Cerebral Cortex, 2023, **33**, 5557–5573
https://doi.org/10.1093/cercor/bhac442
Advance access publication date 5 December 2022
Original Article

Novel objects with causal event schemas elicit selective responses in tool- and hand-selective lateral occipitotemporal cortex

Anna Leshinskaya[1,2,†,*], Mira Bajaj[1,3,†], Sharon L. Thompson-Schill[1]

[1]Department of Psychology, University of Pennsylvania, 425 S. University Ave, Stephen A Levin Building, Philadelphia, PA 19104, United States,
[2]Center for Neuroscience, University of California, Davis, 1544 Newton Court, Room 209, Davis, CA, United States,
[3]The Johns Hopkins University School of Medicine, 733 N Broadway, Baltimore, MD 21205, United States

*Corresponding author: Center for Neuroscience, University of California, Davis, 1544 Newton Court, Room 209, Davis, CA 95618, United States. Email: anna.leshinskaya@gmail.com

†Anna Leshinskaya and Mira Bajaj shared joint first authorship.

Tool-selective lateral occipitotemporal cortex (LOTC) responds preferentially to images of tools (hammers, brushes) relative to non-tool objects (clocks, shoes). What drives these responses? Unlike other objects, tools exert effects on their surroundings. We tested whether LOTC responses are influenced by event schemas that denote different temporal relations. Participants learned about novel objects embedded in different event sequences. Causer objects moved prior to the appearance of an environmental event (e.g. stars), while Reactor objects moved after an event. Visual features and motor association were controlled. During functional magnetic resonance imaging, participants viewed still images of the objects. We localized tool-selective LOTC and non-tool-selective parahippocampal cortex (PHC) by contrasting neural responses to images of familiar tools and non-tools. We found that LOTC responded more to Causers than Reactors, while PHC did not. We also measured responses to images of hands, which elicit overlapping responses with tools. Across inferior temporal cortex, voxels' tool and hand selectivity positively predicted a preferential response to Causers. We conclude that an event schema typical of tools is sufficient to drive LOTC and that category-preferential responses across the temporal lobe may reflect relational event structures typical of those domains.

Key words: category selectivity; causality; fMRI; lateral temporal lobe; tool-selective responses.

Introduction

Among the most classic findings in cognitive neuroscience is the observation of an area in lateral occipitotemporal cortex (LOTC) that responds preferentially to images of tools (hammers, rakes, brushes) relative to other object kinds, including artifacts like clocks and shoes (Martin et al. 1996; Chao et al. 1999; Valyear et al. 2007; Vingerhoets et al. 2009; Anzellotti et al. 2011; Bracci et al. 2011; Simmons and Martin 2011; Garcea and Mahon 2014). Convergently, lesions to LOTC impair the recognition and naming of tools (Damasio et al. 1996, 2004; Tranel et al. 1997; Brambati et al. 2006; Campanella et al. 2010; Buxbaum et al. 2014).

However, the functional role of Tool LOTC remains a puzzle. It is not characterized by a profile of specific shape selectivity. For example, functionally localized Tool LOTC also responds to point-light displays of mechanical motion more than biological motion (Beauchamp et al. 2003). Most strikingly, Tool LOTC responds not only to tools but also to images of hands, which are highly dissimilar in form (Bracci et al. 2011, 2016, 2017). It responds both to elongated and non-elongated tools more than non-tools (Chen et al. 2018). Neither is its role specific to visual recognition. Its preferential response to tools persists with auditorily (Noppeney et al. 2006) and visually (Peelen et al. 2013) presented names as well as item-elicited sounds (Lewis et al. 2005; Doehrmann et al. 2008) and is conserved in individuals born blind (Mahon et al. 2009, 2010; Peelen et al. 2013; Mattioni et al. 2020).

One hypothesis is that this preferential response is driven by motor interaction experience. After participants learn to physically interact with novel objects, the response of areas close to Tool LOTC increases relative to viewing objects with no action experience (Weisberg et al. 2007). LOTC response is also larger to familiar objects one tends to interact with (e.g. forks) relative to less interactive objects, such as picture frames (Magri et al. 2021). However, its role is unlikely in motor execution per se. Damage to areas that anatomically span Tool LOTC affects the recognition of tool-based actions but less so the execution of such actions across a large meta-analysis of lesion data (Tarhan et al. 2016). In anatomically similar areas, responses to observed actions are similar across diverse physical manners of accomplishing a similar goal (Oosterhof et al. 2010; Wurm and Lingnau 2015; Wurm et al. 2017; Vannuscorps et al. 2018). Furthermore, participants born without hands maintain a preferential response to hands and tools in LOTC (Striem-Amit et al. 2017), suggesting that the overlap between tool and hand responses is not due to their coordination in motor experience. Overall, this work suggests that the preferential responses in LOTC are not simply driven by motor association.

We hypothesize that an important role of Tool LOTC is the retrieval of an event schema, a representation capturing the relations among objects and events that are typical of tools. Such tool-canonical event schemas are ones in which the tool object is

Received: December 6, 2021. **Revised:** October 10, 2022. **Accepted:** October 11, 2022

used to affect its surroundings (influence some outcome event). These schemas are relational and abstract because what matters is how the object and other events are related, while the specific visual and motor features of the objects and the events can vary.

Behavioral findings demonstrate that learning how to interact with man-made objects entails learning schemas about object-outcome relations: Beyond their shapes or manner of interaction, functional roles are central to the way even young children classify novel artifacts (Gopnik and Sobel 2000; Kemler Nelson et al. 2000; Nazzi and Gopnik 2003; Truxaw et al. 2006; Oakes and Madole 2008; Hernik and Csibra 2009; Futó et al. 2010; Träuble and Pauen 2011). Unlike the function of non-tool artifacts like a book or a shoe, the function of tools like hammers or paintbrushes involves a direct effect on a target. We hypothesize that Tool LOTC is sensitive to such tool-specific schemas. In support of a related idea, Bracci and Peelen (2013) found that Tool LOTC responds weakly to objects one acts "on" (like tennis balls or musical instruments) relative to familiar tools, which one acts "with."

Here we test one basic prediction of our hypothesis that tool-canonical event schemas drive Tool LOTC: This area should respond more to objects that move "prior to" an event in their surroundings, rather than "following" the same event. The former temporal event structure is consistent with a relationship in which the object causes the event, as in the canonical tool schema, while the latter is consistent with a relationship in which the object responds to it, which is not typical of tools. Causality entails more than temporal structure, but we based this operationalization on our prior research showing that adults' judgments of objects' causal properties are reliably influenced by the temporal structure of events associated with an object (Leshinskaya and Thompson-Schill 2019). Here we tested whether such a schema manipulation also affects the response of functionally localized Tool LOTC.

Materials and methods

Overview

Prior to functional magnetic resonance imaging (fMRI) scanning, participants were taught about 8 novel shapes that each belonged to 1 of 4 event schema conditions (Fig. 1). We demonstrated the event schemas belonging to the objects by systematically embedding each object in a specific animated sequence consisting of an object movement event (e.g. the object rotating) and an ambient event (e.g. stars appearing in the environment). The principal manipulation was whether the object's role in the animated sequence was best described as a Causer or as a Reactor: Causer objects moved prior to the ambient event, whereas Reactors moved afterwards, keeping the component events otherwise identical (Fig. 1). We orthogonally manipulated whether the object movement was self-generated (Self-Mover) or elicited by an animated hand (Hand-Mover) in order to additionally test the sensitivity of LOTC to the involvement of hand manipulation. The alternative hypothesis that motor interaction drives LOTC would predict an overall greater response to the hand-manipulated objects regardless of their role as Causers or Reactors. A third possibility is an interaction whereby both Causal event schemas and manipulation with a hand are both important in explaining LOTC response. Participants memorized the animated sequence that went with each object, then were scanned with fMRI while viewing the objects and recalling their associated animations from memory.

We identified Tool LOTC using a separate category localizer in which we compared responses to images of familiar tools (hammers, paintbrushes, etc.) and non-tool artifacts (books, shoes, etc.; Fig. 2). This design is in line with the above-described findings of a stronger response in LOTC to objects one "acts with" rather than "acts on" (Bracci et al. 2011). We then assessed whether Tool LOTC responded to the novel object conditions according to the structure of their associated event schema, having controlled for factors including object shape, association with another event, association with a hand movement, familiarity, and motor experience.

Code accessibility and preregistration materials

The experimental and analytical methods used in this experiment were preregistered using the Open Science Framework repository (Leshinskaya et al. 2018). Deviations from the preregistered procedures are noted in the manuscript. Custom code for generating and implementing the experimental procedure is also available in the repository. Custom code for neuroimaging analysis is available upon request.

Participants

For the primary experiment, we report the results of 32 participants; to achieve this planned sample size, we scanned 45 participants. Of these, 8 were excluded due to excessive head motion and 2 ended their scan early due to discomfort. Prior to analysis of the primary experimental data, localizer data were individually inspected for quality. In 3 participants, localizer data were exceptionally noisy with no clear activation for the main contrast of interest (tools vs. non-tools) and were thus excluded on that basis. This resulted in the exclusion of 13 participants, who were replaced to fulfill the counterbalancing design (described below) and meet the pre-planned sample size. This resulted in the final sample of 32 participants (19 female), with a mean age of 22.66 (SD = 3.02, range 18–29).

We also report the results of a behavioral pilot study ($n = 32$), which included 22 female participants with a mean age of 21.03 (range 18–28). An additional group participated in an fMRI follow-up experiment ($n = 12$). Of the participants scanned, one was excluded for excessive motion and one ended the scan early due to discomfort. The final sample had $n = 10$ and a mean age of 22.40 (range 19–27) and 8 were female.

We recruited the above participants from the University of Pennsylvania community using the Experiments@Penn website. Procedures were approved by the Institutional Review Board at the University of Pennsylvania and all participants provided written informed consent prior to participation. Participants were required to be between 18 and 35 years of age, right-handed, and eligible to be scanned with MRI (for fMRI studies). Compensation for participating in fMRI studies was $60 total and behavior $10 per hour.

An additional behavioral follow-up study recruited 93 participants from Amazon's Mechanical Turk website. Procedures were approved by the Institutional Review Board at the University of California, Davis. Participants were required to pass a high accuracy threshold on the task and the included sample had 32 participants with a mean age of 38 years (range 23–62), 15 female and 17 male. Participants were compensated with a $3.00 base pay with an accuracy-based bonus of up to $6.

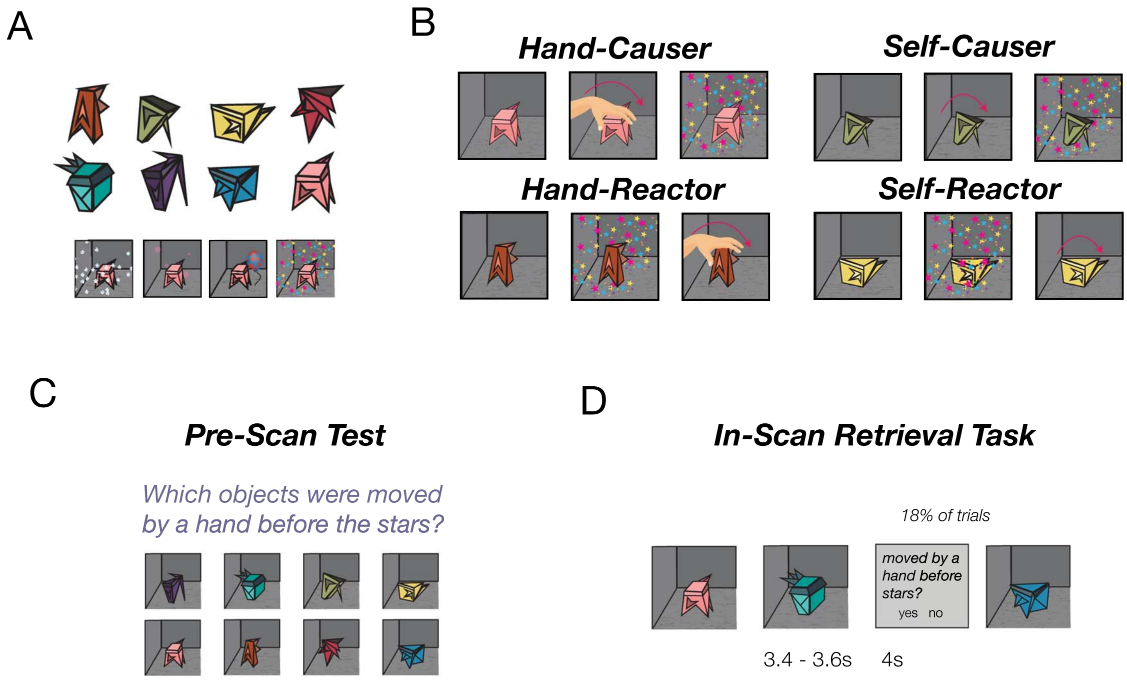

Fig. 1. A) The 8 novel object shapes (top) and still frames from the 4 kinds of ambient events (bottom): snow, bubbles, balloon, and stars. B) Example of the four animation conditions in the experiment. A unique object shape appeared in each distinct event sequence. The Hand-Causer object was moved by a hand prior to an ambient event; the Hand-Reactor object was moved after an ambient event; the Self-Causer object moved on its own prior to an ambient event, and the Self-Reactor moved after an ambient event. The object was present during the entire sequence of events. Shape assignments were counterbalanced across participants and each participant saw 2 instances of each condition, each with one of 2 ambient events selected from the set of 4 shown in A. C) Prior to scanning, participants watched the animated event sequences featuring each object, followed by test questions probing their knowledge, an example of which is shown here. D) During fMRI scanning, still images of each object were shown in randomized order (3.4–3.6 s ISI); 18% of trials were followed by a question probing participants' knowledge of the event sequence that went with the object they had just seen.

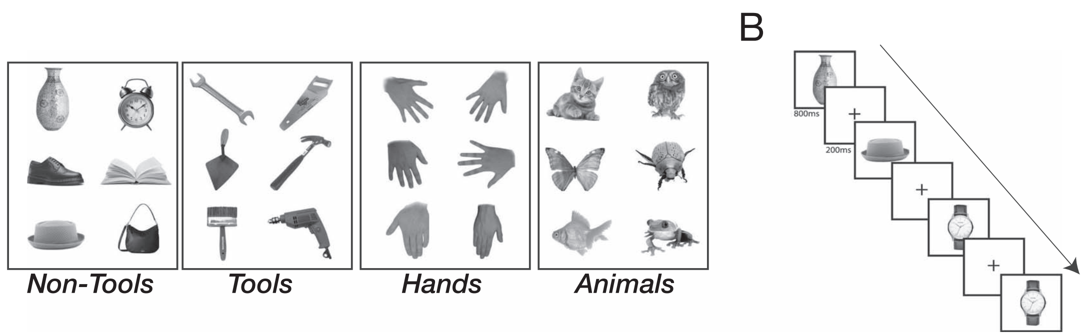

Fig. 2. A) Example images belonging to each of the 4 conditions in the localizer task: Non-Tools, Tools, Hands, and Animals. B) Localizer task procedure, in which images were shown blocked by condition and in randomized order, with a 1 s ISI. Participants hit a button when an identical image repeated twice in a row, as in the last trial pictured here.

Stimuli and materials

Training task

Prior to fMRI scanning, participants learned about the 8 different animated event sequences, each involving a specific, novel, distinctly shaped, and colored geometrical object (Fig. 1). All animated sequences involved a movement of the object and an ambient event (something appearing and disappearing in the background, such as stars or bubbles; Fig. 1), with the object continually present during both events. They belonged to 1 of 4 conditions created by crossing 2 factors: Movement Type and Causality. Movement Type was either Self, in which the object moved on its own (it rotated back and forth) or Hand, in which an animated hand reached in, grabbed the object, and rotated it back and forth. Causality affected the order of events: In the Causer condition, the object moved prior to the ambient event, whereas in the Reactor condition, the object moved afterwards. This was manipulated by reordering the same frames. Crossing these 2 factors created 4 conditions: Self-Causer, Self-Reactor, Hand-Causer, and Hand-Reactor. These 4 conditions had 2 sets of specific animations, which varied in the nature of the ambient event involved (e.g. in set 1, the ambient event was bubbles; in set 2, stars). Each of the 8 animations involved a uniquely shaped object that was systematically mapped to its condition for each participant but counterbalanced across participants (see below).

Animations were created by hand-drawing static frames using Adobe Illustrator and concatenating them into GIFs for each separable component event (i.e. each movement type and each ambient event for each object shape). These were then recombined to make the full event sequences for each condition, ensuring that the components were identical. GIFs were created from individual frame files using Matlab (MathWorks).

All animations took place against a gray background resembling an empty room. Each sequence began with the object smoothly entering the room (13 100-ms frames, total length 1,300 ms). In Causal sequences, the movement event was shown first, followed by the ambient event; in Reactor sequences, this order was reversed. In all cases, each movement and ambient event segment comprised 29 100-ms frames (2,900 ms). The Self-movement event involved the object rotating 10° forward and back; the Hand-movement events involved an animated hand reaching in, grasping the object, and tilting it forward and back, then leaving; both movement events were of equal total duration. Following the movement and ambient events, the object then stayed still for 800 ms. It then repeated the movement and ambient pair and the 800 ms pause 2 more times, for a total of 3, and then smoothly exited the room (2,900 ms). This full set of events, including the repetition, composed the "animated event sequence" for each object shown in the training task.

To control condition comparisons for shape, the 8 geometric objects were assigned to the 4 conditions in counterbalanced fashion with the following procedure: The 8 objects were split into 2 sets of 4, A–D and E–H. Within each set, we ensured that each object appeared in each condition equally often, creating 16 assignment options. We additionally ensured that for each planned comparison (e.g. Causers > Reactor within Movement Type), pairs were perfectly counterbalanced for shape (i.e. equal numbers of comparisons in which A is the Causer and C is the Reactor as vice versa). The 2 sets were themselves paired such that all pair combinations appear equally often in each condition (e.g. A&E, A&F, A&G, and A&H are members of each condition equally often, and so on for the pairings of B, C, and D in place of A). This yielded 16 counterbalancing conditions that were repeated twice.

There were 4 kinds of ambient events: pink bubbles floating across the scene ("bubbles"), snowflakes sweeping from left to right ("snowflakes"), a balloon floating across the screen ("balloon"), and multicolored stars sweeping from left to right ("stars"). Each type of event was created in identical fashion in the context of each object shape. The selection of these events was randomized across participants, but each event appeared equally often in each condition. For each participant, 2 ambient event types were selected (out of the set of 4) and assigned to each shape set. That is, one ambient event was shown with shape set A–D, and the other with E–H. The combination of event and shape counterbalancing conditions was randomized. There were 12 ways to select 2 ambient events from the set of 4 and assign them to the 2 object sets; one of these options was selected randomly for each participant, independently of the object-condition assignments. Overall, these counterbalancing and randomization efforts ensured that the experimental conditions did not systematically differ in terms of the object shapes or ambient events shown in their animated sequences.

Category localizer

Participants performed a category localizer task during fMRI in which they saw grayscale images of objects from 4 categories: Tools (hammers, drills, wrenches, spades, paintbrushes, and saws), Hands, Animals (bats, birds, butterflies, reptiles, small dogs, fish, beetles, salamanders, owls, mice, lobsters, and kittens), and Non-Tools (hats, bags, clocks, watches, books, vases, and shoes); examples of stimuli are shown in Fig. 2. Non-Tools were thus all small, inanimate artifacts which do not have a direct effect on the environment when used in typical fashion. We did not obtain quantitative measures of this factor. Studies of tool-preferential responses do not have a standard "non-tool" comparison condition, but our design choice was designed to avoid both animacy and real-world size confounds, known to separately influence temporal lobe organization (Konkle and Caramazza 2013). Recent approaches typically compare tools to other inanimate objects, such as furniture (Bracci et al. 2011; Striem-Amit et al. 2017); however, we sought to also control for real-world objects size and thus selected small man-made objects. In a related approach, Mahon et al. (2007) defined tools as objects with specific manners of manipulation and compared them with small objects that are arbitrarily manipulated (e.g. wallet, book), but did not find activation differences, only differences in repetition suppression, between these categories. Our operationalization was based on the notion that tools exert effects in the environment and aligns with the findings of Bracci and Peelen (2013) comparing objects one acts "with" rather than "on." For specific stimuli, we used hand and tool images from Bracci et al. (2011) and non-tool images from the small inanimate category from Konkle and Caramazza (2013).

Procedure

Overview

Participants first performed the training task, in which they saw all of the 8 animated sequences and learned to associate each sequence with their participating object shape. They then entered the fMRI scanner, where they reviewed the sequences, performed a retrieval task with the novel objects, and saw the category localizer.

Training task

Participants began the training task 1.5 h prior to scanning. They were instructed to memorize the animated event sequence belonging to each object. The task was implemented using JavaScript and presented in a web browser. They saw each of the 8 animated sequences (featuring each of the 8 objects) one at a time, 15 times over 5 blocks or until criterion (see below). The order of presentation was randomized across blocks but constrained such that each sequence appeared at least twice but not more than 7 times in each block. Each block thus contained 20–28 animated sequences.

After each block, participants were asked 12 questions to assess their learning (Fig. 1C). Each question posed a query regarding the animations they had seen and asked participants to select the objects of which it was true. For example, "which object(s) were moved by a hand before the stars?" Images of all 8 objects were shown in randomized positions and participants responded by selecting any number of the objects. Eight of these questions were specific to both condition and the specific ambient event, such that there was only one correct answer (e.g. "which object(s) tilted on their own before the bubbles?"). The other 4 questions asked more generally about the movement or ambient event types involved, such that there were 4 correct answers ("which object(s) tilted on their own?", "which object(s) were moved by a hand?", "which object(s) involved the bubbles in their videos?", "which object(s) involved the stars in their videos?"). Questions were presented in a randomized order.

At the end of each question set, participants were given an overall score (e.g. "You answered 8/12 questions correctly"). Each block took approximately 5–7 min to complete.

The training task was designed to achieve ceiling performance on all conditions. If participants achieved a score below perfect on the fifth block, they were asked to complete another block of training. Following this additional training, all participants reached criterion (92%); we report obtained average accuracies in Results.

In a separate group of behavioral-only pilot participants ($n = 32$) who only completed the training task, we assessed performance following the first block of training to see if there were any differences by condition, finding the following means: Self-Causer: $M = 80.78\%$, $SD = 12.35$; Hand-Causer: $M = 78.23\%$, $SD = 13.77$; Self-Reactor: $M = 76.08\%$, $SD = 13.04$; Hand-Reactor: $M = 75.27\%$, $SD = 13.39$. Pairwise tests between conditions showed that participants had a higher accuracy for the Self-Causer objects compared to the Self-Reactor objects on the questions following the first block of training ($t(30) = 2.09$, $P = 0.045$). We saw a similar trend in the first 13 fMRI participants. We thus adjusted the training task to show more of the Self-Reactor objects earlier in the training sequence (and fewer later). Thus, the Self-Reactor objects were shown 4 times in the first block, whereas the Self-Causer objects were shown 2 times; this was balanced out in the fourth block, where the Self-Reactor objects were shown twice and the Self-Causer 4 times. As we report in Results, this allowed us to equate the conditions on learning performance in the first block while maintaining equal overall amount of training in all of the conditions.

After they completed the training task, participants were given a sheet of printed object images and asked to write above each image a description of its corresponding events (e.g. "moved by a hand before stars appeared"). If they were unsure about any object, they were asked to complete additional training blocks until they felt comfortable. Overall, 8 participants completed additional training blocks (4 participants completed 1 extra block, 1 participant completed 2 extra blocks, and 3 participants completed 3 extra blocks).

fMRI acquisition parameters

MRI data were collected on Siemens Tim-Trio 3-T scanner at the University of Pennsylvania Center for Neuroimaging. A 64-channel coil was used. Structural scans were acquired with an MPRAGE sequence with 0.8 mm isotropic voxels, a field of view of 256 mm, and matrix size of 320 mm × 320 mm × 320 mm. Functional scans were acquired with an interleaved multiband echo-planar imaging (EPI) sequence (173 volumes/346 s, 72 slices of 2 mm isotropic voxels; no gap; time repetition = 2,000 ms; time echo = 30.0 ms; multiband acceleration factor = 3; flip angle = 75°; field of view = 220 mm; matrix size = 110 × 110 mm).

Anatomical scan and training task review

The scan session began with a structural scan. During this time, participants saw a review of the animations they had seen in the training. Each of the 8 animated sequences was shown 3 times in randomized order, in similar manners except that the object paused for additional time before its animation occurred; participants were instructed to visualize the animations of each object once it appeared. This review was repeated twice during the ~6 min of anatomical scanning.

Retrieval task

In the retrieval task, participants saw still images of each novel object, i.e. with no animations, one at a time (Fig. 1D). This was designed to emulate the circumstance in which images of tools are shown and their action properties are associatively retrieved. Each image appeared on screen for a jittered duration of 3.4, 3.5, or 3.6 s (3 of each duration per run). Here, participants were also instructed to try to explicitly recall each object's corresponding event sequence while viewing the image, in preparation to answer questions which appeared following image presentation on 18% of trials ("catch" trials). Questions were shown for 4 s, fast enough to encourage participants to be prepared for them by retrieving their knowledge ahead of time. Each object image was shown 72 times in total, 9 times per run over 8 runs.

There were 8 types of catch trial questions. Four types were specific to one object (e.g. "Moved by a hand before the stars?"), and 4 applied to 4 objects (e.g. "Moved by a hand?"). Each question type was shown twice for each object (for a total of 16 per object). Each of these 16 questions was assigned randomly among that object's 72 presentation trials across the experiment (the number of questions per run could vary). No feedback was given. Null trials (a fixation cross shown for 4 s) were also distributed through the scan and shown subsequently to 2 of the object image trials, chosen at random. Runs concluded with the presentation of a fixation screen of variable duration to reach a total exact length of 346 s, to accommodate any variance in exact duration due to varying numbers of questions.

Category localizer

Each trial showed one object image on a white background for 800 ms followed by 200 ms of fixation (Fig. 2B). There were 36 images per condition, each shown twice. Trials were blocked by condition, with 6 trials per block and 12 blocks per condition per run. During image presentation, participants performed a one-back exact image repetition detection task. Repetitions were added randomly (1 or 2 per block, or 12.5% of trials), but their total was equated across conditions. There were 2 runs of 426 s duration each.

Post-scan ratings

After the scan, participants rated each animated sequence on its level of perceived causality and tool-like quality ("tool-like-ness"). Participants saw each animation in a randomized order with both questions below. Causality questions asked to what extent the first event seemed to cause the second event (e.g. "To what extent did the hand tilting the object seem to cause the snow?"). Initial instructions clarified that this was "as opposed to events simply following one another." Tool-likeness questions asked, "To what extent did the object seem tool-like?" Responses were on a scale of 1–5, with the answer choices being: "1 = Not at all," "2 = Very little," "3 = Somewhat," "4 = Quite a bit," and "5 = Very much so."

Experimental design and statistical analysis

Preprocessing

Preprocessing was performed using AFNI software (Cox 1996). The data were slice-time corrected using the function 3dTshift and then high-pass filtered (using 3dFourier) with a cutoff of 0.008 Hz. The data were motion-corrected to the 10th volume of the first functional run with 3dVolreg. Linear and polynomial trends up to the third level were removed with 3dDetrend. The localizer runs were smoothed with a 6.0 mm Gaussian filter and the retrieval runs were smoothed with a 3.0 mm Gaussian filter for multivariate analyses and 6.0 mm for univariate analyses. Participants with an overall head displacement of >3.5 mm were excluded and replaced. Linear modeling was done using 3dDeconvolve; details of each model are described below.

Linear modeling: retrieval task

A linear model was fit to the data with a regressor for each of the 8 objects, one for all of the question periods, and derivatives of the 6 motion realignment parameters. For whole-brain univariate analyses, t-tests were used to analyze the main effect of Causality (Causer vs. Reactor) and Movement Type (Hand vs. Self) as well as planned comparisons between Self-Causer vs. Self-Reactor, and Hand-Causer vs. Hand-Reactor. Whole-brain statistical maps were corrected for multiple comparisons using a cluster-size permutation. Null distributions of maximal cluster sizes were generated by shuffling the labels in the linear model within each subject (generating 10 noise maps each) and performing 1,000 iterations of the group analyses. At each iteration, one noise map was chosen from the set generated for each subject and entered into group analyses performed exactly as in the non-shuffled data.

For ROI analyses (see ROI definition, below), linear model fits (t-values) for each participant in all voxels in the ROI were extracted and averaged across voxels. These were entered into an ANOVA using the R function "aov" with block 1 training accuracy as a regressor of non-interest (see Behavioral Results) and Movement and Causality as regressors of interest (all within-subject). Pairwise t-tests were used for planned comparisons between Self-Causer versus Self-Reactor and Hand-Causer versus Hand-Reactor.

Linear modeling: category localizer

Each of the 4 image conditions (Tools, Non-Tools, Animals, Hands) and fixation were included as regressors in a general linear model along with derivatives of the 6 motion realignment parameters. Planned contrasts for Tools > Non-Tools, and exploratory contrasts between other categories (see below), were performed using t-tests.

ROI definition

Category localizer data were used to define functional ROIs according to the following general procedure. For a given contrast, a group statistical map (t-test) was used to define group-level contiguous clusters above a specific voxel-level threshold. Cluster significance was established with permutation analysis over cluster sizes (as described above). Clusters were selected to serve as ROI boundaries if they passed a corrected significance value of $P < 0.05$ and were localized consistently with prior reports for a similar contrast and/or specified in the preregistration. There was one clear candidate for left LOTC and one for left IPS that was selected unambiguously. We then defined subject-specific ROIs within these group clusters using individual functional data, by finding the n contiguous voxels that showed the strongest effect for that individual, where n is either pre-specified or a free parameter (see below). If no voxels above 0 were found in this area for that individual, no data for that ROI for that participant were included. This approach was preregistered; however, 2 parameters were set on the basis of the observed localizer data (not the test data from the novel objects): the initial pre-cluster group-level voxel-wise threshold and the value of n. These are described more in the specific ROI definitions described below.

Tools versus Non-Tools

Our major planned contrast was between Tools and Non-Tools. T-tests across the group were used to compare responses to Tools and Non-tools and thresholded voxel-wise at $P < 0.0001$[1]. Permutation analysis indicated that group clusters over 24 voxels in size were significant at $P < 0.05$ given a pre-cluster threshold at this threshold. For Tools > Non-Tools, this revealed 2 areas, left LOTC (535 voxels) and left IPS (286 voxels), highly consistent with prior findings (Garcea and Mahon 2014; Bracci et al. 2016; Striem-Amit et al. 2018). For Non-Tools > Tools, we identified left and right parahippocampal cortex (left PHC: 60 voxels; right PHC: 363 voxels) and bilateral occipital lobe (right OC: 905 voxels; left OC: 408 voxels), consistent with prior findings on large and non-manipulable inanimate objects (Konkle and Caramazza 2013; Magri et al. 2021). One participant showed no positive effects within the lPHC ROI and was thus excluded from analyses involving lPHC. We selected the more anterior PHC ROIs for analysis as we expected that these were more likely to show high-level effects and sensitivity to learning. We report data from the occipital ROIs in the supplement.

For the value of n (individual ROI size in terms of the number of most selective voxels), our preregistration had specified 200. As follow-up analyses to determine the spatial specificity of the results, we additionally we performed a sweep of this parameter value, including 25, 50, 100, and 200, and corrected for multiple comparisons. These tests were largely nonindependent because the larger ROIs included the voxels in the smaller ROIs, so we estimated the effective number of tests using the *mef* package in R (Derringer 2018), which estimates their non-independence using correlation among the datasets. We then used this number to obtain the corrected significance threshold. In Non-Tool preferring left PHC, the group cluster only had 60 voxels and thus used $n = 25$ and 50 only. For follow-up/exploratory analyses using other ROIs, we used $n = 25$ and 50 to support stronger arguments regarding specificity.

Tool and Hand conjunction

To define Tool and Hand preferring LOTC (Tool and Hand LOTC), we created conjunction maps by intersecting 4 contrasts: Tools > Non-Tools, Tools > Animals, Hands > Non-Tools, and Hands > Animals. Each of the 4 maps was first thresholded voxel-wise at $P < 0.01$. For voxels passing this threshold for all maps, the minimum of the 4 P-values was taken to reflect combined significance. This combined P-value map was in turn thresholded at $P < 0.0001$ to identify regions of interest. This identified 3 areas: left LOTC (297 voxels), left IPS (1183 voxels), and right IPS (375 voxels). To define individual ROIs within each area, we selected the n most selective voxels for each individual as follows: We subtracted a participant's t-values for each contrast (Tools – Non-tools, Tools – Animals, Hands – Non-tools, Hands – Animals) and selected voxels in which each value was greater than 0, then summed their values. The individual ROI was selected as the n contiguous voxels with the maximal summed value.

Whole-brain contrast for Causers versus Reactors

T-tests across the group were used to compare responses to Causers and Reactors and thresholded voxel-wise at $P < 0.001$. Permutation analysis indicated that group clusters over 127 voxels in size were significant at $P < 0.05$ given this pre-cluster threshold.

Ventral stream preference modeling

To query the relationship between category selectivity and novel-object preferences in a more bottom-up fashion, we identified all visual ventral stream voxels and entered their responses to the 4 localizer categories and the 4 novel object conditions into a general linear model for each participant. To define voxels belonging to the ventral stream, we used the overall magnitude of

[1] Our pre-registration planned to use $P < 0.01$ for this threshold, but it was found to be too liberal, including most of ventral temporal cortex as well as LOTC, rather than being specific to LOTC.

any category preference from the category localizer GLM (see Linear Modeling: Category Localizer). We took the cross-participant mean response to each of the 4 localizer categories and identified the maximum of these values. We then thresholded this group-average, category selectivity magnitude at $t > 1.75$, which revealed a swath across bilateral occipitotemporal cortex spanning areas commonly reported in studies of the large-scale organization of the visual ventral stream (Konkle and Caramazza 2013, 2016; Long et al. 2018). We separately defined the left and right ventral stream areas (right: 6,058 voxels; left: 5,500 voxels). Our objective was to predict the magnitude of the Causer versus Reactor difference for novel objects on the basis of 4-way category preference among the real-world, localizer objects. We thus defined a linear model for each participant which predicted the Causer–Reactor difference from the response (t-value) to each of the 4 categories in the localizer data.

Follow-up fMRI experiment

We performed a follow-up experiment with 10 new participants to evaluate neural responses to perceiving the animated event sequences, as opposed to recalling information about them from still images. To this end, this group of participants performed the training task rather than a retrieval task during scanning. They viewed the same 8 animations as before and intermittently answered questions about the animation they just saw using the same questions as in the retrieval task in the main experiment. They were instructed to try to learn which object went with which animation and were probed on their knowledge between runs using the same tests as the training task in the main experiment. fMRI trials had a jittered ITI of between 2.9 and 3.3 s, and each condition was shown 45 times over the course of 5 runs in the scan. Each animation type was followed by each of 8 question types once, their order randomized. Runs contained 72 trials each and lasted 5.83 min. Participants viewed 2 training blocks prior to scanning as a practice. They also performed a category localizer task just as above, in order to identify functional ROIs. Data from the novel objects training task were analyzed with a GLM modeling whether participants were viewing an ambient or an object-based (self-movement or hand-movement) event; question periods were modeled as a single separate regressor. Another GLM modeled each of the 8 conditions during viewing and a regressor for question periods. To define Tool LOTC and hand and Tool LOTC ROIs, we used the group clusters from the main experiment and defined individual ROIs within those areas using the fMRI data collected here, in the same manner as above. We report the results of this follow-up experiment at the end of the Results section, which otherwise concerns the primary, memory retrieval experiment.

Follow-up behavioral experiment

We sought to rule out that conditions differed in how vividly they were recalled. A separate group of online participants performed an identical training task (as described in Training Task) with the same materials, but additionally followed by vividness questions after completing all training blocks. Here they saw each object individually and were directed to "Try to recall the pattern of events that took place with this object. How vividly can you picture it?" They responded on a sliding response scale with values from 0 to 10. Participants were included in analyses if they responded correctly to 7/12 questions by the last block of training for each of the 8 unique objects. Their vividness ratings were then analyzed as a function of condition.

Results

Training task performance

Participants responded to questions following each of 5 (or more) blocks of training. Following their final block of training, participants reached ceiling accuracy in all conditions (Self-Causer: $M = 99.61\%$, $SD = 1.63$; Hand-Causer: $M = 99.61\%$, $SD = 1.63$; Self-Reactor: $M = 100\%$, $SD = 0.00$; Hand-Reactor: $M = 99.87\%$, $SD = 0.74$). There was no main effect of movement-type ($F(1,31) = 1.00$, $MSE = 0.00$, $P = 0.325$) or causality ($F(1,31) = 1.98$, $MSE = 0.05$, $P = 0.169$) on participants' accuracy, and no interaction ($F(1,31) = 1.00$, $MSE = 0.00$, $P = 0.325$). Pairwise tests between conditions showed no significant effects ($P > 0.161$). Pre-final blocks of training likewise did not show accuracy differences, except for the first block (Self-Causer: $M = 81.25\%$, $SD = 15.23$; Hand-Causer: $M = 77.60\%$, $SD = 14.23$; Self-Reactor: $M = 80.60\%$, $SD = 13.86$; Hand-Reactor: $M = 76.30\%$, $SD = 15.57$), in which we saw a significant main effect of Movement Type ($F(1,31) = 6.00$, $MSE = 7.27$, $P = 0.020$, partial $\eta^2 = 0.162$) such that the Self moving objects had higher accuracy than the Hand moved objects. Given this significant behavioral difference, which could reflect differences in the nature of the learning process or memory, we included first-block training accuracy as a within-subject regressor in ROI analyses. There were no other effects, including no main effect of Causality ($F(1,31) = 0.30$, $MSE = 0.44$, $P = 0.588$) and no interaction ($F(1,31) = 0.02$, $MSE = 0.05$, $P = 0.881$). Pairwise tests between individual conditions showed no significant effects ($P > 0.11$).

In-scan task performance

During the in-scan retrieval task, participants were highly accurate in each condition (Self-Causer: $M = 85.01\%$, $SD = 18.38$; Hand-Causer: $M = 85.13\%$, $SD = 16.84$; Self-Reactor: $M = 85.24\%$, $SD = 15.28$; Hand-Reactor: $M = 85.25\%$, $SD = 15.82$). There was no main effect of condition, either for Movement Type ($F(1,31) = 0.00$, $MSE = 0.00$, $P = 0.965$) or Causality ($F(1,31) = 0.02$, $MSE = 0.00$, $P = 0.901$) and no interaction ($F(1,31) = 0.00$, $MSE = 0.00$, $P = 0.957$). No pairwise tests between conditions showed a significant effect (all $P > 0.895$).

Causality and tool-likeness ratings

After completing the scan, participants rated each of the objects on two 1–5 scales, Causality and Tool-likeness. The Causality scale probed to what extent participants perceived the first event in the sequence to have caused the second event in the sequence. We note that all animations were expected to be rated as somewhat causal, as even in the Reactor sequences, the ambient event could be seen as influencing subsequent the object movement. Accordingly, causality ratings revealed that participants rated most of the sequences as "somewhat" causal (Self-Causer $M = 3.17$, $SD = 1.12$; Hand-Causer $M = 3.00$, $SD = 1.11$; Self-Reactor $M = 2.89$, $SD = 1.27$; Hand-Reactor $M = 2.70$, $SD = 1.13$). There was, however, a main effect of Causality ($F(1,31) = 6.31$, $MSE = 2.67$, $P = 0.017$, partial $\eta^2 = 0.169$), demonstrating that participants' ratings were higher for the Causer objects than for the Reactor objects. There was no main effect of Movement Type ($F(1,31) = 1.45$, $MSE = 1.03$, $P = 0.237$) and no interaction ($F(1,31) = 0.00$, $MSE = 0.00$, $P = 0.952$). Pairwise tests showed that the Self-Causer animations were seen as more causal than the Hand-Reactor animations ($t(31) = 2.65$, $P = 0.012$). No other effects were significant (all $P > 0.065$).

The Tool-likeness scale rated how "tool-like" participants perceived each object to be. Across conditions, objects were seen as "very little" or "somewhat" tool-like (Self-Causer $M = 2.41$,

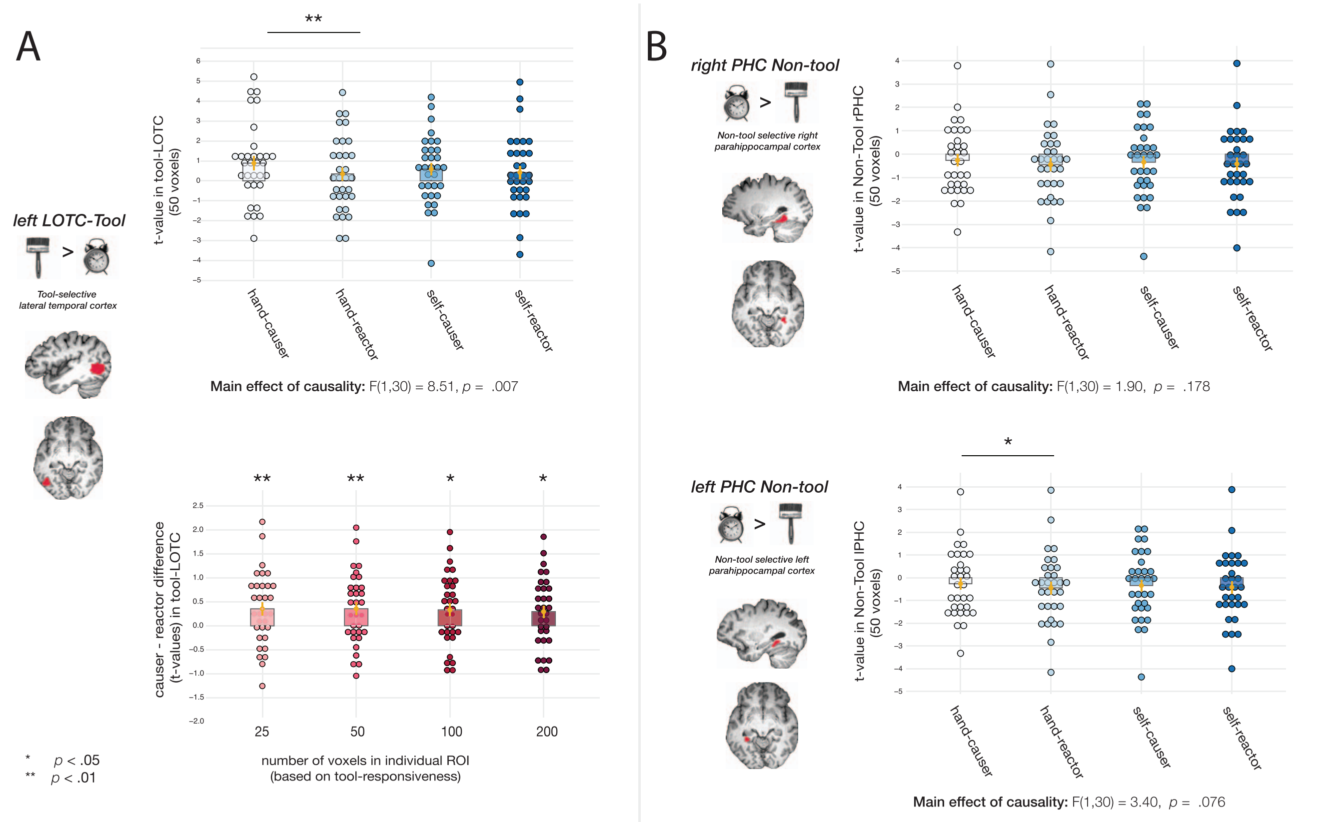

Fig. 3. A) Results in Tool-preferring left LOTC. Left: The significant group cluster defined using the contrast Tools > Non-Tools in the localizer experiment, from which individual ROIs were selected. Top, t-values reflecting the activation in response to each of the novel object conditions relative to baseline, revealing a main effect of Causality and a significant pair-wise comparison between Hand-Causer and Hand-Reactor conditions. This is shown when left Tool LOTC is defined using the top 50 voxels in each individual. Bottom, The overall Causer–Reactor difference as a function of the number of voxels in each individual ROI. B) Results in Non-Tool preferring right PHC and left PHC, defined using the top 50 voxels in each individual. Left, The significant group clusters defined using the contrast Non-Tools > Tools in the localizer experiment, from which individual ROIs were selected. Top, t-values reflecting the activation in response to each of the novel object conditions relative to baseline, in right PHC. There were no significant differences among conditions. Bottom, t-values reflecting the activation in response to each of the novel object conditions relative to baseline, in left PHC. There was a trend towards a main effect of causality and a significant difference between the Hand-Causer and Hand-Reactor conditions. *$P < 0.05$.

SD = 1.04; Hand-Causer $M = 2.72$, SD = 1.23; Self-Reactor $M = 2.69$, SD = 1.22; Hand-Reactor $M = 2.69$, SD = 1.06). There was no main effect of Movement Type ($F(1,31) = 2.49$, MSE = 0.78, $P = 0.125$) or Causality ($F(1,31) = 1.15$, MSE = 0.50, $P = 0.292$) and no interaction ($F(1,31) = 1.54$, MSE = 0.78, $P = 0.224$). Pairwise tests showed one significant difference, between the Self-Causer and Hand-Causer ($t(31) = 2.21$, $P = 0.035$), with Hand-Causer seen as more-tool like. No other effects were significant (all $P > 0.071$). While we had expected stronger condition differences in the perceptions of tool-likeness, it is likely that adults typically use a constellation of features to judge whether something should be called a tool, likely including canonical shape and transparency of physical mechanism (Malt and Johnson 1992; Bechtel et al. 2013).

Tool and Non-Tool preferring areas

We defined Tool LOTC using the familiar object localizer data: activation in response to familiar Tools versus Non-Tools was contrasted. Consistently with past findings, areas in left LOTC and left IPS were significantly more responsive to Tools than Non-Tools; left LOTC is shown in Fig. 3 and left IPS in Supplementary Fig. S1. We defined individual ROIs using each participant's top n tool-preferring voxels within these significant group-level areas; n was initially set to 200 per our preregistration.

Our critical analysis was to test the response of Tool LOTC to the different novel object conditions in the main experiment. Block one training accuracy was used as a within-subject regressor of non-interest. A 2 (Movement Type: Self vs. Hand) × 2 (Causality: Causer vs. Reactor) ANOVA revealed a main effect of Causality ($F(1,30) = 6.15$, MSE = 3.05, $P = 0.019$, partial $\eta^2 = 0.170$), such that activation in LOTC was higher for Causer objects than Reactor objects (Self-Causer: $M = 0.61$, SD = 1.48; Hand-Causer: $M = 0.80$, SD = 1.75; Self-Reactor: $M = 0.44$, SD = 1.64; Hand-Reactor: $M = 0.38$, SD = 1.58). There was no main effect of Movement Type ($F(1,30) = 0.00$, MSE = 0.00, $P = 0.968$) and no interaction ($F(1,30) = 1.05$, MSE = 0.51, $P = 0.313$). Planned comparisons between Causer and Reactor within Movement Type revealed no effect of Causality in the Self-mover objects ($F(1,30) = 1.24$, MSE = 0.45, $P = 0.274$), but a significant effect of Causality in the Hand-moved objects ($F(1,30) = 7.67$, MSE = 3.55, $P = 0.010$, partial $\eta^2 = 0.204$).

We next investigated whether the Causer versus Reactor effect was observed in the most highly tool-preferring voxels, as an exploratory follow-up to the planned analysis. We varied n by taking either the 100, 50, or 25 most tool-preferring voxels for each individual LOTC. Results across values of n are displayed in Table 1 and Fig. 3, which show that the main effect of Causality remained robustly significant when restricting analyses to these smaller, more tool-preferring ROIs (100 voxels: $F(1,30) = 7.36$, MSE = 3.88, $P = 0.011$, partial $\eta^2 = 0.197$; 50 voxels: $F(1,30) = 8.51$, MSE = 4.45, $P = 0.007$, partial $\eta^2 = 0.221$; 25 voxels: $F(1,30) = 8.20$,

Table 1. Results from Movement × Causality ANOVA in tool areas.

	LOTC								
	Self-Causer	Self-Reactor	Hand-Causer	Hand-Reactor	Causality	Movement	Movement × Causality	Causality within Self	Causality within Hand
	M (SD)	M (SD)	M (SD)	M (SD)	Main effect	Main effect	Interaction	Simple effect	Simple effect
25 voxels	0.55 (1.69)	0.40 (1.84)	0.85 (2.02)	0.29 (1.91)	$F=8.20$ $P=0.008$ $\eta_p^2=0.215$	$F=0.05$ $P=0.817$	$F=2.68$ $P=0.112$	$F=0.87$ $P=0.360$	$F=11.55$ $P=0.002$ $\eta_p^2=0.278$
50 voxels	0.59 (1.70)	0.42 (1.87)	0.89 (1.97)	0.35(1.88)	$F=8.51$ $P=0.007$ $\eta_p^2=0.221$	$F=0.13$ $P=0.723$	$F=2.38$ $P=0.134$	$F=1.08$ $P=0.306$	$F=11.62$ $P=0.002$ $\eta_p^2=0.279$
100 voxels	0.58 (1.66)	0.41 (1.79)	0.84 (1.91)	0.35(1.78)	$F=7.36$ $P=0.011$ $\eta_p^2=0.197$	$F=0.05$ $P=0.826$	$F=1.79$ $P=0.191$	$F=1.19$ $P=0.283$	$F=9.90$ $P=0.004$ $\eta_p^2=0.248$
200 voxels	0.61 (1.48)	0.44 (1.64)	0.80 (1.75)	0.38(1.58)	$F=6.15$ $P=0.019$ $\eta_p^2=0.170$	$F=0.00$ $P=0.968$	$F=1.05$ $P=0.313$	$F=1.24$ $P=0.274$	$F=7.67$ $P=0.010$ $\eta_p^2=0.204$
IPS									
25 voxels	0.12 (2.01)	0.19 (2.03)	0.34 (2.21)	0.13 (2.38)	$F=0.58$ $P=0.454$	$F=0.25$ $P=0.619$	$F=1.81$ $P=0.189$	$F=0.22$ $P=0.646$	$F=2.35$ $P=0.136$
50 voxels	0.18 (2.01)	0.20 (1.97)	0.39 (2.15)	0.20 (2.31)	$F=0.80$ $P=0.379$	$F=0.63$ $P=0.435$	$F=1.15$ $P=0.293$	$F=0.03$ $P=0.861$	$F=2.03$ $P=0.164$
100 voxels	0.21 (1.92)	0.19 (1.88)	0.38 (2.04)	0.20 (2.18)	$F=1.62$ $P=0.213$	$F=0.45$ $P=0.508$	$F=0.70$ $P=0.410$	$F=0.04$ $P=0.8^{49}$	$F=2.08$ $P=0.159$
200 voxels	0.13 (1.78)	0.10 (1.79)	0.27 (1.85)	0.02 (1.90)	$F=2.94$ $P=0.097$	$F=0.01$ $P=0.930$	$F=1.21$ $P=0.279$	$F=0.07$ $P=0.797$	$F=3.60$ $P=0.067$

Note: All ANOVA degrees of freedom are (1,30).

MSE = 4.53, $P=0.008$, partial $\eta^2=0.215$). Multiple comparison correction adjusting for non-independence among tests at each level of n yielded a correct P-value of 0.020, at which all effects remained significant. No other effects were observed within these ROIs (uncorrected P's > 0.11). We did not observe any significant correlations between this effect and causality ratings across participants.

In tool-preferring left IPS, we did not find differences between the novel object conditions at $n=200$ (Self-Causer: $M=0.13$, SD = 1.78; Hand-Causer: $M=0.27$, SD = 1.85; Self-Reactor: $M=0.10$, SD = 1.79; Hand-Reactor: $M=0.02$, SD = 1.90). A 2 × 2 ANOVA revealed no main effect of Causality ($F(1,30)=2.94$, MSE = 0.65, $P=0.097$), no main effect of Movement Type ($F(1,30)=0.008$, MSE = 0.00, $P=0.930$), and no interaction ($F(1,30)=1.21$, MSE = 0.39, $P=0.279$). Planned comparisons between Causer and Reactor within Movement Type revealed no effect of Causality in the Self-moving objects ($F(1,30)=0.07$, MSE = 0.02, $P=0.797$) nor in the Hand-moving objects ($F(1,30)=3.60$, MSE = 1.09, $P=0.067$). This pattern of findings was consistent across values of n; detailed results are shown in Table 1.

We also defined Non-Tool-preferring ROIs, which included areas in the left and right parahippocampal cortex (rPHC and lPHC); these areas are shown in Fig. 3. In right PHC with $n=200$, we found similar levels of activation across conditions (Self-Causer: $M=-0.31$, SD = 1.29; Hand-Causer: $M=-0.19$, SD = 1.26; Self-Reactor: $M=-0.42$, SD = 1.31; Hand-Reactor: $M=-0.42$, SD = 1.33). A 2 ×2 ANOVA revealed a marginal main effect of Causality ($F(1,30)=3.74$, MSE = 0.89, $P=0.063$, partial $\eta^2=0.111$), no main effect of Movement Type ($F(1,30)=0.02$, MSE = 0.01, $P=0.890$), and no interaction ($F(1,30)=0.29$, MSE = 0.13, $P=0.593$). Planned comparisons between Causer and Reactor conditions within Movement Type revealed no effect of Causality in the Self-moving objects ($F(1,30)=0.67$, MSE = 0.15, $P=0.420$), nor in the Hand-moving objects ($F(1,30)=2.90$, MSE = 1.09, $P=0.099$). Table 2 shows that at more selective values of n, there were again no significant or marginal effects.

We next compared Non-Tool right PHC with Tool LOTC directly, expecting an interaction. A 2 (ROI: LOTC vs rPHC) × 2 (Causality) × 2 (Movement Type) ANOVA was performed at each level of n, with results shown in Fig. 4 and Table 3. We did not observe an ROI by Causality interaction at $n=200$. However, we did observe such an interaction at each of the more selective ROI sizes ($n=100$, 50, and 25), perhaps because these most category-preferential voxels were also more different in their responses to the novel object conditions. The strongest effect was found at $n=25$, the most specific ROI, where we saw a main effect of ROI ($F(1,31)=6.75$, MSE = 46.03, $P=0.014$, partial $\eta^2=0.179$), indicating greater activation in the LOTC overall, and a main effect of Causality ($F(1,30)=5.18$, MSE = 3.60, $P=0.030$, partial $\eta^2=0.147$), indicating overall greater activation to Causers than Reactors, and a significant ROI × Causality interaction ($F(1,31)=6.31$, MSE = 1.08, $P=0.017$, partial $\eta^2=0.169$), indicating a greater effect of Causality in LOTC than in rPHC. Multiple comparison correction adjusting for non-independence among the tests at each level of n yielded a correct P-value threshold of 0.021, and thus, the interaction results at $n=25$ and $n=50$ remain significant. Overall, these findings suggest that preferential responses to Causal novel objects are seen more reliably in Tool-preferring voxels than Non-Tool-preferring ones, but only in the most-preferring voxels.

The Non-Tool preferring left PHC group level cluster (Fig. 3) was only 60 voxels in size and thus less strongly selective than the right PHC for Non-Tools (its P-value is higher according to our cluster-size based statistics). We could thus only use $n=25$ and $n=50$. At $n=25$, we found significant effects among the novel object conditions in the left PHC (Self-Causer: $M=-0.44$, SD = 1.41; Hand-Causer: $M=-0.27$, SD = 1.35; Self-Reactor:

Table 2. Results from Movement × Causality ANOVA in Non-Tool areas.

	Self-Causer *M* (SD)	Self-Reactor *M* (SD)	Hand-Causer *M* (SD)	Hand-Reactor *M* (SD)	Causality Main effect	Movement Main effect	Movement × Causality Interaction	Causality within Self Simple effect	Causality within Hand Simple effect
rPHC									
25 voxels	−0.30 (1.48)	−0.35 (1.53)	−0.26 (1.50)	−0.40 (1.49)	F = 0.01 P = 0.939	F = 0.17 P = 0.686	F = 0.12 P = 0.730	F = 0.10 P = 0.752	F = 1.19 P = 0.284
50 voxels	−0.34 (1.47)	−0.42 (1.51)	−0.26 (1.46)	−0.45 (1.55)	F = 1.90 P = 0.178	F = 0.01 P = 0.923	F = 0.18 P = 0.672	F = 0.22 P = 0.644	F = 1.71 P = 0.201
100 voxels	−0.33 (1.39)	−0.39 (1.40)	−0.25 (1.34)	−0.44 (1.46)	F = 2.12 P = 0.156	F = 0.06 P = 0.803	F = 0.30 P = 0.589	F = 0.18 P = 0.676	F = 2.02 P = 0.166
200 voxels	−0.31 (1.29)	−0.42 (1.31)	−0.19 (1.26)	−0.42 (1.33)	F = 3.74 P = 0.063	F = 0.02 P = 0.890	F = 0.29 P = 0.593	F = 0.67 P = 0.420	F = 2.90 P = 0.099
lPHC									
25 voxels	−0.44 (1.41)	−0.57 (1.26)	−0.27 (1.35)	−0.59 (1.37)	F = 4.35 P = 0.046 $\eta_p^2 = 0.130$	F = 0.56 P = 0.460	F = 0.64 P = 0.430	F = 0.60 P = 0.443	F = 4.89 P = 0.035 $\eta_p^2 = 0.144$
50 voxels	−0.54 (1.34)	−0.64 (1.25)	−0.43 (1.17)	−0.70 (1.21)	F = 3.40 P = 0.076	F = 0.07 P = 0.795	F = 0.62 P = 0.438	F = 0.40 P = 0.531	F = 4.29 P = 0.047 $\eta_p^2 = 0.129$

Note: All ANOVA degrees of freedom are (1,30) in rPHC and (1,29) in lPHC.

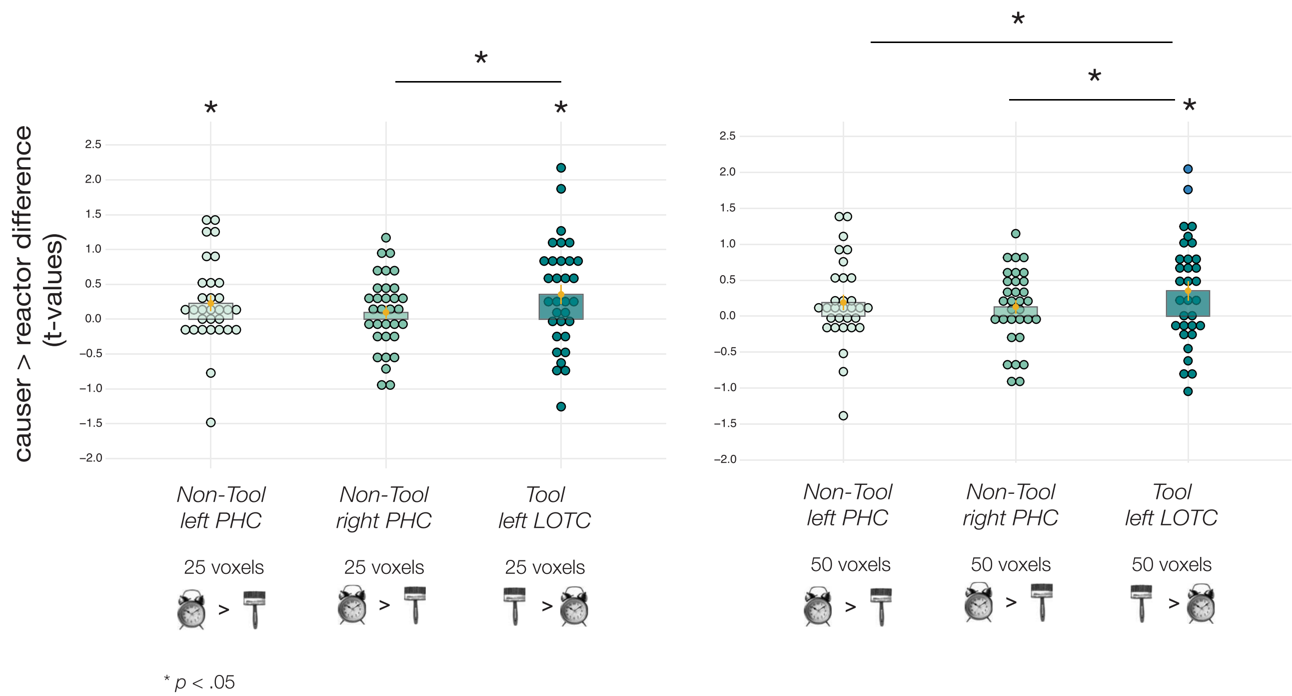

Fig. 4. Main effect of Causality (Causer–Reactor mean difference in *t*-values) in the novel object conditions (same data as Fig. 3), plotted across 3 ROIs: Non-Tool Left PHC, Non-Tool Right PHC, and Tool Left LOTC. Left: Results when ROIs are defined using the 25 most selective voxels in each individual. Right: Results when ROIs are defined using the 50 most selective voxels in each individual. Results reveal a greater effect of Causality in Tool Left LOTC relative to Non-Tool Right PHC at both ROI sizes, and a relative to Non-Tool Left PHC at 50 voxels. *$P < 0.05$.

$M = -0.57$, $SD = 1.26$; Hand-Reactor: $M = -0.59$, $SD = 1.37$) as shown in a 2 × 2 ANOVA, which revealed a main effect of Causality ($F(1,29) = 4.35$, $MSE = 1.68$, $P = 0.046$, partial $\eta^2 = 0.130$), no main effect of Movement Type ($F(1, 29) = 0.56$, $MSE = 0.23$, $P = 0.460$), and no interaction ($F(1, 29) = 0.64$, $MSE = 0.26$, $P = 0.430$). Comparisons between Causer and Reactor within Movement Type revealed no effect of Causality in the Self-moving objects ($F(1,29) = 0.60$, $MSE = 0.26$, $P = 0.443$), but there was a main effect of Causality in the Hand-moving objects ($F(1,29) = 4.89$, $MSE = 1.69$, $P = 0.035$, partial $\eta^2 = 0.144$). There was no interaction with left LOTC, as tested by a 2 (ROI: LOTC vs left PHC) × 2 (Causality) × 2 (Movement Type) ANOVA. This instead revealed a main effect of ROI ($F(1,30) = 11.94$, $MSE = 65.13$, $P = 0.002$, partial $\eta^2 = 0.285$), indicating greater activation in the LOTC overall, and a main effect of Causality ($F(1,29) = 7.34$, $MSE = 6.01$, $P = 0.011$, partial $\eta^2 = 0.202$), such that there was overall greater activation to Causers than Reactors, but no other effects ($P > 0.159$).

However, at $n = 50$, the pattern changed, showing no significant main effect of Causality (Table 2; Fig. 3) but an effect of Causality within the Hand-movers ($F(1,29) = 4.29$, $MSE = 1.27$, $P = 0.047$, partial $\eta^2 = 0.129$). We also observed a significant ROI × Causality

Table 3. Results from ROI × Movement × Causality ANOVA.

rPHC × LOTC							
	rPHC M (SD)	LOTC M (SD)	ROI Main effect	Causality Main effect	ROI × Causality Interaction	ROI × Movement Interaction	ROI × Causality × Movement Interaction
25 voxels	−0.33 (1.48)	0.52 (1.86)	F = 6.75 P = 0.014 $\eta_p^2 = 0.179$	F = 5.18 P = 0.030 $\eta_p^2 = 0.147$	F = 6.31 P = 0.017 $\eta_p^2 = 0.169$	F = 0.60 P = 0.444	F = 2.01 P = 0.167
50 voxels	−0.37 (1.48)	0.56 (1.85)	F = 8.86 P = 0.006 $\eta_p^2 = 0.222$	F = 6.07 P = 0.020 $\eta_p^2 = 0.168$	F = 5.20 P = 0.030 $\eta_p^2 = 0.144$	F = 0.55 P = 0.466	F = 1.49 P = 0.232
100 voxels	−0.35 (1.38)	0.55 (1.78)	F = 11.19 P = 0.002 $\eta_p^2 = 0.265$	F = 5.56 P = 0.025 $\eta_p^2 = 0.156$	F = 4.58 P = 0.040 $\eta_p^2 = 0.129$	F = 0.50 P = 0.486	F = 0.89 P = 0.353
200 voxels	−0.34 (1.29)	0.56 (1.60)	F = 16.15 P = 0.003 $\eta_p^2 = 0.343$	F = 5.52 P = 0.026 $\eta_p^2 = 0.155$	F = 2.29 P = 0.097	F = 0.02 P = 0.889	F = 0.49 P = 0.491

lPHC × LOTC							
	lPHC M (SD)	LOTC M (SD)	ROI Main effect	Causality Main effect	ROI × Causality Interaction	ROI × Movement Interaction	ROI × Causality × Movement Interaction
25 voxels	−0.47 (1.34)	0.56 (1.87)	F = 11.94 P = 0.002 $\eta_p^2 = 0.285$	F = 7.34 P = 0.011 $\eta_p^2 = 0.202$	F = 2.09 P = 0.159	F = 0.00 P = 0.998	F = 0.84 P = 0.367
50 voxels	−0.58 (1.23)	0.59 (1.87)	F = 16.86 P < 0.001 $\eta_p^2 = 0.360$	F = 6.71 P = 0.015 $\eta_p^2 = 0.188$	F = 4.42 P = 0.044 $\eta_p^2 = 0.128$	F = 0.35 P = 0.560	F = 0.74 P = 0.395

Note: All ANOVA degrees of freedom are (1,31) in rPHC and (1,30) in lPHC.

interaction with Tool LOTC ($F(1,29) = 4.42$, MSE = 0.48, $P = 0.044$, partial $\eta^2 = 0.128$; Table 3, Fig. 4). Thus, the effects in left PHC were not stable across voxel selection parameters. Nonetheless, there do appear to be some effects of causality in this area. Its weaker Non-Tool preference and partial sensitivity to Causality are likely related and may be an outcome of its shared left lateralization with tool responses. As the data at $n = 25$ and 50 were highly correlated, the estimated effective number of comparisons was 1 and the corrected P-value threshold was 0.05.

Neither of the occipital Non-Tool-preferring ROIs (left OCC or right OCC) showed any significant or marginal effects in the novel object conditions and each of them interacted significantly with Tool LOTC, at every setting of n (see Supplementary Tables S1 and S2). Thus, the effects of the novel object conditions appear to be limited to higher-level temporal areas.

Hand and Tool LOTC

Prior research has reliably reported overlap between the response to tools and the response to hands in LOTC (Bracci et al. 2011, 2016, 2017; Bracci and Peelen 2013). This hand-and-tool preferring area is preserved in those born without hands (Striem-Amit et al. 2017) and exhibits a stronger response to objects one acts "with"—such as canonical tools—than to objects one acts "on," such as musical instruments (Bracci and Peelen 2013). A response to both tools and hands may thus be an important predictor of voxels with a strong response to tool-like event schemas. We tested these predictions with follow-up analyses that were not specifically pre-registered but were driven by theory based on this prior literature.

We used a conjunction analysis to define Hand-and-Tool LOTC (Hand–Tool LOTC), specifically by contrasting Tools > Non-Tools and Hands > Animals, then using the minimum statistic from these 2 contrasts to define the statistic for the conjunction. We selected the n (25 and 50) most significant voxels in the conjunction statistic for each individual within a significant group cluster for the same analysis, identified in LOTC (Fig. 5). We then probed the pattern of responses in this area to the novel object conditions. A 2 (Movement Type: Self vs. Hand) × 2 (Causality: Causer vs. Reactor) ANOVA with block one training accuracy as a regressor of non-interest revealed highly reliable effects of Causality in this area at $n = 25$ ($F(1,30) = 9.50$, MSE = 4.43, $P = 0.004$, partial $\eta^2 = 0.240$) and $n = 50$ ($F(1,30) = 8.78$, MSE = 3.86, $P = 0.006$, partial $\eta^2 = 0.226$), such that activation in Hand–Tool LOTC was higher for Causers than Reactors ($n = 25$: Self-Causer: $M = 0.37$, SD = 1.75; Hand-Causer: $M = 0.73$, SD = 2.11; Self-Reactor: $M = 0.24$, SD = 1.93; Hand-Reactor: $M = 0.16$, SD = 2.08; $n = 50$: Self-Causer: $M = 0.34$, SD = 1.77; Hand-Causer: $M = 0.66$, SD = 2.13; Self-Reactor: $M = 0.21$, SD = 1.93; Hand-Reactor: $M = 0.15$, SD = 2.06). No other effects were observed in these ROIs ($P > 0.093$). Analyses within Movement Type revealed an effect of Causality within the Hand-moving objects at $n = 25$ ($F(1,30) = 13.72$, MSE = 6.37, $P = 0.001$, partial $\eta^2 = 0.314$) and $n = 50$ ($F(1,30) = 12.37$, MSE = 5.31, $P = 0.001$, partial $\eta^2 = 0.292$). This effect was not observed in the Self-moving objects ($P > 0.369$). Because the data at $n =$ 25 and $n = 50$ were highly correlated, an estimate of the number of comparisons indicated no correction was needed, but notably all effects would pass correction for 2 comparisons.

We next examined whether effects of Causality were stronger in Hand–Tool LOTC than in Non-Tool PHC. Similarly to the comparison with Tool LOTC, Causality effects in Hand–Tool LOTC were stronger than those in right PHC but not left PHC (Table 4). For right PHC at $n = 25$, we found a main effect of

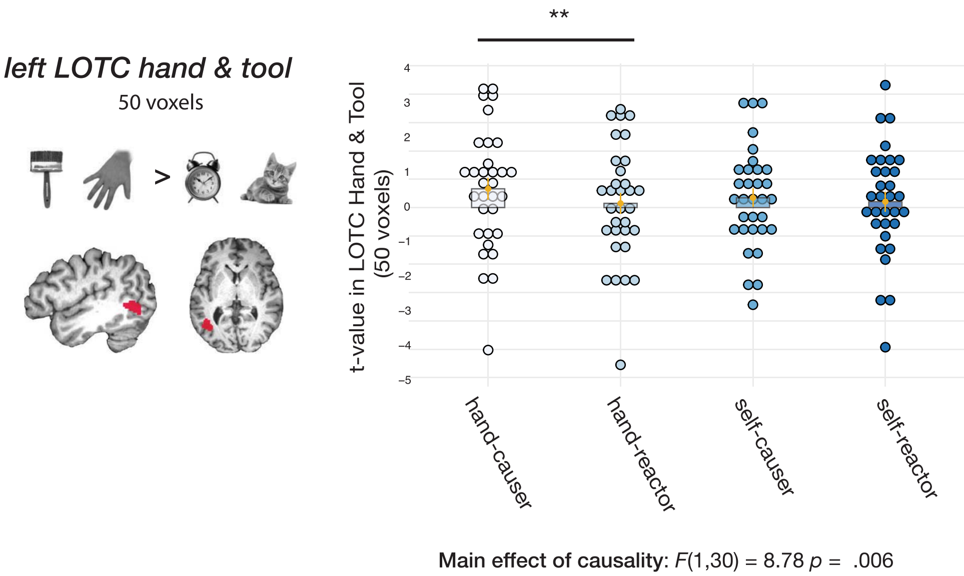

Fig. 5. Results in Hand and Tool left LOTC, defined using the 50 most selective voxels in each individual. Left, The significant group cluster showing stronger responses to Tools and Hands relative to Non-Tools and Animals, from which individual ROIs were selected. Right, t-values reflecting the activation in response to each of the novel object conditions relative to baseline. Results show a main effect of Causality and a significantly stronger response to the Hand-Causer than the Hand-Reactor conditions. $^{**}P < 0.01$.

ROI ($F(1,31) = 4.68$, MSE = 31.50, $P = 0.038$, partial $\eta^2 = 0.131$), indicating greater activation in Hand–Tool LOTC overall, as well as a main effect of Causality ($F(1,30) = 5.55$, MSE = 3.54, $P = 0.025$, partial $\eta^2 = 0.156$). We also observed a significant ROI × Causality interaction ($F(1,31) = 6.41$, MSE = 1.00, $P = 0.017$, partial $\eta^2 = 0.171$), indicating a greater effect of Causality in Hand–Tool LOTC than in rPHC. No other significant effects were observed ($P > 0.183$). For left PHC at $n = 25$, we also found a main effect of ROI ($F(1,30) = 8.03$, MSE = 43.19, $P = 0.008$, partial $\eta^2 = 0.211$), indicating greater activation in Hand–Tool LOTC overall, as well as a main effect of Causality ($F(1,29) = 8.36$, MSE = 6.10, $P = 0.007$, partial $\eta^2 = 0.224$). However, we found no significant ROI × Causality interaction ($F(1,30) = 2.05$, MSE = 0.297, $P = 0.163$) and no other significant effects ($P > 0.194$). Finally, Causality effects in Hand–Tool LOTC were not stronger than those in Tool LOTC (ROI – Causality interaction, $P > 0.40$).

Outside of these ROIs, we did not find any areas with significantly differential responses to any of our novel object conditions when using whole-brain contrasts.

Ventral stream preference modeling

The above pattern of findings suggests that areas with strong Tool preferences and Tool & Hand preferences also exhibit stronger responses to novel Causer objects than Reactor objects. This appears less the case in areas or voxels with weaker or other category preferences, painting an overall picture in which there is a relationship across the temporal lobe between a preference for real-world categories and a preference to a certain pattern of event relations, as captured in our novel object conditions. To bolster this finding with a more bottom-up and broader test that doesn't rely on particular method of ROI selection, we used an exploratory analysis in a fuller swath of the visual ventral stream. We extracted the responses to each condition across ventral stream voxels separately in each hemisphere and measured the correlation between the responses to real-world objects and the responses to the novel object conditions. A within-participant regression model tested whether the Causer–Reactor response difference across ventral stream voxels varied as a function of how those voxels responded to each of the 4 real-world object categories during the localizer.

To test the relationship between category preference and the Causer–Reactor difference, we ran within-participant linear regression to predict the Causer–Reactor difference in each voxel on the basis of those voxels' response to Animals, Non-Tools, and an average of the response to Hands and Tools, then compared the beta coefficients of these predictors across participants.

In the left ventral stream, we found that Hand–Tool responses were significantly predictive of the Causer–Reactor difference; $M = 0.0725$, SD = 0.167, $t(31) = 2.460$, $P = 0.0197$, CI [0.0124, 0.133], $d = 0.442$. Animal responses were not significantly predictive, $M = -\ 0.009$, SD = 0.171, $P = 0.772$, but were not significantly less predictive than Hand and Tool responses ($P = 0.124$). Non-Tool responses approached significance in the reverse direction, $M = -0.051$, SD = 0.162, $t(31) = -1.777$, CI [−0.109, 0.007], $P = 0.085$, $d = 0.319$, indicating that if anything, they predicted a stronger response to Reactors than Causers. As such, Non-Tool coefficients were significantly lower than Hand–Tool coefficients; $t(31) = 2.496$, CI = [0.023, 0.224], $P = 0.018$, $d = 0.448$. This confirms that the Causer–Reactor effect was significantly stronger in Hand and Tool preferring voxels than Non-Tool preferring voxels, as found in the ROI analyses.

In the right ventral stream, we found a similar pattern. The strength of Hand–Tool responses significantly predicted the Causer–Reactor difference, $M = 0.086$, SD = 0.434, $t(31) = 2.42$, CE [0.0135, 0.159] $P = 0.022$, $d = 0.434$. Animal responses were not significantly predictive, $M = -0.001$, SD = 0.214, $P = 0.0989$, but not differently from the Hand–Tool predictor ($P = 0.204$). Non-Tool responses were significantly predictive in the reverse direction, $M = -0.068$, SD = 0.152, $t(31) = -2.55$, CE [−0.123,-0.0138] $P = 0.016$, $d = 0.459$, and were significantly different from the Hand–Tool coefficients, $t(31) = 3.118$, CI = [0.053, 0.256], $P = 0.004$, $d = 0.560$. Thus, the interaction between Tool and Non-Tool predictors remained strong in both hemispheres of the ventral stream, despite the absence of focal regions of Tool selectivity in the right hemisphere. Results collapsing across the hemispheres are shown in Fig. 6. These findings overall support the conclusion that there

Table 4. Results from ROI × Movement × Causality ANOVA in Hand-and-Tool LOTC.

rPHC × LOTC							
	rPHC *M* (SD)	LOTC *M* (SD)	ROI Main effect	Causality Main effect	ROI × Causality Interaction	ROI × Movement Interaction	ROI × Causality × Movement Interaction
25 voxels	−0.33 (1.48)	0.37 (1.96)	$F = 4.68$ $P = 0.038$ $\eta_p^2 = 0.131$	$F = 5.55$ $P = 0.025$ $\eta_p^2 = 0.156$	$F = 6.41$ $P = 0.017$ $\eta_p^2 = 0.171$	$F = 1.33$ $P = 0.259$	$F = 2.69$ $P = 0.111$
50 voxels	−0.37 (1.48)	0.34 (1.96)	$F = 5.09$ $P = 0.031$ $\eta_p^2 = 0.141$	$F = 6.08$ $P = 0.020$ $\eta_p^2 = 0.168$	$F = 4.02$ $P = 0.054$	$F = 0.82$ $P = 0.373$	$F = 1.85$ $P = 0.184$
lPHC × LOTC							
	lPHC *M* (SD)	LOTC *M* (SD)	ROI Main effect	Causality Main effect	ROI × Causality Interaction	ROI × Movement Interaction	ROI × Causality × Movement Interaction
25 voxels	−0.47 (1.34)	0.37 (1.99)	$F = 8.03$ $P = 0.008$ $\eta_p^2 = 0.211$	$F = 8.36$ $P = 0.007$ $\eta_p^2 = 0.224$	$F = 2.05$ $P = 0.163$	$F = 0.07$ $P = 0.787$	$F = 0.86$ $P = 0.362$
50 voxels	−0.58 (1.23)	0.34 (1.99)	$F = 9.68$ $P = 0.004$ $\eta_p^2 = 0.244$	$F = 7.22$ $P = 0.012$ $\eta_p^2 = 0.199$	$F = 3.06$ $P = 0.091$	$F = 0.49$ $P = 0.492$	$F = 0.61$ $P = 0.440$

Note: All ANOVA degrees of freedom are (1,31) in rPHC and (1,30) in lPHC.

is a reliable relationship between category-selective responses to familiar objects and differential responses to Causers and Reactors, such that Tool and Hand preferring voxels show a stronger response to Causer than Reactor objects, while Non-tool preferring voxels do not.

Follow-up fMRI experiment

A follow-up fMRI experiment with 10 participants was run to rule out an alternative possible explanation of our results. In the retrieval phase of main study, participants were encouraged to recall the event sequence associated with each object. If they visualized this sequence in its typical order, they may have visualized the first part of the event more reliably than the second (e.g. if they ran out of time). In this case, stronger responses to the Causer objects than the Reactor objects in LOTC could reflect an overall preference to object or hand movement events over ambient events. Even though both ambient and movement events are dynamic, LOTC responds more to tool motion relative to animate motion (Beauchamp et al. 2003), and it is not known how it would respond to our ambient event stimuli. We thus scanned new participants with fMRI while they viewed the same animated event sequences directly, rather than retrieving them from memory. This allowed us to obtain the perceptual responses to viewing these stimuli. We then compared responses to the 2 kinds of events, ambient (averaging all ambient events) and object-based (averaging hand-moved and self-moved events). We found a reliable preference for the ambient events over object or hand movements in Tool LOTC at $n = 25$ voxels: Ambient, $M = 2.97$, $SD = 1.76$; Object-Based Movement, $M = -2.07$, $SD = 1.24$; $t(9) = 5.65$, $P < 0.001$; and at $n = 50$ voxels: Ambient, $M = 2.93$, $SD = 1.62$; Object-Based Movement, $M = -2.06$, $SD = 1.13$; $t(9) = 6.03$, $P < 0.001$. In Hand–Tool LOTC at $n = 25$ and at $n = 50$, we saw very similar effects: $t(9) = 5.71$, $P < 0.001$ and $t(9) = 4.89$, $P < 0.001$. Altogether, this runs counter to the predictions of the alternative hypothesis, helping to rule it out. We speculate that the reason for the significant preference to the ambient events could be that they are vividly colored and often more dynamic than object-based movements.

We also compared the stimuli when categorized as Causer and Reactor sequences, but we did not observe reliable differences between these in either Tool or Hand–Tool ROIs. We suspect that this difference was absent because when these are viewed as presented, then event sequences in both conditions exhibit causal relations between 2 entities: In the Causer conditions, the central object moving causes the ambient event, while in the Reactor condition, the event causes the object to move. Thus, both conditions always had a Causer and a Reactor in this perceptual condition. In contrast, the retrieval task in the main experiment showed an isolated object that had reliably served either and only as a Causer or a Reactor. These results may suggest that LOTC responses to Causers could be elicited by diverse stimuli.

Follow-up behavioral experiment

We furthermore aimed to rule out the possibility that our conditions differed in how vividly they were recalled. While it is not possible to know post hoc how vividly our specific fMRI participants recalled their event sequences, we pursued this question with a separate group of participants, of equal sample size ($n = 32$), who were given identical training materials. We only included participants with high accuracy on the training task, out of 12: Self-Causer $M = 10.72$, Hand-Causer $M = 10.36$, Self-Reactor $M = 10.67$, and Hand-Reactor $M = 10.47$. Comparisons revealed that Self conditions had higher accuracy than Hand conditions, $t(31) = 2.097$, $P = 0.044$, but there were no other differences, all $P > 0.20$. Vividness ratings (0–10 scale) were also similar across conditions, Self-Causer $M = 8.33$, Hand-Causer $M = 8.08$, Self-Reactor, $M = 8.09$, Hand-Reactor $M = 7.72$. Self conditions were more vivid than Hand conditions, $t(31) = 2.16$, $P = 0.039$, but there were no other differences, all $P > 0.10$. Numerically, the ratings were higher for Causers than Reactors, however, and it is thus possible that differences in imagery vividness could be detected in a larger sample. As it is, we believe the effect is too small to account for our fMRI effects. If driven by vividness, we would also

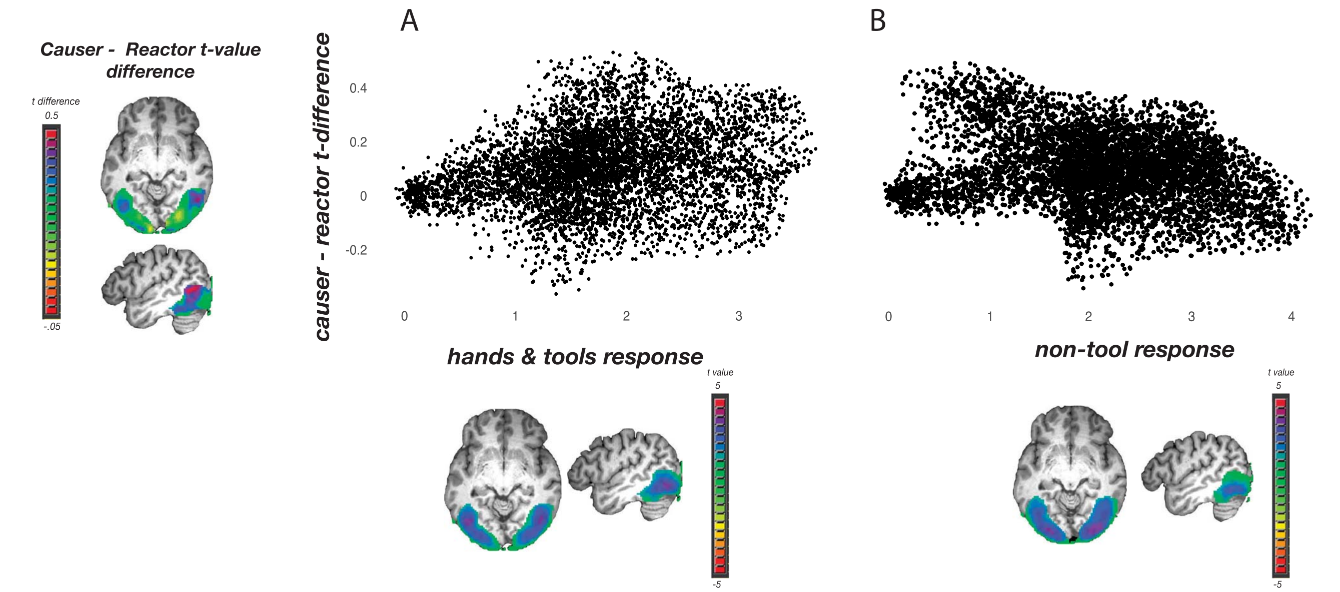

Fig. 6. Results of the ventral stream preference analysis. Left: Main effect of Causality (t-value difference between Causers and Reactors) within the ventral stream mask. Bottom, The Hands & Tools and Non-Tools response within the ventral stream mask. Scatterplot (A) shows the magnitude of each voxel's Causer–Reactor contrast as a function of its response to the Hands & Tools average response. Scatterplot (B) shows the Causer–Reactor difference as a function of the Non-Tool response.

have expected to see a main effect of Self versus Hand in the fMRI data, which was the strongest vividness effect. A vividness difference would also not explain the correlation between the effect of Causality and the Hand and Tool preference among voxels in the ventral visual stream. Nonetheless, future work should obtain vividness ratings from fMRI participants directly.

Discussion

Our findings support the hypothesis that preferential responses in Tool LOTC can be elicited by the retrieval of an event schema that has a structure consistent with that of real-world tools. Novel Causer objects appeared to influence the appearance of a surrounding event, while Reactor objects appeared to respond to it, as conveyed solely by the order of animated events. After learning about these objects, participants viewed static images of them during fMRI. We found that functionally localized Tool LOTC responded more strongly to Causers than Reactors, revealing that the different structure of their associated event schemas was sufficient to drive preferential responses in this area, without any shape or motor differences. In contrast, this preference was absent and significantly weaker in right PHC, an area that responds more to familiar non-tools than tools. Across the inferior temporal lobe, real-world category preferences to both tools and hands were positively predictive of a preference to Causers over Reactors, and non-tool preferences negatively so. We thus suggest that domain-preferential organization in the temporal lobe is partially reflective of the typical event structures characterizing different domains.

This result is not in line with the view that preferential responses to tools must be due to their greater association with certain kinds of sensory-motor information, such as visual motion or motor action (Martin 2007; Patterson et al. 2007; Lambon Ralph et al. 2017), because Causers and Reactors did not vary in these factors but their difference in event structure was sufficient to drive differential responses. In contrast, our findings are in line with research emphasizing the importance of causal and functional information in children's and adults' categorization of artifacts (Keil et al. 1998; Gopnik and Sobel 2000; Kemler Nelson et al. 2000; Truxaw et al. 2006; Kelemen and Carey 2007; Oakes and Madole 2008; Hernik and Csibra 2009; Futó et al. 2010; Yee et al. 2010, 2011; Träuble and Pauen 2011; Lombrozo and Rehder 2012; Bechtel et al. 2013) and work showing the importance of relational structure in semantic memory more generally (Gopnik and Meltzoff 1997; Markman and Stilwell 2001; Jones and Love 2007; Chatterjee 2008; Goodman et al. 2008; Carey 2009; Kemp et al. 2010). We argue that such properties are equally important as factors in the neural organization of object knowledge, providing one possible explanation for the preservation of tool-selective responses in congenitally blind individuals (Mahon et al. 2009, 2010; Peelen et al. 2013; Mattioni et al. 2020) and in those born without hands (Striem-Amit et al. 2018).

Our findings contribute to a broader picture of how relational structure influences responses across the temporal lobe. Isik et al. (2017) found that responses in posterior superior temporal sulcus (a region important in social understanding) also exhibited a preferential response to point-light walkers that interact relative to those that act independently, suggesting that this area is influenced by spatiotemporal relations. Relatedly, Abassi and Papeo (2020) found that the extrastriate body area of the ventral stream responds more strongly to 2-body dyads that are facing each other than those that are facing away, suggesting this area is tuned to spatial configurations of bodies. Together with these findings, our work reinforces the conclusion that relational information is an important part of the content and organization of domain-selective temporal areas.

How event relations come to be learned from experience and attributed to objects is active area of research (Schulz and Gopnik 2004; Kelemen and Carey 2007; Schulz et al. 2008; Waismeyer et al. 2014; Liu et al. 2019). We operationalized tool-canonical event schemas as temporally ordered event relations whose structure was conditional on the particular object involved, based in part on our prior research investigating how adults judge the causal properties of novel objects (Leshinskaya and Thompson-Schill 2019). This work found that participants made use of a hierarchical structure of event relations: they judged an object to cause an event so long as that event appeared contingently on the object's motion, but not if it simply occurred frequently in its presence. This allowed us to formalize the differences between our conditions in terms of observable event properties while controlling for aspects like associated event frequency and visual attributes.

Our findings also speak directly to the function of LOTC. A closely related proposal suggests that LOTC is tuned to objects that extend the body by serving as an effector used with the hand/arm to influence another object or surface (Bracci and Peelen 2013). Bracci and Peelen obtained ratings indicating to what extent specific familiar objects extend the body; items like forks, tennis rackets, and combs had higher ratings than doorknobs, books, and violins. These ratings were in turn strong predictors of the extent of LOTC activation in response to them, whereas ratings on motor association and graspability were not. On this account, the schema influencing LOTC involves an object that exerts a causal effect on its surroundings and it is used physically by the body to achieve this. Our notion of causal influence is broader, only having the first requirement. These 2 accounts lead to slightly different predictions. We would predict that some items categorized as low on body extension in Bracci and Peelen's study (such as light switches) would exhibit stronger responses in LOTC than items with less of a causal influence overall (e.g. books); however, their data did not separate such items explicitly. For our data, a body extension account would predict that the effect of causality would be strongest in hand-moved objects, an interaction we did not see robustly. The failure to observe this interaction may be due to a lack of power and future work should more deeply investigate the exact structure of event schemas that best drive LOTC. Nonetheless, both of these theories take event relations as critical aspects of the function of LOTC.

We additionally replicated the result that LOTC is preferentially activated not just by images of tools, but also of hands (Bracci et al. 2011, 2016, 2017, Striem-Amit et al. 2017) and found that this common area also exhibited a strong preference for novel Causer versus Reactor objects. The reason for this spatial overlap and functional similarity remains a matter of debate. The body-extension account suggests that it is driven by the fact that tools and hands are functionally interrelated, in that the former extend the latter (Bracci and Peelen 2013). A causal account would argue it is driven by the fact that tools and hands are both causally effective and explains why novel Causers also engaged similar areas. This latter view is supported by computational work showing that learning to visually identify hands is facilitated by tracking objects that are "movers" of other objects (Ullman et al. 2012). Using this "mover" principle to learn to visually recognize hands would naturally lead to grouping hands with tools in the visual recognition system, regardless of whether the objects act in concert with the hands or alone.

Finally, our work is also broadly complementary to neuropsychological and neuroimaging evidence that a variety of nearby loci in lateral temporal cortex are involved in understanding actions, events, tools, and how they relate to one another (Leshinskaya et al. 2020; Wurm and Caramazza 2021), as well as explicit representation of relational structure and thematic association in semantic memory (Wu et al. 2007; Kalénine and Buxbaum 2016; Xu et al. 2018; Leshinskaya and Thompson-Schill 2020; Wang et al. 2020).

Limitations

We note a few limitations of our work. While we reported a main effect of Causality, we did not observe an interaction with hand manipulation, which is either due to a lack of power or because Tool LOTC is driven broadly by event schemas denoting a causal role of an object, regardless of whether it is used by a hand. If the latter, then the event schema driving LOTC may be less specific than the event schemas associated with tools per se. As we argued above, however, this may be the reason that tools and hands both engage LOTC—they are both "causers," broadly construed. Behaviorally, we also found that causer objects were not always rated as more "tool-like" by participants. While Hand-manipulated Causers we rate as more tool-like than Self-moving Causers, these ratings did not fully go in line with Tool LOTC activation. It is thus possible that the use of the lexical label "tool" is not perfectly aligned with the factors driving cortical organization of object representations.

We also note that functional localizers used to define Tool LOTC vary across different studies. Whereas we used a contrast of tools to non-tool objects similar to recent approaches (Bracci et al. 2011), others use contrasts with animals or differently selected inanimate objects (Chao and Martin 2000; Mahon et al. 2007). The lack of a standard localizer across studies limits the certainty with which we can integrate findings across them and we encourage future work that hones a more standard approach.

Conclusion

Prior findings have reported a tool and hand preferring region in LOTC whole profile is not easily reduced to any particular shape preference or motor association. Here we showed that tool and hand-LOTC can be engaged by the retrieval of an event schema associated with a novel object, if the object featured in that schema appears to affect other events rather than respond to them. In the context of other work, we argue that domain-preferential organization in the temporal lobe may reflect differences not only in sensory-motor features but in the event structures typical of their preferred domains.

Acknowledgments

We thank Stefania Bracci and Talia Konkle for providing visual stimulus materials for the category localizer and early feedback on the experiment.

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Data availability statement

Data are available on request.

Funding

This work was supported by National Institutes of Health grants P30 NS45839 to the Center for Functional Neuroimaging at the University of Pennsylvania, and R01DC015359 and R01DC009209 to SLT-S.

Conflict of interest statement: The authors declare no competing financial interests.

References

Abassi E, Papeo L. The representation of two-body shapes in the human visual cortex. *J Neurosci*. 2020:40(4):852–863.

Anzellotti S, Mahon BZ, Schwarzbach J, Caramazza A. Differential activity for animals and manipulable objects in the anterior temporal lobes. *J Cogn Neurosci*. 2011:23(8):2059–2067.

Beauchamp MS, Lee KE, Haxby JV, Martin A. FMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci*. 2003:15(7):991–1001.

Bechtel S, Jeschonek S, Pauen S. How 24-month-olds form and transfer knowledge about tools: the role of perceptual, functional, causal, and feedback information. *J Exp Child Psychol*. 2013:115(1): 163–179.

Bracci S, Peelen MV. Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *J Neurosci*. 2013:33(46):18247–18258.

Bracci S, Cavina-Pratesi C, Ietswaart M, Caramazza A, Peelen MV. Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J Neurophysiol*. 2011:107:1443–1456.

Bracci S, Cavina-Pratesi C, Connolly JD, Ietswaart M. Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia*. 2016:84:81–88.

Bracci S, Daniels N, Op de Beeck HP. Task context overrules object- and category-related representational content in the human parietal cortex. *Cereb Cortex*. 2017:27(1):310–321.

Brambati SM, Myers D, Wilson A, Rankin KP, Allison SC, Rosen HJ, Gorno-Tempini ML. The anatomy of category-specific object naming in neurodegenerative diseases. *J Cogn Neurosci*. 2006: 18(10):1644–1653.

Buxbaum LJ, Shapiro AD, Coslett HB. Critical brain regions for tool-related and imitative actions: a componential analysis. *Brain*. 2014:137(7):1971–1985.

Campanella F, D'Agostini S, Skrap M, Shallice T. Naming manipulable objects: anatomy of a category specific effect in left temporal tumours. *Neuropsychologia*. 2010:48(6):1583–1597.

Carey S. *Origin of concepts*. Oxford: Oxford University Press; 2009

Chao LL, Martin A. Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*. 2000:12(4):478–484.

Chao LL, Haxby JV, Martin A. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci*. 1999:2(10):913–919.

Chatterjee A. The neural organization of spatial thought and language. *Semin Speech Lang*. 2008:29(3):226–238.

Chen J, Snow JC, Culham JC, Goodale MA. What role does "elongation" play in "tool-specific" activation and connectivity in the dorsal and ventral visual streams? *Cereb Cortex*. 2018:28(4):1117–1131.

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996:29(3): 162–173.

Damasio H, Grabowski T, Tranel D, Hichwa RD, Damasio AR. A neural basis for lexical retrieval. *Nature*. 1996:380:499–505.

Damasio AR, Tranel D, Grabowski T, Adolphs R. Neural systems behind word and concept retrieval. *Cognition*. 2004:92:179–229.

Derringer J. *A simple correction for non-independent tests*. PsyArXiv. 2018. https://psyarxiv.com/f2tyw/.

Doehrmann O, Naumer MJ, Volz S, Kaiser J, Altmann CF. Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia*. 2008:46(11):2776–2786.

Futó J, Téglás E, Csibra G, Gergely G. Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition*. 2010:117(1):1–8.

Garcea FE, Mahon BZ. Parcellation of left parietal tool representations by functional connectivity. *Neuropsychologia*. 2014:60: 131–143.

Goodman, ND, Tenenbaum, JB, Griffiths, TL, Feldman, J. 2008. Compositionality in rational analysis: In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford University Press, USA. 377–406.

Gopnik A, Sobel DM. Detecting blickets: how young children use information about novel causal powers in categorization and induction. *Child Dev*. 2000:71(5):1205–1222.

Gopnik A, Meltzoff AN. *Words, thoughts, and theories*. Cambridge, MA: MIT Press; 1997

Hernik M, Csibra G. Functional understanding facilitates learning about tools in human children. *Curr Opin Neurobiol*. 2009:19(1): 34–38.

Isik L, Koldewyn K, Beeler D, Kanwisher NG. Perceiving social interactions in the posterior superior temporal sulcus. *Proc Natl Acad Sci U S A*. 2017:114(43):E9145–E9152.

Jones M, Love BC. Beyond common features: the role of roles in determining similarity. *Cogn Psychol*. 2007:55(3):196–231.

Kalénine S, Buxbaum LJ. Thematic knowledge, artifact concepts, and the left posterior temporal lobe: where action and object semantics converge. *Cortex*. 2016:82:164–178.

Keil FC, Smith WC, Simons DJ, Levin DT. Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge. *Cognition*. 1998:65(2–3):103–135.

Kelemen D, Carey S. The essence of artifacts: developing the design stance. In: Laurence S, Margolis E, editors. *Creations of the mind: artifacts and their representation*. Oxford: Oxford University Press; 2007. pp. 415–449

Kemler Nelson DG, Frankenfield A, Morris C. Young children's use of functional information to categorize artifacts: three factors that matter. *Cognition*. 2000:77:133–168.

Kemp C, Tenenbaum JB, Niyogi S, Griffiths TL. A probabilistic model of theory formation. *Cognition*. 2010:114(2):165–196.

Konkle T, Caramazza A. Tripartite organization of the ventral stream by animacy and object size. *J Neurosci*. 2013:33(25):10235–10242.

Konkle T, Caramazza A. The large-scale organization of object-responsive cortex is reflected in resting-state network architecture. *Cereb Cortex*. 2016:27:4933–4945.

Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT. The neural and computational bases of semantic cognition. *Nat Rev Neurosci*. 2017:18(1):42–55.

Leshinskaya A, Bajaj M, Thompson-Schill SL. *ToolCausality*. 2018. https://doi.org/10.17605/OSF.IO/67XBR.

Leshinskaya A, Thompson-Schill SL. From the structure of experience to concepts of structure: how the concept "cause" applies to streams of events. *J Exp Psychol*. 2019:148(4):619–643.

Leshinskaya A, Thompson-Schill SL. Transformation of event representations along middle temporal gyrus. *Cereb Cortex*. 2020:30(5): 3148–3166.

Leshinskaya, A, Wurm, MF, Caramazza, A. 2020. Concepts of actions and their objects. In D. Poeppel, G. Mangun, & M. S. Gazzaniga (Eds.), *The Cognitive Neurosciences*. The MIT Press, Cambridge MA. 757–765.

Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. Distinct cortical pathways for processing tool versus animal sounds. *J Neurosci*. 2005:25(21):5148–5158.

Liu S, Brooks NB, Spelke ES. Origins of the concepts cause, cost, and goal in prereaching infants. *Proc Natl Acad Sci U S A*. 2019:116(36): 17747–17752.

Lombrozo T, Rehder B. Functions in biological kind classification. *Cogn Psychol*. 2012:65(4):457–485.

Long B, Yu CP, Konkle T. Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc Natl Acad Sci U S A*. 2018:115(38):E9015–E9024.

Magri C, Konkle T, Caramazza A. The contribution of object size, manipulability, and stability on neural responses to inanimate objects. *NeuroImage*. 2021:237:118098.

Mahon BZ, Milleville SC, Negri GL, Rumiati RI, Caramazza A, Martin A. Action-related properties shape object representations in the ventral stream. *Neuron*. 2007:55(3):507–520.

Mahon BZ, Anzellotti S, Schwartzbach J, Zampini M, Caramazza A. Category-specific organization in the human brain does not require visual experience. *Neuron*. 2009:63:297–405.

Mahon BZ, Schwarzbach J, Caramazza A. The representation of tools in left parietal cortex is independeant of visual experience. *Psychol Sci*. 2010:21(6):764–771.

Malt BC, Johnson EC. Do artifact concepts have cores? *J Mem Lang*. 1992:31(2):195–217.

Markman AB, Stilwell CH. Role-governed categories. *J Exp Theor Artif Intell*. 2001:13(4):329–358.

Martin A. The representation of object concepts in the brain. *Annu Rev Psychol*. 2007:58:25–45.

Martin A, Wiggs CL, Ungerleider LG, Haxby JV. Neural correlates of category specific knowledge. *Nature*. 1996:379(15 February): 649–652.

Mattioni S, Rezk M, Battal C, Bottini R, Mendoza KEC, Oosterhof NN, Collignon O. Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *elife*. 2020:9:1–33.

Nazzi T, Gopnik A. Sorting and acting with objects in early childhood: an exploration of the use of causal cues. *Cogn Dev*. 2003:18(3): 299–317.

Noppeney U, Price CJ, Penny WD, Friston KJ. Two distinct neural mechanisms for category-selective responses. *Cereb Cortex*. 2006:16(3):437–445.

Oakes LM, Madole KL. *Function revisited: how infants construe functional features in their representation of objects*. Vol. 36; 2008. pp. 135–185

Oosterhof NN, Wiggett AJ, Diedrichsen J, Tipper SP, Downing PE. Surface-based information mapping reveals crossmodal vision-action representations in human parietal and occipitotemporal cortex. *J Neurophysiol*. 2010:104(2):1077–1089.

Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*. 2007:8(December):976–989.

Peelen MV, Bracci S, Lu X, He C, Caramazza A, Bi Y. Tool selectivity in left occipitotemporal cortex develops without vision. *J Cogn Neurosci*. 2013:25(8):1225–1123.

Schulz LE, Gopnik A. Causal learning across domains. *Dev Psychol*. 2004:40(2):162–176.

Schulz LE, Goodman ND, Tenenbaum JB, Jenkins A. Going beyond the evidence: abstract laws and preschoolers' responses to anomalous data. *Cognition*. 2008:109:211–223.

Simmons WK, Martin A. Spontaneous resting-state BOLD fluctuations reveal persistent domain-specific neural networks. *Soc Cogn Affect*. 2011:7(4):467–475.

Striem-Amit E, Vannuscorps G, Caramazza A. Sensorimotor-independent development of hands and tools selectivity in the visual cortex. *Proc Natl Acad Sci U S A*. 2017:114(18):4787–4792.

Striem-Amit E, Vannuscorps G, Caramazza A. Plasticity based on compensatory effector use in the association but not primary sensorimotor cortex of people born without hands. *Proc Natl Acad Sci U S A*. 2018:115(30):7801–7806.

Tarhan LY, Watson CE, Buxbaum LJ. Shared and distinct neuroanatomic regions critical for tool-related action production and recognition: evidence from 131 left-hemisphere stroke patients. *J Cogn Neurosci*. 2016:27(12):2491–2511.

Tranel D, Damasio H, Damasio AR. A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*. 1997:35(10):1319–1327.

Träuble B, Pauen S. Cause or effect: what matters? How 12-month-old infants learn to categorize artifacts. *Br J Dev Psychol*. 2011: 29(3):357–374.

Truxaw D, Krasnow MM, Woods C, German TP. Conditions under which function information attenuates name extension via shape. *Psychol Sci*. 2006:17(5):367–371.

Ullman S, Harari D, Dorfman N. From simple innate biases to complex visual concepts. *Proc Natl Acad Sci U S A*. 2012:109(44): 18215–18220.

Valyear KF, Cavina-Pratesi C, Stiglick AJ, Culham JC. Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*. 2007:36:T94–T108.

Vannuscorps G, Wurm MF, Striem-Amit E, Caramazza A. Large-scale organization of the hand action observation network in individuals born without hands. *Cereb Cortex*. 2018:29(8):3434–3444.

Vingerhoets G, Acke F, Vandemaele P, Achten E. Tool responsive regions in the posterior parietal cortex: effect of differences in motor goal and target object during imagined transitive movements. *NeuroImage*. 2009:47(4):1832–1843.

Waismeyer A, Meltzoff AN, Gopnik A. Causal learning from probabilistic events in 24-month-olds: an action measure. *Dev Sci*. 2014:18(1):175–182.

Wang WC, Hsieh LT, Swamy G, Bunge SA. Transient neural activation of abstract relations on an incidental analogy task. *J Cogn Neurosci*. 2020:33(1):77–88.

Weisberg J, van Turennout M, Martin A. A neural system for learning about object function. *Cereb Cortex*. 2007:17(3):513–521.

Wu DH, Waller S, Chatterjee A. The functional neuroanatomy of thematic role and locative relational knowledge. *J Cogn Neurosci*. 2007:19(9):1542–1555.

Wurm MF, Caramazza A. Two "what" pathways for action and object recognition. *Trends Cogn Sci*. 2021:23(2):103–116.

Wurm MF, Lingnau A. Decoding actions at different levels of abstraction. *J Neurosci*. 2015:35(20):7727–7735.

Wurm MF, Caramazza A, Lingnau A. Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *J Neurosci*. 2017:37(3):562–575.

Xu Y, Wang X, Wang X, Men W, Gao JH, Bi Y. Doctor, teacher, and stethoscope: neural representation of different types of semantic relations. *J Neurosci*. 2018:38(13):2562–2517.

Yee E, Drucker DM, Thompson-Schill SL. fMRI-adaptation evidence of overlapping neural representations for objects related in function or manipulation. *NeuroImage*. 2010:50(2):753–763.

Yee E, Huffstetler S, Thompson-Schill SL. Function follows form: activation of shape and function features during object identification. *J Exp Psychol*. 2011:140(3):348–363.