Experimental Brain Research (2019) 237:363–375
https://doi.org/10.1007/s00221-018-5426-9

RESEARCH ARTICLE

The dominant role of functional action representation in object recognition

Long Ni[1,2] · **Ye Liu**[1,2] · **Wenyuan Yu**[1,2]

Received: 30 May 2018 / Accepted: 1 November 2018 / Published online: 9 November 2018

Abstract

Action representation of manipulable objects has been found to be involved in object recognition. Recently, studies have indicated the existence of two distinct action systems: functional action specifying how to use an object and structural action concerning how to grasp an object. Despite evidence revealing the systems' anatomical and functional differences, few preceding studies have dissociated their respective roles in object recognition. The present study aimed to tease apart their roles in the recognition of manipulable objects with a priming paradigm. Specifically, we used static stimuli (photos, Experiments 1 and 2) and dynamic stimuli (video clips, Experiments 3 and 4) depicting functional and structural action hand gestures as primes and measured the magnitude of functional and structural action priming effect in object recognition. We found that static and dynamic priming stimuli induced a robust action priming effect only for functional action prime-target pairs. Naming latencies of the target objects were shorter when functional action representations of the prime and target were congruent than when they were incongruent. Moreover, as compared to static priming photos, dynamic priming stimuli induced a larger functional action priming effect. By contrast, neither static nor dynamic priming stimuli elicited a structural action priming effect. Behavioral data from our four experiments provide consistent evidence of the dominant role of functional action representation in the recognition of manipulable objects, suggesting that action knowledge regarding how to use rather than grasp an object is more likely an intrinsic component of objects' conceptual representation.

Keywords Functional action · Structural action · Priming effect · Object recognition

Introduction

Ventral and dorsal visual streams, since Goodale and his colleagues first pointed out their distinction, are considered to be two separate visual processing pathways (Goodale and Milner 1992; Ungerleider and Haxby 1994). Over the past two decades, accumulating evidence has indicated that the two pathways are interconnected (e.g., Almeida et al. 2010, 2018; Cloutman 2013; Kristensen et al. 2016; Mahon et al. 2013). One important line of evidence shows that action knowledge, considered to be subserved by the dorsal stream, is involved in the ventral-stream-engaged processing of object recognition (e.g., Chao and Martin 2000; Harris et al. 2012; Helbig et al. 2006; Lee et al. 2018; Myung et al. 2006; Vainio et al. 2008).

A number of behavioral studies have indicated that activation of action knowledge induced by a priming stimulus affects the processing of object recognition. Specifically, recognition of the target object could be facilitated when the action knowledge of the prime and target matched, and adversely affected when they did not match (Borghi et al. 2007; Borghi and Riggio 2015; Bub et al. 2013; Campanella and Shallice 2011; Harris et al. 2012; Helbig et al. 2006, 2010; Vainio et al. 2008). The involvement of action knowledge in object recognition has been supported by neurophysiological and neuroimaging studies. For instance, Wolk, Coslett, and Glosser (2005) reported a patient with a high-level visual recognition deficit who was impaired in naming

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00221-018-5426-9) contains supplementary material, which is available to authorized users.

✉ Ye Liu
liuye@psych.ac.cn

[1] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Rd., Chaoyang District, Beijing 100101, People's Republic of China

[2] Department of Psychology, University of the Chinese Academy of Sciences, Beijing, People's Republic of China

living items while preserving the ability to name nonliving items. Moreover, the patient's naming performance was significantly more reliable for the nonliving items whose sensory-motor information was available (i.e., items that could be physically grasped, such as pliers) than those for which such information was not available. Indirect evidence from neuroimaging data shows that action knowledge is automatically evoked upon naming (Chao and Martin 2000; Chouinard and Goodale 2010), categorizing (Gerlach et al. 2002) or even passively viewing the manipulable objects (Creem-Regehr et al. 2007; Grèzes et al. 2003; Vingerhoets 2008; Wadsworth and Kana 2011). Taken together, this evidence suggests that action knowledge is one of the essential parts of object representation and is heavily involved in object recognition.

In recent years, accumulating evidence has suggested the existence of two types of object-related action knowledge: structural action knowledge that concerns how to grasp an object, and functional action knowledge that concerns how to use an object for a purpose (Binkofski and Buxbaum 2013; Bub and Masson 2006, 2012; Bub et al. 2008; Jax and Buxbaum 2010, 2013; Lee et al. 2018; Masson et al. 2008). Researchers have observed the dissociation between structural and functional actions in patients with apraxia who lost their ability to functionally use an object but were still capable of grasping objects (Sirigu et al. 1995), and in patients with optic ataxia who maintained the ability to use familiar objects but manifested a deficit in grasping objects (Jeannerod et al. 1994). A deficit in grasping objects is linked to brain lesions primarily in the left superior parietal lobule (SPL) (Rueschemeyer et al. 2010; Sirigu et al. 1995), whereas a deficit in using objects is associated with brain damage in left inferior parietal lobule (IPL) (Jeannerod et al. 1994).

This dissociation suggests a distinct neural basis for the functional and structural action systems. More recently, Buxbaum and Kalénine (2010) proposed the *Two Action System* Model (2AS), suggesting the existence of two distinct action pathways: structural and functional action systems. The structural action system is modulated by the dorso-dorsal visual stream, initiating from V3a to V6 to V6a and the medial intraparietal area (MIP) in the SPL and ending up in the dorsal premotor cortex, whereas the functional action system is subserved by the ventro-dorsal system, running from the medial superior temporal (MST) area to the IPL (see also Rizzolatti and Matelli 2003). Functionally, the structural action system is specialized for the processing of objects' spatial properties (including shape, size, and orientation) and thus for the online control of grasping action, whereas the functional action system, with its stored representation of functional action features in long-term memory, is specialized for skilled using action (Binkofski and Buxbaum 2013; Buxbaum and Kalénine 2010; Watson and Buxbaum 2014). Thus, under the 2AS theory, functional action knowledge is more likely to be an intrinsic component of objects' conceptual representation as it requires activation of long-term representations. In contrast, structural action relies on online processing of objects' visual properties, thus engaging in sensorimotor memory.

Involvement of semantic representation in functional action is evidenced by neuroimaging studies showing that its neural substrate *ventro-dorsal stream* is functionally related to the ventral visual pathway responsible for objects' conceptual representation (Almeida et al. 2013; Kristensen et al. 2016). Specifically, the left IPL, a key region of the ventro-dorsal pathway and presumed to store the functional action knowledge (Binkofski and Buxbaum 2013; Buxbaum and Kalénine 2010; Evans et al. 2016; Ishibashi et al. 2011), was found to receive visual input from the ventral pathways during the processing of tools (Almeida et al. 2013; Kristensen et al. 2016; Mahon et al. 2013).

Despite evidence suggesting the coexistence of two action systems of object-hand interactions, most of the preceding studies did not take their difference into account when testing the role of action knowledge in object recognition. For instance, Helbig et al. (2006) observed the action priming effect in pairs of prime-target objects that share the same action knowledge. However, the fact that the prime and target objects afforded both structural and functional actions (e.g., a plier and a nutcracker afford the same structural action "power grip" and functional action, "clutching the handles"), precludes us from understanding which of the two action congruencies contributes to the action priming effect (see also Myung et al. 2006). Some studies tested exclusively the role of structural (Borghi et al. 2005; Vainio et al. 2008) or functional action knowledge (Harris et al. 2012; Helbig et al. 2010) in object recognition, thus avoiding the question of which action system is a more essential part of object representation.

Among the few studies attempting to disentangle the two systems' contributions to object recognition, McNair and Harris (2012) found that the action priming effect occurred when the prime and target objects afforded the same structural but not functional action. The authors concluded that structural rather than functional action knowledge is automatically accessed upon viewing a manipulable object and thus is involved in object representation. However, two potential issues in their study compromise the authors' conclusion. First, they used object photos as both primes and targets, which were not well controlled for the confounding variable of global shape. Specifically, the paired prime-target objects that were matched in both functional and structural actions had higher global shape similarities than the paired prime-target objects that were matched only in functional or structural action. Second, the prime object was presented only for 167 ms in their study. This duration might be too

short to elicit the activation of functional action representation as evidence has indicated that functional action activation is evoked more slowly, but is sustained longer, than structural action activation (Binkofski and Buxbaum 2013; Jax and Buxbaum 2010; Kiefer et al. 2011; Lee et al. 2013, 2018). By recording event-related potentials (ERPs), Lee et al. (2018) reported that activation of structural action representation started in the early time window of 0–150 ms (P1 component) and continued to the window of 150–300 ms (N1 and P2 components). In contrast, activation of functional action representation revealed itself in a later time window of 300–500 ms, as indicated by the N400s component, which is an index of semantic processing. Thus, it is highly possible that increasing the duration of priming stimuli can give rise to functional action activation, consequently facilitating recognition of the target objects that share the same functional action.

In the present study, we used photos and video clips depicting hand action gestures as priming stimuli, thus avoiding the potential confounding variables (e.g., global shape) possibly introduced by object photos (see also Borghi et al. 2007; Bub et al. 2013; Vainio et al. 2008). By using action gestures, we were able to create priming stimuli that match more precisely in structural and functional action with the target objects. In addition, given the temporal dynamics of the two types of actions, we presented the priming stimuli with longer durations (1000 ms in Experiments 1 and 2, and 2000 ms in Experiments 3 and 4). With these carefully-designed experiments, we aimed to disentangle the relative roles of structural and functional actions in object recognition.

We predicted that functional and structural action priming effects could be induced by both static and dynamic priming stimuli. We expected the magnitude of action priming effect to be higher for functional action prime-target pairs than for structural action prime-target pairs, as functional action knowledge is suggested to be an essential component of object representation. Dynamic priming stimuli that were presented longer and visually more compelling were also expected to induce a larger priming effect compared to the static priming stimuli.

Experiment 1

In Experiment 1, we aimed to dissociate the roles of structural and functional action representation in object recognition. We adopted the priming paradigm from Borghi et al. (2007) and used static photo stimuli displaying structural and functional action hand gestures as primes. Thus, by measuring and comparing the magnitude of the priming effect induced by functional and structural action primes, we were able to see which type of action information had a larger impact on object recognition.

Method

Participants

Thirty-two undergraduate and graduate students (average age 22.4, 15 females, 17 males) from Beijing Forestry University participated in Experiment 1. All participants had normal or corrected-to-normal vision, and all but two participants were right-handed. They gave informed consent before the experiment and were paid after the experiment for their participation. The study adhered to the Declaration of Helsinki and was approved by the institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences.

Stimuli and design

Stimuli were displayed on a CRT screen with a resolution of 1024 × 768 and a refresh rate of 100 Hz. The priming stimuli consisted of 25 photos of hand gestures, including 20 functional action gestures (e.g., trigger, palm, and squeeze) and 5 structural action gestures (e.g., horizontal power grasp and vertical power grasp). We photographed the hand gestures against a white background and set the photo size to 691 × 461 pixels. The target stimuli consisted of 53 colored object photos (size: 500 × 500 pixels) selected from the China Image Set (Ni et al., 2018, *under review*). Target objects depicted in the photos were clearly segmented from the white background and were presented with an average visual angle of approximately 8.12° vertically and 8.07° horizontally.

Of the 53 target objects, 20 afforded structural action only (e.g., an apple) (*S* objects), 5 afforded functional action only (e.g., a faucet that could only be rotated) (*F* objects), and the remaining 28 afforded both structural and functional actions (e.g., a calculator) (*FS* objects). As a result, we had a total of 48 target objects that could afford structural action (28 *FS* objects + 20 *S* objects) and a total of 33 target objects that could afford functional action (28 *FS* objects + 5 *F* objects).

Pairs of prime hand gestures and target objects were created such that half of the prime–target pairs were action congruent and the other half were action incongruent. Specifically, each *S* target object was paired with a congruent and an incongruent structural action hand gesture, and each *F* target object was paired with a congruent and an incongruent functional action hand gesture. In addition, each *FS* target object was paired with two structural action gestures (congruent vs. incongruent) and two functional action gestures (congruent vs. incongruent) (see Fig. 1 for examples). (See Supplementary Materials for the complete set of prime-target pairs

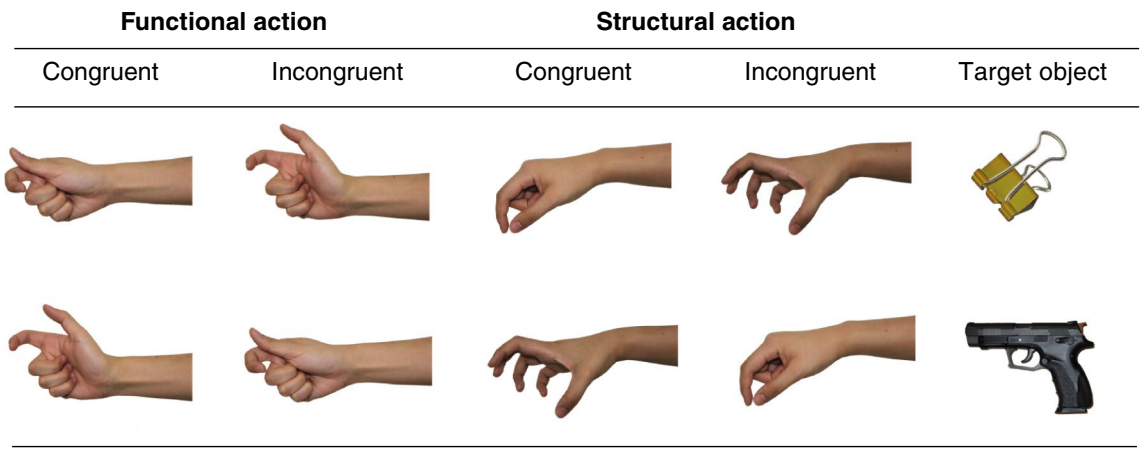

Fig. 1 Exemplars of the *FS* objects and their corresponding congurent vs. incongruent functional and structrual action hand gestures

used in Experiment 1). Thus, we had a total of 96 structural prime-target pairs (20 $S\times2+28$ $FS\times2$) that were equally divided into 2 blocks (structural action block) and a total of 66 functional prime-target pairs (5 $F\times2+28$ $FS\times2$) that were equally divided into 2 blocks as well (functional action block). In each block, half of the target objects were combined with congruent structural/functional action gestures and the other half with incongruent structural/functional action gestures. Each of the prime–target pairs in a block was randomly selected on a single trial. Moreover, half of the participants conducted the two structural action blocks first and the other half conducted the two functional action blocks first, with the running order of the two structural/functional blocks balanced across the participants.

Action congruency ratings We instructed a group of 21 participants who did not participate in Experiment 1 to rate functional and structural action congruency between the prime hand gestures and target objects. Participants rated action congruency on a 6-point Likert scale, where 1 indicates the hand gesture depicted in the photo matches poorly the functional/structural action afforded by the target object and 6 indicates a perfect match in functional/structural action between the hand gesture and the target object. As expected, the rating was significantly higher for congruent (4.95 ± 0.71) than incongruent (1.94 ± 0.31) functional action pairs [$t(20)=22.89$, $p<0.001$] and was also significantly higher for congruent (4.54 ± 0.83) than incongruent (1.65 ± 0.36) structural action pairs [$t(20)=17.67$, $p<0.001$].

Procedure

The experimental procedure was programed with E-prime 2.0. Each trial started with a fixation cross on the screen center for 500 ms, followed by a prime hand gesture that was presented for 1000 ms and immediately replaced by a 250-ms blank screen. The blank screen was followed by a target object that remained on screen for 80 ms (see Fig. 2 for the trial procedure). Participants were asked to name the target object as quickly and accurately as possible. They were told to refrain from making meaningless sounds (e.g., "um" and "en") that could trigger the onset of name recording. Reaction times of photo naming were recorded via voice key. The experimenter monitored participants' responses in the same room and recorded naming accuracy with a button box that was connected to the computer. Participants sat in

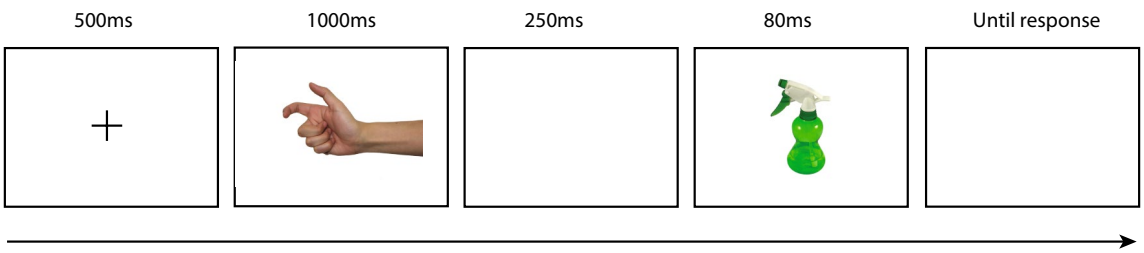

Fig. 2 Trial procedure of Experiment 1. Each trial started with a fixation cross for 500 ms, followed by a priming hand gesture presented for 1000 ms and then replaced by a blank screen for 250 ms. A target object was then presented for 80 ms. Participants were required to speak out the name of the target object as rapidly and accurately as possible once it appeared on screen

a dim room in front of the computer monitor with a viewing distance of approximately 70 cm.

Before the experiment session, participants practiced naming the target objects twice and were provided feedback by the experimenter to correct their naming errors. To familiarize participants with the trial procedure, the experiment session started with another 12 practice trials in which randomly selected nontarget objects were used. No feedback was provided in the experiment session. The whole experiment took approximately 30 min to complete.

Results and discussion

For each participant's data, we first removed trials with reaction times (RT) below and above three standard deviations, and trials with RT smaller than 100 ms that usually resulted from the recording of meaningless sounds. In total, we removed 1.12% of trials across all participants. Our analysis only focused on the RT, as naming accuracy reached a ceiling (over 95%) for all participants. We conducted two analyses on RT: one by-participant analysis (with participants as random variable), and one by-item analysis (with items as random variable).

We first performed separate analyses on the 48 target objects that could afford structural action (i.e., 28 *FS* objects + 20 *S* objects) and the 33 target objects that could afford functional action (i.e., 28 *FS* objects + 5 *S* objects). Paired *t* tests revealed a significant action priming effect for functional action pairs [by-participant analysis: $t(31) = -3.08$, $p < 0.01$; by-item analysis: $t(32) = -3.03$, $p < 0.01$] but not for structural action pairs [by-participant analysis: $t(31) = -1.58$, $p = 0.12$; by-item analysis: $t(47) = -1.503$, $p = 0.14$]. The results indicated that naming latency was significantly shorter when the functional action implied by the hand gesture matched that of the target object.

We then conducted a two-way (action type × action congruency) repeated measures ANOVA on the 28 *FS* target objects affording both actions. ANOVA results revealed only a main effect of action congruency [by-participant analysis: $F(1, 31) = 9.36$, $p < 0.01$, $\eta^2 = 0.23$; by-item analysis: $F(1, 27) = 14.23$, $p < 0.001$, $\eta^2 = 0.35$], indicating that response latency was shorter when the action implied by the static hand gesture matched that of the target object (see Fig. 3). There was no main effect of action type or interaction effect [by-participant analysis: $Fs < 1.70$, $p < 0.05$; by-item analysis: $Fs < 1.26$, $ps > 0.05$]. The absence of interaction effect indicated that both functional and structural actions contributed to the priming effect, though the magnitude of functional action priming effect (31 ± 56) appeared to be larger than that of structural action prime effect (16.2 ± 53).

Note that the number of objects that afford a specific functional action (e.g., to poke) is limited, whereas the number of objects that afford a structural action (e.g., precision grip) can be very large (i.e., many small objects afford the same precision grip). Thus, it is possible that the shorter response latency we observed in the functional action pairs resulted from a semantic effect rather than an action priming effect, i.e., it is easier for participants to predict the target object associated with a functional action.

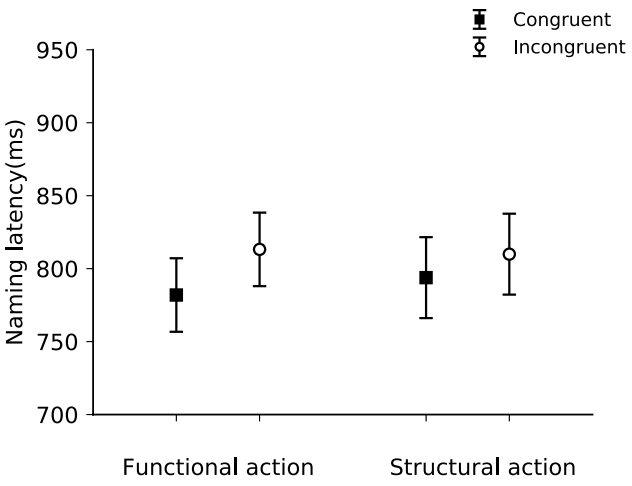

Fig. 3 Mean naming latency in the four conditions averaged across the 32 participants in Experiment 1. Error bars indicate 1 standard error above and below the means

To exclude this possibility, we recruited a group of 23 participants (mean age 23.2, 14 males, 9 females) who did not participate in Experiment 1 to recognize the functional and structural hand gestures. The procedure was identical to the Experiment 1 session except that the target object was replaced by a mask created by scrambling all photos of the targets. As in Experiment 1, the participants practiced naming the target objects twice before performing the gesture recognition task. On each trial, after viewing the hand gesture, they were encouraged to speak out the name of an object they thought was associated with the hand gesture. If they had difficulty coming up with a specific object name, they were asked instead to describe 'the general purpose of the hand gesture' (i.e., what is the hand gesture supposed to do?). When they were neither able to give a specific name nor describe the general purpose of the hand gesture, they were allowed to say "I don't know". They were required to respond in a time window of 3500 ms, a duration longer than the maximum valid naming latencies reported in Experiment 1. The time constraint was introduced to render the procedure comparable to that of Experiment 1, in which the participants were encouraged to respond as accurately and quickly as possible.

We adapted both strict and liberal criteria to define the correct responses. Under the strict criterion, a response was considered correct only when it was the recognizable name of one of the target objects that afforded the action congruent

with the prime hand gesture. Under the liberal criterion, a response was considered correct whenever it revealed the exact action corresponding to the action gesture or was the name of a nontarget object affording the same action. For instance, to the action gesture of "power grasp", the response "apple" (one of the target objects) was considered correct with the strict criterion, and the responses "grasp something big" or "mango" (not a target object) were also considered correct with the liberal criterion.

Using the strict criterion, 13.5% of participants on average recognized the functional action gestures, whereas 9.2% of participants on average recognized the structural action gestures. Using the liberal criterion, however, 50% of participants on average recognized the structural action gestures, as compared to 18.5% of participants on average who recognized the functional action gestures. The data suggested that the functional action gestures were not much more recognizable than the structural action gestures. To nullify the potential facilitating effect of semantic priming on naming latency, we removed the pairs in which the functional action gestures were recognized by more than 10% of participants under the liberal criterion. Note that this was an extremely strict inclusion criterion by which all structural prime-target pairs should be discarded. We then reanalyzed the remaining functional action prime-target pairs containing 8 functional action gestures that were recognized by less than 2% of the participants on average (by the liberal criterion). The results still revealed a significant action priming effect for functional action prime-target pairs [$t(7) = -2.2$, $p = 0.03$, one tailed], thus ruling out the possible role of semantic priming.

Taken together, the results of Experiment 1 suggested that though both functional and structural action representations are involved in object recognition, the role of functional action representation is more robust and dominant.

Experiment 2

In Experiment 1, we presented functional and structural action prime-target pairs in separate blocks. It should be pointed out that we had a smaller number of unique structural actions (i.e., only five different structural action gestures in Experiment 1) than the number of unique functional action primes. Although there is no way to completely control for this variable (i.e., grasping actions are not as diverse as skilled-use actions), we should be aware of one potential problem that could accompany this issue. Repeatedly presenting a small set of structural action primes could blunt observers' neural sensitivity to them, leading to decreased activation of structural action representation. Therefore, the unbalanced number of unique structural and functional action gestures might contribute to the small magnitude of structural action priming effect observed in Experiment 1.

To mitigate this concern, we reran Experiment 1 with the change that functional and structural action prime-target pairs were mixed in the each block. Thus, the observers were required to switch back and forth between ventro-dorsal and dorso-dorsal streams that are considered to be the neural substrates of functional and structural action representation.

Method

Participants

Thirty-two undergraduate and graduate students (average age 22.6, 15 females, 17 males) from Beijing Forestry University participated in Experiment 2. All participants had normal or corrected-to-normal vision, and all but two participants were right-handed and. They gave informed consent before the experiment and were paid after the experiment for their participation.

Stimuli and procedure

The priming stimuli consisted of the same set of static hand gestures from Experiment 1. The target stimuli consisted of only the 28 *FS* objects (i.e., affording both structrual and functional actions) used in Experiment 1. Every target object was paired with two structural action gestures (congruent vs. incongruent) and two functional action gestures (congruent vs. incongruent). Thus, we ended up with 112 (28 FS objects × 4) prime-target pairs that were equally divided into four blocks. In each block, half of the target objects were combined with congruent (or incongruent) functional (or structural) action gestures and the other half were combined with incongruent (or congruent) structural (or functional) action gestures. Each of the prime–target pairs in a block was randomly selected on a single trial. The running order of the four blocks were balanced across the participants. The trial procedure was identical to that in Experiment 1. The whole experiment took approximately 20 min to complete.

Results and discussion

For each participant's data, we removed trials with reaction times (RT) below and above three standard deviations, and trials with RT smaller than 100 ms that most likely resulted from the recording of meaningless sounds. As a consequence, we removed 9.48% of the trials across all participants. We conducted both by-participant and by-item analyses, as in Experiment 1.

We plotted the mean naming latencies in each of the four conditions across the 32 participants (see Fig. 4). A two-way repeated measures ANOVA revealed that there was a significant interaction between action type (functional vs. structural) and action congruency (match vs. mismatch)

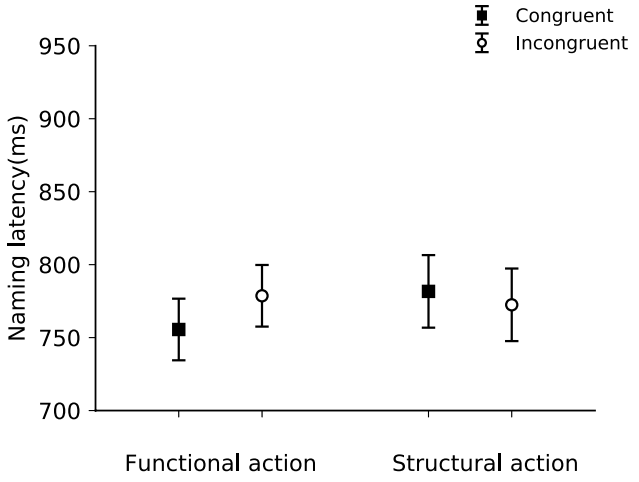

Fig. 4 Mean naming latency in the four conditions averaged across the 32 participants in Experiment 2. Error bars indicate 1 standard error above and below the means

[by-participant analysis: $F(1, 31) = 7.79$, $p < 0.01$, $\eta^2 = 0.20$; by-item analysis: $F(1, 27) = 5.92$, $p < 0.05$, $\eta^2 = 0.16$] but no main effects of action type and action congruency [by-participant analysis: $Fs < 2.35$, $ps > 0.05$; by-item analysis: $Fs < 1.58$, $ps > 0.05$]. Planned contrasts showed that naming latency was shorter when functional actions implied by the hand gestures matched those of the target objects [by-participant analysis: $t(31) = -2.85$, $p < 0.01$; by-item analysis: $t(27) = -2.16$, $p < 0.05$]. In contrast, no action priming effect was observed for the structural action pairs [by-participant analysis: $t(31) = 1.02$, $p = 0.31$; by-item analysis: $t(27) = 1.24$, $p = 0.23$].

Experiment 2 revealed only a significant functional action priming effect, which was different from Experiment 1 in which structural action, albeit to a lesser extent, was also found to contribute to the priming effect. The results from Experiment 2 suggested that it is the functional rather than structural action representation that is more fundamentally involved in the processing of object recognition.

Experiment 3

In the preceding two experiments, we observed neither a small nor no action priming effect for the structural action prime-target pairs. One possible reason is that the static hand gestures are not "visually compelling" enough to induce strong activation of structural action representation and, therefore, fail to facilitate recognition of the target objects. This assumption is supported by the finding from Vainio et al. (2008) that a stronger priming effect on object recognition was induced with dynamic priming stimuli displaying a series of hand gesture photos. To test this possibility, we thus used video clips depicting dynamic action gestures as primes in Experiment 3. We expected a stronger action priming effect to appear for both functional and structural actions. Another benefit of using dynamic priming stimuli is that we could represent some functional actions that are not easily displayed with static hand photos (e.g., hand gestures to use a lighter or a pencil sharpener).

Method

Participants

Thirty-two undergraduate and graduate students (average age 21.9, 17 females, 15 males) from Beijing Forestry University participated in Experiment 3. All participants were right-handed and had normal or corrected-to-normal vision. They gave informed consent before the experiment and were paid after the experiment for their participation.

Stimuli and procedure

The priming stimuli consisted of 14 movie clips, 12 of which depicted functional action gestures (e.g., using a nail clipper) and the remaining 2 of which depicted structural action gestures (e.g., precision gripping a nail clipper). We instructed a female hand model to mimic functional and structural actions toward target objects that were not present. To obtain precise and recognizable dynamic gestures, we allowed the hand model to practice manipulating real objects before the filming. We filmed only her hand movements against a white background. Each of the 14 movie clips had a duration of 2000 ms and an equal resolution of 691 × 461 pixels.

The target stimuli consisted of 41 object photos selected from Experiment 1. Among the 41 target objects, 20 afforded structural action only (*S* objects), 5 afforded functional action only (*F* objects), and 16 afforded both structural and functional actions (*FS* objects), Thus, we ended up with a total of 36 target objects that could afford structural action (16 *FS* + 20 *S*) and a total of 21 target objects that could afford functional action (16 *FS* + 5 *F*).

As in Experiment 1, we paired each *S* object with a congruent and an incongruent structural action hand gesture, each *F* target with a congruent and an incongruent functional action hand gesture, and each *FS* target with two structural action hand gestures (congruent vs. incongruent) and two functional action hand gestures (congruent vs. incongruent). (see Supplementary matetial for the complete set of prime-target pairs used in Experiment 3). As a result, we had a total of 72 structural action pairs (16 *FS* × 2 + 20 *S* × 2) that were equally divided into 2 blocks, and a total of 42 functional action pairs (16 *FS* × 2 + 5 *F* × 2) that were also equally divided into two blocks. The design within each block was the same as in Experiment 1.

The trial procedure was identical to that of Experiment 1 except that static hand gestures were replaced with video clips as primes that lasted 2000 ms. The whole experiment took approximately 35 min to complete.

Action congruency ratings We asked 22 participants who did not participate in Experiment 3 to rate functional and structural action congruency between the hand gestures depicted in the video clips and the actions afforded by the target objects. As expected, functional action congruency was significantly higher in congruent (5.31 ± 0.44) than incongruent (1.49 ± 0.28) prime-target pairs [$t(21) = 37.68$, $p < 0.001$], and structural action congruency was significantly higher in congruent (4.85 ± 0.80) than incongruent (1.77 ± 0.35) prime-target pairs [$t(21) = 17.83$, $p < 0.001$].

Results and discussion

For each participant's data, we removed trials with reaction times (RT) below and above three standard deviations, and trials with RT smaller than 100 ms that usually resulted from the recording of meaningless sounds. In total, we removed 2.88% of the trials across all participants. We conducted both by-participant and by-item analyses on the RT.

As in Experiment 1, we first performed separate analyses on the 36 target objects that could afford structural action (i.e., 16 *FS* objects + 20 *S* objects) and the 21 target objects that could afford functional action (i.e., 16 *FS* objects + 5 *F* objects). Paired *t* tests revealed a significant priming effect for functional action prime-target pairs [by-participant analysis: $t(31) = -3.31$, $p < 0.01$; by-item analysis: $t(20) = -2.74$, $p < 0.05$] but not for structural action prime-target pairs [by-participant analysis: $t(31) = 0.66$, $p = 0.51$; by-item analysis: $t(20) = 0.51$, $p = 0.61$].

We then conducted a two-way repeated measures ANOVA on the 16 *FS* target objects (Fig. 5). Statistical results showed that there was a significant main effect of action congruency [by-participant analysis: $F(1, 31) = 8.22$, $p < 0.01$, $\eta^2 = 0.21$; by-item analysis: $F(1, 15) = 9.66$, $p < 0.01$, $\eta^2 = 0.39$;] and a significant interaction between action type and action congruency [by-participant analysis: $F(1, 31) = 5.63$, $p < 0.05$, $\eta^2 = 0.15$; by-item analysis: $F(1, 15) = 6.76$, $p < 0.05$, $\eta^2 = 0.31$;] but no main effect of action type [by-participant analysis: $F(1,31) = 1.11$, $p = 0.30$; by-item analysis: $F(1,15) = 1.81$, $p = 0.20$]. Planned contrasts further indicated that the shorter naming latency in action congruent conditions primarily resulted from the action priming effect in functional action pairs [by-participant analysis: $t(31) = -3.43$, $p < 0.01$; by-item analysis: $t(15) = -3.25$, $p < 0.01$]. In contrast, there was no significant action priming effect in structural action pairs [by-participant analysis: $t(31) = -0.46$, $p = 0.65$; by-item analysis: $t(15) = -0.63$, $p = 0.54$].

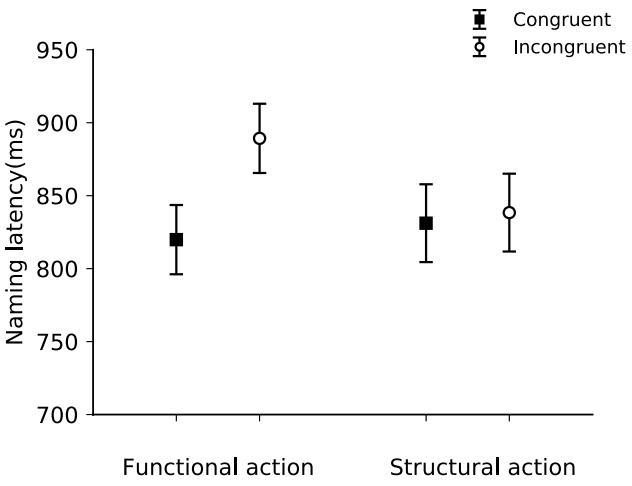

Fig. 5 Mean naming latency in the four conditions across the 32 participants in Experiment 3. Error bars indicate 1 standard error above and below the means

The use of dynamic priming stimuli did not give rise to the priming effect in structural action prime-target pairs. As expected, the magnitude of the functional action priming effect induced by the dynamic priming stimuli (69.6 ± 114.9) was significantly and marginally significantly stronger than those induced by the static priming stimuli in Experiments 1 [23.8 ± 47.2, $t(42) = 2.16$, $p < 0.05$] and 2 [31.3 ± 56.1, $t(42) = 1.69$, $p = 0.052$], respectively. However, it is noted that it might be easier for participants to predict the target objects upon viewing the dynamic action gestures unfolding over time. Thus, the stronger action priming effect could be the result of a higher semantic priming effect. To test this possibility, we asked the same group of 23 participants who performed the hand gesture recognition task in Experiment 1 to recognize the dynamic action gestures, following the same procedure as in Experiment 1. We found that 25% and 13% of participants on average recognized the functional and structural action hand gestures, respectively, by the strict criterion, both of which were higher than reported in Experiment 1. The results suggested that the semantic priming effect more or less contributed to the short naming latencies observed in Experiment 3. However, we should be cautious with the interpretation for the reasons that will be discussed in detail shortly.

Consistent with the results from Experiment 2, we observed a significant priming effect only for functional action pairs. Taken together, the results indicate a robust role of functional action representation in object representation.

Experiment 4

In Experiment 3, we observed that the dynamic hand gestures induced a significant functional but not structural action priming effect. However, in Experiment 3 there were more target objects that could afford structural than functional action (36 vs. 21 target objects). To remove this confounding variable, we conducted Experiment 4 in which only *FS* target objects (i.e., objects affording both structural and functional actions) were used, as in Experiment 2. Moreover, to confirm the action priming effect observed previously, we employed a new set of 21 *FS* target objects in Experiement 4.

Method

Participants

Thirty undergraduate and graduate students (average age 22.5, 13 females, 17 males) from Beijing Forestry University participated in Experiment 4. All participants had normal or corrected-to-normal vision, and all but one participant were right-handed. They gave informed consent before the experiment and were paid after the experiment for their participation.

Stimuli and procedure

The priming stimuli consisted of 14 movie clips, including 12 depicting functional action gestures and 2 depicting structural action gestures. The standardization process for the movies clips was the same as in Experiment 3. The target objects consisted of 21 color photos depicting objects that afforded both structural and functional actions (*FS* objects).

We paired each of the 21 object targets with two dynamic functional action hand gestures (congruent vs. incongruent) and two dynamic structural action hand gestures (congruent vs. incongruent). (See Supplementary material for the complete set of prime-target pairs used in Experiment 4). Thus, we had a total of 84 (21 *FS* objects × 4) prime-target pairs that were equally divided into four blocks. The design with each block was the same as in Experiment 2. The experimental procedure was identical to that in Experiment 3. The whole experiment took approximately 25 min to complete.

Action congruency ratings We asked 22 participants who did not participate in Experiment 4 to rate structural/functional action congruency between the priming hand gestures and those afforded by the target objects. Functional action congruency was rated significantly higher for congruent (5.00 ± 0.69) than incongruent (1.92 ± 0.42) functional action pairs [$t(21) = 25.81$, $p < 0.001$], and structural action congruency was rated significantly higher for congruent (4.05 ± 1.00) than incongruent (2.09 ± 0.62) structural action pairs [$t(21) = 7.27$, $p < 0.001$].

Results and discussion

For each participant's data, we removed trials with reaction times (RT) below and above three standard deviations, and trials with RT smaller than 100 ms that usually resulted from the recording of meaningless sounds. In total, we removed 2.62% of the trials across all participants. We performed both by-participant and by-item analyses.

As depicted in Fig. 6, the action priming effect appeared to occur for functional action prime-target pairs, whereas there was no evident difference between congruent and incongruent conditions for structural action prime-target pairs. A two-way repeated measures ANOVA confirmed the observation by revealing a significant interaction between action type and action congruency [by-participant analysis: [$F(1, 29) = 4.61$, $p < 0.05$, $\eta^2 = 0.14$; by-item analysis: $F(1, 20) = 6.97$, $p < 0.05$, $\eta^2 = 0.26$]. There was neither main effect of action type nor action congruency [by-participant analysis: $Fs < 1.67$, $ps > 0.05$; by-item analysis: $Fs < 1.44$, $ps > 0.05$]. Planned contrasts showed that a significant priming effect was present for functional action pairs, i.e., faster naming response in action congruent (804 ± 21 ms) than incongruent condition [by-participant analysis: $t(20) = -2.35$, $p < 0.05$; by-item analysis: 837 ± 26 ms, $t(20) = -2.91$, $p < 0.01$], but not for structural action pairs [by-participant analysis: $t(29) = 0.86$, $p = 0.40$; by-item analysis: $t(29) = 0.96$, $p = 0.34$].

As in Experiments 1 and 3, the same group of participants who performed the hand gesture recognition task in

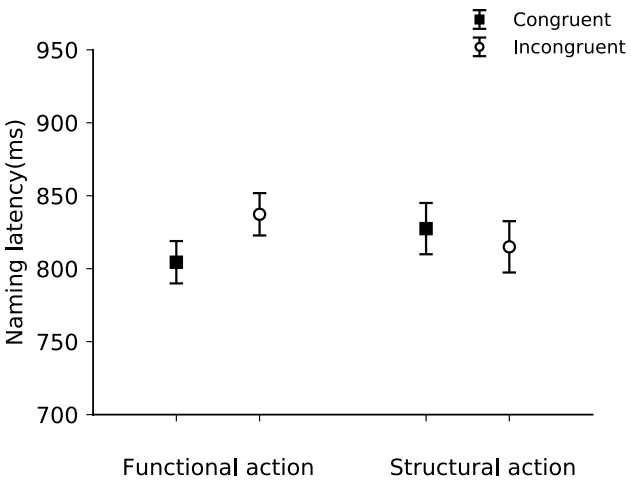

Fig. 6 Mean naming latency in four conditions across the 30 participants in Experiment 4. Error bars indicate 1 standard error above and below the means

Experiment 1 was asked to recognize the dynamic structural and functional action gestures. We found that 18.2% and 8.3% of participants on average recognized the functional and structural action hand gestures, respectively, by the strict criterion, which were higher than reported in Experiment 1 but lower than reported in Experiment 3.

In line with the preceding experiments, we only observed a significant priming effect for functional action prime-target pairs in Experiment 4, thus confirming that functional action representation plays a dominant role in the object recognition processing.

General discussion

In all of the four experiments, we observed an action priming effect for functional action prime-target pairs. Specifically, we found that naming latency is significantly shorter when functional actions of the prime and target object are congruent than when they are incongruent. Comparably, using video clips depicting hand action gestures, Helbig et al. (2010) also reported a significant priming effect for functional action prime-target pairs. Note that although they did not distinguish the two types of actions, most of the prime-target pairs matched only in functional actions (e.g., screwing with a screwdriver).

In contrast, by using photos depicting static hand gestures as primes, we observed a small structural action priming effect in Experiment 1 and a complete absence of structural action priming effect in Experiment 2. The absence of structural action priming effect appeared to be inconsistent with the findings from Borghi et al. (2007). However, they observed a congruency effect for structural action pairs only after a preliminary motor training phase was introduced.

In Experiments 3 and 4, we employed video clips depicting dynamic action gestures as priming stimuli but again failed to observe an action priming effect for the structural action pairs. This contradicts with the findings from Vingerhoets (2008) showing that dynamic hand gestures (i.e., a sequence of structural action hand gestures) elicited an action priming effect, consequently facilitating categorization of the target objects. However, it should be pointed out that in their study, the priming hand gesture (the last one in a sequence) was imposed on the transparent target object when the response was made, thus providing a stronger visual congruency/incongruency effect compared to our condition, in which the prime and target were presented sequentially. The setup they used may explain why they observed a significant action priming effect for structural action pairs.

Our results also stand in contrast with the findings from McNair and Harris (2012), who reported a significant action priming effect only for structural action prime-target pairs. Specifically, they compared naming accuracy in four prime-target pair conditions: the prime and target matched in both functional and structural actions (G+A+), matched in only structural action (G+A−), matched in only functional action (G−A+), and matched in neither of them (G−A−). Naming accuracy was found significantly higher in G+A+ and G+A− conditions than in G−A+ and G−A− conditions. They concluded that it is the knowledge regarding how to grasp rather than use an object that is automatically accessed upon identifying a manipulable object. As we mentioned in the Introduction, several potential issues in their experimental design might compromise their conclusions. First, they used object photos as both primes and targets but left the global shape between the prime and target objects uncontrolled across the four conditions, i.e., there was a higher global shape similarity in the G+A+ condition than in other conditions. Second, though there was no significant difference in the average functional action similarity between the G+A+ and G−A+ conditions, that is not obviously the case for some of prime-target pairs. For instance, among a total of 8 sets of stimuli pairs, the pairs of *peg* and *dropper, hammer* and *racquet*, and *staple gun* and *holepunch* in the G+A+ condition appear to have higher functional action similarity than the corresponding pairs of *peg* and *plier, hammer* and *stamp*, and *staple gun* and *clip* in the G−A+ condition. Since McNair and Harris (2012) did not report rating scores on each of the pairs in the two conditions, we are not able to confirm our observation. But if this difference exists as we believe, it could dilute the overall functional action priming effect induced by other pairs that have higher functional action similarity in the G−A+ condition. Third, and also more certainly, the short duration of priming stimuli (167 ms) in their experiment may not have been long enough to induce the activation of functional action. Studies examining the temporal dynamics of structural and functional action activation have indicated that activation of functional action is evoked later but sustained longer, whereas activation of structural action is faster rising but also faster decaying (Jax and Buxbaum 2010; Kiefer et al. 2011; Lee et al. 2013, 2018). So it is quite possible that the absence of functional action priming effect was due to the insufficient activation of functional action representation within the short duration. The temporal dynamics of structural action representation also help to explain why we did not observe a significant priming effect for structural action. As the stimuli duration increases, activation of structural action representation decays quickly.

One potential confounding variable in our series of experiments is the semantic priming effect that is possibly involved in the naming processing. Specifically, observers might be able to predict the target object upon viewing the priming action for a certain period of time. If this is the case, the functional action priming effect is then simply the result of preactivation of the semantic representation

of the target object. However, this possibility is strongly invalidated by the results of Experiment 1. By analyzing the prime-target pairs in which the functional action primes were recognized by less than 2% of participants under the liberal criterion, the functional action priming effect was still present in Experiment 1. Note that under the liberal criterion, we should exclude from analysis all the structural action prime-target pairs, as their structural action primes were recognized by more than 10% of participants on average. Besides, we suspect that participants in our experiments would not intentionally attempt to "guess" the target objects upon viewing the prime actions, for two reasons. First, in all four experiments, participants were not asked to attend to the prime actions. Therefore, a "smart" participant could afford to ignore the priming stimuli entirely and just focus on naming the target objects. This stands in contrast with the action recognition task in which we explicitly forced participants to "guess" the target objects. Thus, the proportion of recognized hand gestures reported in the action recognition task is very likely to be overestimated. Second, there is little, if any, benefit for participants to deliberately "guess" the target objects. As each priming hand gesture was paired with equal number of congruent and incongruent target objects, allocating cognitive rescources to guess the target objects could fail in half of the total trials at least (i.e., taking into account the wrong guesses). Thus, there is no good reason for our participants to adopt such an ineffective "guessing" strategy throughout the experiment. However, we cannot rule out the possibility that there might be a few trials in which the participant attempts to "guess" the targets when they are too revealing to them.

Recall that the dynamic hand gestures depicted in the video clips from Experiments 3 and 4 were more recognizable than those depicted in the photos from Experiment 1, which suggests that preactivation of the objects' conceptual representation was the cause of the stronger functional action priming effect observed in Experiment 3. Given the above reasons, however, we should be cautious with this interpretation. Although we cannot completely rule out the role of preactivation (especially in Experiments 3 and 4, in which dynamic action movies were used as primes), it is clear that the action priming effect is the primary source of shorter naming latencies observed in the congruent functional action prime-target pairs.

Across the four experiments, we found that functional action had a robust and dominant role in object recognition, suggesting that the knowledge concerning how to use is more likely to be an intrinsic part of the object's conceptual representation than how to grasp the object. Our conclusion is consistent with the theory of the *Two Action System* proposed by Buxbaum and Kalénine (2010) (see also Binkofski and Buxbaum 2013). Under 2AS, functional action knowledge, assumed to be subserved by the ventro-dorsal pathway, is stored in long-term memory. Its retrieval thus requires access to semantic knowledge and entails more processing time (Buxbaum and Kalénine 2010; Binkofski and Buxbaum 2013). In contrast, structural action, modulated by the dorso-dorsal stream, relies on online processing of visual information of the object (e.g., shape, size, orentation) and thus more engages in "a rapidly decaying sensorimotor memory" (Buxbaum and Kalénine 2010). The assumption that functional action relies more on semantic representation helps to explain the difference in temporal dynamics of the two action systems observed by Kiefer et al. (2011). By recording event-related potentials (ERPs), Kiefer et al. (2011) found an early ERP effect of action priming on the central scalp of approximately 100 ms and a later ERP component of N400 known to reflect semantic integration (see also Sim et al. 2014). More recently, Lee et al. (2018) confirmed that the early ERP components (P1, P2, and N2) were induced by the structural action, whereas the later component (N400s), an index of semantic processing, was evoked by functional action.

Note that, albeit to a lesser extent, structural action might also require activation of long-term representation (Bub and Masson 2006, 2012; Bub et al. 2008; Osiurak et al. 2013). Bub and Masson (2006, 2012), Bub et al. (2008) reported that structural and functional actions were evoked by both pictures and names of manipulable objects, as revealed by the subsequent primed production of the corresponding action. More recently, Osiurak et al. (2013) found that observers took more time to grasp and then move an object to someone (a receiver) than to grasp and use it, suggesting that a grasp-to-move action requires activation of long-term representation of some physical features of the object (e.g., weight). However, we should be cautious with the interpretation, as a grasp-to-transport action might involve a different mechanism from a grasp-alone action.

We conclude that recognition of manipulable objects is facilitated when they are preceded by action hand gestures implying the similar functional action but not the structural action. The functional action priming effect suggests that functional action knowledge is a more stable and thus an intrinsic part of objects' conceptual representation. With a series of behavioral experiments, our study provides strong evidence for the robust and dominant role of functional action in object recognition.

Acknowledgements This research was supported by the grant of the National Natural Science Foundation of China (Grant nos. 61632004 and 61773379) and the grant of the German Research Foundation (DFG) and the National Natural Science Foundation of China in project Crossmodal Learning (NSFC61621136008/DFGTRR-169).

Author Contributions LN, YL, and WYY designed the experiments. LN collected and analyzed the data. LN, YL, and WYY wrote the paper.

Compliance with ethical standards

Conflict of interest The authors declare no competing financial and non-financial interests.

References

Almeida J, Mahon BZ, Caramazza A (2010) The role of the dorsal visual processing stream in tool identification. Psychol Sci 21(6):772–778

Almeida J, Fintzi AR, Mahon BZ (2013) Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. Cortex 49(9):2334–2344

Almeida J, Amaral L, Garcea FE, Aguiar de Sousa D, Xu S, Mahon BZ, Martins IP (2018) Visual and visuomotor processing of hands and tools as a case study of cross talk between the dorsal and ventral streams. Cogn Neuropsychol 35(5–6):288–303

Binkofski F, Buxbaum LJ (2013) Two action systems in the human brain. Brain Lang 127(2):222–229

Borghi AM, Riggio L (2015) Stable and variable affordances are both automatic and flexible. Front Hum Neurosci 9:351–351

Borghi AM, Bonfiglioli C, Lugli L, Ricciardelli P, Rubichi S, Nicoletti R (2005) Visual hand primes and manipulable objects. In: COGSCI2005. XXVII annual conference of the cognitive science society. Lawrence Erlbaum Associates Inc., Mahwah, NJ, pp 322–327

Borghi AM, Bonfiglioli C, Lugli L, Ricciardelli P, Rubichi S, Nicoletti R (2007) Are visual stimuli sufficient to evoke motor information? Studies with hand primes. Neurosci Lett 411(1):17–21

Bub D, Masson M (2006) Gestural knowledge evoked by objects as part of conceptual representations. Aphasiology 20(9):1112–1124

Bub DN, Masson ME (2012) On the dynamics of action representations evoked by names of manipulable objects. J Exp Psychol Gen 141(3):502–517

Bub DN, Masson ME, Cree GS (2008) Evocation of functional and volumetric gestural knowledge by objects and words. Cognition 106(1):27–58

Bub DN, Masson ME, Lin T (2013) Features of planned hand actions influence identification of graspable objects. Psychol Sci 24(7):1269–1276

Buxbaum LJ, Kalénine S (2010) Action knowledge, visuomotor activation, and embodiment in the two action systems. Ann N Y Acad Sci 1191(1):201–218

Campanella F, Shallice T (2011) Manipulability and object recognition: is manipulability a semantic feature? Exp Brain Res 208(3):369–383

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

Chouinard PA, Goodale MA (2010) Category-specific neural processing for naming pictures of animals and naming pictures of tools: an ALE meta-analysis. Neuropsychologia 48(2):409–418

Cloutman LL (2013) Interaction between dorsal and ventral processing streams: where, when and how? Brain Lang 127(2):251–263

Creem-Regehr SH, Dilda V, Vicchrilli AE, Federer F, Lee JN (2007) The influence of complex action knowledge on representations of novel graspable objects: evidence from functional magnetic resonance imaging. J Int Neuropsychol Soc 13(6):1009–1020

Evans C, Edwards MG, Taylor LJ, Ietswaart M (2016) Perceptual decisions regarding object manipulation are selectively impaired in apraxia or when tDCS is applied over the left IPL. Neuropsychologia 86:153–166

Gerlach C, Law I, Paulson OB (2002) When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects. J Cogn Neurosci 14(8):1230–1239

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15(1):20–25

Grèzes J, Tucker M, Armony J, Ellis R, Passingham RE (2003) Objects automatically potentiate action: an fMRI study of implicit processing. Eur J Neurosci 17(12):2735–2740

Harris IM, Murray AM, Hayward WG, O'callaghan C, Andrews S (2012) Repetition blindness reveals differences between the representations of manipulable and nonmanipulable objects. J Exp Psychol Hum Percept Perform 38(5):1228–1241

Helbig HB, Graf M, Kiefer M (2006) The role of action representations in visual object recognition. Exp Brain Res 174(2):221–228

Helbig HB, Steinwender J, Graf M, Kiefer M (2010) Action observation can prime visual object recognition. Exp Brain Res 200(3–4):251–258

Ishibashi R, Ralph MAL, Saito S, Pobric G (2011) Different roles of lateral anterior temporal lobe and inferior parietal lobule in coding function and manipulation tool knowledge: evidence from an rTMS study. Neuropsychologia 49(5):1128–1135

Jax SA, Buxbaum LJ (2010) Response interference between functional and structural actions linked to the same familiar object. Cognition 115(2):350–355

Jax SA, Buxbaum LJ (2013) Response interference between functional and structural object-related actions is increased in patients with ideomotor apraxia. J Neuropsychol 7(1):12–18

Jeannerod M, Decety J, Michel F (1994) Impairment of grasping movements following a bilateral posterior parietal lesion. Neuropsychologia 32(4):369–380

Kiefer M, Sim E-J, Helbig H, Graf M (2011) Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. J Cogn Neurosci 23(8):1864–1874

Kristensen S, Garcea FE, Mahon BZ, Almeida J (2016) Temporal frequency tuning reveals interactions between the dorsal and ventral visual streams. J Cogn Neurosci 28(9):1295–1302

Lee C, Middleton E, Mirman D, Kalénine S, Buxbaum LJ (2013) Incidental and context-responsive activation of structure-and function-based action features during object identification. J Exp Psychol Hum Percept Perform 39(1):257–270

Lee C, Huang H-W, Federmeier KD, Buxbaum LJ (2018) Sensory and semantic activations evoked by action attributes of manipulable objects: evidence from ERPs. NeuroImage 167:331–341

Mahon BZ, Kumar N, Almeida J (2013) Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. J Cogn Neurosci 25(6):862–871

Masson ME, Bub DN, Newton-Taylor M (2008) Language-based access to gestural components of conceptual knowledge. Quart J Exp Psychol 61(6):869–882

McNair NA, Harris IM (2012) Disentangling the contributions of grasp and action representations in the recognition of manipulable objects. Exp Brain Res 220(1):71–77

Myung J, Blumstein SE, Sedivy JC (2006) Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. Cognition 98(3):223–243

Osiurak F, Roche K, Ramone J, Chainay H (2013) Handing a tool to someone can take more time than using it. Cognition 128(1):76–81

Rizzolatti G, Matelli M (2003) Two different streams form the dorsal visual system: anatomy and functions. Exp Brain Res 153(2):146–157

Rueschemeyer S-A, van Rooij D, Lindemann O, Willems RM, Bekkering H (2010) The function of words: distinct neural correlates for words denoting differently manipulable objects. J Cogn Neurosci 22(8):1844–1851

Sim EJ, Helbig HB, Graf M, Kiefer M (2014) When action observation facilitates visual perception: activation in visuo-motor areas contributes to object recognition. Cereb Cortex 25(9):2907–2918

Sirigu A, Cohen L, Duhamel J-R, Pillon B, Dubois B, Agid Y (1995) A selective impairment of hand posture for object utilization in apraxia. Cortex 31(1):41–55

Ungerleider LG, Haxby JV (1994) `What'and `where'in the human brain. Curr Opin Neurobiol 4(2):157–165

Vainio L, Symes E, Ellis R, Tucker M, Ottoboni G (2008) On the relations between action planning, object identification, and motor representations of observed actions and objects. Cognition 108(2):444–465

Vingerhoets G (2008) Knowing about tools: neural correlates of tool familiarity and experience. Neuroimage 40(3):1380–1391

Wadsworth HM, Kana RK (2011) Brain mechanisms of perceiving tools and imagining tool use acts: a functional MRI study. Neuropsychologia 49(7):1863–1869

Watson CE, Buxbaum LJ (2014) Uncovering the architecture of action semantics. J Exp Psychol Hum Percept Perform 40(5):1832–1848

Wolk DA, Coslett HB, Glosser G (2005) The role of sensory-motor information in object recognition: evidence from category-specific visual agnosia. Brain Lang 94(2):131–146