

Understanding the dynamics of functional and volumetric action representations when prepared
for immediate execution

by

Duo Wang
BSc, University of Victoria, 2016

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

© Duo Wang, 2018
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other
means, without the permission of the author.

Supervisory Committee

Understanding the dynamics of functional and volumetric action representations when prepared
for immediate execution

by

Duo Wang
BSc, University of Victoria, 2016

Supervisory Committee

Dr. Daniel Bub (Department of Psychology)
Co-Supervisor

Dr. Michael Masson (Department of Psychology)
Co-Supervisor

Abstract

This study examines the state of competing affordances when an action is prepared for immediate production. More specifically, we investigated the nature of motor representations evoked by distinct action intentions, with a special interest in functional (grasp to use) and volumetric (grasp to lift) actions. With just two objects available, participants were asked to prepare an action on a particular object (e.g., preparing to lift the cellphone), and when signaled, either to perform this original action plan or to switch to executing an alternative one, either on the same or different object. By manipulating cueing methods used for indicating the preparatory and the target action plans, we found distinct patterns in the effect of preserving either object (different action on the same object) or action (same grasp type on a different object) on action execution. Changing either component of the action-object pairing incurred a cost in response time. In Experiment 1, a cost was observed when a prepared action was switched to an alternate action on the same object. For example, preparing to lift a cellphone but switching instead to a use action on the same object, incurred a cost. A further cost was found when subjects prepared a functional action to one object (e.g., use the cellphone) but switched to the same class of action on the alternate object (e.g., use the spray can). Both these effects were found to operate at the motor level. No costs were observed when subjects switched from a planned action to naming the target object (Experiment 3). Crucially, it was found that the nature of the cueing method instructing subjects to switch from a planned to an alternate action impacted the effect of action congruency. The cost observed in Experiment 1 when subjects switched to an alternate action on the same object occurred when the switch from a planned to an alternate action was cued by a verb-noun combination (e.g., *use cellphone*). No such cost occurred when the action was cued by a verb (e.g., *use*) and the target object was spatially cued by an arrow pointing to its location

(Experiment 2). The cost of switching from a planned action type (e.g., a use action) to the same action type carried on the alternate object also depended on whether the planned action was verbally or spatially cued. These results provide new evidence on the nature of action representations associated with different motor intentions, as well as of the nature of action-object pairings.

Table of Contents

Supervisory Committee.....	ii
Abstract	iii
Table of Contents.....	v
List of Figures.....	vi
Acknowledgments	vii
Introduction	1
Experiment 1	5
1.1 Method.....	5
1.2 Results	11
1.3 Discussion.....	17
Experiment 2	21
2.1 Method.....	21
2.2 Results	23
2.3 Discussion.....	27
Experiment 3	30
3.1 Method.....	30
3.2 Results	32
3.3 Discussion.....	35
General discussion	38
Action-object pairing	39
Action-object-feature pairing.....	39
Congruency effect of object	42
Congruency effect of action	44
References	47

List of Figures

<i>Figure 1.</i> Function, volumetric and touch action hand gesture definitions.....	6
<i>Figure 2.</i> Response apparatus used in Experiment 1	7
<i>Figure 3.</i> Experimental material layout	8
<i>Figure 4.</i> Example of possible trial conditions in Experiment 1.....	9
<i>Figure 5.</i> The sequence of events in Experiment 1.. ..	11
<i>Figure 6.</i> Green dot response time in Experiment 1	13
<i>Figure 7.</i> Mean response time by switch conditions examining object congruency effect in Experiment 1.....	15
<i>Figure 8.</i> Mean response time of four switch conditions in Experiment 1.....	16
<i>Figure 9.</i> Mean response time by switch conditions examining action congruency effect in Experiment 1	17
<i>Figure 10.</i> Green dot response time in Experiment 2.....	23
<i>Figure 11.</i> Mean response time by switch conditions examining object congruency effect in Experiment 2.....	25
<i>Figure 12.</i> Mean response time of four switch conditions in Experiment 2.....	26
<i>Figure 13.</i> Mean response time by switch conditions examining action congruency effect in Experiment 2.....	27
<i>Figure 14.</i> Green dot response time in Experiment 3.....	33
<i>Figure 15.</i> Mean response time by switch conditions examining object congruency effect in Experiment 3.....	34
<i>Figure 16.</i> Mean response time by switch conditions examining action congruency effect in Experiment 3.....	35
<i>Figure 17.</i> Two processing routes involved in programming action-object-feature pairings.....	41
<i>Figure 18.</i> Example status of processing routes in Experiment 1.	43
<i>Figure 19.</i> Example status of processing routes in Experiment 2.	44

Acknowledgments

I would express my very great appreciation to:

Dr. Daniel Bub and Dr. Michael Masson for their great team work providing patient guidance, enthusiastic encouragement and useful critiques;

Marnie Jedynak for making all the experiments come alive and working smoothly;

Research assistants for their great help in the process of data collection;

Colleagues in the lab and the department for being an amazing support group;

And my parents, sister, and friends for their unconditional love and companionship throughout my studies.

Introduction

Interactions with objects and tools are essential for day-to-day living. Depending on our goals, the action applied to an object varies; for instance, we apply a vertical power grasp to lift and transport a spray can. Alternatively, we use the spray can by applying a precision grip with an extended forefinger to depress the nozzle.

Action plans can occur at various levels of abstraction (Pacherie, 2008).

A *proximal level* of representation deals with an immediate goal of the agent, and the situation he or she currently perceives. A proximal intention is sufficiently abstract that it can be accessed through language. The phrase “use cellphone” for example, refers to the proximal goal of carrying out the typical function of the object. Below this level are procedures that represent the actual movements needed to reach for and grasp the target object. The shape and positioning of our hands on an object will vary depending on the nature of our proximal intentions. Grasping to use an object like a cellphone requires a different set of hand postures than grasping to lift it. Action representations related to lifting and moving an object will be termed *Volumetric* actions; whereas action representations driven by the intention to carry out the proper function of the target object will be called *Functional* actions (Bub, Masson, & Cree, 2008). Neuropsychological evidence indicates that tasks evoking volumetric and functional action representations elicited discrepant cortical activation patterns (Buxbaum, Kyle, Tang, & Detre, 2006; Creem-Regehr & Lee, 2005). Johnson-Frey (2004) further corroborated these findings by showing brain lesion patients exhibiting selective behavioral deficits of lifting and using actions, signaling differential neural substrates involved in activating the two types of motor representations. Moreover, behavioral data revealed that under certain task demands, these two types of action

representations exhibited differences in time course (Bub, Masson, & van Mook, 2018; Jax & Buxbaum, 2010; Osiurak, Roche, Ramone, & Chainay, 2013)

To date, several studies have dealt with the nature of functional and volumetric action representations evoked under various task demands and experimental designs. It has been suggested that semantic processing of words referring to manipulable objects (e.g., *cellphone*) embedded in a sentence would evoke multiple affordances including both the functional and volumetric actions (Bub et al., 2008; Masson, Bub, & Newton-Taylor, 2008), while contextual information provided by the sentence facilitates the selection of particular action (Bub & Masson, 2010). Furthermore, when presented with the object (either object pictures or real manipulable objects), volumetric motor representations are evoked through a rapid visuomotor route guided by the global shape and structural features of the target object; whereas the activation of functional actions requires accessing stored knowledge of the object identity and previous experience interacting with it (Bub et al., 2008; Osiurak & Badets, 2016).

Recently, researchers have also examined the switch cost of functional and volumetric actions with a constantly visible array of familiar objects. In particular, Bub et al. (2018) investigated switch costs while controlling and manipulating the state of motor representations based on switches that could occur when the preceding action plan was executed, repeated or postponed. For example, in one of the experiments, Bub and colleagues asked participants responding to short imperative sentences (e.g., *use the cellphone*) by acting on a target object with the indicated grasp on consecutive trials. The imperative sentences invited either functional or volumetric actions on an object in specific locations. From trial to trial, the action classes (functional or volumetric) and the selected object varied, and the cost of switching from a finished motor representation to a newly activated one was examined. Interestingly, the

researchers found that there was a significant cost when switching between action classes on a particular object, but no effect when maintaining the type of action while switching between objects (Bub et al., 2018). However, the switch cost from a completed action to another examines the carry-over effect of the previously integrated motor representation on the programming of a new one. Little is known about the cost of switching from functional and volumetric motor representations that are ready for immediate action execution but not yet performed.

To appreciate the importance of this issue, consider the following scenario. You are about to pick up your cellphone which has dropped on the floor. While you are reaching out with the palm facing down, you receive a new message on this device. You would then automatically rotate your wrist orientation to grasp the cellphone from the side in order to use the keyboard and reply to the text. In this case, the original goal of interacting with the cellphone was to lift it, but now has been aborted, and instead must give way to a new goal of using it.

Clearly, adaptive action control is fundamental to successful interaction with manipulable objects given the dynamic environment we live in (Pezzulo & Cisek, 2016). For this reason, it is important to investigate cognitive processes underlying the reprogramming of action plans for a better understanding of action representations. This current study, therefore, aims at expanding our understanding of cognitive mechanisms involved in selecting between competing motor representations before action execution. In the set of experiments, we created a condition where subjects prepared for the immediate production of an action, and completed the action when presented with a green dot. This occurred on 25% of the trials. On the remaining trials, subjects were cued to abort the original action and switch to an alternative hand action afforded by either the same or a different object. We examined the relationship between an actively prepared action

and an alternative target action that is executed instead of the prepared action. It will be seen that the relationship between object-action pairings plays an important role in selecting between competing motor representations.

Experiment 1

With a particular interest in motor representations invoked by distinct intentions, Experiment 1 examines how a fully planned action on a particular object can have an impact on programming an alternative motor representation with congruent or incongruent action plan elements. In addition to the functional and volumetric actions that were discussed above, we introduce another action class in this experiment, *touch*. Notice, unlike functional and volumetric actions motivated by intentions that are meaningful and common in daily life, touch is rarely associated with any meaningful goals, though occasionally touch is considered a hand gesture used for checking surface temperature. So far, literature concerning touch actions mostly focused on tactile sensitivity instead of its motor properties (Pacherie, 2008). It is assumed that a touch action can be automatically triggered since it requires no prior knowledge, is affected by physical features to a minimal extent, and relies predominantly on the object location.

Therefore, our expectation is that the inclusion of touch actions will enhance our understanding of the nature of its motor representation, as well as serving as an important comparison with respect to the functional and volumetric grasp actions.

1.1 Method

Participants

Thirty-eight undergraduate students (females = 30, age range = 18-37, mean age = 20.97, right-handedness = 33) from the University of Victoria were recruited through the research participation system SONA. Participants received bonus credits toward their undergraduate psychology courses. Informed consent was given by all participants at the beginning of the experiment. The experiment was approved by the University of Victoria Human Research Ethics Committee.

Material

Two response elements simulating day-to-day manipulable objects (cellphone, spray can) were selected, and participants were instructed to consider the elements as real-life objects when acting on them. The chosen response elements possessed strongly contrasting physical features so that each of the two objects was associated with unique functional, and volumetric grasps. For consistency, we had specific hand gesture definitions of each action type on the two response elements for a total of six possible actions (see Figure1).







Object	Action		
	Functional	Volumetric	Touch
Cellphone			
Spray can			

Figure 1. Function, volumetric and touch action hand gesture definitions of the cellphone and spray can for a right-handed participant

The response elements were configured on an apparatus panel such that the centerlines of the response elements were symmetrical to each other across the midline of the panel base with a horizontal distance of approximately 15 cm apart (see Figure 2). The position of the response elements (spray can-cellphone vs. cellphone-spray can) was counterbalanced between subjects.

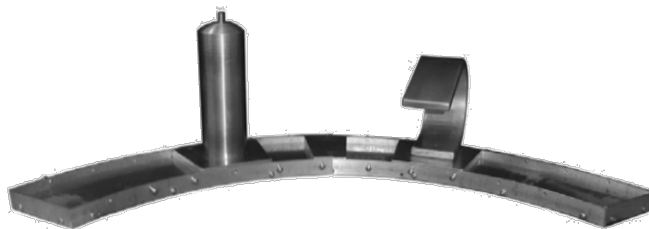


Figure 2. Response apparatus used in Experiment 1. Two response elements modeling real life objects mounted onto the apparatus panel (Left: spray can; Right: cellphone). The location of elements was counterbalanced across subjects.

Responses were made by a combination of key presses on a six-key button box (keys horizontally arranged), and reach-and-grasp actions on the response apparatus. The button box allowed measuring of key lift-off response time, whereas the response apparatus was equipped with touch-sensitive detectors that enabled the recording of movement/reach-out response time (see Bub et al., 2008, for more detailed information of the response device).

Two types of cueing methods with visual and audio stimuli were used for this experiment. The visual stimuli were printed on a G3 Macintosh computer monitor screen, and the audio output was delivered through a Logitech USB computer headset. The first type of cueing method was a visual language cue (Phrase cue), consisting of a capitalized verb (i.e., USE, LIFT, or TOUCH) and a capitalized noun (i.e., CELLPHONE, or SPRAY CAN). An example phrase cue would be USE CELLPHONE. The phrase cues were centered in a 360x360 pixel image. The other cue used was a joint audio-visual stimulus (Audio/Arrow cue). A recorded English-speaking female voice of either *use*, *lift*, or *touch* was first given through the headphone, and immediately after which, an arrow centered in a 159x210 pixel image was presented vertically on the screen, offsetting either to the left or to the right 300 pixels away from

the vertical midline, pointing down to one of the response elements. In addition, a green dot centered in a 360x360 pixel image appeared in some trials serving as a go signal. Figure 3 shows the layout of the materials used.

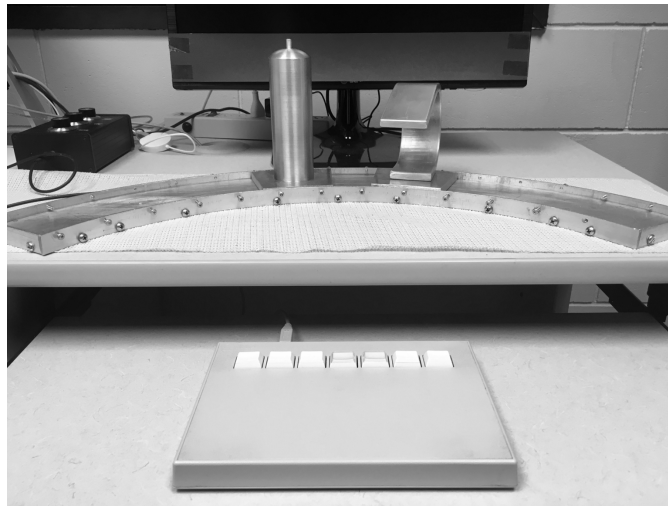


Figure 3. Experimental material layout (rear to near): monitor screen, response apparatus, and button box.

Design

Participants were instructed to prepare an action plan indicated by an Audio/Arrow cue and respond to an execution stimulus. Subjects carried out this particular action plan prepared when a green dot appeared; when they saw a printed phrase instead, they switched to executing the action plan suggested by the phrase. The green dot trials were introduced in an attempt to ensure full preparation of the Audio/Arrow action plan by creating a responding time limit. Whereas in the switch condition, a 3 (PREP: action type prepared) x 3 (ACT: action type executed) x 2 (Congruency of object: location of the action for the preparatory and execution phases) factorial design was adopted. See Figure 4 for an example of possible trial conditions given a prepared action plan.

The experiment consisted of two training blocks, one practice block, and seven critical blocks. There were rest breaks between blocks. During the two training blocks, participants first learned to make hand gestures indicated by Audio/Arrow cues, and then practiced responding to Phrase cues. Each of the six action-object combinations was repeated twice in each training session (24 trials). In the critical trials, 25% of the trials were green dot trials, and the remaining 75% were switch trials. Following a practice block, which contained 16 trials randomly selected out of the 48 possible critical trials, the six possible action plans were repeated 14 times in the green dot condition (84 trials), and in the switch condition (252 trials), an equal number of trials was presented for each level of the design randomly across blocks.

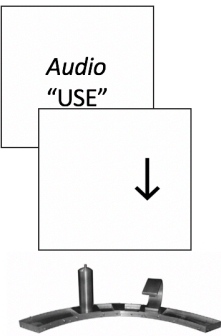
Action type and location		
Prepared action plan	Cued action plan	Conditions
	<i>Green dot trials</i>	<i>Green dot trials</i>
	●	Same
	<i>Switch trials</i>	<i>Switch trials</i>
	Use cell phone	Same
	Use spray can	Object change only
	Touch cell phone	Action change only
	Lift spray can	Both object and action change

Figure 4. Example of possible trial conditions for the prepared action plan *use the cellphone* in Experiment 1.

Procedure

The participants were tested in individual sessions with a research assistant in a quiet lab room. They would sit comfortably in a chair approximately 50 cm away from the monitor screen, with their dominant hand resting on the furthest key of the button box aligned with the

dominant hand. The two response elements were mounted on the response apparatus placed halfway between the participant and the monitor, without obscuring the monitor screen.

Both written and verbal instructions were given for each testing block of the experiment. Participants first, in the training phase, learned to make hand gestures associated with the Audio/Arrow stimuli. Participants put on the headset, and when they saw a fixation cross in the center of the screen, used their dominant hand to press down the far left or far right button which is on the same side as their dominant hand to hear the audio verb and see the arrow pointing at one of the response elements. Once they were ready to act, the participants released the button and mimicked the indicated hand gesture with their dominant hand by grasping the appropriate response element. Similarly, they were trained to make the actions indicated by the Phrase cues.

Once the training blocks were completed, the experiment proceeded to the critical condition (see Figure 5 for the sequence of events in Experiment 1). The subjects kept holding down the same button as before with their dominant hand when seeing the fixation cross. After a blank delay of 250 ms, the Audio/Arrow cue would appear. The participants were told to hold down the button as long as they need to fully prepare this action, then lift off and press down the button again when ready. Following another 250 ms blank delay after the button was pressed, either a green dot or a printed word phrase would show up, and would remain on the screen until the button was released for action execution. For the purpose of ensuring a state of readiness for the prepared Audio/Arrow action plan, there was a deadline of 800 ms applied to the lift-off time in the green dot trials. A warning message, “Too slow”, would appear on the screen if the participant failed to respond within the deadline. The program would document the response once contact was made on the response element. The response time was recorded starting from the onset of the green dot or the printed phrase cue, till the contact with the response element

being registered by the computer. Prior to commencing the next trial, the experimenter manually scored responses as either *correct*, *incorrect*, or *spoiled*. Trials involving technical or procedural issues would be spoiled. For example, if the participant sneezed while the hand was in flight, it would be deemed invalid and thus would be spoiled. A verbal debriefing as well as a performance summary was provided immediately after the completion of the experiment.

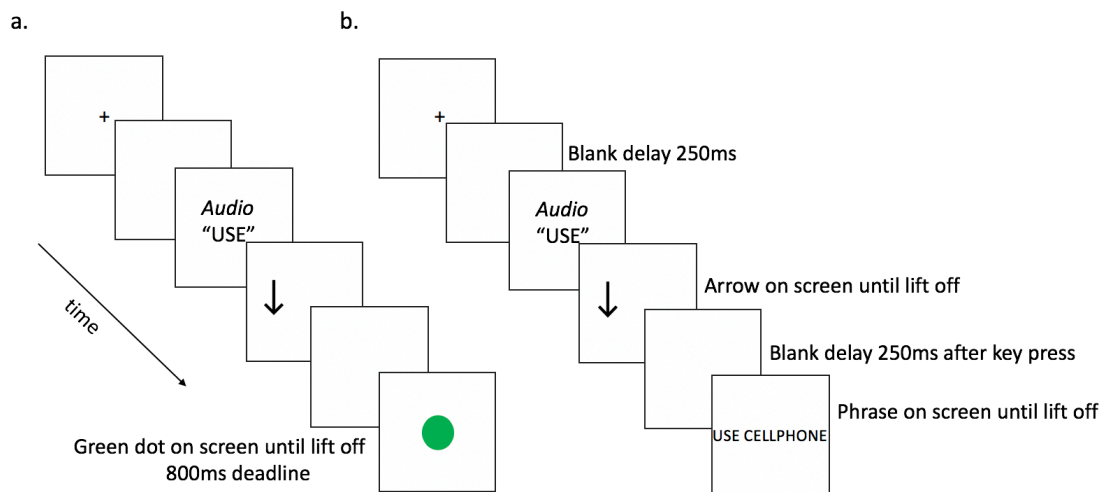


Figure 5. The sequence of events in Experiment 1. a: green dot trials with 800 ms lift-off time limit, occurring 25% of the time; b: switch trials, occurring 75% of the time.

1.2 Results

It is crucial for the experiment that the participants are fully prepared for the initially cued action so that they are able to act immediately. The green dot trials with a deadline, therefore, served as an important indication and feedback on preparation performance. Two subjects were excluded from the data due to low probability of meeting the green dot time limit (< 50% of the time). Data of response time that was faster than 200 ms or slower than 2200 ms (.364% of trials) were removed prior to analysis allowing maximum .5% correct responses to be

excluded from the data (Ulrich & Miller, 1994). Averaged error rates were very low for both the green dot condition (1.06%) as well as the switch condition (1.15%) suggesting mastery of the task; thus, no meaningful analysis can be done on these error data and only the response time data were reported and examined for this and following experiments with a similar design.

The term congruency was used, when constructing the data, in reference to the component of the executed action plan remaining the same as the prepared action plan. Two major components of an action plan that are of interest in these series of experiments are the action class and the associated object. The effect of object congruency would compare responses made on the congruent object with those on the incongruent object; namely, examine if there is a benefit of preserving the object information when acting on the same object as prepared. Similarly, the effect of action congruency refers to the effect of keeping the action class the same while undergoing a switch of action plans. For example, by asking a participant to prepare a use action on the cellphone and switch to lifting the cellphone, we are examining the congruency effect of object; whereas switching from *use cellphone* to *use spray can* would reveal the effect of action congruency. Furthermore, there are two ways of examining the congruency effect comparing different trial conditions (shown in Figure 4): (1) both levels congruent versus one level congruent; and (2) one level congruent versus both levels incongruent. More specifically, comparisons between the *same* condition and the *object change only* condition, as well as between the *action change only* condition and the *both change* condition provide a basis for examining the congruency effect of the object; comparisons between the *same* condition and the *action change only* condition, together with those between the *object change only* condition and the *both change* condition allow us to evaluate the effect of action type congruency.

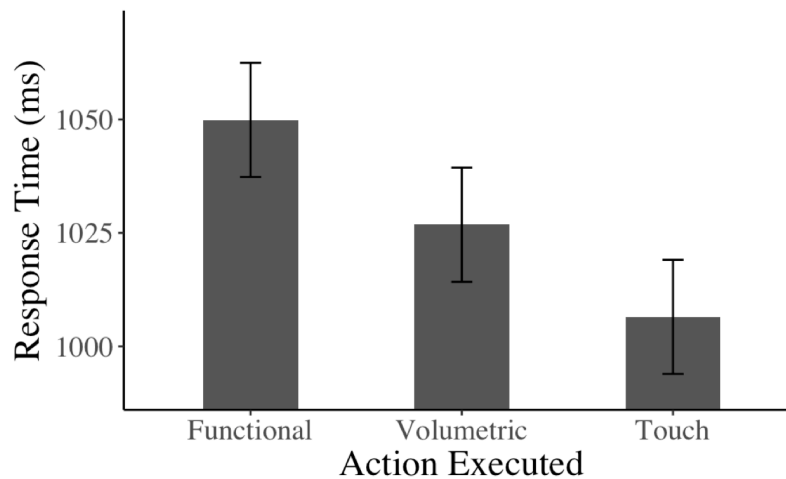


Figure 6. Green dot response time of each action type, averaged across locations in Experiment 1. Error bars represent 95% within-subject confidence intervals.

The mean response time of each action class for the green dot condition was presented in Figure 6. The data suggest that functional actions (1049.9 ms) were slower than volumetric actions (1026.8 ms), which were slower than touch actions (1006.5 ms).

With the main interest of understanding the state of the motor system during the switch, a 3 (PREP: functional, volumetric, touch) x 3 (ACT: functional, volumetric, touch) x 2 (Congruency of object: congruent, incongruent) repeated-measure analysis of variance (ANOVA) was applied to the data in order to have a closer examination of each action class and the effect of object congruency. A significant main effect of the action class performed was found, $F(2, 70) = 8.748$, $MSE = 5989$, $p < .001$, suggesting that the three action classes were significantly different from each other in the switch trials. Additionally, a three-way interaction was found between the three factors, $F(4, 140) = 36.46$, $MSE = 1607$, $p < .001$.

More specifically, further statistical tests analyzing the effect of object congruency for each action class revealed an interesting pattern as shown in Figure 8. When the action class prepared and performed were the same, there was a significant congruency effect of object for all three action classes. When performing a functional action with a prepared functional action, performance on the congruent object was faster than on the incongruent object, $F(1, 35) = 33.54$, $MSE = 2330$, $p < .001$. For volumetric actions prepared, volumetric actions performed on the congruent object lead to faster responses than performances made on an incongruent object, $F(1, 35) = 34.85$, $MSE = 2136$, $p < .001$. The same pattern was found for touch actions, $F(1, 35) = 9.495$, $MSE = 2896$, $p < .005$. Surprisingly, once the action class performed was different from the prepared action class, a different pattern was observed. When the preparation involved a functional action, the congruency effect of object disappeared for performing a volumetric action, $F(1, 35) = 1.477$, $MSE = 2331$, $p > .05$; and performing a touch action on the prepared object was slower than touching the non-prepared object, suggesting a reversed congruency effect of object, $F(1, 35) = 7.54$, $MSE = 895$, $p < .01$. For volumetric actions prepared, a significant reversed congruency effect of object was revealed for both functional and touch actions performed: $F(1, 35) = 8.529$, $MSE = 1363$, $p < .01$, $F(1, 35) = 12.45$, $MSE = 2524$, $p < .005$, respectively. As to touch actions prepared, it yielded the same pattern as functional actions prepared such that there was no benefit for volumetric actions performed, $F(1, 35) = 2.766$, $MSE = 1850$, $p > .1$; and a reversed congruency effect of object for functional actions performed, $F(1, 35) = 6.453$, $MSE = 1870$, $p < .05$.

If we now turn to what was shown in Figure 8, the data were aggregated across specific action classes. The reaction time for each switch condition was shown. When the participants were instructed to switch to executing the exact same action plan, it was significantly faster than

to switch to performing the same action on a different object, demonstrating the effect of preparation, providing evidence of the preparatory stage, $F(1, 35) = 36.68$, $MSE = 4687$, $p < .01$. Moreover, the inhibitory effect of object congruency was again evident once the action class was different, $F(1, 35) = 17.54$, $MSE = 1468$, $p < .01$.

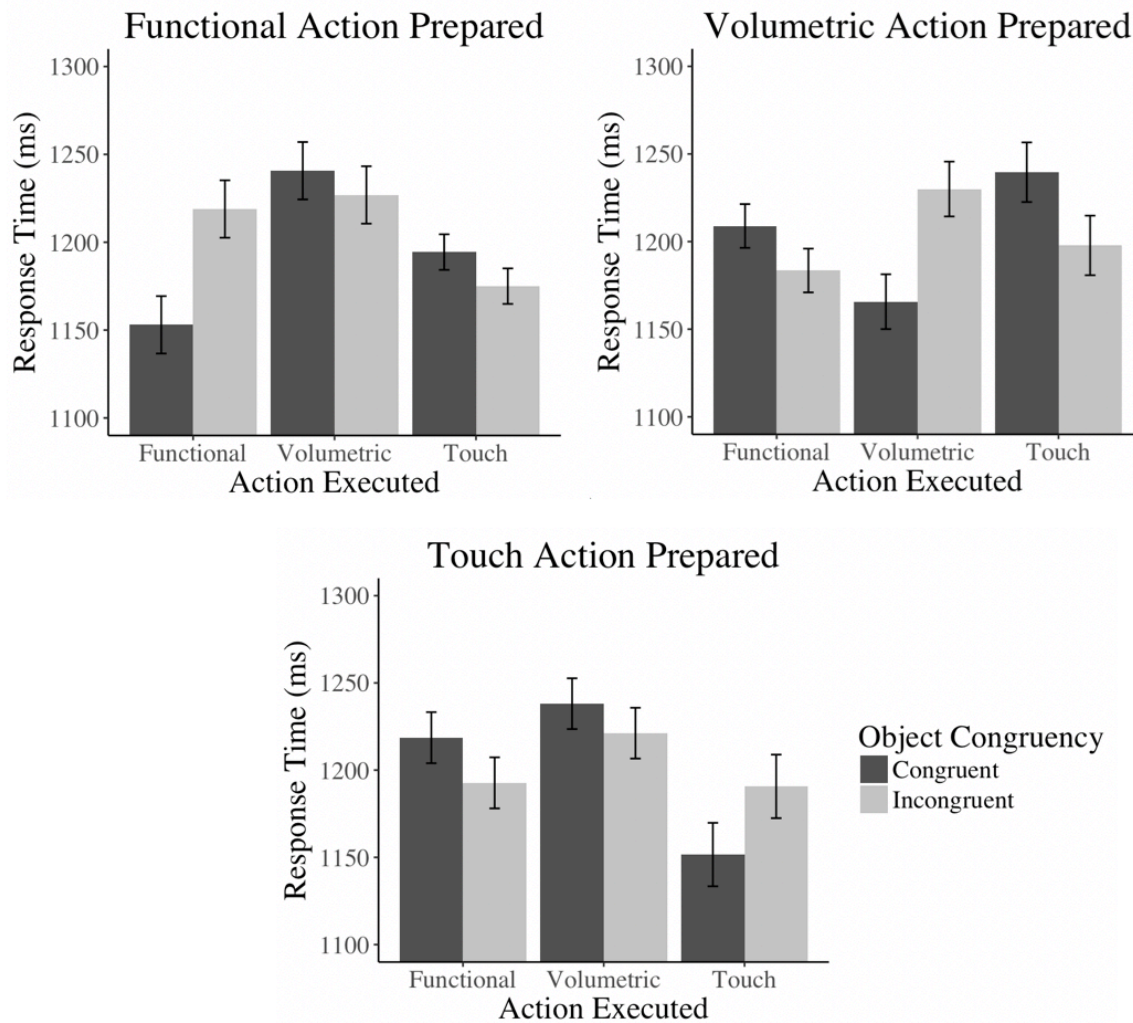


Figure 7. Experiment 1: Response time as a function of action type prepared, action type performed, and congruency of object. Error bars represent 95% within-subject confidence intervals suitable for evaluating congruency effects (Loftus & Masson, 1994; Masson &

Loftus, 2003). Upper left: Participants prepared a functional action. Upper right: Participants prepared a volumetric action. Bottom: Participants prepared a touch action.

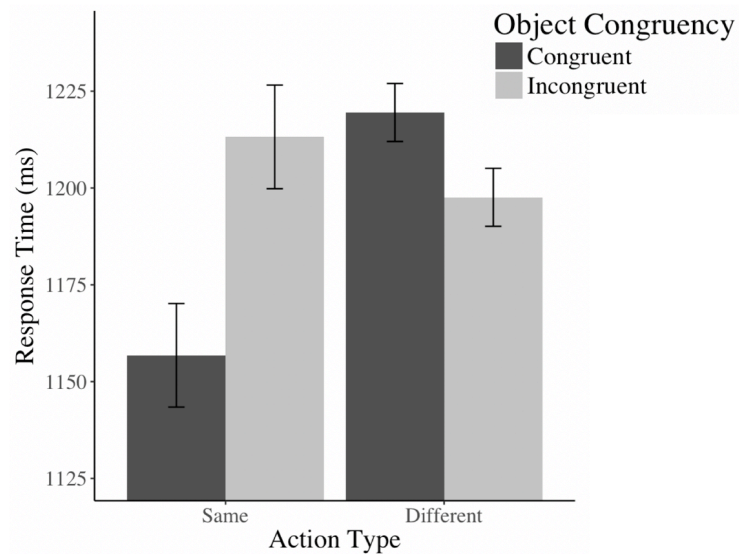


Figure 8. Mean response time of the four switch conditions in Experiment 1. Error bars indicate 95% within-subject confidence intervals.

The second set of analysis aimed at examining the effect of action congruency on performance. As discussed at the beginning of the Results section, we conducted a repeated measure ANOVA allowing the following comparisons between a) the repeat action and the action change only condition; and b) the object change only and the both change conditions (see Figure 9). The former set of comparison examines the congruency effect of action when acting on the congruent object, whereas the latter examines the effect when switching to the non-prepared object (i.e., incongruent object). When the objects were congruent, congruent action conditions were faster than incongruent action conditions for all three action classes: functional actions [$F(1, 35) = 42.38$, $MSE = 1554$, $p < .001$]; volumetric actions [$F(1, 35) = 70.58$, $MSE = 1380$, $p < .001$]; and touch actions [$F(1, 35) = 30.68$, $MSE = 1712$, $p < .001$]. In the incongruent object condition, responses were significantly slower when the action type prepared and

performed was functional actions, $F(1, 35) = 16.14$, $MSE = 1049$, $p < .001$, meaning that there was a cost of preserving the functional actions. However, this pattern was not found when the action preserved being volumetric ($F < 1$) or touch actions, $F(1, 35) = 1.671$, $MSE = 1083$, $p = .205$.

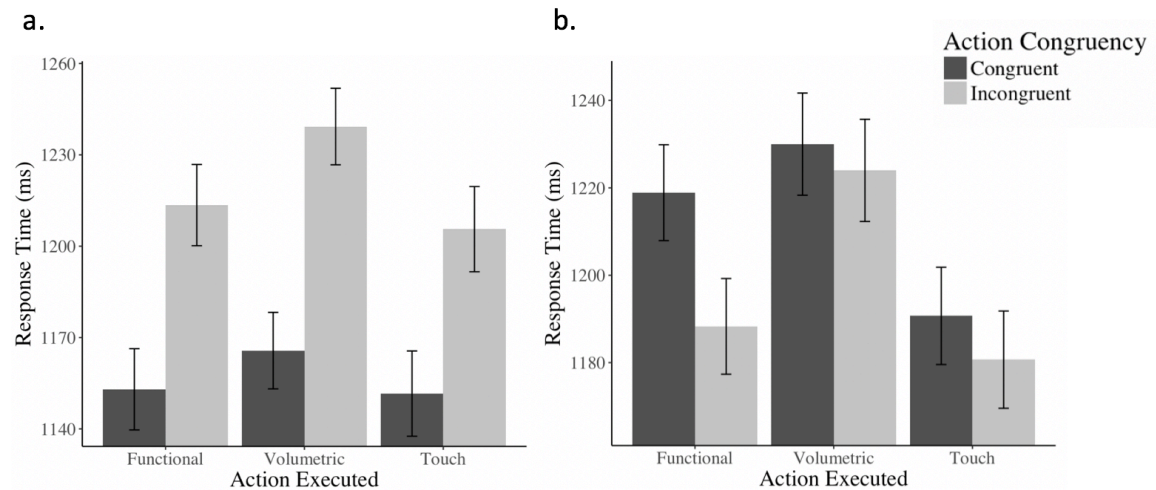


Figure 9. Experiment 1: Response time as a function of type of action performed, and congruency of action type between the preparatory and execution stage when acting on the incongruent object. Error bars represent 95% within-subject confidence intervals. a. Congruent object; b: Incongruent object.

1.3 Discussion

Experiment 1 was designed to examine the activation state of competing affordances when the selected action plan was ready for execution. Participants were asked to prepare an action plan indicated by an Audio/Arrow cue, and carry out an action depending on the subsequent cue presented. Upon receiving a green dot, the participants carried out this fully prepared action plan. Alternatively, when receiving a Phrase cue, they executed the action plan denoted by the new cue. Green dot condition performance indicates that functional action

responses were significantly slower than volumetric responses, with touch actions being the fastest among the three action classes. This result was consistent with the literature that lift responses are faster than the production of use actions under specific experimental designs (e.g., Osiurak & Badets, 2016). This result can be explained by the differences in inherent action complexities. Functional actions often involve more precise grips than do the other two classes. Also, the programming of a touch action relied less on visuomotor information than the grasp actions, as evident not only in the simplicity of the specific hand posture denoting touch in the current study, but also the nature of the goal motivating the touch gestures.

Another significant finding suggested by the data is that when the action type is changed, the congruency of object negatively impacts response time performance. Surprisingly, there was no asymmetrical relationship between the three action classes. We infer that this generic reversed congruency effect has to do with the inhibition of competing motor representations. There are several studies that consider the role of competition between affordances in the selection of an action (Cisek, 2007; Cisek & Kalaska, 2010). It has been proposed that potential affordances are processed in parallel upon receiving sensory information, and compete with each other until an action incurs an activation level above an execution threshold and is performed. Furthermore, there is neurophysiological evidence suggesting that potential affordances are processed and activated simultaneously. Specifically, actions sharing a similar degree of movement parameters mutually excite each other and dissimilar ones compete against each other through inhibition. Based on this theory, the inhibition effect observed can be explained by a modified version of the affordance competition theory as follows.

With the two potential target objects present in view at all times, all six possible actions (three types of actions on each object) are available for selection. When asked to prepare a

specific action, the affordances on the target object would compete with each other such that the non-target actions on the target object would be inhibited, whereas the non-target object and its affordances become irrelevant. Considering the level of neural activation, the prepared target action is the highest, with the alternative actions on the target object being the lowest since they are inhibited, and the actions on the non-targeted object rest in the middle. Therefore, when instructed to switch to an alternative action, it is slower to act on the previously prepared object than to act on the other object. This competition between affordances accounts for the inhibitory effect of object congruency when the action class was changed observed in the current experiment.

Further evidence demonstrated that for functional actions, preserving the *action class* during the switch yielded slower performance when compared to conditions where the action class was changed. Contrary to the effect of object congruency, this pattern was specific to *functional* actions, and there was no effect of action congruency on performance for volumetric or touch actions. We theorize that the specific effect obtained for functional actions occurred because these actions invariably demand access to conceptual knowledge of object identity.

The logic behind this conjecture is as follows. Abundant evidence indicates that preparing a use action on an object always requires access to stored knowledge of its function, whereas a lift action can be based either directly on the object's perceived shape or on a stored representation of the object's structural properties (see Bub et al., 2018, for a review). Preparing a lift or touch action on a target object is most easily accomplished by directly attending to its globe shape. For example, a vertical power grasp is afforded by the shape of the spray can, while a horizontal power grasp is afforded by the cellphone. Preparing a use action requires a more

conceptual level of representation that includes a linguistic description of the intended action in working memory (e.g. use the cell phone). On switch trials, competition between use actions on different objects occurs because they are categorized as serving the *same* type of intention. Lift and touch actions, in contrast, are planned by attending to the shape of the target object. The targeted action invokes no associated conceptual/linguistic representation in working memory, and no competition occurs when switching to the alternate object entailing the same class of action.

Experiment 2

The second experiment was conducted to examine the validity of effects and patterns found in Experiment 1. The task demand was once again to place subjects in a state where they were ready to execute an action immediately, and then when the prepared action plan was ready for execution, instruct them to switch to an alternative action plan on a percentage of trials. Two different modalities of cueing stimuli were used to distinguish between the preparatory and the switch phases. As in Experiment 1, the prepared action was cued by an auditory verb and a pointing arrow, and the action plan switched to was indicated by an imperative phrase consisting of a verb and a noun printed on the screen. In both cases, the action class is suggested by language. However, the differences lie in the object representation. The phrase cue denotes the object via language (i.e., object name), whereas the Audio/Arrow cue signals the location using a strong spatial cue (i.e., arrow). An object name always requires conceptual processing, but an arrow can be used to produce an action on an object based directly on its shape. Thus, in Experiment 2 we reversed the order of cueing methods: the initially prepared action was cued by the Phrase and the switch conditions were cued by the Audio/Arrow cue. The same pattern should be observed if the two cueing stimuli evoke similar motor representations.

2.1 Method

Participants

Forty-four subjects (females = 30, age range = 17-51, mean age = 20.77, right-handedness = 41) were recruited from the same pool as in Experiment 1. None of them have participated in Experiment 1. The experiment was approved by the University of Victoria Human Research Ethics Committee.

Material

Same materials were used as in Experiment 1.

Design

The designs and measures were similar to the ones used in Experiment 1, but with the following exceptions. First, the order of the cueing methods was reversed. Second, the reaction time was recorded starting from the onset of the arrow picture on the screen until the grasp being registered by the system, omitting the presentation time of the audio verb. Thus, it would not be surprising if the averaged response times for each condition were smaller than those observed in Experiment 1. Even though the processing time for the audio cue was not fully recorded, the response time course difference was considered an additive effect that would not affect the general pattern observed. In addition, the green dot responses were prompted using an 800 ms deadline applied to the lift-off response time in Experiment 1. It was reported that some participants had the tendency to lift off the button very rapidly to meet the deadline even though they were not ready for executing the action, and instead programmed the action while their hand was already in motion. Therefore, to avoid such situations confounding the results, the 800 ms deadline of the green dot was modified to be applied to the total response time instead of the lift-off time in Experiment 2. A training block for green dot was introduced (12 trials) familiarizing the participants with the procedure.

Procedure

A similar procedure was adopted in the current experiment. The study started with a training phase for the Phrase cue, and then a green dot training block. In the green dot training block, the participants prepared for an action indicated by the Phrase cue and when they were ready, pressed and held down the button, then carried out the action as soon as the green dot appeared. After which, the participants completed the final training phase for the Audio/Arrow

cue, and a practice block before starting the critical trials. A performance summary was shown at the end of the experiment upon completion.

2.2 Results

Data from eight subjects were excluded from the analysis: three subjects withdrew from the experiment; three subjects had averaged green dot response time over 900 ms given an 800 ms deadline; two other subjects were rejected due to over 15% green dot errors. Response time beyond the range of 200 ms and 1500 ms were excluded (.441% of critical trials).

The mean response time in the green dot condition was presented in Figure 10. Once more, it was shown that functional actions were the slowest (744.1 ms), and touch actions were the fastest (726.2 ms).

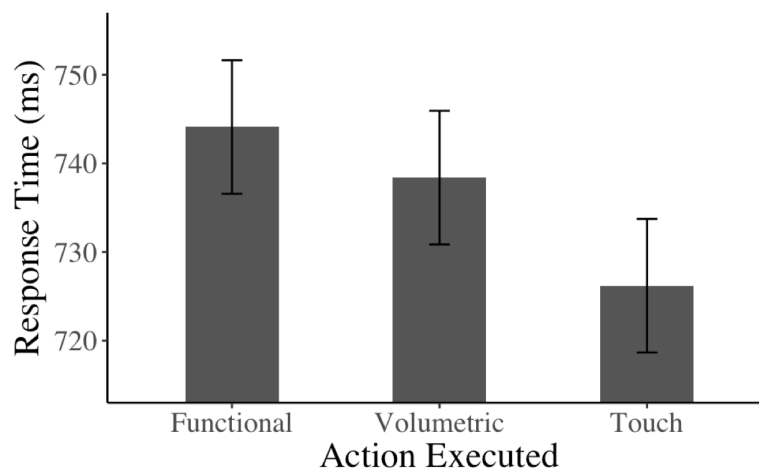


Figure 10. Green dot response time of each action type, averaged across locations in Experiment 2. Error bars represent 95% within-subject confidence intervals.

A repeated measure ANOVA of the response time was computed. The analysis revealed significant main effects of all three factors: PREP, $F(2, 70) = 66.83$, $MSE = 4037$, $p < .001$; ACT, $F(2, 70) = 5.513$, $MSE = 808$, $p < .01$; and Congruency of object, $F(1, 35) = 4.799$, $MSE = 5863$, $p < .05$. In addition, there was a significant three-way interaction between the factors, F

(4, 140) = 28.85, $MSE = 5989$, $p < .001$. No other significant effects were found. More specifically, when segregating the data based on individual action class performed, a slightly different pattern was found when compared to Experiment 1 (see Figure 11). During the switch trials, if the action class was kept the same, there was a significant congruency effect of object for all three action types consistent with Experiment 1: functional actions, $F(1, 35) = 36.47$, $MSE = 1906$, $p < .001$; volumetric actions, $F(1, 35) = 28.21$, $MSE = 2227$, $p < .001$; and touch actions, $F(1, 35) = 16.52$, $MSE = 1739$, $p < .001$. When the action class was changed during the switch, unlike Experiment 1, no reversed congruency effect of object was found across action classes, except for the condition where a prepared functional action was switched to a touch action, $F(1, 35) = 9.062$, $MSE = 808$, $p < .005$.

The condition means of the switch conditions were shown in Figure 12, aggregating the three action classes. The effect of object congruency obtained demonstrated that when the action class was the same, performance on a congruent object would be faster than that on an incongruent object, $F(1, 35) = 39.44$, $MSE = 3951$, $p < .001$; and executing a different action class on the same object prepared showed no effect of object congruency, $F(1, 35) = 1.747$, $MSE = 1830$, $p = .195$.

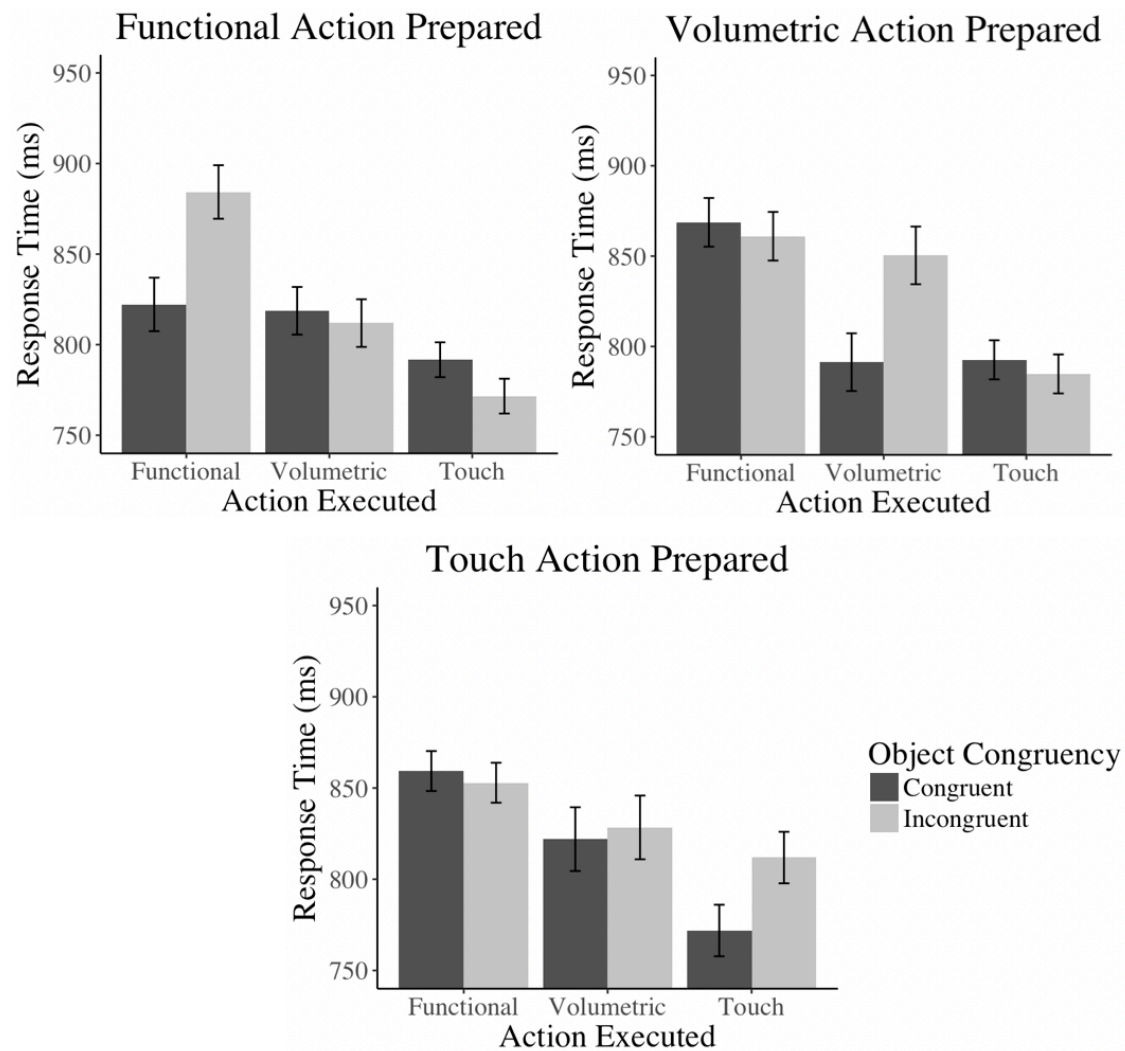


Figure 11. Experiment 2: Response time as a function of action type prepared, action type performed, and congruency of object. Error bars represent 95% within-subject confidence intervals. Upper left: Participants prepared a functional action. Upper right: Participants prepared a volumetric action. Bottom: Participants prepared a touch action.

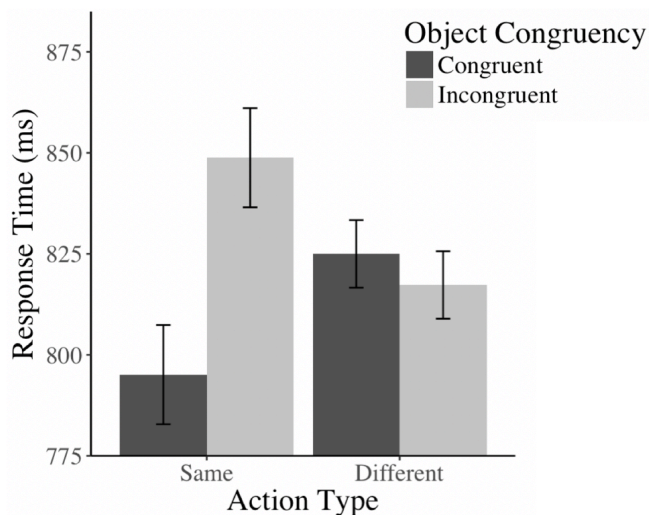


Figure 12. Mean response time of the four switch conditions in Experiment 2. Error bars indicate 95% within-subject confidence intervals.

To investigate the effect of action congruency on response latency, two ANOVA were done first for the congruent objects, and then for the incongruent object condition (see 1.2 Results section for a detailed rationale). In the congruent object condition (Figure 13 a.), again, significant congruency effect of action was observed for all three action classes: functional actions [$F(1, 35) = 39.1, MSE = 767, p < .001$]; volumetric actions [$F(1, 35) = 19.08, MSE = 756, p < .001$]; and touch actions [$F(1, 35) = 12.45, MSE = 601, p < .001$]. When examining the incongruent object condition (Figure 13 b.), there was a significant main effect of action congruency on response time that response times were longer for congruent trials collapsing across action classes, $F(1, 35) = 58.5, MSE = 920, p < .001$. Interestingly, the inhibitory effect of action congruency was found for all three action types: functional actions, $F(1, 35) = 26.53, MSE = 546, p < .001$; volumetric actions, $F(1, 35) = 21.2, MSE = 914, p < .001$; and touch actions, $F(1, 35) = 29.59, MSE = 695, p < .001$. This suggests that there is a generic cost of preserving the action type. No other significant effect was found.

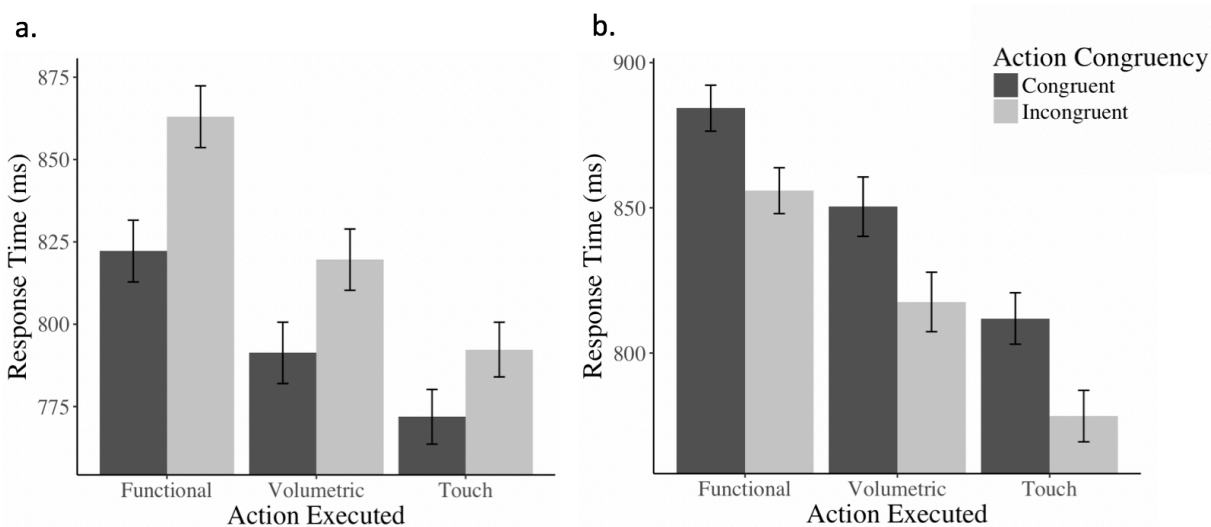


Figure 13. Experiment 2: Response time as a function of type of action performed, congruency of action type. Error bars represent 95% within-subject confidence intervals. a. Congruent object; b: Incongruent object.

2.3 Discussion

Experiment 2 was a replication of Experiment 1, investigating any potential impact the cueing methods might pose on the patterns observed in the previous experiment. The order of cueing methods used was reversed: participants were switching from a Phrase cued preparatory state to an action plan designated by an Audio/Arrow cue. A similar pattern of response latencies for the three action classes was observed in the green dot condition where subjects executed the particular action plan prepared. The effect of object congruency presented in Experiment 2 was in accordance with what was found in Experiment 1, but surprisingly no consistent inhibitory effect was observed of the competing affordances on the congruent object. In Experiment 1, we inferred that the inhibitory effect observed was a result of affordance competition. However, little is known about the nature of competing affordances, such as the processing stages where the competition occurs. Competition could occur either at the motor level where specific actions

(i.e., hand gestures) are selected, or it could operate at the level where specific features of an object are selected. A useful example of the selection of object features is: when intending to lift a cellphone, we pay more attention to the edges of the object; and when using a cellphone, it the surface of the object that is of interest. The absence of inhibitory effect in the current experiment provides a more nuanced understanding of where the affordances compete. Instead of inhibiting specific actions, it seems that the inhibition is evident in a higher level of processing where an action is planned depending on relevant object properties.

Cueing the switch conditions using an arrow directed attention to the shape and location of the alternate target object. Actions are rapidly produced, and the competing influence of the planned action on the same object is no longer apparent. Strikingly, despite the absence of any such competition, Experiment 2 suggests that a generic cost of preserving action type was evident for all three action types when switching from the prepared action to execute the same type of action on an incongruent object. This finding differs from Experiment 1, where only functional actions demonstrated the slowing of responses when preserving action class during the switch. We proposed in the first experiment that the functional action demonstrated a slower performance due to its requirement of conceptual knowledge. It is reasonable to speculate that in Experiment 2, the manipulation of cueing methods encouraged accessing conceptual knowledge so that all three action classes presented a similar pattern.

The reason for this claim is as follows: all action types were prepared via the encoding of a linguistic cue (i.e., use/touch/lift cellphone/spray can). Competition occurred whenever the same action type was repeated on switching to an alternate object. In Experiment 1 the prepared action on a target object was cued by means of a verb denoting the action class and an arrow indicating the location of the object. We conjectured that a conceptual level of representation

was held in working memory only for a use action, while lift and touch actions were based on the shape and location of the target object. In Experiment 2, all actions were prepared via language codes (verb/noun combinations). Switching between objects while maintaining the same type of action incurred a cost because all prepared actions involved conceptual levels of representation in working memory.

Experiment 3

In Experiment 3, we wish to further advance our understanding of the cognitive mechanisms involved in switching between a planned and an alternate action. Are the effects observed in the last two experiments evoked specifically when different grasp actions on objects are required, or do they also occur when the target objects are classified without the intention to use or lift them? We, therefore, introduced a perceptual identification task into the experimental design adopted in Experiment 1 and 2. In the current experiment, participants fully prepared either a functional or volumetric action, and when instructed would switch to an alternative grasp action or naming the object indicated by the switch cue. We were particularly interested in assessing the effect of object congruency on the naming task.

Note that even though naming does not require any action to be performed on the object, in the current study naming was referred to as an action class, similarly to the functional and volumetric classes of action. Moreover, since we have found no distinct results for touch actions when compared to functional and volumetric actions, touch was not included in this experiment.

3.1 Method

Participant

Forty-two new participants (females = 33, age range = 17- 63, mean age = 21.66, right-handedness = 38) were recruited from the same research pool. None of them have participated in Experiment 1 or Experiment 2.

Material

The same material was used as in Experiment 1 and 2. A new headset, Mpow H2 Bluetooth headphones with Mic, was used in Experiment 3 enabling the recording of verbal responses.

Design

A similar design was adopted in this experiment as in Experiment 2 with a few modifications. Only two grasp actions (i.e., use and lift) were involved in the preparatory stage in this experiment, and we introduced naming to the switch conditions. Furthermore, due to technical constraints, only lift-off response time was recorded for grasp actions. Therefore, the green dot deadline was adjusted to 500 ms applied to the lift-off response time.

The participants were instructed to fully prepare an action (either use or lift) indicated by a Phrase cue (e.g., use cellphone). On 33% of the trials, the participant executed this prepared action upon seeing a green dot. On the remaining 67% of the trials, participants switched to an alternative action plan indicated by an audio verb (i.e., use, lift, and name) and a printed arrow (Audio/Arrow cue). An equal number of trials were assigned to each switch conditions. The response time of the grasp actions (in both green dot and switch trials) was measured from the onset of the arrow until the button lift-off. Vocal responses were classified as a valid response input when at the first detection of sound, the microphone detected more than 25 ms of activity above a sensitivity threshold set during a Mic test conducted at the beginning of the experiment. Thus, the response time of the naming trials was measured from the onset of the arrow cue until the computer detected the first sound that passed the microphone limit to be classified as a response.

A 2 (PREP: functional, volumetric) x 3 (ACT: functional, volumetric, name) x 2 (Congruency of object: congruent, incongruent) factorial design was adopted. Three training blocks each consisting of 12 trials were conducted where participants practiced responding to the Phrase cues, Phrase cues following green dots, as well as the Audio/Arrow cues. The use and lift actions for each object were practiced three times in each training block. Twelve trials were randomly generated from the total of 36 possible critical trials (24 switch trials and 12 green dot

trials) in the practice trial. In the critical condition, 324 trials were shown and 108 of those were green dot trials.

Procedure

Before commencing the experiment, a brief microphone test was done for each subject to ensure that the threshold set for vocal inputs was sufficient for registering naming responses. Participants all went through two training blocks and one practice block for the critical condition. In the critical trials, participants pressed and held down the far right or far left button with their dominant hand when seeing a fixation cross. After a blank delay of 250 ms, a phrase cue appeared on the monitor screen and the participant could take as long as they need to prepare for this action. When they were ready, they lifted off and pressed down the button one more time to indicate that they were fully prepared to act. After a 250 ms blank delay, either a green dot or an Audio/Arrow cue appeared. The participant carried out the prepared action when they saw a green dot. If the participant lifted off the button later than 500 ms past the onset of the green dot, a “Too slow” message was printed on the screen. When the participants received the Audio/Arrow cue instead, and the audio verb was either use or lift, they switched to grasp the response element indicated by this cue. When the participant was asked to name, they kept pressing down the button and at the same time named the object pointed by the arrow. The experimenter would manually score the correctness of the responses. A brief performance summary was generated and shown, and the participants were debriefed.

3.2 Results

Four subjects withdrew from the experiment and were excluded from the analysis, and two subjects were rejected due to green dot response time over 550 ms. Responses faster than

100 ms or slower than 1600 ms were excluded as outliers from the analysis so that only .493% correct responses were excluded.

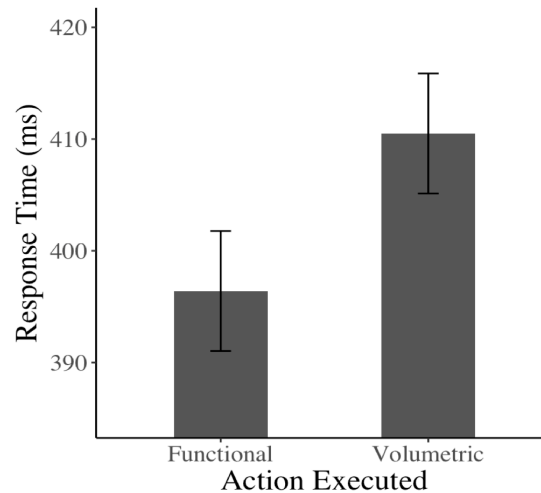


Figure 14. Green dot response time of each action type in the preparatory stage, averaged across locations in Experiment 3. Error bars represent 95% within-subject confidence intervals.

Mean response time of the green dot condition was shown in Figure 14, showing a different pattern from the last two experiments that functional actions were faster than volumetric actions (396.4 ms vs. 410.5 ms).

Data with a 2 (PREP) x 3 (ACT) x 2 (Congruency of object) factorial design was analysed in a repeated measure ANOVA. There was a significant main effect of object congruency with faster responses on the congruent object, $F(1, 35) = 7.902$, $MSE = 7770$, $p < .01$. There was also a main effect of action type performed, $F(2, 70) = 153.3$, $MSE = 13045$, $p < .001$, indicating differences between action classes. A significant three-way interaction between factors was reported, $F(2, 70) = 7.931$, $MSE = 1120$, $p < .001$. Explicitly, when the action class was preserved through the switch, there was a significant congruency effect of object for both

functional and volumetric actions: $F(1, 35) = 8.459$, $MSE = 3320$, $p < .01$; $F(1, 35) = 8.229$, $MSE = 2236$, $p < .01$, respectively. When the prepared action class was changed to the other class, there was no effect of object congruency for both prepared functional ($F < 1$) and volumetric actions ($F < 1$). Importantly, the data showed that when switching from a grasp action to naming, there was a significant benefit of object congruency in both preparation conditions (Figure 15). When switching from a prepared functional action on a particular object to naming an object, the response time was shorter for congruent objects, $F(1, 35) = 10.87$, $MSE = 1849$, $p < .01$; the same pattern held for a prepared volumetric action, $F(1, 35) = 6.334$, $MSE = 2404$, $p < .05$.

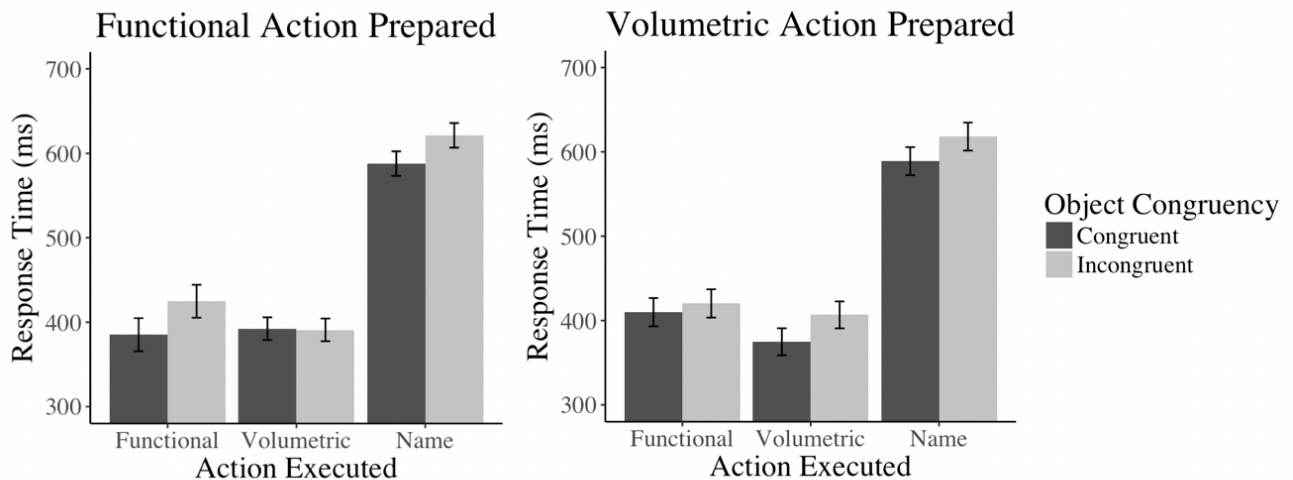


Figure 15. Experiment 3: Response time as a function of action type prepared, action type performed, and congruency of object. Error bars represent 95% within-subject confidence intervals. Left: Participants prepared a functional action. Right: Participants prepared a volumetric action.

The analysis testing the effect of action congruency was done on a modified data file where naming trials were excluded. The within-subject repeated measure ANOVA suggests that when the objects were congruent (Figure 16 a.), both action classes demonstrated benefit of preserving action class, but the effect was of a smaller magnitude than the previous experiments: functional actions [$F(1, 35) = 10.65, MSE = 1031, p < .01$]; and volumetric actions [$F(1, 35) = 5.946, MSE = 927, p < .05$]. For incongruent objects (Figure 16 b.), a significant reversed action congruency effect was found for volumetric actions [$F(1, 35) = 5.944, MSE = 824, p < .05$], but not for functional actions ($F < 1$).

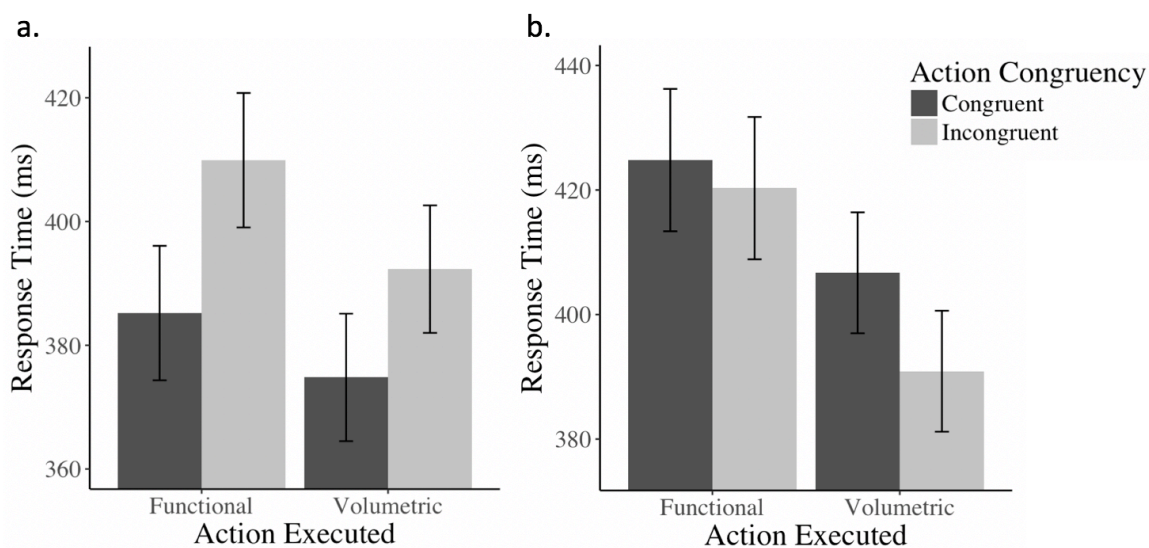


Figure 16. Experiment 3: Response time as a function of type of action performed, congruency of action type. Error bars represent 95% within-subject confidence intervals. a. Congruent object; b. Incongruent object.

3.3 Discussion

This final experiment was a replication of Experiment 2 while introducing a naming task with the aim of assessing the effect of object congruency at both the motor level and at the level

of a task that does not require the planning of a grasp action. Instead of switching from a prepared hand gesture to another hand gesture as in the first two experiments, participants were sometimes instructed to name the particular object they prepared the functional or volumetric actions on.

It is interesting to notice that in Experiment 3, unlike the first two experiments, the functional actions were faster than the volumetric actions in green dot trials. A possible explanation for this might be that by introducing naming as a potential switch condition, the subjects would be more aware of the functional properties of an object (see for example, Bub & Masson, 2012 on the role of functional knowledge in the conceptual representation of an object), and therefore more prepared for functional actions.

In Experiment 1 and 2, it was found that when the action class was changed in the switch condition, no facilitating effect of object congruency was observed. Indeed, switching to an alternate action on the same object either incurred a cost (Experiment 1) or no benefit (Experiment 2). Clearly, these effects occurred at the level of planning grasp actions. Switching from a planned grasp action to naming an object yielded *faster* responses times when the target object remained the same on switch trials. Note that although the action congruency effect replicated the pattern in Experiment 2, no reversed congruency effect occurred in the incongruent object condition for functional actions. With only two different actions required for each object, the discriminability between actions was high, and thus it was hard to find an effect of maintaining action types during an object switch. Also, another possible account could be that there was an uneven pairing between the prepared actions and the executed action with a ratio of two (i.e., functional and volumetric) to three (i.e., functional, volumetric, and naming).

A further point is that the naming task was not incorporated in the preparation set. As described in earlier sections, the prepared actions were indicated by a phrase explicitly indicating the action class and the name of the object. A resulting naming response would, in this case, entail mere repetition of the word, and does not require directing attention to the objects (i.e., spatial location), and is not compatible or consistent with the purpose behind Experiment 3. With the findings of this current replication experiment as supporting evidence, future studies could manipulate the cueing methods to include naming into the preparatory stage to create a balanced relationship between prepared and performed response.

General discussion

The present study was designed to compare the state of motor representations evoked by distinct intentions when an action was prepared for immediate execution. The critical manipulation between Experiment 1 and 2 was the cueing methods used to indicate the prepared action and the final target action on switch trials. In the first experiment, participants prepared an action indicated by an audio verb/arrow combination, and switched to an action suggested by an imperative phrase cue. In the second experiment, the participants switched from an action cued by a phrase to an audio verb/arrow cued target action.

Comparing the effect of object congruency, the first two experiments revealed a similar pattern: executing the prepared action plan was faster than performing the prepared action on a different object; once the action class was changed, there was no facilitatory effect of object congruency. More specifically, in Experiment 1, there was a cost of object congruency such that executing non-prepared actions on the same object was *slower* than abandoning the original plan and programming a completely new one on the alternate object. In Experiment 2, no such result occurred; neither costs nor benefits occurred when a different action was carried out on a prepared object. Results of Experiment 3, however, indicate that there was a benefit in speeded performance when subjects, having prepared an action on a target object, switched to naming it.

In terms of the effect of action congruency, the two experiments demonstrate clearly distinct patterns. Experiment 1 showed a reversed congruency effect of action when maintaining the functional action class, but not when preserving volumetric or touch actions. In contrast, all three types of actions revealed impaired performance when the action class was preserved in Experiment 2. The asymmetrical pattern in Experiment 1 vanished once the order of cueing methods was changed.

Action-object pairing

We now consider in some detail, the nature of differences in the state of motor representations evoked by differential cueing stimuli. After indicating an action class linguistically (e.g., lift), the object is either indicated by a linguistic cue (i.e., object name) or a spatial cue (i.e., arrow). The object name presented linguistically requires semantic processing before localizing the spatial position. In contrast, an arrow indicates a spatial location directly as it points down at the shape, directing attention to the object's visual features. These cueing methods setting up the preparatory stage of an action, lead to different representations of the action plan. We propose that linguistic cues activate a *goal object* representation with associated conceptual knowledge, whereas the audio verb and arrow cues evoke spatial representations in the form of a *goal shape* consisting of a set of physical features. We further infer that when asked to prepare an action, an action-object pairing will be established, and the state of the object representation depends on the cueing stimulus used. For example, an imperative phrase *use cellphone* activates an action-object pairing such that the action goal is to *use*, and the object is *cellphone*. More specifically, the object *cellphone* is a *goal object* with an emphasis on its conceptual properties. This abstract level representation of an action plan is essential in orienting the way in which the participants directing attention to the object.

Action-object-feature pairing

Once an action-object pairing is established, more detailed processing of the action-object relationship is needed for the programming of specific goal postures related to the target object (i.e., action-object-feature pairing). Take the *use-cellphone* action-object pairing as an example. This action representation can be applied to multiple cellphones with distinct shapes or sizes, and thus is in accordance with ultimately an infinite number of gestures. Further integration of target

object's global shape and physical features is crucial to the programming of action representations. We thus propose that two additional processing pathways are recruited simultaneously when programming an action class specific to an object (see Figure 17): a conceptual route and a direct route. The conceptual route is strongly engaged when identifying object identity is needed (through linguistic cues), and selects the associated object features required by the target action. The direct route, on the other hand, is faster than the conceptual route, and is based only on the overall physical features of the object. The two processing routes regulate how attention is paid to the target object and thus determine how an action is programmed. The relative strength of the two routes depends on task demands; namely, the specific cueing method used. For example, an action cued using an object name would activate a strong conceptual route; and actions cued by an arrow would rely more on the direct route.

The evidence that a reversed congruency effect of object was observed in Experiment 1 but not in Experiment 2, suggests that the competition between affordances occurs at a level that deals with each action type and its relationship to a particular object, rather than competition between hand gestures. It seems reasonable to assume that the competition affects the ability to switch between the features of the object attended to, occurring at the level of the conceptual route. For example, when instructed to prepare a use action on the cellphone, attention is paid to the physical features of the cellphone related to the use action, whereas the other features are inhibited in the competition.

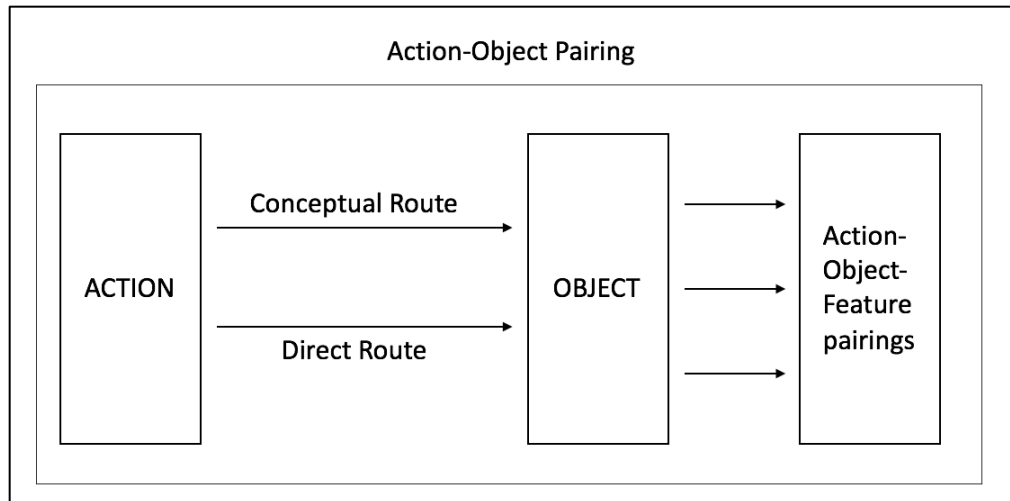


Figure 17. Two processing routes involved in programming action-object-feature pairings.

We have so far proposed two levels of representation of an action plan, namely, the activation of action-object pairings at an abstract level, and the programming of action-object-feature pairings at a lower motor level. It is important to note that the distinction between a *goal object* and a *goal shape* lies in the abstract level of representation (i.e., the action-object pairing level) such that the content of an action plan is represented differently in working memory (e.g., lift cellphone vs. lift this shape in this location, respectively). Moreover, we assume that when the system is allowed enough time to prepare an action to a familiar object, both processing routes in the motor level would be recruited so that the differential qualities of these two object representations would not impact on the motor level of action representation (i.e., the action-object-feature pairing level). For example, when given an audio verb/arrow cue, a proximal goal would be formulated on the abstract level based on a goal shape representation, which requires little or no conceptual knowledge of the object, and the direct route would be quickly recruited in the motor level programming specific hand postures. When given enough processing time, the conceptual route would later also be activated in the motor level requiring access to the

conceptual knowledge of target object identity, so that despite the goal shape representation in the action-object pairing level, the action-object-feature pairing would always entail an object-specific description of the action plan.

Congruency effect of object

Given that two distinct pathways are involved in processing specific action type and relevant object features, and the competition between affordances occurs at the level of object features in relation to the relevant grasp, we can explain the patterns of object congruency observed in Experiment 1 and 2 by examining the nature of the processing route recruited for the prepared action and target action. When asked to prepare an action indicated by an audio verb and an arrow (Experiment 1), subjects were given sufficient time to recruit both conceptual and direct routes to select the target action (see Figure 18). For instance, as shown in Figure 18, when asked to prepare a lift action on the cellphone, the arrow directed attention to the target shape without referencing any conceptual knowledge. Given enough preparation time, features related to a lift action on the cellphone were also selected through the conceptual route, while concurrently the other object features associated with a different grasp were inhibited. When instructed to perform a different action on the target object, the linguistic designation of the object (e.g., lift cellphone) engaged the conceptual route. The competing actions which were inhibited in the preparatory stage would thus be harder to perform, yielding the reversed congruency effect.

On the other hand, when the prepared action was indicated by a linguistic cue (Experiment 2), the conceptual route would be recruited and non-target features of the object would be suppressed (see Figure 19). When instructed to switch to another action on the same object indicated by an arrow cue, the participants were able to quickly respond to the shape

through the direct route, bypassing the inhibition incurred in the conceptual route, and thus no inhibitory effect was observed. This outcome was replicated in Experiment 3.

Naming tasks are not operating at the motor level where grasp actions are programmed, and are not dependent on the specific physical properties of an object. Therefore, the competition in the conceptual route had no impact on naming the prepared object in a switch trial. Indeed, there was a benefit of object congruency when switched to naming (Experiment 3).

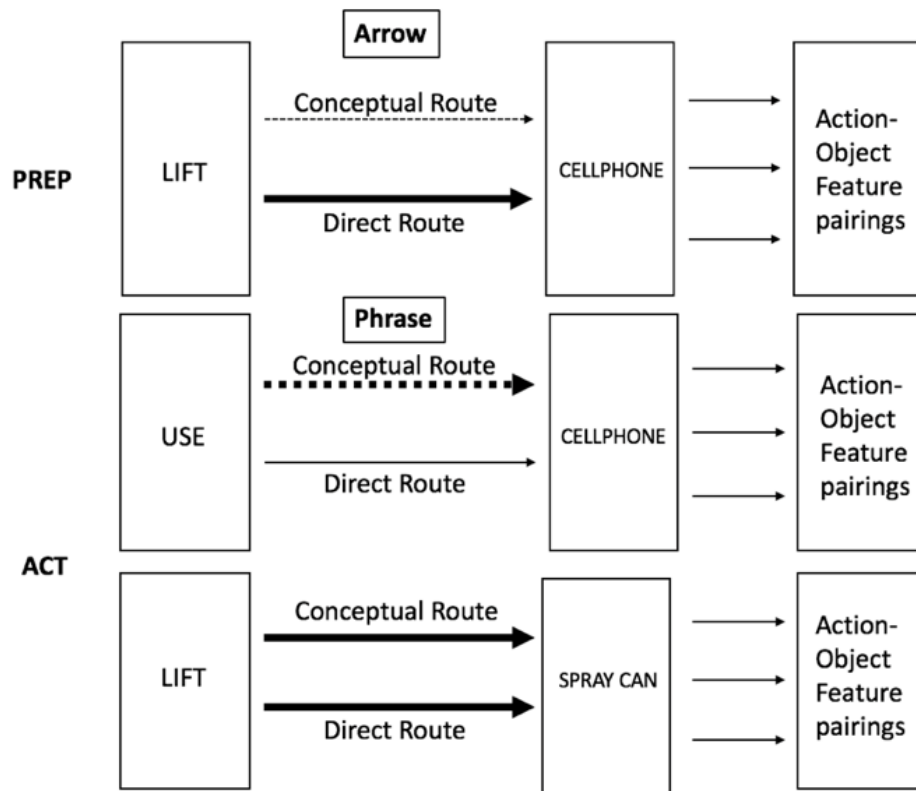


Figure 18. Example status of processing routes in Experiment 1 when prepared to lift the cellphone. The upper panel indicates the prepared action cued by an arrow; the middle panel demonstrates switching to an alternate action on the prepared object cued by a phrase (object name); the lower panel indicates switching to an alternate object cued by a phrase while preserving the action class. Thicker lines represent faster activation; dashed lines refer to competition.

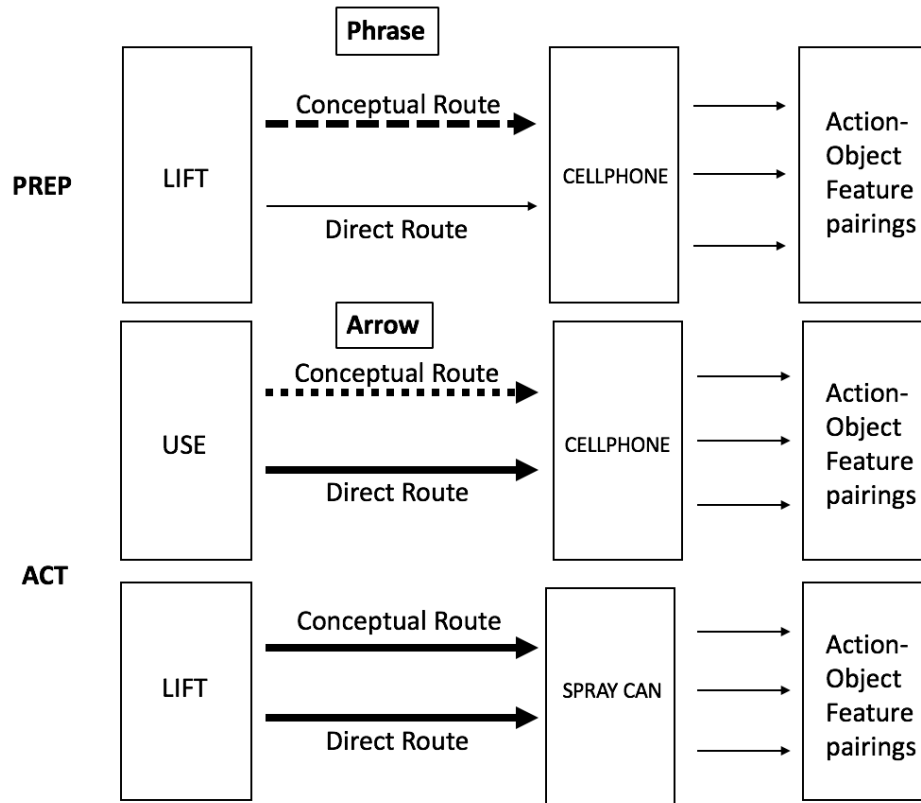


Figure 19. Example status of processing routes in Experiment 2 when prepared to lift the cellphone. The upper panel indicates the prepared action cued by a phrase (object name); the middle panel demonstrates switching to an alternate action on the prepared object cued by an arrow; the lower panel indicates switching to an alternate object cued by an arrow while preserving the action class. Thicker lines represent faster activation; dashed lines refer to competition.

Congruency effect of action

Let us now turn to the reversed congruency effect for functional actions observed in Experiment 1. No such effects occurred for lift and touch actions. Recall that in Experiment 2, the inhibitory effect of action congruency was consistent across all three action classes. Why are functional actions affected in both experiments? Similar to the nature of cueing stimuli, the

differences between motor representations evoked by distinct intentions can be understood using the *goal object*, and *goal shape* status of an object representation. Functional actions are directed by the intended use of an object. Conceptual knowledge of object identity and prior experience of use actions are vital to accomplishing functional grasps. By contrast, volumetric actions are gestures used to manipulate object position or location in space, and touch actions are simple interactions with the object enabling physical contact. Both volumetric and touch actions are dependent solely on the global physical features of an object. Therefore, we infer that the activation of a functional motor representation requires identification of a *goal object*, whereas it is sufficient to have *goal shape* information when eliciting volumetric or touch motor representations.

In terms of the level of action-object pairing, when a goal object presentation is activated, higher level cognitive processes are involved in identifying the object, which would enhance the integration between the action class and object. However, with only the goal shape encoded, an action type depending on the global physical features of the object can be easily dissociated from the identity of the object. Put differently, encoding a goal object (either through linguistic cueing or through preparing a functional action) will involve an integrated object-action pairing, making it harder to switch objects while preserving the action class.

Referring back to Experiment 1, the cueing stimulus used in the preparatory stage encouraged activation of a goal shape representation (cued by an arrow), and the switch cue required execution of an action with a goal object representation (cued linguistically). Experiment 2, in contrast, cued the prepared action linguistically, and defined the object representation at a conceptual level (i.e., goal object). As discussed above, the action-object integration would be stronger when the object is represented as a goal object, and thus would

incur a cost when maintaining the action class and switching the target object, consistent with the pattern found in Experiment 2. No such strong integration would be evident in action-object pairings with a goal shape activation; consequently, the switch should be less costly if the prepared action is cued by an arrow. Given that functional actions naturally prompt identification of goal object, the transitions would be always from *goal object* to *goal object* for functional actions in both experiments, explaining the asymmetrical pattern observed in Experiment 1.

To summarize, we hypothesized that two levels of processing are involved in programming an action representation resulting in the observed pattern of the congruency effect between object and action. The action-object pairing level indicates an abstract goal state, and the action-object-feature processing level recruits two distinct processing routes integrating specific features of an object required by an action class. Competition occurs between different object representations (i.e., *goal object* vs. *goal shape*) at the level of action-object pairing, and also between visual physical features of an object associated with distinct action types at the level of action-object-feature pairing. Our findings provided new evidence on the nature of action representations associated with different motor intentions, as well as of the processes involved in generating motor representations.

References

- Bub, D. N., & Masson, M. E. J. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition, 116*(3), 394–408.
<https://doi.org/10.1016/j.cognition.2010.06.001>
- Bub, D. N., & Masson, M. E. J. (2012). On the dynamics of action representations evoked by names of manipulable objects. *Journal of Experimental Psychology: General, 141*(3), 502–517. <https://doi.org/10.1037/a0026748>
- Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition, 106*(1), 27–58.
<https://doi.org/10.1016/j.cognition.2006.12.010>
- Bub, D. N., Masson, M. E. J., & van Mook, H. (2018). Switching between lift and use grasp actions. *Cognition, 174*, 28–36. <https://doi.org/10.1016/j.cognition.2018.01.013>
- Buxbaum, L. J., Kyle, K. M., Tang, K., & Detre, J. A. (2006). Neural substrates of knowledge of hand postures for object grasping and functional object use: Evidence from fMRI. *Brain Research, 1117*(1), 175–185. <https://doi.org/10.1016/j.brainres.2006.08.010>
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1485), 1585–1599. <https://doi.org/10.1098/rstb.2007.2054>
- Cisek, Paul, & Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience, 33*(1), 269–298.
<https://doi.org/10.1146/annurev.neuro.051508.135409>

- Creem-Regehr, S. H., & Lee, J. N. (2005). Neural representations of graspable objects: are tools special? *Cognitive Brain Research*, *22*(3), 457–469.
<https://doi.org/10.1016/j.cogbrainres.2004.10.006>
- Deschrijver, E., Wiersema, J. R., & Brass, M. (2016). The interaction between felt touch and tactile consequences of observed actions: An action-based somatosensory congruency paradigm. *Social Cognitive and Affective Neuroscience*, *11*(7), 1162–1172.
<https://doi.org/10.1093/scan/nsv081>
- Dijkerman, H. C. (2017). On feeling and reaching: Touch, action, and body space. In A. Postma & I. J. M. van der Ham (Eds.), *Neuropsychology of space: Spatial functions of the human brain*. (pp. 77–122). San Diego, CA: Elsevier Academic Press.
<https://doi.org/10.1016/B978-0-12-801638-1.00003-3>
- 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.
<https://doi.org/10.1016/j.cognition.2010.01.004>
- Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, *8*(2), 71–78. <https://doi.org/10.1016/j.tics.2003.12.002>
- Masson, M. E. J., Bub, D. N., & Newton-Taylor, M. (2008). Language-Based Access to Gestural Components of Conceptual Knowledge. *Quarterly Journal of Experimental Psychology*, *61*(6), 869–882. <https://doi.org/10.1080/17470210701623829>
- Osiurak, F., & Badets, A. (2016). Tool use and affordance: Manipulation-based versus reasoning-based approaches. *Psychological Review*, *123*(5), 534–568.
<https://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*(1), 76–81.
<https://doi.org/10.1016/j.cognition.2013.03.005>
- Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*(1), 179-217. doi:10.1016/j.cognition.2007.09.003
- Pezzulo, G., & Cisek, P. (2016). Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences*, *20*(6), 414–424.
<https://doi.org/10.1016/j.tics.2016.03.013>
- Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General*, *123*(1), 34–80. <https://doi.org/10.1037/0096-3445.123.1.34>