Cognition 174 (2018) 28–36

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Switching between lift and use grasp actions

Daniel N. Bub, Michael E.J. Masson*, Hannah van Mook

Department of Psychology, University of Victoria, Victoria, British Columbia, Canada

ARTICLE INFO

Keywords:
Actions
Distal goal
Object identification

ABSTRACT

Switching between competing grasp postures incurs costs on speeded performance. We examined switch costs between lift versus use actions under task conditions that required subjects to identify familiar objects. There were no asymmetrical interference effects, though reliable costs occurred when the same object required a different action on consecutive trials. In addition, lift actions were faster to objects targeted for a prospective use action than objects irrelevant to this intended goal. The benefit of a lift-then-use action sequence was not merely due to the production of two different actions in short order on the same object; use actions to an object marked for the distal goal of a lift action were not faster than use actions applied to another object. We propose that the intention to use an object facilitates the prior action of lifting it because the motor sequence *lift-then-use* is habitually conscripted to enact the proper function of an object.

1. Introduction

Grasp actions vary depending on whether we wish to use or lift an object. To use an object according to its proper function, manual actions are often directed at structural features that are not the most salient (e.g., depressing the keys of a cellphone to make a call). These actions are guided by stored knowledge, also referred to as *manipulation knowledge*, of how we typically use an object (Osiurak & Badets, 2016). The grasp applied for lifting instead of using an object can be generated directly from the object's global shape. At least in principle, lift actions can be accomplished without prior knowledge of actions linked to an object's identity, and so should be generated more rapidly than use actions.

Under certain task conditions, lift actions are indeed produced faster than use actions on the same objects (Jax & Buxbaum, 2010; Osiurak & Badets, 2016). For example, Jax and Buxbaum presented familiar objects one at a time to subjects whose vision was initially occluded by liquid crystal display goggles. Shortly after a warning tone, the goggles cleared to reveal a single object on a platform. Depending on the instructions, a given block of trials required the subject to apply either a lift or a use grasp action to the revealed object. Irrespective of task order, lift actions were generated more rapidly than use actions (Jax & Buxbaum, 2010; Osiurak, Roche, Ramone, & Chainay, 2013). Furthermore, for objects that required different use and lift hand postures (e.g., a power grasp for lifting a pocket calculator and a closed hand with an extended index finger for using it), production of use actions interfered with the subsequent production of lift actions to the same objects.

Two possible explanations have been proposed for the *use-on-lift* interference effect reported by Jax and Buxbaum (2010). The representation of a use action might remain active long after it has been generated to a particular object. A subsequent lift action would then be delayed if the same object continued to evoke the prior (and competing) use action. Alternatively, repeated production of use actions may induce a task set that entails an overall bias towards using rather than lifting objects. If the task set persists, the motor system will trigger a use action that interferes with the production of a lift action.

In what follows, we re-consider the nature of the motor representations governing use and lift actions under different task conditions. Note that the goal of a use action is typically defined in abstract terms. The action occurs in order to carry out the predetermined function of a tool or utensil, an object property that is necessarily dependent on stored knowledge. By contrast, the intention behind a lift action is usually defined more concretely. We produce these actions "… simply to grasp and move the object from one location to another" (Osiurak & Badets, 2016; p. 538). In fact, though, a variety of distal goals can be satisfied by the lifting of an object. We may reach for and lift an object to rapidly snatch it away from a child, for example, if we perceive the object to be dangerous. We can lift and hand the object to someone else; we can lift and transport the object to a new location. Finally, we may grasp and move an object into our peri-personal space because we intend to use it.

There is good evidence that not all these ways of lifting an object are performed without access to stored knowledge. Osiurak et al. (2013) have reported that lift actions are in fact generated more slowly than

* Corresponding author at: Department of Psychology, University of Victoria, P.O. Box 1700 STN CSC, Victoria, British Columbia V8W 2Y2, Canada.
E-mail address: mmasson@uvic.ca (M.E.J. Masson).

https://doi.org/10.1016/j.cognition.2018.01.013
Received 27 June 2017; Received in revised form 28 January 2018; Accepted 29 January 2018
0010-0277/

use actions when the goal of lifting is to hand the object to another person. These authors suggested that a "lift-to-give" action is determined not only by perceptual information like the object's position and size, but also requires access to long-term knowledge of an object's weight and other non-visual attributes.

Other results confirm that stored motor representations are accessed for grasp actions directed at lifting, as well as using, an object. Gentilucci (2002) has shown that knowledge of how we typically interact with objects has an influence on the kinematics of a grasp action. Herbort and Butz (2011) found that habitual actions determine the grasp chosen to rotate an object, overriding the posture more directly afforded by the intended goal of the movement. We have recently established that depicted objects, even when rotated from their upright orientation, can rapidly trigger constituents of grasp actions based on their canonical (upright) description if task demands draw attention to this stored representation (Bub, Masson, & Kumar, 2018; Chua, Bub, Masson, & Gauthier, in press).

Recent methodological developments in robotics also place emphasis on the role of prior knowledge in the formation of lift actions. As noted by Bohg, Morales, Asfour, and Kragic (2014), many computational approaches rely on a data base of object models associated with a set of stable grasps determined by prior experience. Once an object has been recognized, its position or pose is estimated and a suitable grasp retrieved from an "experience database". For novel objects, it is often possible to generate grasp postures from stored knowledge of familiar objects they resemble. It is only when task demands minimize any dependence on prior experience that candidate grasps are generated via direct consultation of structural data.

To summarize, the role of stored knowledge in the production of lift versus use actions is of considerable interest. Switching between actions types carried out on objects that appear suddenly after occlusion yields evidence consistent with the view that lift actions are produced by a fast visuomotor route that does not rely on stored motor representations. We wish to further evaluate switch costs incurred when producing use versus lift actions to a set of familiar objects that remain continuously in view. Assume these objects are all clearly visible and placed close together, and that a grasp action is produced to one object (e.g., *use the cellphone*), followed by another action carried out on the same (*lift the cellphone*) or a different object (*lift the pencil*), and so on. Because each of the possible targets remains constantly in view, the task requires the programming of various grasp actions to objects that have already been identified, the situation that normally applies to the production of use or lift actions on objects in peri-personal space. The question of interest is the following: what switch costs, if any, occur under these conditions, and what light do they shed on the nature of lift and use actions?

2. Experiment 1

In Experiment 1, subjects were cued on each trial to initiate a lift or use grasp action to an object by the image of a hand grasping that object with the appropriate grip (see Fig. 1). Subjects made their response by reaching and grasping one of three response elements continuously available in front of them, as shown in Fig. 1. A continuous sequence of cued reach-and-grasp responses was executed with the critical manipulation being the relationship between the action performed on the current trial and the action performed on the previous trial. These trial-to-trial transitions allowed for repetition of the same action on consecutive trials and for three types of action switch: different action (different action applied to the same object on two consecutive trials), different object (the object changed across two trials, but the same action type was applied), and both action and object were different across the two trials.

Response times were used to measure the cost of switching one or both components of the object-action configuration across two consecutive trials. The switch costs observed in Experiment 1 should reflect the influence of a completed task on the performance of a newly

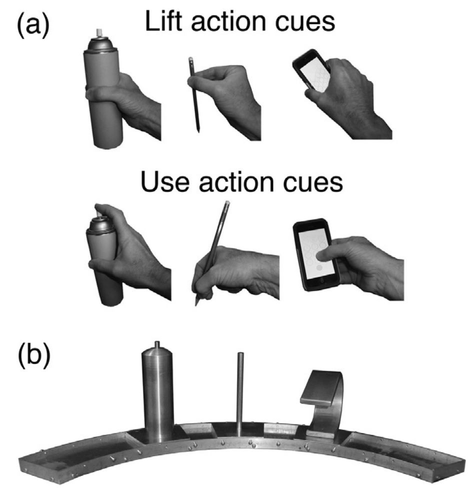

Fig. 1. Action cues (a) and response apparatus (b) used in Experiment 1.

established task and serve as a baseline against which to compare the switch costs in Experiment 2, where a specified action is generated in response to an imperative sentence.

2.1. Method

2.1.1. Subjects

Thirty-two English-speaking students (25 female, age range 18–26 years, median = 20 years) were recruited from undergraduate psychology classes at the University of Victoria. They were given extra credit in their course as an incentive to participate in the experiment. The target sample size ($n = 32$) was commensurate with the goal to detect a small effect size ($d = 0.2$) in a related-samples pairwise comparison assuming a correlation between related samples of between 0.90 and 0.95, power of 0.80, and type I error rate of 0.05. Past research using reach-and-grasp actions in our laboratory have yielded correlations between conditions in that general range. Sample-size estimates were computed using G∗Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009).

2.1.2. Materials

Unique use and lift grasps (one grasp of each type for each object) were identified for three specific objects: cellphone, pencil, and spray can. Digital grayscale photographs were made with each of the three objects posed with a male human right hand demonstrating each of the associated actions (see Fig. 1). A second version of these images was created by making a mirror reversal of each original for use with left-handed subjects. These images were used as cues to indicate to subjects which action to perform on a given trial.

2.1.3. Procedure

Subjects were tested individually in a quiet room. The images serving as action cues were presented on a monitor positioned about 50 cm from the subject. Stimulus presentation and data collection were controlled by an iMac computer. Image presentation was initiated by having the subject use the index finger of their dominant hand to press a button on a button box that was placed on the table in front of the subject. A response was initiated by lifting the response hand from the button, providing a measure of liftoff time. The subject then reached and grasped an element mounted on a response apparatus was positioned between the button box and the computer monitor (see Fig. 1). Three elements were mounted on the apparatus, each to be used for the use and lift actions associated with one of the three objects. Positioning of the elements on the base of the apparatus was counterbalanced across subjects. A weak electrical current was passed through the apparatus which was connected to the computer that controlled the experiment. Contact with a response element broke the circuit, signaling completion of the reach-and-grasp response. The experimenter viewed a second monitor which indicated the target response on each trial. The experimenter observed each response and classified it as correct or incorrect (in case contact was made with the wrong response element or the wrong action was produced).

Subjects were given 32 practice trials and 384 critical trials in which one of the action cues was presented and the corresponding action was executed. They were instructed to respond as quickly and accurately as possible. Trials began with the subject resting his or her dominant hand on a button on the button box. When a fixation cross appeared on the monitor, the subject pressed down on the button and the fixation cross disappeared. After a 250-ms interval, the action cue was presented and the subject used his or her dominant hand to make a speeded reach-and-grasp movement to the relevant element on the response apparatus. Subjects were encouraged to think of each response element as the object in the corresponding action cue and to associate the use and lift actions for that object with the response element.

In the critical trials, each action cue was presented 64 times in a random order. A break was provided after every 64 critical trials. Because the action cues were presented in a random order, the number of trial-to-trial switches of each of the four possible types (same object and same action, same object and different action, different object and same action type, different object and different action type) varied across subjects. The first trial of a block was excluded from analysis because there was a time delay on the order of seconds between that trial and the previous trial. Also, trials on which an error was made or response time was considered an outlier or an anticipation response were excluded. Across the set of subjects who were tested, the minimum number of trials in a switch condition was 5 and the maximum was 76 (mean = 45).

2.2. Results and discussion

Response time for each trial was defined as the time between the onset of the action cue and liftoff from the button box. We also captured the time taken to move from the response box to the response element, but we report only liftoff times because in Experiments 3 and 4 that is the only measure we were able to capture. For the sake of consistency, we report liftoff time across all of the experiments reported here. Nevertheless, the pattern of results for Experiment 1, where the full response time taken to complete a reach-and-grasp action was available, was the same for liftoff time and total response time.

Response times below 200 ms were considered anticipation responses and were excluded. In addition, response times greater than 1300 ms were excluded as outliers (0.35% of correct responses). The upper cutoff was set so that no more than 0.5% of correct observations were excluded (based on recommendations by Ulrich & Miller, 1994). This policy was applied in all of the experiments reported here, resulting in some variation across experiments with respect to the upper cutoff value. Error rates were very low ($M = 0.1\%$), with 25 subjects making no errors, so no analyses of these data are not reported.

The response time data were summarized by computing for each subject the mean correct response time in each of the four switch conditions for each action type (lift and use). The means taken across subjects are shown in Fig. 2. The analyses we report are analyses of variance (ANOVA) computed by collapsing across the observations within a condition for each subject. We also computed linear mixed models analyses of the data for each experiment and these analyses yielded the same pattern of significant and nonsignificant fixed effects as we report for ANOVAs, with one exception noted in Experiment 4. A three-factor analysis of variance with action type (lift, use), object switch (same, different), and action switch (same type, different type) as repeated-measures factors was computed. The type I error rate was set at 0.05. The only significant effect in this analysis was a main effect of object switch, with a longer response time when a different object had been cued on the previous trial relative to when the same object was cued on the previous trial (446 ms vs. 440 ms), $F(1, 31) = 14.81$, $MSE = 196$, $p < .001$, $\eta_p^2 = 0.32$.

Experiment 1 did not produce differences in the time to initiate use and lift actions. Rather, these results suggest that when subjects imitate the actions demonstrated pictorially, use and lift-actions take the same amount of time to initiate (443 ms for both), $F < 1$. This outcome provides a baseline against which to compare our subsequent experiments. Jax and Buxbaum (2010) conducted a similar control study with images and also found no systematic differences in the initiation times to hand postures depicting lift and use actions.

3. Experiment 2

In Experiment 2, we created a situation more closely tailored to everyday life. Rather than responding to images of a hand grasping objects, subjects responded to imperative sentences that demanded a use or lift action. The task involved in carrying out such actions on a set of objects allowed us to evaluate switch costs incurred when generating actions to short sentences like *use the cell phone* or *lift the spray can*. From one trial to the next, a switch could occur either if an alternate grasp action was to be applied to the same object as on the previous trial (e.g., *use the cellphone* followed by *lift the cellphone*), or if the next trial implicated a different object. The generality of use-on-lift interference effects can be tested in two contexts in this case. First, a switch between action types (e.g., use then lift) could occur in response to the same object. The continued activation, for example, of a use action associated with a cellphone would interfere with the subsequent production of a lift action to the same object. Second, the action class (use or lift) could be maintained in switching to another object (e.g., *use the spray can* followed by *use the pencil*). If a task set exists for using or lifting in general, then switch costs should be less pronounced when the action class is maintained across consecutive trials but the object changes, relative to when both the action class and the object are changed.

3.1. Method

3.1.1. Subjects

Thirty-two new English-speaking students (24 female, age range 18–34 years, median = 19.5 years) were recruited from the same pool as in Experiment 1, and none had participated in that experiment. The target sample size was set using the same criteria as in Experiment 1. One subject's data were omitted from the reported analyses because of a persisting tendency to begin a response before the action cue was presented. Data are reported for the remaining 31 subjects.

3.1.2. Materials and procedure

The same hand actions and objects were used as in Experiment 1. Instead of cuing actions with an image of a hand grasping an object, subjects were cued on each trial with an auditory sentence. Six sentences were recorded in English in a female voice, each corresponding to a use or lift action associated with one of the three objects from Experiment 1 (e.g., *use the spray can*). The recordings were edited so that verb-article combination had a duration of 630 ms and the object names were allowed to vary in length. Aside from the use of the auditory sentences to cue the action on each trial, the procedure was the same as in Experiment 1. Across the full sample of subjects, the minimum number of trials in a switch condition for any particular subject was 30 and the maximum was 68 (mean = 47).

3.2. Results

Response time was measured from the onset of the object name to liftoff from the button box, as in Experiment 1. Response times below

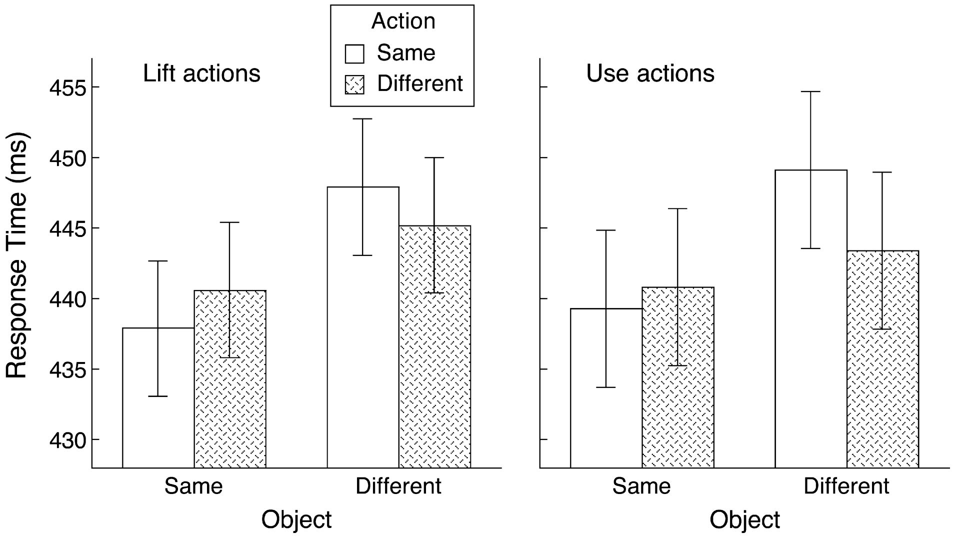

Fig. 2. Mean response time as a function of action type and switch condition in Experiment 1. Error bars are 95% within-subject confidence intervals suitable for comparing switch conditions within an action type (Loftus & Masson, 1994; Masson & Loftus, 2003).

200 ms were excluded as anticipation responses and response times greater than 1300 ms were excluded as outliers (0.30% of correct responses). Error rates once again were very low ($M = 0.2\%$) and 24 subjects made no errors, so we do not report analyses of these data.

Mean correct response time in each condition was computed for each subject and the means computed across subjects are shown in Fig. 3. An ANOVA with action type, action switch, and object switch as repeated-measures factors indicated significant main effects for action switches (a cost of 13 ms), $F(1, 30) = 22.84$, $MSE = 460$, $p < .001$, $\eta_p^2 = 0.43$, and object switches (a cost of 36 ms), $F(1, 30) = 42.59$, $MSE = 1902$, $p < .001$, $\eta_p^2 = 0.59$. There was no significant difference between lift and use actions (708 ms for both), $F < 1$. In addition, there was a significant interaction between the two types of switch, $F(1, 30) = 12.00$, $MSE = 547$, $p < .005$, $\eta_p^2 = 0.29$, with the action switch having minimal impact (3-ms cost) when the object also changed between trials (e.g., *use the stapler* followed by *lift the cellphone*), $F < 1$, and a substantial cost (23 ms) when the object remained the same (e.g., *use the cellphone* followed by *lift the cellphone*)), $F(1, 30) = 21.74$, $MSE = 774$, $p < .001$, $\eta_p^2 = 0.42$.

3.3. Discussion

Our results indicate no reliable difference between use and lift actions, but a reliable switch cost occurs when a different action was carried out on the same object on consecutive trials. Furthermore, we did not observe asymmetrical interference effects when switching between use and lift actions, of the kind reported by Jax and Buxbaum (2010). Thus, switching from a use to a lift action on the same object produced a cost (22-ms) that was no greater than the cost for the reverse sequence of events (a switch from lift to use; 24-ms), $F < 1$. There was also no evidence that the sequential production of actions on different objects implicated a generic use (or lift) task set. Response times of nearly equal magnitude (724 ms vs. 727 ms) occurred regardless of whether the type of action (use versus lift) was maintained between objects (e.g., *use the pencil* followed by *use the cellphone*) or varied (e.g., *lift the pencil* followed by *use the cellphone*), $F < 1$. If persistence of a general task set for use or lift actions had occurred, then pairwise repetition of action type (use/lift) across two different objects should have yielded better performance than a switch from a use (or lift) to a lift (or use) action. No such benefit was observed.

4. Experiment 3

Our findings thus far show no asymmetrical switch costs between use and lift actions, so in Experiment 3 we incorporated a design similar to one used by Jax and Buxbaum (2010). We evaluated switch costs that occurred between blocks of use and lift actions applied to a variety of

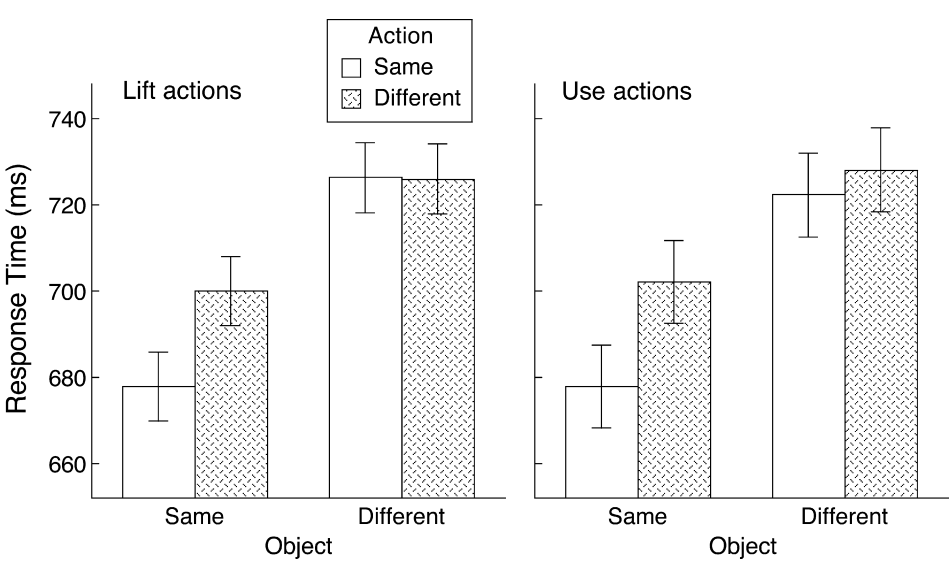

Fig. 3. Mean response time as a function of action type and switch condition in Experiment 2. Error bars are 95% within-subject confidence intervals suitable for comparing switch conditions within an action type.

objects. Switching between two action types separated into blocks of trials may invoke weaker controlled processes than the same two action types occurring in a random order (Tornay & Milán, 2001). Less dependence on control processes responsible for the suppression of a now irrelevant task set might allow us to observe the use-on-lift interference effect reported by Jax and Buxbaum.

4.1. Method

4.1.1. Subjects

Forty-eight English-speaking students (37 female, age range 17–45 years, median = 20 years) were recruited from the same pool as in Experiments 1 and 2. None had taken part in the earlier experiments. Half of the subjects were randomly assigned to each of the two block order conditions (lift then use, use then lift). This sample size was consistent with achieving power of 0.8 to detect a small interaction effect ($f = 0.1$) between a repeated-measures factor and a between-subjects factor, each with two levels, assuming a type I error rate of 0.05 and a correlation between the repeated-measures conditions of 0.90.

4.1.2. Materials

We selected six objects, each with a unique use and lift action, for Experiment 3. Three of the objects were those included in Experiments 1 and 2, and three were new (calculator, lotion bottle, and stapler). A physical instance of each of these objects was affixed to the base of the response apparatus shown in Fig. 1, replacing the three generic response elements used in the previous experiments. The objects were placed on the base in two different orders and half of the subjects were assigned to each of the orders. A digit audio recording in a female voice was made of the name of each of the objects.

4.1.3. Procedure

Subjects were tested under conditions similar to Experiments 1 and 2. They were given four blocks of trials, each of which began with an instruction indicating whether they were to perform lift or use actions in that block of trials. Lift and use actions alternated across the four blocks, with half of the subject performing lift actions in the first block and the remainder starting with use actions. We refer to the first two blocks (one block of each type of trial) as cycle 1 and the second two blocks as cycle 2. No training trials were provided. At the beginning of each trial, a fixation cross appeared at the center of the computer monitor signaling the subject to press a button on the button box. The fixation cross was then erased and 500 ms later the auditory recording of the name of one of the six objects was presented. In response, the subject used his or her dominant hand to make the required action for that block of trials on the indicated object. It was not possible to include the mounted objects in the electrical circuitry of the response apparatus, so response time was measured from the onset of the auditory cue to lift-off from the button box. In each of the four blocks of trials, the names of the six objects were presented six times each in a random order, creating 36 trials in each block. The relevant action instruction (lift or use) was provided at the beginning of each block of trials.

4.2. Results

Response time was measured as in Experiments 1 and 2. Response times shorter than 200 ms were removed as they were considered anticipations, and response times longer than 3500 ms were classified as outliers (0.48% of correct responses). The mean error rate was 0.1% and 42 subjects made no errors, so no analysis of error data is reported.

Mean correct response time in each condition is presented in Fig. 4. An ANOVA of response time data with block order (lift then use, use then lift) as a between-subjects factor and cycle and action type as repeated-measures factors indicated that response times were reliably shorter in the second cycle than in the first (860 ms vs. 892 ms), $F(1, 46) = 8.37$, $MSE = 5856$, $p < .01$, $\eta_p^2 = 0.15$. This difference suggests a general practice effect. The interaction between block order and action type was significant, $F(1, 46) = 5.51$, $MSE = 1927$, $p < .05$, $\eta_p^2 = 0.11$, but was subsumed by the significant three-way interaction, $F(1, 46) = 9.53$, $MSE = 2028$, $p < .005$, $\eta_p^2 = 0.17$. We examined the latter interaction by analyzing the effects of block order and action type separately for the first and the second cycles. There were no reliable effects in the second cycle ($Fs < 1$), but for cycle 1 there was a significant interaction between block order and action type, $F(1, 46) = 10.07$, $MSE = 2909$, $p < .005$, $\eta_p^2 = 0.18$. Additional analyses indicated that among subjects who performed lift actions first, the response-time difference (22 ms) between lift and use actions was not significant, $F(1, 46) = 2.77$, $MSE = 2093$, $p > .10$, $\eta_p^2 = 0.11$. For subjects who performed use actions first, lift actions were carried out significantly faster than use actions (48-ms difference), $F(1, 46) = 7.38$, $MSE = 3724$, $p < .05$, $\eta_p^2 = 0.24$.

4.3. Discussion

The consistent production of use actions did not yield interference on a subsequent block of lift actions. On the contrary, lift actions occurred more rapidly after use actions made during a previous block of trials. This facilitation is not just the result of a practice effect; the reverse order (a block of use actions carried out after a block of lift actions) produced a reliably smaller advantage for use over lift actions. The selective benefit of use on lift actions was observed in Cycle 1 but dissipated by Cycle 2, presumably on account of the build-up of practice effects.

The facilitating effect of prior use on lift actions under these task conditions is noteworthy. To better understand this surprising outcome, we invoke a distinction between the distal and proximal representation of a grasp action (cf. Pacherie, 2008). Distal intentions refer to higher-level goals that are accomplished via an ordered sequence of proximal actions. Within this sequence are what we and others have termed “use actions” (for example, the keys on the phone are depressed with a finger). Clearly, though, the distal intention of carrying out an object’s proper function also extends to lift actions. To use a spray can, for example, we may lift the object from a shelf, followed by a dexterous change in grasp posture that allows the index finger to depress the nozzle while the object remains firmly held against the palm of the hand. Although the first part of the sequence depends on a lift action and the second part on a use action, both components form part of an integrated sequence at the level of the distal goal (i.e., spraying with spray can).

Evidence indicates that the representation of a grasp action is “shaded” (Botvinick, 2007) by the distal goal. For example, neurons in the inferior parietal lobe of the monkey code the proximal act of lifting an object differently, depending on the final goal of the action (Fogassi et al., 2005). Indeed, movement kinematics reflect the modulatory effect of context. Parameters like grip aperture and wrist velocity differ between lift-to-pour and lift-to-drink grasp actions (Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016). It is even the case that observers can make use of kinematic features to classify the distal intentions of others (Becchio, Manera, Sartori, Cavallo, & Castiello, 2012).

Consider a block of trials in which subjects are asked to place their hand on each object as though to use it. We assume that the task set invokes both distal and proximal levels of a motor hierarchy; a grasp posture is generated through the intention to use an object. Now consider a switch to a second block of trials in which subjects are asked to place their hand on each object so as to lift it. What influence on grasp-to-lift actions remains from the previous block of use actions? Vallacher and Kaufman (1996) have argued that proximal actions are inherently unstable (i.e., the motor representation of a proximal act rapidly dissipates after being generated), and that distal goals are typically emphasized when people conceptually represent their own or another’s

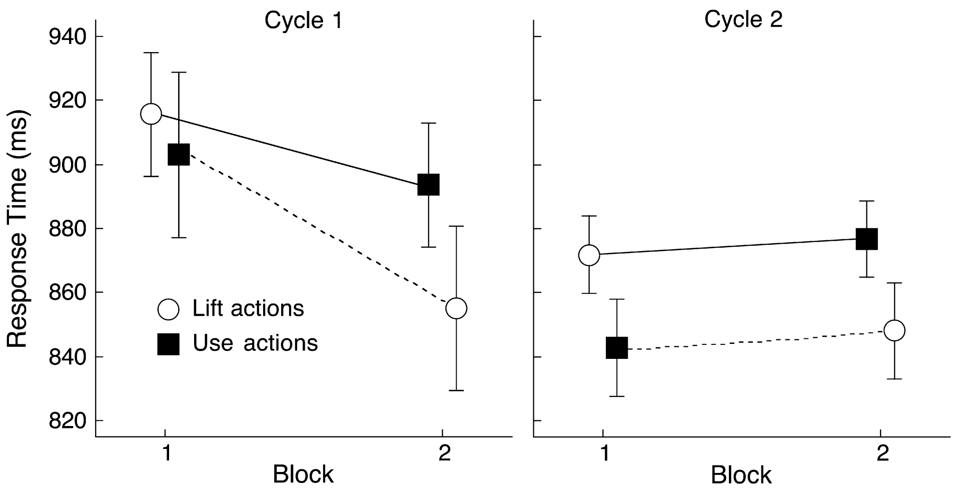

Fig. 4. Mean response time as a function of action type and trial block in Experiment 3. Subjects performed one lift block and one use block in each cycle and within-subject means for each block are connected by lines. Error bars are 95% within-subject confidence intervals suitable for comparing one action type against another (connected by a line) within a cycle.

actions (for additional commentary see Masson, Bub, & Lavelle, 2013). We make the reasonable assumption that a use task set, when sustained after a block of trials, is represented in terms of distal intentions rather than proximal actions. To be clear, it is not a use action like poking the keys of a calculator that we assume is sustained over time, but the more abstract intention to make use of the calculator. We conjecture that this higher level influence was responsible for the benefit of prior use on lift actions.

5. Experiment 4

We tested the idea that the intention to carry out the proper function of an object facilitates the production of a lift action on the same as opposed to some other object. Consider the following task: Subjects are told that they must plan to carry out a use action on object A. Just prior to this action they must produce a lift action on the same object or on a different object B. Faster lift actions to object A than B would confirm the idea that the distal intention to use an object in working memory has a facilitating effect on an intervening lift action to the same object. No such benefit would occur if the prior intention is to lift object A and subjects first produce a use action on the same or a different object. In other words, although a use task set should facilitate lift actions, no such effect should be found when the intention to lift an object provides the task context for a use action.

5.1. Method

5.1.1. Subjects

A new sample of 31 subjects (26 female, age range 17–25 years, median = 20 years) was drawn from the same pool as in the earlier experiments. None had taken part in the earlier experiments. The target sample size ($n = 32$) was established using the same criteria as in Experiments 1 and 2, although we were able to test only 31 subjects before access to our subject pool ended for the academic term.

5.1.2. Materials

The six objects and associated lift and use actions from Experiment 3 were used. An imperative sentence was recorded in a female voice specifying the use or lift action for each object (e.g., *use the cellphone*). These sentences were used to specify an action that the subject was to retain in working memory and to execute at the end of a trial. In addition, a digital grayscale photograph of a male right hand gripping each of the objects in either a use or lift posture was produced, similar to the images used in Experiment 1. Mirror image versions of these photographs were generated for use with left-handed subjects. These visual cues were used to specify an action that the subject was to make immediately upon seeing the cue. The response apparatus with mounted objects from Experiment 3 was used.

For the experimental task, subjects were presented one auditory sentence and one visual cue on each trial. All possible pairs of the 12 sentences and 12 visual cues were used, excluding pairs that designated the same action (e.g., picture of a hand pressing a button on a cellphone, *use the cellphone*). The resulting combinations comprised six conditions, defined by the type of action indicated by the visual cue (lift, use) and the by the type of switch between visually cued action and action indicated by the imperative sentence (same object and different action: action switch; different object and same action type: object switch; different object and different action type: both switch).

5.1.3. Procedure

Subjects were tested under conditions similar to Experiment 3. They were given two training tasks, the first requiring them to make the reach-and-grasp responses indicated by the 12 visual cues. Each cue was presented twice in a random order. In the second task, the imperative sentences were presented twice each in a random order and subjects executed the described action in response to each cue. The experimental task was then presented. Each trial began with a fixation cross at the center of the computer monitor, which signaled the subject to press a button on the button box with the index finger of his or her dominant hand. The fixation cross was then erased and an auditory sentence was presented. Subjects were instructed to hold the described action in working memory as they would be cued to perform that action at the end of the trial. Following the sentence and a 500-ms delay, a visual action cue was presented and the subject was to immediately carry out that action with his or her dominant hand. After the action was completed the experimenter used the computer keyboard to classify the response as correct or incorrect (with any error classified according to whether the response involved an incorrect action, incorrect object, or both). Once the experimenter's input was entered, the subject again pressed the button on the button box and an orange disk appeared on the monitor signaling the subject to execute the action that had been described in the auditory sentence at the start of that trial. The experimenter classified that response as correct or incorrect and the trial was then complete.

Subjects were first given 36 practice trials, with six trials representing each of the six conditions defined by visually cued action type (lift, use) and switch condition (action, object, both). A series of 180 critical trials was then presented with a rest break provided after every 30 trials. Across the 180 trials, each of the 12 visual cues was presented 15 times. Of the 15 occurrences of each visual cue, five were assigned to each of the three switch conditions. For the action-switch condition, the same object was implicated by the visual cue and the auditory sentence (e.g., a hand lifting the stapler as the visual cue, *use the stapler* as the auditory sentence) and each particular pairing was presented five times. For the object-switch condition the same action class (lift or use) was designated by the visual cue and the auditory sentence, but different objects were involved (e.g., a hand lifting the stapler, *lift the cellphone*). Each of the five presentations of a particular visual cue in this condition was paired with an auditory sentence specifying a different one of the five available alternative objects. The both-switch condition was structured just like the object-switch condition, except that the visual cue and auditory sentence specified different actions (e.g., a hand lifting the stapler, *use the cellphone*). Across the full set of critical trials, then, each visual cue and each auditory sentence was presented equally often.

5.2. Results

Our primary interest was in responses to the visual cues, made while holding a different action in working memory (as specified by an auditory sentence). Response time was measured from the onset of the visual cue to liftoff from the button box. As in Experiment 3, we were not able to include reach time in our measure. Response times shorter than 200 ms were considered to be anticipations and were excluded. Response times longer than 2500 ms were excluded as outliers (0.46% of responses). Unlike the earlier experiments, subjects performed the visually cued action while holding a competing action in working memory. This dual task requirement led to a higher error rate than was observed in the first three experiments (1.4% overall), and so we present an analysis of error rates for this experiment.

For completeness, the mean response time for each condition in the experiment is shown in the upper section of Fig. 5. As described in the introduction to this experiment, however, our primary interest was in two specific contrasts and so we restrict our statistical tests to those cases.[1] The first contrast tested the prediction that holding the distal goal of using an object would privilege the corresponding lift action for that object. This influence should be made apparent by a shorter response time for lift actions when a use action involving the same object is held in working memory (action switch) relative to when a use action applied to a different object is being held for future execution (both switch). This contrast yielded a significant effect of 32 ms, $F(1, 30) = 5.14$, $MSE = 3097$, $p < .05$, $\eta_p^2 = 0.15$.

The second contrast was complementary to the first and was aimed at performance of use actions in the action-switch and both-switch conditions. In this case, we anticipated no reliable effect because the standard action hierarchy has the functional (use) action as the ultimate distal goal and it subsumes the proximal lift action. It is this relationship that was expected to produce the effect seen with lift actions. But the reverse relationship does not hold, so no comparable effect was expected for use actions. Indeed, the contrast between action- and both-switch conditions for use actions was not significant (a 3-ms effect, $F < 1$). A Bayesian test of this effect using the *ttestBF* function in the *BayesFactor* package in **R** indicated that the data supported the null model over a model that included an effect of switch condition with a Bayes factor of 4.9 (indicating positive evidence in favor of the null model).

The mean error percent for each condition is shown in the lower section of Fig. 5. Again, our primary interest was in the two planned contrasts, so we report significance test for those specific cases. Interestingly, for lift actions, the error rate was significantly higher in the action-switch condition than in the both-switch condition (a 2.1% effect), $F(1, 30) = 13.71$, $MSE = 4.91$, $p < .001$, $\eta_p^2 = 0.31$. No effect was found for use actions (a 0.5% difference, $F < 1$; Bayes factor of 3.4 favoring the null model). The effect on error rates for lift actions might

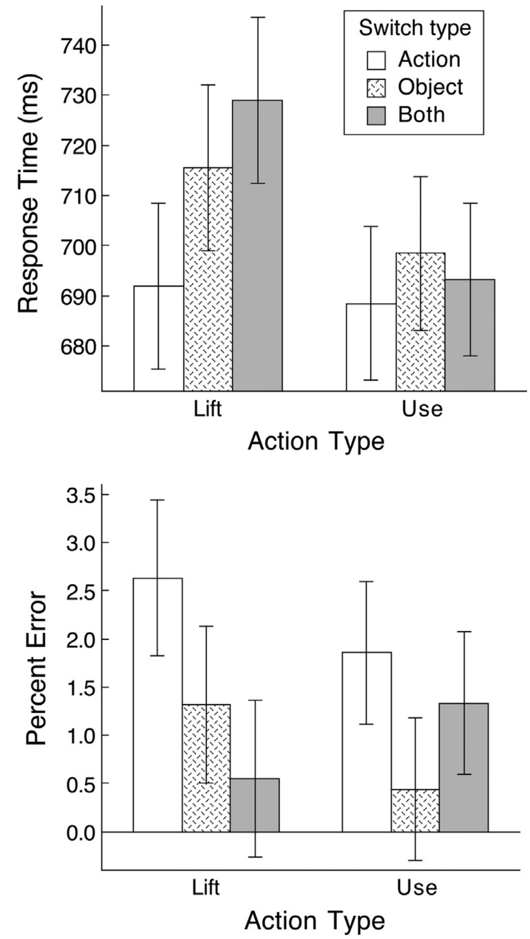

Fig. 5. Mean response time and percent error as a function of action type and switch condition in Experiment 4. Error bars are 95% within-subject confidence intervals suitable for comparing switch conditions within an action type.

initially be taken as a speed-accuracy tradeoff, but there are a number of aspects of the data that we believe support a different interpretation. First, no such tradeoff occurred for use actions and it is not clear why a tendency to make fast, error-prone responses would be associated with one type of action and not another, particularly as no tendency of that kind was apparent in any of the previous experiments. Second, the error made when subjects were cued to perform a lift action in the action-switch condition consisted, in every instance, of making the use action instead of the designated lift action to the indicated object. The intrusion of the use action occurred at a significantly higher rate than in the baseline condition (both-switch) for the lift actions, but the complementary effect was not seen for use actions. This asymmetry, both in response times and in action-intrusion errors, suggests a special status for contexts in which observers make a lift action while planning to later make a use action with the same object.

To ensure that subjects were reliably maintaining in working memory the action described by the imperative sentence on each trial, we assessed the accuracy of the responses to those cues. On average, subjects correctly performed the action described by the auditory sentence on 98.2% of trials. An analysis of errors on this task with action type and switch type as factors revealed no significant effects ($Fs < 1.7$).

5.3. Discussion

The final goal of an act like grasping an object – whether the intention, for example, is to eat the object or to place it in a container – determines how the action is represented in the motor cortex (Fogassi

[1] An ANOVA with switch type (action only, both) and action type (lift, use) as factors did not find a significant interaction, $F(1, 30) = 3.43$, $MSE = 1894$, $p < .10$, although the linear mixed models analysis found that interaction to be significant, $t = 2.21$, $p < .05$.

et al., 2005). We obtained striking evidence that this principle applies to lift actions on familiar objects. A lift-then-use action is an organized motor sequence frequently produced to enact the proper function of an object. Accordingly, lift actions are faster when a use action on the same object (versus a different object) is maintained prospectively in working memory. No such contextual effect occurs when a use action is produced in the context of a lift action as a distal goal.

It is noteworthy that our results offer no support for the view that use actions interfere with subsequent lift actions. We will have more to say about this discrepancy in the final discussion. For now, we emphasize that previous evidence for use-on-lift interference was observed when objects appeared suddenly after being occluded, and lift actions were in general produced more rapidly than use actions (Jax & Buxbaum, 2010). Beyond the fact that in our experiments objects were continuously viewed, a number of other methodological differences may have prevented any reliance on a fast visuomotor route for lift actions. We utilized few objects (3 or 6) evoking a limited set of grasp actions, and the repeated exposure to these target items could have induced a greater reliance on stored motor representations to generate lift actions.

A second consideration is the role of language or more generally, conceptual levels of task set in modulating the nature of lift actions. We either used imperative sentences (Experiments 2 and 4) or concrete nouns (Experiment 3) to designate an action on a particular object. It seems plausible (and to us, quite likely) that the production of grasp actions to labelled objects is invariably mediated by accessing stored knowledge. This raises questions about the nature of lift actions evoked in tasks like picture-word matching or in any context determined by abstract conceptual codes. We relied on verbal instructions to evoke the distal intention of using an object in Experiment 4. Remarkably, this active representation in working memory produced effects on lift actions cued by a visual image (the depiction of a hand grasping an object). Without the modulating influence of a distal goal, motor responses to such images yielded neither costs nor benefits in switching between lift and use actions (see Experiment 1). Given the distal intention to use an object, however, the same kind of visual image generated lift actions that were encapsulated within a motor hierarchy. In the context of a prior intention to use an object, an intervening lift action – even when triggered by a visual image – is now represented as part of an ordered sequence habitually connected to the object's proper function.

6. General discussion

We have examined switch costs between use and lift actions on objects in a continuously visible array. In a first experiment, actions were cued randomly by means of a picture demonstrating a lift or use action with respect to a particular object. Repeated actions on the same object were faster, but lift actions were not produced more rapidly than use actions, and there were no switch costs. This limited result occurred for actions that were largely imitative, cued by the picture of a hand acting on the target object. In the next experiment, subjects responded to imperative sentences (e.g., *lift the spray can*) rather than to the image of a depicted action on an object. Switch costs were observed but they again did not show any difference between lift and use actions. As in Experiment 1, repeating an action on the same object yielded the fastest responses. Switching to a different action on the same object induced a cost for both lift and use actions. A further cost was incurred when a switch took place between objects. Actions cued by imperative sentences are thus affected by (a) switch costs induced when a new action must be produced on the same object, and (b) a further cost when the sentence demands the switching of attention to another object. We found no evidence that lift actions were produced more rapidly than use actions, nor for a persistent task set induced by use actions. Switch costs were the same regardless of whether the action type was maintained between objects or changed from a use to a lift action.

In Experiment 3, we implemented a blocked design that was a closer approximation to conditions yielding support for the claim that different motor systems underlie the production of lift and use actions. Objects remained continuously in view, but a block of lift actions was followed by a block of use actions, or vice versa. The production of use and lift actions separated into blocks of trials may place less demand on controlled processes than random switches, allowing the impact of a change in task set to reveal itself more clearly.

Interestingly, a novel result emerged. Lift actions did not occur more rapidly than use actions; there was no difference between the speed of lift actions made first in block 1 and use actions occurring second in block 2. However, when subjects performed use actions first, subsequent lift actions were carried out more rapidly than use actions. In our view, the benefit of use actions on a block of lift actions occurs because the distal goal of using an object can be sustained over time. Lift actions after a block of use actions are modulated by this higher-level context.

The proposed facilitation of use on lift actions was tested directly in Experiment 4 by requiring subjects to hold in working memory the distal goal of carrying out a use action on an object. Before producing this action, they were cued to make another action on the same or a different object. Lift actions were faster to objects targeted for a subsequent use action than objects irrelevant to the distal goal. This result was not simply due to carrying out two different actions in short order on the same object; use actions to an object marked for the distal goal of a lift action were not faster than intervening use actions applied to another object. We propose that the intention to use an object facilitates the prior action of lifting it because the motor sequence – lift then use – is habitually conscripted to enact the proper function of an object.

Recently, we have obtained additional support for this conjecture (Masson & Bub, 2016). With just three objects (a cellphone, a pencil, and a spray can) continuously visible in an array, subjects were asked to prepare a lift or use action by means of a sentence (e.g., *use the cell phone*). In the interim, they executed a speeded action cued by the image of a hand representing the goal posture of a use or lift action. As predicted, the distal intention to produce a use action facilitated an intervening lift action on the same object relative to a lift action on a nontarget object. The prior intention to lift an object (say, the pencil) not only conferred no advantage to an intervening use action but in fact, caused the action to be produced more slowly on the target object (the pencil) than to another object (the cellphone or spray can).

Clearly, a distal goal provides a contextual influence on proximal actions. A lift grasp enacted on an object, governed by the intention to use it, is a well-practiced motor event. In contrast, the distal intention to lift an object is incompatible with a proximal use action on the same object. In other words, a lift-to-use action has a well-defined hierarchical representation in the motor system, but no such coherent organization exists for "using-to-lift" an object.

We conclude with the following additional remarks on the difference between use and lift actions. There is good evidence that under certain task conditions, when immediate visuomotor transformations are called upon without enlisting conceptual levels of representation, an object can rapidly trigger a grasp action determined only by its perceived structure (Rossetti, Pisella, & Vighetta, 2003). For example, lift actions are much faster than use actions if an object appears suddenly, instructions are to produce a rapid grasp action, and the task context discourages any higher-level constraints on motor intentions (Jax & Buxbaum, 2010; Osiurak et al., 2013).

A fundamental question that remains, however, is the role of knowledge- versus structurally-based representations in generating lift versus use actions under conceptually based task conditions, such as in language tasks like picture-word matching (Lee, Middleton, Mirman, Kalénine, & Buxbaum, 2013). Do features of action play a role in the semantic classification of objects even when observers have no intention to engage in reach-and-grasp actions?

This issue is undoubtedly contentious (e.g., Proctor & Miles, 2014).

In some conceptual tasks, objects evoking different lift and use actions (like a spray can) have been shown to yield slower classification responses than objects like a baseball, evoking very similar use and lift actions (e.g., Kalénine, Wamain, Decroix, & Coello, 2016). Interpretation of this outcome is based on the assumption that conflict occurs between a rapidly triggered lift action driven by the perceived structure of the object, and slower use actions generated from stored knowledge of the object's functional properties. We should note that the mere presence of a discrepancy in the speed of responding to "conflict" versus "non-conflict" objects is not necessarily indicative of the presence of competing motor representations. The latter class of object – affording the same or very similar use and lift actions – tends to include exemplars that are structurally simpler than conflict objects. For instance, a baseball, various kinds of bottles or containers, a vase (these examples of "non-conflict" objects are all taken from Kalénine et al.) have few if any parts, whereas a remote control, scissors, tape dispenser and calculator (examples of "conflict" objects) are visually more complex. Visual complexity, as pointed out by Gerlach and Marques (2014), produces different effects on performance, depending on the level of object discrimination needed to perform a task. If a perceptual decision can be made on the basis of coarse-grained information (global shape) without the need to determine the identity of an object, visual complexity can enhance performance, because there are fewer objects with similar shapes that compete for selection. When finer-grained discrimination is required, perceptual decisions are harder for visual objects that have more complex structural representations.

Task demands do indeed appear relevant to understanding the different effects induced by conflict and non-conflict objects on performance (Roche & Chainay, 2017). When subjects were asked to produce speeded judgements on whether an object could be reached with a grasp action, Kalénine et al. (2016) reported that conflict objects in peripersonal space yielded slower "yes" responses than non-conflict objects. Objects that were outside this space, demanding a "no" response, showed if anything, the reverse pattern; responses to conflict objects were faster than responses to non-conflict objects. These results make sense if "yes" responses to an object in peripersonal space were based on fine-grained analysis of shape, yielding slower responses to visually more complex objects. In contrast, "no" responses to objects in extra-personal space were mainly determined by their global structure. Consistent with this interpretation, a speeded classification task that always required detailed classification of object identity (whether the object was typically found in the kitchen) yielded slower performance for conflict than non-conflict objects, regardless of whether they occurred in peri- or extrapersonal space. Although competing motor representations may contribute to the difference between conflict and non-conflict objects when reach-and-grasp actions are produced (e.g., Watson & Buxbaum, 2015), visual complexity appears to determine the difference between these two classes of object under task conditions that do not invoke object-related action representations.

When objects appear suddenly in peripersonal space after visual occlusion, lift actions following a block of use actions are slower for conflict than non-conflict objects (Jax & Buxbaum, 2010). One interpretation of this result is that use actions persist over trials, generating interference with competing lift actions to conflict objects; no such competition affects responses to objects that invite non-conflicting use and lift actions. We suggest an alternative account; the sudden appearance of an object can trigger a grasp action determined by its overall shape, without access to higher level conceptual knowledge. Use actions, by contrast, invariably depend on stored motor representations. The repeated production of use actions will result in a persistent reliance on object identity when the task set demands the production of lift grasps, pre-empting the contribution of the fast visuomotor route. Because the objects in our version of this experiment (Experiment 3) were constantly in view, no such contrast between lift actions performed before versus after use actions was in effect.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2018.01.013.

References

Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: From thought experiments to empirical evidence. *Frontiers in Human Neuroscience, 6*, 117. http://dx.doi.org/10.3389/fnhum.2012.00117.

Bohg, J., Morales, A., Asfour, T., & Kragic, D. (2014). Data-driven grasp synthesis–A survey. *IEEE Transactions on Robotics, 30*, 289–309.

Botvinick, M. M. (2007). Multilevel structure in behaviour and in the brain: A model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society B, 362*, 1615–1626.

Bub, D. N., Masson, M. E. J., & Kumar, R. (2018). Time course of motor affordances evoked by pictured objects and words. *Journal of Experimental Psychology: Human Perception and Performance, 44*, 53–68.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports, 6*, 37036. http://dx.doi.org/10.1038/srep37036.

Chua, K.-W., Bub, D. N., Masson, M. E. J., & Gauthier, I. (in press). *Grasp representations depend on knowledge and attention*. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science, 308*, 662–667.

Gentilucci, M. (2002). Object motor representation and reaching-grasping control. *Neuropsychologia, 40*, 1139–1153.

Gerlach, C., & Marques, J. F. (2014). Visual complexity exerts opposing effects on object categorization and identification. *Visual Cognition, 22*, 751–769.

Herbort, O., & Butz, M. V. (2011). Habitual and goal-directed factors in (everyday) object handling. *Experimental Brain Research, 213*, 371–382.

Jax, S. A., & Buxbaum, L. J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition, 115*(2), 350–355.

Kalénine, S., Wamain, Y., Decroix, J., & Coello, Y. (2016). Conflict between object structural and function affordances in peripersonal space. *Cognition, 155*, 1–7.

Lee, C., Middleton, E., Mirman, D., Kalénine, S., & Buxbaum, L. J. (2013). Incidental and context-responsive activation of structure- and function-based action features during object identification. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 257–270.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Masson, M. E. J., & Bub, D. N. (2016). *Action representations evoked by imperative sentences*. Paper presented at the annual meeting of the Psychonomic Society, Boston, MA.

Masson, M. E. J., Bub, D. N., & Lavelle, H. (2013). Dynamic evocation of hand action representations during sentence comprehension. *Journal of Experimental Psychology: General, 142*, 742–762.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

Osiurak, F., & Badets, A. (2016). Tool use and affordance: Manipulation-based versus reasoning based approaches. *Psychological Review, 123*(5), 534–668. http://dx.doi.org/10.1037/rev0000027.

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

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*, 179–217.

Proctor, R. W., & Miles, J. D. (2014). Does the concept of affordance add anything to explanations of stimulus-response compatibility effects? In B. H. Ross (Vol. Ed.), *The psychology of learning and motivation: Vol. 60*, (pp. 227–266). San Diego: Elsevier.

Roche, K., & Chainay, H. (2017). Is there a competition between functional and situational affordances during action initiation with everyday tools? *Frontiers in Psychology, 8*, 1073. http://dx.doi.org/10.3389/fpsyg.2017.01073.

Rossetti, Y., Pisella, L., & Vighetta, A. (2003). Optic ataxia revisited: Visually guided action versus immediate visuomotor control. *Experimental Brain Research, 153*, 171–179.

Tornay, F. J., & Milán, E. G. (2001). A more complete task-set reconfiguration in random than in predictable task switch. *Quarterly Journal of Experimental Psychology, 54A*, 785–803.

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General, 123*, 34–80.

Vallacher, R. R., & Kaufman, J. (1996). Dynamics of action identification: Volatility and structure in the mental representation of behavior. In P. M. Gollwitzer, & J. A. Bargh (Eds.). *The psychology of action: Linking cognition and motivation to behavior* (pp. 260–282). New York: Guilford Press.

Watson, C. E., & Buxbaum, L. J. (2015). A distributed network critical for selecting among tool-directed actions. *Cortex, 65*, 65–82.