LOC and ATL, with their division of labor weighted by behavioral intent: When the task required discerning on-screen size regardless of object identity, perturbing the LOC attenuated the congruency effect while perturbing the ATL left the effect intact. By contrast, when the task demanded discerning real-world size (necessitating explicit object recognition), disrupting the LOC or ATL both eradicated the otherwise reliable effect. By including both a control site (the vertex) and control stimuli (numbers), our paradigm allowed us to rule out nonspecific confounding factors. Specifically, including control stimuli ensured that the disruptive impact selectively occurred to object stimuli rather than a blanket effect inhibiting neural processing for all stimuli, whereas including a control site ensured that the effect can only be ascribed to sites causally relevant to a cognitive function. Our results support the anatomical and functional specificity of TMS and highlight the distinct profiles of LOC and ATL in reaction to different task-based intents.

To dissect the distinct functional profiles, we first consider the representational composition of size knowledge. To access the canonical size of a particular object (e.g., knowing a pea is small), two representations are entailed: object identity and modality-based size attributes (typically extracted from visual statistical structures, such as a pea extends a tiny range of visual angle from an usual viewing distance, although tactile information can be another source). This visually based size attribute should not be conflated with sensory-retinal size that reflects the area a stimulus spans across the retina and is coded in the EVC. Instead, it represents statistical regularity of frequently co-occurring features (e.g., objects small/large in volume usually cover foveal/peripheral vision and are rounded/boxy in shape) (Konkle and Oliva, 2011; Nasr et al., 2014). Two lines of evidence suggest that the occipitotemporal cortex underpins such visual regularities: First, the “lateral-small versus medial-large” size preference map (the zone our LOC site of stimulation falls in) differentiates object stimuli based on their canonical sizes, even when the task did not require explicit retrieval of size information (e.g., detecting repetition of images) (Konkle and Oliva, 2012a; Konkle and Caramazza, 2013). This profile of neural response is akin to the behavioral congruency effect: as we demonstrated in Experiment 1, when exposed to object stimuli, participants showed involuntary access to canonical sizes, even when the context provided little incentive to recognize those objects (see also Konkle and Oliva, 2012b). Second, neural signatures of acquiring statistical regularities of visual features have been identified in the LOC (Turk-Browne et al., 2009). However, when a context requires establishing linkages between size attributes and explicitly iden-

tified objects (Experiment 2), the ATL becomes an indispensable node of the network that maintains the congruency effect. This is consistent with the ATL’s hypothesized role as the culmination of object processing stream where abstract identity is distilled (Kravitz et al., 2013) and cooperation between the ATL and modality-based areas to generate coherent transmodal concepts (Patterson et al., 2007; Lambon Ralph, 2014).

Our experimental design allowed exploration of the balance between the ATL and LOC under different circumstances: The extent of ATL involvement varies with task demands: its contribution is crucial when the context stipulates access to object identity and fine-grained size comparison at a conceptual level, as required by Experiment 2 (e.g., is a raspberry smaller than a strawberry?). The ATL’s contribution is significantly reduced (such that targeted neurostimulation has no effect) when the task only requires size information at a “shallower” perceptual level. By marked contrast, the LOC contributes to the computation of object size regardless of whether the task is perceptually (Experiment 1) or conceptually focused (Experiment 2). Overall, these suggest an adaptive network in which the LOC automatically computes configuration regularity to derive canonical size in the presence of object stimuli, whereas the ATL is flexibly summoned to secure conceptual mapping between objects and their canonical sizes when required by the task. This ATL-LOC synergistic relationship and their distinct functional responses to top-down signal resonate with recent findings concerning the interplay between “bottom-up” and “top-down” systems in object recognition (Harel et al., 2014; McKee et al., 2014). These recent studies have moved away from classical feedforward theory of visual circuitry in which a hierarchical system reacts to the input of a preceding stage in a passive and bottom-up fashion. Instead, these contemporary accounts emphasize the importance of top-down modulation from the high-level cortical zones, such as the anterior temporal and prefrontal cortices.

It is noteworthy that the origin of top-down signals may well be the prefrontal cortex, given its established role in generating cognitive control and modulatory messages (e.g., Lee and D’Esposito, 2012). Interestingly, in response to (presumably) prefrontal-mediated modulatory feedback signals, different sections along the occipital-to-temporal stream seem to have differential characteristics. The caudal sections (e.g., the LOC), albeit still sensitive to modulation, respond in a more rigidly “stimulus-driven” fashion to visual input, not caring whether the task-relevant dimension is sensory-visual size or abstract-conceptual size. On the contrary, the more rostral regions (e.g., the ATL) form a partnership with posterior areas when a context demands conceptual-level processing.

Some tantalizing clues arise from the neurocognitive effect that ATL and LOC stimulation causes to the whole neural system. While bearing in mind that the interaction resulted from cortical stimulation modulating the absence/presence of congruency effects rather than direct differences between stimulation sites, we observed some consistent patterns across Experiments 1 and 2 (Fig. 3). Relative to the vertex baseline, LOC stimulation interfered with the congruency effect via shortening incongruent RTs. In contrast, ATL stimulation operated via prolonging congruent RTs. A future direction that may give insight into these processes would be to investigate functional/effective connectivity between the ATL and LOC in response to congruent and incongruent displays and to examine how connectivity patterns may be related to behavioral performance.

In a broader sense, the present findings and the growing consensus of rapid top-down interactivity concord with contempo-

rary hypotheses about the nature of semantic representation. Specifically, the hub-and-spokes hypothesis (for a recent review, see Lambon Ralph, 2014) suggests that the synergy between modality-specific “spokes” stored in association cortices and a transmodal “hub” housed in the ATL is key to the genesis of coherent, generalizable concepts. Convergent evidence from patients, neuroimaging, and neurostimulation indicates that the trans-modality computation of this representational hub is implemented in the ventrolateral portions of the ATL (Lambon Ralph, 2014). The locus of our ATL stimulation falls within this ventrolateral region, indicating that visual cognition benefits from interaction with this semantic-related region when the task demands explicit retrieval of visually based semantic attributes. This conclusion accords with three lines of inquiries. First, intracranial-electrode recording has revealed rapid activation of the ATL (~130 ms after stimulus), which drives semantically related feedback to posterior visual cortices (Chan et al., 2011). Second, computational instantiations of the hub-and-spokes model demonstrate that the responses of visual-feature units are shaped through feedback from a higher transmodal-hub layer (Rogers et al., 2004). Third, perhaps most strikingly, the performance of patients with semantic dementia due to atrophy centered on the ATL mirror the current cTBS data (Ikeda et al., 2006; Lambon Ralph et al., 2010): the patients remain able to match visually presented items when the stimuli vary purely in size but become severely impaired when the task requires object recognition (e.g., matching across different exemplars from the same category).

References

- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Chan AM, Baker JM, Eskandar E, Schomer D, Ulbert I, Marinkovic K, Cash SS, Halgren E (2011) First-pass selectivity for semantic categories in human anteroventral temporal lobe. *J Neurosci* 31:18119–18129. [CrossRef Medline](#)
- Chiou R, Sowman PF, Etschell AC, Rich AN (2014) A conceptual lemon: theta burst stimulation to the left anterior temporal lobe untangles object representation and its canonical color. *J Cogn Neurosci* 26:1066–1074. [CrossRef Medline](#)
- Collins JA, Olson IR (2014) Beyond the FFA: the role of the ventral anterior temporal lobes in face processing. *Neuropsychologia* 61:65–79. [CrossRef Medline](#)
- Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 7:555–562. [CrossRef Medline](#)
- Halpern CH, Glosser G, Clark R, Gee J, Moore P, Dennis K, McMillan C, Colcher A, Grossman M (2004) Dissociation of numbers and objects in corticobasal degeneration and semantic dementia. *Neurology* 62:1163–1169. [CrossRef Medline](#)
- Harel A, Kravitz DJ, Baker CI (2014) Task context impacts visual object processing differentially across the cortex. *Proc Natl Acad Sci U S A* 111:E962–E971. [CrossRef Medline](#)
- Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005) Theta burst stimulation of the human motor cortex. *Neuron* 45:201–206. [CrossRef Medline](#)
- Ikeda M, Patterson K, Graham KS, Ralph MA, Hodges JR (2006) A horse of a different colour: do patients with semantic dementia recognise different versions of the same object as the same? *Neuropsychologia* 44:566–575. [CrossRef Medline](#)
- Kaufmann L, Koppelstaetter F, Delazer M, Siedentopf C, Rhomberg P, Golaszewski S, Felber S, Ischebeck A (2005) Neural correlates of distance and congruity effects in a numerical Stroop task: an event-related fMRI study. *Neuroimage* 25:888–898. [CrossRef Medline](#)
- Konkle T, Caramazza A (2013) Tripartite organization of the ventral stream by animacy and object size. *J Neurosci* 33:10235–10242. [CrossRef Medline](#)
- Konkle T, Oliva A (2011) Canonical visual size for real-world objects. *J Exp Psychol Hum Percept Perform* 37:23–37. [CrossRef Medline](#)
- Konkle T, Oliva A (2012a) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74:1114–1124. [CrossRef Medline](#)
- Konkle T, Oliva A (2012b) A familiar-size Stroop effect: real-world size is an automatic property of object representation. *J Exp Psychol Hum Percept Perform* 38:561–569. [CrossRef Medline](#)
- Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci* 17:26–49. [CrossRef Medline](#)
- Lambon Ralph MA (2014) Neurocognitive insights on conceptual knowledge and its breakdown. *Philos Trans R Soc Lond B Biol Sci* 369:20120392. [CrossRef Medline](#)
- Lambon Ralph MA, Sage K, Jones RW, Mayberry EJ (2010) Coherent concepts are computed in the anterior temporal lobes. *Proc Natl Acad Sci U S A* 107:2717–2722. [CrossRef Medline](#)
- Lee TG, D’Esposito M (2012) The dynamic nature of top-down signals originating from prefrontal cortex: a combined fMRI-TMS study. *J Neurosci* 32:15458–15466. [CrossRef Medline](#)
- McKee JL, Riesenhuber M, Miller EK, Freedman DJ (2014) Task dependence of visual and category representations in prefrontal and inferior temporal cortices. *J Neurosci* 34:16065–16075. [CrossRef Medline](#)
- Nasr S, Echarria CE, Tootell RB (2014) Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *J Neurosci* 34:6721–6735. [CrossRef Medline](#)
- Patterson K, Nestor PJ, Rogers TT (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci* 8:976–987. [CrossRef Medline](#)
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442. [CrossRef Medline](#)
- Pobric G, Jefferies E, Ralph MAL (2007) Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. *Proc Natl Acad Sci USA* 104:20137–20141. [CrossRef Medline](#)
- Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K (2004) Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychol Rev* 111:205–235. [CrossRef Medline](#)
- Rotshtein P, Henson RN, Treves A, Driver J, Dolan RJ (2005) Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8:107–113. [CrossRef Medline](#)
- Sandrini M, Umiltà C, Rusconi E (2011) The use of transcranial magnetic stimulation in cognitive neuroscience: a new synthesis of methodological issues. *Neurosci Biobehav Rev* 35:516–536. [CrossRef Medline](#)
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK (2009) Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* 21:1934–1945. [CrossRef Medline](#)
- Visser M, Lambon Ralph MA (2011) Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J Cogn Neurosci* 23:3121–3131. [CrossRef Medline](#)