Psychonomic Bulletin & Review

Visual context modulates potentiation of grasp types during semantic object categorization

Journal:	*Psychonomic Bulletin & Review*
Manuscript ID:	PBR-BR-13-183.R1
Manuscript Type:	Brief Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Kalénine, Solène; IRHIS-CNRS UMR 8529, Université Lille Charles de Gaulle Lille 3 Shapiro, Allison; Moss Rehabilitation Research Institute, Flumini, Andrea; Università di Bologna, Psychology Borghi, Anna; Università di Bologna, Buxbaum, Laurel; Moss Rehabilitation Research Institute,
Keywords:	stimulus-response compatibility, concepts and categories, visual selective attention, embodied cognition

Visual context modulates potentiation of grasp types during semantic object categorization

Solène Kalénine[1,2], Allison D. Shapiro[1], Andrea Flumini[3], Anna M. Borghi[3,4], and Laurel J. Buxbaum[1]

[1] Moss Rehabilitation.Research Institute, Elkins Park, PA, USA

[2] Lille Nord de France University, Lille, France

[3] University of Bologna, Bologna, Italy

[4] ISTC-CNR, Rome, Italy

Word count (excl. title, abstract, and figures) : 4692

Corresponding Author:

Solène Kalénine

IRHIS-CNRS UMR 8529

Université Charles-de-Gaulle Lille III

Domaine universitaire du Pont de Bois

BP 149 - 59653 Villeneuve d'Ascq Cedex

France

Email: solene.kalenine@gmail.com

Phone: +33 (0)3 20 41 73 25

Abstract

Substantial evidence suggests that conceptual processing of manipulable objects is associated with potentiation of action. Such data have been viewed as evidence that objects are recognized via access to action features. Many objects, however, are associated with multiple actions. For example, a kitchen timer may be clenched with a power grip to move it, but pinched with a precision grip to use it. The present study tested the hypothesis that action evocation during conceptual object processing is responsive to the visual scene in which objects are presented. Twenty-five healthy adults were asked to categorize object pictures presented in different naturalistic visual contexts that evoke either move- or use-related actions. Categorization judgments (natural vs. artifact) were performed by executing a move- or use-related action (clench vs. pinch) on a response device, and response times were assessed as a function of contextual congruence. Although the actions performed were irrelevant to the categorization judgment, responses were significantly faster when actions were compatible with the visual context. This compatibility effect was largely driven by faster pinch responses when objects were presented in use- compared to move-compatible contexts. The present study is the first to highlight the influence of visual scene on stimulus-response compatibility effects during semantic object processing. These data support the hypothesis that action evocation during conceptual object processing is biased toward context-relevant actions.

Introduction

Evidence from numerous behavioral studies suggests that conceptual processing of manipulable objects is associated with potentiation of action (e.g., Craighero, Bello, Fadiga, & Rizzolatti, 2002; Ellis & Tucker, 2000; Girardi, Lindemann, & Bekkering, 2010; Tucker & Ellis, 1998, 2001). Many of these studies show that conceptual processing of a visually-presented object is facilitated when the motor response required for the task is compatible with the action typically associated with that object, even when that action is task-irrelevant. For example, participants are faster to categorize a small, “pinchable” object (such as a strawberry) as a natural rather than manufactured object when they indicate their categorization choice by performing a precision (pinch) grip compared to a power (clench) grip on an experimental apparatus (Tucker & Ellis, 2001). Such stimulus-response compatibility effects have been taken as evidence that conceptual object representations are composed in part of sensorimotor features associated with object manipulation (e.g., Barsalou, 2008).

Many manipulable objects, however, are associated with several actions. For example, a kitchen timer may be clenched with a power grip to move it, but pinched with a precision grip to use it. Recent studies have shown that object processing may recruit both of these action types (Bub, Masson, & Cree, 2008; Lee, Middleton, Mirman, Kalénine, & Buxbaum, 2013). In one such study, for example, participants were first trained to associate different actions with distinct colors, then viewed objects whose color signaled the action to be performed on an experimental device. Despite the apparent irrelevance of the motor response to the object identification task, responses that were congruent with using or moving the objects (e.g. poking-calculator; clenching-spray bottle) were executed faster than incongruent actions (Bub et al., 2008).

More recently, Jax and Buxbaum (2010, 2013) demonstrated that use- and move-related actions may compete with each other within single objects. In particular, initiation of use actions is slower for objects associated with distinct move-related actions (hereafter, “conflict” objects, e.g. calculator) as compared to objects for which use- and move-related actions are similar (“non-conflict” objects, e.g. drinking glass). This, and associated data indicating that initiation of move-related actions is no slower for conflict- than non-conflict objects, suggests that move-related activations may be relatively rapid, thus interfering with planning of use-related actions. Jax and Buxbaum (2010) proposed that the intention to act on an object triggers a race-like competition between functional and structural responses during action selection. Only functional responses require activation of long-term conceptual representations; thus, structural responses can be activated more quickly than functional responses.

The evidence for two classes of actions associated with a given object raises questions about the factors that may influence the strength and time course of their activation. One possibility is that both types of action are invariably activated during object recognition. Alternatively, and more likely in our view, action activation may be responsive to task goals and context (see Buxbaum & Kalénine, 2010). In support of this latter possibility, a recent eye-tracking study demonstrated that activation of move- and use-related competition between objects in a visual array may be accelerated by congruent verbal context (Lee et al., 2013). For instance, cueing of target identity with action sentences such as “he picked up the calculator” or “he used the calculator” accelerated competition between the target (calculator) and distractor objects that are picked up or used similarly, respectively. These data suggest that verbal context may influence the activation of both of these classes of action.

To our knowledge, the question of whether visual scene context may modify activation of move- and use-related actions has not previously been addressed. In the present study, we tested the hypothesis that evocation of move- or use-related actions is indeed responsive to the congruence of the visual context in which objects are presented. To this aim, we used a stimulus-response compatibility paradigm first developed by Tucker and Ellis (2001) and presented conflict objects in move-compatible or use-compatible visual scenes.

Methods

Participants

Twenty-five healthy adults (10 females, mean age = 62, SD = 6.4, mean education = 15.5 years, SD = 2.7 years) took part in the study. All participants were recruited from the Moss Rehabilitation Research Institute Research Registry (Schwartz, Brecher, Whyte, & Klein, 2005), Philadelphia, USA. They had no history of traumatic brain injury, neurologic disorders, alcohol or drug abuse, or history of psychosis, and achieved a score of at least 27 on the Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975). They gave informed consent according to guidelines of the Institutional Review Board of Albert Einstein Healthcare Network and were paid $15 for their participation.

Materials and procedure

The study included a baseline experiment designed to control for individual grasping time differences and a main experiment designed to test the influence of visual context on action activation during object semantic processing. Critical stimuli were only involved in the main experiment and were selected from a preliminary study (see Supplementary Materials online). They were colored pictures of 20 manufactured objects associated with different move and use hand postures (e.g., kitchen timer). Objects were presented in either a MOVE environment, in which the visual scene was a context in which the object would be clenched with a power grip (e.g., kitchen timer in drawer) or a USE environment in which the object would be pinched with a precision grip (e.g. kitchen timer on countertop, with food). The association between the MOVE and USE scenes and the gestures evoked by the conflict objects (clench or pinch) was confirmed in the norming study (Supplementary Materials). There were 40 photographs corresponding to the two visual contexts for each of the 20 conflict objects (see example in Figure 1 and list in Supplementary Materials). The scenes represented an office, kitchen, or bathroom. In addition to the critical conflict objects, each scene also contained 4 distractor objects, both man-made and natural (e.g., fruit, vegetables, plants, flowers). A subset of these distractors was used as target objects on filler trials. Thirty natural and 10 man-made distractor objects appeared in both MOVE and USE context pictures. The other natural and man-made distractors objects only appeared in one picture. Distractor objects could afford either power or precision grips or both/none (e.g. plants). For each conflict object, we ensured that the different affordances were represented in equivalent proportions between use and move contexts. For instance for the kitchen timer (Figure 1), all distractor objects would be grasped with a clench, expect for broccoli (use context) and lime (move context) that may afford both clench and pinch grips.

Figure 1. Example of conflict object (kitchen timer) presented in a MOVE (left) or a USE (right) scene.

Sound files corresponding to category labels “natural?” and “man-made?” were recorded by a female native speaker of American English.

The response apparatus consisted of a 4-inch long by 1-inch diameter cylinder that afforded both a power grip by clenching the whole cylinder and a precision grip by pinching the tip of the cylinder. The response device was programmed in E-prime to record reaction times when participants squeezed the cylinder (Figure 2).

Figure 2. Experimental set up using the response device allowing clench and pinch grasps.

Baseline experiment

The goal of the baseline experiment was to provide individual mean reaction times for clenching and pinching the device without visual stimuli or a semantic task. Participants reached to and grasped the apparatus with either a pinch or a clench in response to “YES” and “NO” verbal cues (see Supplementary Materials).

Main experiment

On each trial, a fixation cross appeared in the center of the screen. Participants began each trial by pressing and holding the middle key of the response box with the index finger of their left hand[1]. The mobility of the right limb was limited with an arm sling. Immediately after pressing the key, the scene picture appeared on the screen. After a 1250ms delay, a red box appeared around the target or one of the four distractors. Location of target and distractors was randomized. Simultaneously, they heard an auditory cue, either “natural?” or “man-made?”. Participants then indicated whether the category label matched the object in the box by using the response device to indicate a “YES” or “NO” response. This was accomplished by releasing the response box key and reaching to grasp the cylinder with either a clench or a pinch. The picture disappeared when the start button was released. Participants in Group 1 clenched the device to respond “YES” and pinched it to respond “NO”, whereas participants in Group 2 performed the opposite mapping. They were instructed to respond as quickly and accurately as possible. Movement initiation and transport times were recorded automatically in E-Prime. Accuracy was coded online by an experimenter (c=clench, p=pinch, n=none). Gesture videotaping was used for offline accuracy checking. Participants performed 12 practice trials with feedback on accuracy, using pictures that were not displayed in the experiment. The experiment contained 120 trials.

Each of the 40 scenes was presented 3 times in randomized order resulting in 120 experimental trials. On 40 critical trials, the red box appeared around the conflict object in the scene. For the remaining filler trials, the box appeared around a natural distractor object on 60 trials and around a man-made distractor on 20 trials. Thus, the target object was natural

[1] Participants were always asked to respond with their left hand while their right arm was immobilized for future comparison with left hemisphere stroke patients. Left hemisphere stroke patients frequently have reduced right arm mobility.

and man-made on an equal number of trials. Each scene was repeated 3 times: once with the conflict object as target and twice with a distractor object as target. Since each conflict object was the target twice, once in the MOVE and once in the USE scene, the number of repetition of distractor objects as target was varied among filler trials so that overall, object category, object repetition across pictures, and target repetition were not informative in predicting which object in the scene would be the target on a given trial.

On half of the trials, the target object was coupled with the label "natural?" and on the other half coupled with the label "manmade?". Repeated target objects could be associated with the same label or a different label on both occurrences. Hence, when a given object was the target for the second time, the likelihood of hearing a repeated or new label was equivalent.

Data Analysis

In the baseline experiment, individual initiation times[2] for pinch and clench were calculated and used to reduce between-subject variability in the data from the main experiment (see below; also see Supplemental Materials for additional detail).

In the main experiment, data were trimmed and adjusted as follows. First, participants who were at chance level in at least one condition (accuracy < 75% according to binomial probability) were excluded from further analysis (N=3). One participant was particularly slow in baseline initiation times (3SD below the group mean) and was also excluded. Thus, the final data set included 21 participants. Second, analyses on initiation

[2] Our analyses focused on movement initiation times since object action-related features have been shown to affect grasp planning prior to movement execution (e.g., Bub, Masson, & Cree, 2008; Jax & Buxbaum, 2010; Girardi, Lindemann, & Bekkering, 2010). Nonetheless, note that we did not observe any effect of the variables of interest on transport times (all p's > .25).

times were conducted after removing incorrect trials (4% data) (No trials were excluded for being shorter than 200ms or longer than 3 standard deviations from the group mean in the corresponding condition). Finally, adjusted initiation times were computed at the individual level in each condition by subtracting initiation baseline times for pinch and clench from the respective initiation times in the main experiment.

A 2*2 Analysis of Variance was conducted on mean adjusted initiation times from critical trials with Gesture (pinch, clench) and Context (MOVE, USE) as within-subject (F_1) or within-item (F_2) factors. Distribution normality and variance homogeneity were verified. Errors were extremely rare: of the total of 840 trials run by all subjects in the experiment, only 35 trials had errors. Error distribution was highly skewed and not suited to a similar analysis as the one conducted on initiation times. Nevertheless, proportions of correct responses between conditions were compared using chi-square.

Results

Initiation times

There was no main effect of Gesture [$F1_{(1,20)}$= 0.31, R^2= 0.02 , p= .58; $F2_{(1,19)}$= 1.08, R^2= 0.05 , p= .31] or Context [$F1_{(1,20)}$= 1.15, R^2= 0.05 , p = .29; $F2_{(1,19)}$= 3.18, R^2=0.14 , p = .09]. Critically, the Gesture x Context interaction was significant in both the by-subject [$F1_{(1,20)}$= 4.8, R^2= 0.19, p = .04] and by-item [$F2_{(1,19)}$= 6.31, R^2=0.25 , p =.02] analyses. As shown in Figure 3, there was a greater advantage of the use context compared to the move context in the pinch gesture condition compared to the clench gesture condition.

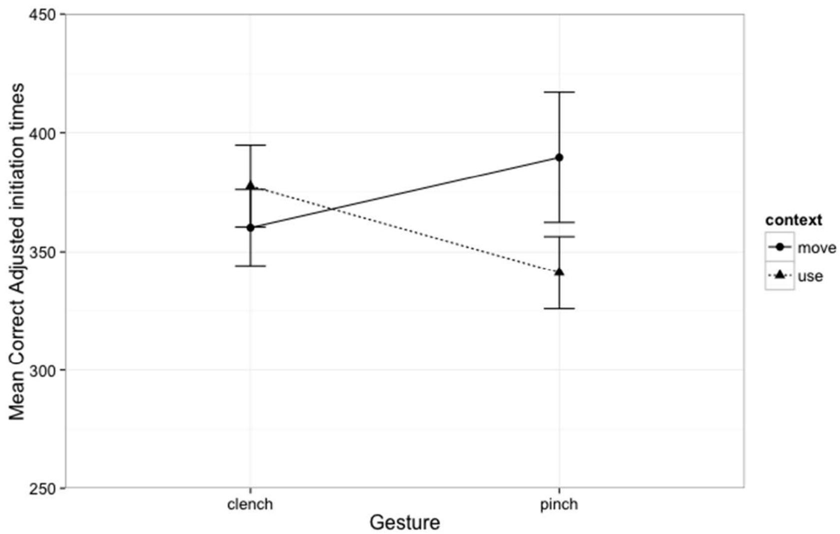

Figure 3. Mean correct adjusted initiation times (and standard errors) for clench and pinch categorization responses as a function of context (MOVE, USE).

Post-hoc comparisons of the by-item analysis indicated that the interaction between Gesture and Context was likely due to shorter initiation times in the use than in the move context for pinch ($t = 2.74$, $p = .01$), whereas there was no difference between use and move contexts for clench ($t = -0.2388$, $p = 0.81$). None of the post-hoc tests reached significance in the by-subject analysis, though the results were consistent with those demonstrated in the by-item analysis ($t=1.6891$, $p = .10$ between move and use contexts for pinch; $t=-0.5447$, $p = .59$ between move and use contexts for clench).

Correct responses

Chi-square test on accuracy data did not show any significant difference in proportion of correct responses between the four Gesture x Context conditions ($\chi^2 = 6.65$, $p= .08$). As can be seen in Table 1, the number of correct responses was numerically inferior for pinch responses in the use context, but this was anecdotal considering the absence of significant difference between conditions and the very limited number of errors. Consequently, accuracy data will not be further discussed.

Table 1: Number and proportion of correct responses in the different Context x Gesture conditions.

Context	Gesture	Number of correct responses	Proportion of correct responses
Move	Clench	203	96.2%
Use	Clench	204	96.6%
Move	Pinch	202	97.6%
Use	Pinch	196	92.9%

Discussion

We report context-dependent compatibility effects between the motor responses performed during object semantic categorization and the action evoked by the object in a given visual context. Prior demonstrations indicate that action evocation during object processing may be modulated by verbal context (Costantini, Ambrosini, Scorolli, & Borghi, 2011; Lee et al., 2013), affordances of distractor objects (Caligiore, Borghi, Parisi, Ellis, Cangelosi, et al., 2013; Ellis, Tucker, Symes, Vainio, 2007; Pavese & Buxbaum, 2002; Tipper, Howard, & Jackson, 1997), and relationships to other objects or agents (Borghi, Flumini, Natraj, & Wheaton, 2012; Ellis et al., 2013; Girardi et al., 2010; Yoon, Humphreys, & Riddoch, 2010). The present data extend such findings by demonstrating that activation of move- and use-related gestures during semantic object processing may additionally be modulated by the visual environment in which objects are presented. The visual environments used here were composed of 5 objects naturally displayed on a furnished room background. The fact that we observed compatibility effects with complex visual contexts provides additional ecological validity to action evocation phenomena during object processing and reinforces the idea that affordances are flexibly activated in natural environmental conditions. In addition, the data suggest that the contextual modulation observed in the present study is the outcome of a global visual processing of the scene that can be distinguished from the influence of single object affordances. Although distractor objects may have also activated the actions associated with them, their affordances were equivalent between contextual conditions. Thus, the context-dependent compatibility effects reported here are likely related to the meaning conveyed by the array and by the action intention that emerges from the visual scene.

The existence of such effects raises the challenge of identifying when and how visual context influences compatibility effects in the cascade of perceptual and motor processes. It is well-recognized that preparation of a motor response orients attention towards action-relevant features and may facilitate visual processing of stimuli that are congruent with that action (the "motor-visual attention" effect, e.g., Allport, 1987; Bekkering & Neggers, 2002; Botvinick, Buxbaum, Bylsma, & Jax, 2009; Craighero, Fadiga, Rizzolatti, & Umiltà, 1999; Hannus, Cornelissen, Lindemann, & Bekkering, 2005; Pavese & Buxbaum, 2002). Preparing a clench or a pinch may facilitate processing of distinct conflict object features (e.g., the entire kitchen timer vs. the timer dial, respectively). Consequently, faster object processing may be observed when the features highlighted by response preparation are compatible with one of the actions evoked by the object. At the same time, visual object processing appears to activate action representations, even in tasks not involving a motor response (e.g., Kalénine, Mirman, Middleton, & Buxbaum, 2012; Lee et al., 2013; Myung et al., 2010). Additionally, visual context influences object processing (e.g., Gronau, Neta, & Bar, 2008; Mudrik, Lamy, & Deouell, 2010). In objects associated with more than one action, such as the conflict objects presented here, we may speculate that the visual context serves to amplify the action associated with it (e.g., Wurm, von Cramon, & Schubotz, 2012). In an iterative manner, this "bottom up" facilitation of an object-related action by the context may resonate with the intention-driven facilitation of action by the planned action (see Chambon et al., 2011; Shen & Paré, 2011 for related accounts). Further investigations of context-dependent compatibility effects could potentially employ variations in the timing of experimental perceptual and motor events to specify how environment-based and intention-based processes interact during object processing.

Another main issue concerns the stage of object processing at which the observed context-relevant action effects emerge. While most studies on effect of context on action evocation from objects have induced "deep" object processing by using semantic decision tasks, a few studies have contrasted different processing levels and showed that affordances are not activated when the task requires shallow object processing (e.g., color judgments; Pellicano, Iani, Borghi, Rubichi, & Nicoletti, 2010; Tipper, Paul, & Hayes, 2006). One possibility is that context-relevant action modulation arises before conceptual object processing is completed, perhaps on the basis of associations between the target object, context, and actions. Context-dependent activation of object affordances could then impact semantic processing while emerging from earlier (pre-conceptual) stages of perceptual processing. Alternatively, object-related actions might be automatically evoked during early processing stages (Goslin, Dixon, Fischer, Cangelosi, & Ellis, 2012) and context modulation might arise later on during conceptual processing. Context could work as a late filter, which would enhance relevant action features and turn off irrelevant ones. Regardless, results overall suggest that all action features are not systematically integrated to object concepts and that context and goals play a decisive role in this integration.

The compatibility effects observed in the present study were largely driven by faster initiation of use-related actions when the object was presented in a use-compatible context compared to a move-compatible context. In contrast, initiation of move-related actions did not appear sensitive to visual context. This asymmetry could have been related to the fact that participants were required to respond with their left hand. Indeed, manual asymmetries have been reported in visually primed grasping (Vainio, Ellis, Tucker & Symes, 2006). However, manual differences were observed in the opposite direction, with an absence of object size-grip type compatibility effects when precision grip responses were performed with the left hand. A reduction of affordance effects has also been recently observed when right-handed participants used their left hand to execute memorized instructions on objects with handles that were spatially congruent or incongruent with the dominant hand (Apel, Cangelosi, Ellis, Goslin, & Fisher, 2013), suggesting that compatibility effects may be more difficult to observe when responses are performed with the left hand. Moreover, while manual asymmetries could possibly account for a main effect of grip type in the present paradigm (which we did not observe), they could not explain the observed context effects on precision grips.. If compatibility effects are overall enhanced/reduced for precision grips depending on the response hand, this should affect move and use context conditions equally. Thus, reasons for the asymmetry reported here remain uncertain but several potential explanations can be formulated. First, pinch grasps might be more context-specific than clench grasps. For instance, pinch might be more associated with opening a bottle with a corkscrew than clench is associated with moving this item. Second, use-related actions are often preceded by move-related actions, particularly in naturalistic environments. For example, one must first pick up a corkscrew with a clench prior to using it with a pinch. Accordingly, clenches may be equally triggered by use- and move-compatible contexts while pinches would be more strongly activated in use-compatible contexts. Finally, at the action planning level, one could consider the clench hand posture less specified than the pinch hand posture. In other words, the first phase of any grasping movement (pinch or clench) could start in some cases with a clench-like posture, and the position of the different fingers that are opposed to the thumb could require further determination. This possibility accords with neurophysiological data showing additional fronto-parietal recruitment for the control of precision grips compared to power grips (Ehrsson et al., 2001). Hence, clench action initiation would be as relevant for use-compatible and move-compatible environments and context would show little influence on clench responses.

In summary, the present study is the first to highlight the influence of visual scene on stimulus-response compatibility effects during semantic object processing. This finding brings additional support to action models that consider both action subtypes and context as key determinants for understanding interactions between object and action processing (e.g., Buxbaum & Kalénine, 2010). Moreover, our finding may have strong implications for object processing in naturalistic tasks where objects are perceived in their natural visual environments.

Disclosure

The authors report no competing interests.

Ackowledgements

We thank Rob Ellis and Mike Tucker for providing us with the response apparatus.

This research was supported by NIH RO1-NS065049 and James S. McDonnell Foundation #220020190 grant awards to Laurel Buxbaum.

References

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). NJ: Hillsdale: Erlbaum.

Apel JK, Cangelosi A, Ellis R, Goslin J, Fischer MH. (2012) Object affordance influences instruction span. *Experimental Brain Research, 223(2),* 199-206.

Barsalou, L. W. (2008). Grounded cognition. *Annual review of psychology*, *59*, 617–45.

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological science*, *13*(4), 370–4.

Borghi, A. M., Flumini, A., Natraj, N., & Wheaton, L. A. (2012). One hand, two objects: emergence of affordance in contexts. *Brain and cognition*, *80*(1), 64–73.

Botvinick, M. M., Buxbaum, L. J., Bylsma, L. M., & Jax, S. a. (2009). Toward an integrated account of object and action selection: a computational analysis and empirical findings from reaching-to-grasp and tool-use. *Neuropsychologia*, *47*(3), 671–83.

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.

Buxbaum, L. J., & Kalénine, S. (2010). Action knowledge, visuomotor activation, and embodiment in the two action systems. *Annals of the New York Academy of Sciences*, *1191*, 201–18.

Caligiore, D., Borghi, A.M., Parisi, D., Ellis, R., Cangelosi, A., Baldassarre, G. (2013). How affordances associated with a distractor object affect compatibility effects: A study with the computational model TRoPICALS. *Psychological Research, 77*(1), 7-19.

Chambon, V., Domenech, P., Pacherie, E., Koechlin, E., Baraduc, P., & Farrer, C. (2011). What are they up to? The role of sensory evidence and prior knowledge in action understanding. *PloS one*, *6*(2), e17133.

Costantini, M., Ambrosini, E., Scorolli, C., & Borghi, A. M. (2011). When objects are close to me: affordances in the peripersonal space. *Psychonomic bulletin & review*, *18*(2), 302–8.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, *40*(5), 492–502.

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: a motor-visual attentional effect. *Journal of experimental psychology. Human perception and performance*, *25*(6), 1673–92.

Ehrsson, H. H., Fagergren, A., Forssberg, H., Mosier, K., Lau, C., Wang, Y., Venkadesan, M., et al. (2001). Differential Fronto-Parietal Activation Depending on Force Used in a Precision Grip Task□: An fMRI Study. *Journal of Neurophysiology*, *85*(6), 2613–2623.

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, *91*(4), 451–471.

Ellis R, Tucker M, Symes E, Vainio L. (2007). Does selecting one visual object from several require inhibition of the actions associated with nonselected objects? J*ournal of Experimental Psychology: Human Perception Performance, 33(3)*, 670-91.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189–98.

Girardi, G., Lindemann, O., & Bekkering, H. (2010). Context effects on the processing of action-relevant object features. *Journal of experimental psychology. Human perception and performance*, *36*(2), 330–40.

Goslin, J., Dixon, T., Fischer, M. H., Cangelosi, A., & Ellis, R. (2012). Electrophysiological examination of embodiment in vision and action. *Psychological science*, *23*(2), 152–7.

Gronau, N., Neta, M., & Bar, M. (2008). Integrated contextual representation for objects' identities and their locations. *Journal of cognitive neuroscience*, *20*(3), 371–88.

Hannus, A., Cornelissen, F. W., Lindemann, O., & Bekkering, H. (2005). Selection-for-action in visual search. *Acta psychologica*, *118*(1-2), 171–91.

Jax, S. a, & Buxbaum, L. J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition*, *115*(2), 350–5.

Jax, S. A., & Buxbaum, L. J. (2013). Response interference between functional and structural object-related actions is increased in patients with ideomotor apraxia. *Journal of neuropsychology*, *7*(1), 12–8.

Kalénine, S., Mirman, D., Middleton, E. L., & Buxbaum, L. J. (2012). Temporal dynamics of activation of thematic and functional knowledge during conceptual processing of manipulable artifacts. *Journal of experimental psychology. Learning, memory, and cognition*, *38*(5), 1274–95.

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*(1), 257–70.

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object-scene processing. *Neuropsychologia*, *48*(2), 507–17.

Myung, J., Blumstein, S. E., Yee, E., Sedivy, J. C., Thompson-Schill, S. L., & Buxbaum, L. J. (2010). Impaired access to manipulation features in Apraxia: evidence from eyetracking and semantic judgment tasks. *Brain and language*, *112*(2), 101–12.

Pavese, A., & Buxbaum, L. J. (2002). Action matters: The role of action plans and object affordances in selection for action. *Visual Cognition*, *9*(4), 559–590.

Pellicano, A., Iani, C., Borghi, A. M., Rubichi, S., & Nicoletti, R. (2010). Simon-like and functional affordance effects with tools: the effects of object perceptual discrimination and object action state. *Quarterly journal of experimental psychology (2006)*, *63*(11), 2190–201.

Schwartz, M. F., Brecher, A. R., Whyte, J., & Klein, M. G. (2005). A patient registry for cognitive rehabilitation research: a strategy for balancing patients' privacy rights with researchers' need for access. *Archives of physical medicine and rehabilitation*, *86*(9), 1807–14.

Shen, K., & Paré, M. (2011). Neural basis of feature-based contextual effects on visual search behavior. *Frontiers in behavioral neuroscience*, *5*, 91.

Tipper, S. P., Paul, M. A., & Hayes, A. E. (2006). Vision-for-action: the effects of object property discrimination and action state on affordance compatibility effects. *Psychonomic bulletin & review*, *13*(3), 493–8.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of experimental psychology. Human perception and performance*, *24*(3), 830–46.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, *8*(6), 769–800.

Vainio L, Ellis R, Tucker M, Symes E (2006). Manual Asymmetries In Visually Primed Grasping. Experimental Brain Research, 173 (3), 395–406.

Wurm, M. F., von Cramon, D. Y., & Schubotz, R. I. (2012). The context-object-manipulation triad: cross talk during action perception revealed by fMRI. *Journal of Cognitive Neuroscience, 24(7)*, 1548-59.

Yoon, E. Y., Humphreys, G. W., & Riddoch, M. J. (2010). The paired-object affordance effect. *Journal of experimental psychology. Human perception and performance*, *36*(4), 812–24.

Supplementary materials

Normative study

The 40 critical scene pictures were assessed for their ability to evoke a particular gesture, either pinch or clench, when the subject was explicitly asked to pantomime a gesture appropriate for the conflict object as depicted in the scene. Sixteen additional healthy volunteers (11 female / 5 male) participated in the norming experiment. On each trial, the scene picture appeared on the screen. After 1000 ms, a red box appeared around the conflict item in this photograph for 750 ms and then disappeared. Participants' task was to pantomime with their left hand how they would interact with the highlighted object in that particular context. The mobility of the right limb was limited with an arm sling. Subjects were given explicit instructions to take note of the context as it would inform them about how they might interact with that object in real life. They were also instructed to respond as quickly as possible so as to reflect the most natural and immediate gesture evoked by the object in that environment.

A response was coded as a clench if pantomimed contact with an imaginary object included the palm of the hand, had a rounded aperture, and used more than 3 active fingers. A response was coded as a pinch if the subject pantomimed contact with the object with only the thumb and the index finger or only the thumb, index, and middle finger. It was also coded as a pinch if the subject gestured with more than 3 fingers but pantomimed contacting the object with the tips of the fingers and/or formed a hand posture with a flat aperture. Responses were recorded by video camera and coded offline by one of 2 experimenters, who demonstrated 90% inter-rater reliability. To prevent coding biases experimenters did not have knowledge of the scene or object to which the subject was pantomiming. The norming experiment contained 40 trials, presented randomly. There were two scenes for each of 20 conflict objects: one scene depicted the conflict object in a USE context and the other scene depicted the conflict object in a MOVE context.

Norming data confirmed that each conflict object received more pinch gestures in the USE compared to the MOVE context while it received more clench gestures in the MOVE compared to the USE context. Overall, in the MOVE context there were 175 clench and 142 pinch gestures, whereas in the USE context there were 103 clench and 210 pinch gestures.

Baseline Experiment

A fixation cross appeared on the screen and participants began the trial by using the index finger of the left hand to press the middle key of a response box positioned in front of them. The mobility of the right limb was limited with an arm sling. Participants were asked to keep the key depressed until presentation of a verbal cue. After a variable delay of between 1500 and 3000ms, the word "YES" or "NO" was delivered through speakers. As soon as they heard the word, participants released the key and grasped the response device, which was a cylinder mounted vertically in a wooden support, located 13 inches from the response box (Figure 2). Participants were randomly assigned to one of 2 groups. Participants in Group 1 had to clench the cylinder when they heard "YES" and pinch it when they heard "NO", whereas those in Group 2 had to perform the opposite mapping. They were instructed to respond as quickly and accurately as possible. Initiation times (word offset to liftoff) and transport times (liftoff to cylinder contact) were recorded automatically in E-Prime. Accuracy was coded online by the experimenter (c= clench, p=pinch, n=none). Gesture videotaping was used for offline accuracy checking. There were 10 practice trials with reaction time and accuracy feedback (5 "YES" and 5 "NO" in random order), followed by 24 baseline trials where no feedback was provided (12 "YES" and 12 "NO" in random order).

Correct movement initiation times were computed as a function of the Gesture performed on the device in response to "yes" and "no" labels (pinch or clench). After the 10 practice trials, participants were 100% correct on the 24 baseline trials. One participant was particularly slow in initiation times (3SD below the group mean) and was excluded from further analysis. For the remaining 24 participants, initiation times that were either shorter than 200ms or longer than 3 standard deviations below the mean of the group in the pinch and clench conditions were considered outlier trials and removed from the data (1.5%).

Conflict item list with their corresponding MOVE and USE scenes

Conflict item	Move scene	Use scene
Book	On bookshelf	Open on desk
Playing cards	Stacked in drawer	Stacked on card table
Binder clip	In supply drawer	Affixed to paper stack on desk
Cookie jar	On pantry shelf	On counter, lid slightly ajar
Corkscrew	In kitchen drawer	In corked wine bottle
Cheese grater	In kitchen drawer	In bowl on kitchen counter
Jewelry box	On bathroom shelf	On bathroom counter, slightly ajar
Keys	In desk drawer	Inserted to desk lock
Lamp	On supply shelf	On desk, angled toward magazine
Pencil sharpener	On bathroom shelf	On desk, pencil inserted
Pin cushion	On bathroom shelf	On bathroom counter
Post-it	On bathroom shelf	On desk, top note written on
Pot lid	On dish rack	On pot on stove
Soda can	On pantry shelf	On counter next to glass
Measuring spoons	In kitchen drawer	On kitchen counter, inserted in backing soda
Tape dispenser	In desk drawer	On desk, with gift wrapping supplies
Kitchen timer	In kitchen drawer	On kitchen counter
Tissue	On bathroom shelf	On bathroom counter, slightly ajar
Toilet paper	On bathroom shelf	On roll next to toilet
Tupperware	On dish rack	On kitchen counter with food inside, slightly ajar