



Learning to use novel objects: A training study on the acquisition of novel action representations

M. van Elk^{a,b,*}, M. Paulus^a, C. Pfeiffer^{a,b}, H.T. van Schie^c, H. Bekkering^a

^a Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

^b Laboratory of Cognitive Neuroscience, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Switzerland

^c Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 6 November 2009

Keywords:

Novel objects
End location
Grip
Action representations
Action observation

ABSTRACT

Many studies have suggested that the motor system is organized in a hierarchical fashion, around the prototypical end location associated with using objects. However, most studies supporting the hierarchical view have used well-known actions and objects that are highly over-learned. Accordingly, at present it is unclear if the hierarchical principle applies to learning the use of novel objects as well. In the present study we found that when learning to use a novel object subjects acquired an action representation of the end location associated with using the object, as evidenced by slower responses in an action observation task, when the object was presented at an incorrect end location. By showing the importance of knowledge about end locations when learning to use a novel object, the present study suggests that end locations are a fundamental organizing feature of the human motor system.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Imagine yourself wandering around in a museum of modern art, in which all types of different colored objects are placed at tables. A sign on the wall encourages you to grasp the objects to find out what happens. When you grasp the first object, you notice that a tiny melody comes out of the object and you move the object closer to your ear. Your movement towards the next object elicits a breeze of rose smell and you move the object towards your nose to find out whether it is the object that produces the odor. As the example nicely illustrates, the end location to which objects are directed plays an important role in understanding what the object can be used for. For example, we put headphones in our ears to listen to music or we grasp a comb to straighten our hair before dinner.

Developmental studies suggest that the ability to grasp and to play with objects precedes infant's knowledge about what an object can be used for (Barrett, Davis, & Needham, 2007). Within the first year of life infants gradually learn how to grasp and manipulate objects according to their physical properties (Bourgeois, Khawar, Neal, & Lockman, 2005), but it is only during the second year of life that children develop the ability to effectively grasp and use objects in a meaningful fashion (e.g. grasping a spoon to eat; McCarty, Clifton, & Collard, 1999). For example, when presented with a spoon with the handle oriented to the left or the right, 9-month old infants always grasped the spoon with their dominant hand, irrespective of the object's orientation, ending up with an odd grip when approaching the mouth. In contrast, older infants anticipated the desired end location (i.e. bringing the spoon to the mouth) and adjusted their manner of grasping the spoon depending on the object's orientation.

* Corresponding author. Address: Laboratory of Cognitive Neuroscience, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Station 19, AI 2101, 1015 Lausanne, Switzerland. Fax: +41 21 693 1770.

E-mail address: m.vanelk@epfl.ch (M. van Elk).

The finding that older infants learn to adjust their manner of grasping according to the desired end location is an instance of the more general principle that the way in which people approach and grasp objects is largely determined by their consecutive use. Studies in the domain of motor control show for instance that during grasping movements people automatically select the handgrip in a way that they will end the movement in a comfortable end-position, which is known as the end-state comfort effect (Rosenbaum, Vaughan, Barnes, & Jorgensen, 1992). These findings are in line with the view that our motor system is organized in a hierarchical fashion and that the specific grip that is used for grasping an object is selected in accordance with a specific end location (Grafton & Hamilton, 2007; Majdandzic et al., 2007; van Schie & Bekkering, 2007). In support of this view, it has been found that neurons in the inferior parietal lobe of the monkey respond differently if the same grasping action is part of an action sequence directed towards the mouth (e.g. grasping to eat) or towards the table (e.g. grasping to place), suggesting that different movement elements are linked to their final end location (Fogassi et al., 2005). In addition, it has been found that the processing of goal- and grip-related information during action planning calls upon different brain networks (Majdandzic et al., 2007; van Schie & Bekkering, 2007). Furthermore, recent studies suggest that conceptual knowledge about objects may be organized in a hierarchical fashion as well around the prototypical end location associated with objects (Lindemann, Stenneken, van Schie, & Bekkering, 2006; van Elk, van Schie, & Bekkering, 2008b). That is, because we have a lot of experience bringing objects to specific end locations (e.g. we typically move a cup towards our mouth in order to drink) strong associations exist between an object and the object's associated end location (e.g. cup-mouth).

However, most studies discussed thus far involved actions and objects that were highly familiar and over-learned. Accordingly, at present it is unclear if the dominance of end locations over grips in our interaction with objects reflects a general feature of the human motor system, or is limited to well-known or familiar actions. The aim of the present study was to investigate if the dominance of end locations over grips applies to the acquisition of novel action representations concerning the object's use as well. To investigate this question we employed an object training paradigm. Subjects first trained to use novel objects in a specific fashion. In the first experiment subjects learned to use objects that elicited a specific action effect (i.e. producing a sound or odor when brought towards the ear or nose) and in the second experiment subjects trained with objects without an action effect. The rationale for using objects that elicited an action effect in the first experiment was to mimic naturalistic object use as close as possible. That is, in daily life most of our actions do not consist of the arbitrary movement of objects towards end locations, but the movements have a specific function (e.g. grasping a headphone to bring it towards the ear serves the function of hearing). To replicate the findings from the first experiment and to examine the importance of object function for learning end locations, in a second experiment subjects practiced with objects that did not elicit an action effect.

After subjects trained with the novel objects, an object recognition task was administered, in which subjects were first presented with a photograph of a trained or untrained object, which was followed by an action snapshot. The picture of the single object served as an implicit cue for the retrieval of action information associated with that object (e.g. the object is brought towards the ear). Previous studies have indicated that action information about objects is automatically activated when objects are perceived (e.g. Bub, Masson, & Cree, 2008) and that an important aspect of this knowledge consists of the object's prototypical end locations (van Elk, van Schie, & Bekkering, 2009).

The object picture was followed by an action snapshot, in which subjects observed an actor performing an action with the object that was correct or incorrect with respect to how subjects were trained to use the objects. Subjects were instructed to decide if the objects that were presented in the two subsequent pictures (single object and the action snapshot) involved the same object or different objects (two-choice response task). A comparable action observation paradigm has been used successfully to assess the functional organization of conceptual knowledge about well-known objects (van Elk, van Schie, & Bekkering, 2008a). By manipulating the correctness of the action context in which an object was presented, a stronger interference of end location-violations (e.g. cup presented near ear) than of grip-violations (e.g. cup presented with odd grip) was found, suggesting that conceptual knowledge for understanding other's actions is organized primarily around end locations. In a similar fashion, in the present study the end location and the grip that the actor applied to the objects were independently manipulated. Although the action-object relation was only implicit and irrelevant to the subjects' task (same/different decision), a facilitation in object identification was expected if the observed action was appropriate for the object. Trials with incorrect or incompatible actions, on the other hand, were expected to delay object identification, and to result in slower reaction times. As a consequence, any effect of the action correctness on reaction times could be regarded as an implicit effect of activated action representations on the processing of observed actions.

2. Experiment 1

2.1. Methods

In the first experiment subjects performed goal directed actions with objects that elicited a sound-effect or an odor-effect when brought towards the ear or nose respectively. Accordingly, the objects had a clearly defined function that was directly related to the specific way in which the object was grasped and moved towards the face. After training, subjects conducted an object recognition task, deciding whether two subsequent pictures (single object and the action snapshot) represented

the same object. Training and test blocks were presented in an alternating fashion, to minimize the possible influence of repeatedly being confronted with pictures representing objects being used in an incorrect fashion on acquired object representations.

2.2. Subjects

In the first experiment, 16 right-handed subjects (four males, mean age: 19.1 years) participated. All subjects were students at the Radboud University Nijmegen and received course credits or 8 euro's for participation. Handedness was assessed through informal verbal inquiry.

2.3. Stimuli

Four novel objects were constructed for which in daily life no specifically known function exists. The four objects were divided in two object-sets. Each object-set consisted of one object that could be brought towards the ear (ear-object) and one object that could be brought towards the nose (nose-object). The ear-object produced a sound-effect when brought towards the ear and the nose-object produced an odor (shampoo smell). As a sound-effect a bell sound of 2000 ms duration was used. The volume was adjusted for each individual subject to ensure that the sound effect was heard when the object was brought towards the ear. For the nose-object informal verbal enquiry confirmed that participants smelled the object when brought towards the nose. One of the ear objects could be grasped with a pronated powergrip in order to move the object subsequently towards the ear (see Fig. 1 upper panel). The other ear object could be grasped with a supinated powergrip to move it towards the ear (see Fig. 1 lower panel). One of the ear objects could be grasped at the handle, with the thumb opposite to the fingers and this object had to be inverted in order to bring it towards the nose. The other ear object could be grasped at the handle with a supinated precision grip in order to move it towards the nose. Subjects always trained only with one object-set (target objects) and did not train with the other objects (no-go objects). The training of different object-sets was counterbalanced between participants.

For each object four different action pictures were taken, in which the correctness of the end location and the grip applied to an object were independently manipulated with respect to the way in which subjects trained to use that object. Accordingly, objects could be categorized with respect to their designated use as: (1) correct end location, correct grip (e.g. ear object directed towards ear and grasped in a proper way), (2) correct end location, incorrect grip (e.g. ear object directed towards ear and grasped in incorrect way), (3) incorrect end location, correct grip (e.g. ear object directed towards nose, but grasped in a proper way), or (4) incorrect end location, incorrect grip (e.g. ear object directed towards nose, and grasped in an incorrect way). In addition, a single picture was taken of each object to be used as a task cue in the training phase (described below) and as an implicit recall cue in the test phase.

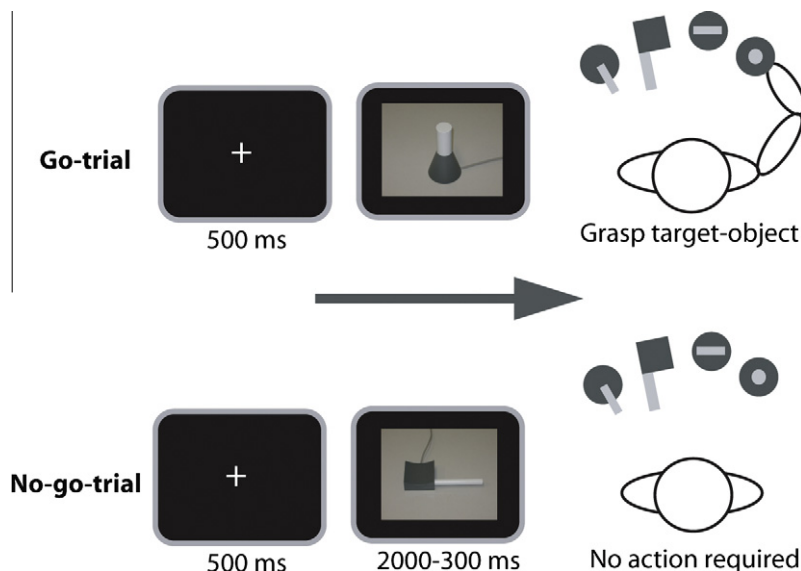


Fig. 1. Experimental procedure during training blocks. If a picture on the screen represented a target-object (upper panel), subjects responded by releasing the starting button, grasping the object and bringing the object to the end location (ear or nose). The next trial was initiated when the subject returned to the starting button. If the picture on the screen was a no-go object (lower panel), no subsequent action was required and subjects held their hand at the starting button until the next trial was presented.

2.4. Setup

Subjects were seated comfortably in front of a table on which a computer display, four novel objects and a response-device were placed (see Fig. 1). Pictures were presented on a 19" computer display at a resolution of 1024×768 pixels. At a viewing distance of 80 cm this resulted in pictures with a visual angle of approximately 13° and the size of the objects in the picture closely resembled real object size. The training phase, in which each participant acquired action experience with two out of four novel objects, was divided in three training blocks, alternated with three test blocks, in which participants performed a visual object recognition task. During training blocks movement kinematics of the right thumb and index finger were recorded, using the MiniBird-motion tracking system (miniBIRD 800, Ascension Technology Corporation). Spherical areas of 5 cm around each object and 10 cm around the ear and the nose were defined in order to control task performance online and to present a sound-effect if the ear object was transported towards the ear. Error messages were presented, if a wrong object was grasped and if an object was brought to a wrong end location. The experiment was controlled using Presentation 11.0.03 (neurobehavioral systems, Albany, CA).

2.5. Training phase

Before the experiment started, subjects were instructed how to grasp the two target-objects in a proper way and how to bring the objects to their designated end location. During each training block, subjects trained to use the objects in the designated way. In order to be able to directly compare performance in visual object recognition for trained and untrained objects, in the training phase participant had to perform specific actions with two out of four novel objects (target objects), whereas no action was required for the remaining two novel objects (no-go objects) on the onset of a visual cue. The assignment of objects to go- and no-go sets was randomized between participants, thereby counterbalancing object-specific influences on visual object recognition on the level of the group. The timeline for a single training trial is represented in Fig. 1. A fixation cross was presented for 500 ms, followed by a picture of a single object. The subject's task was to perform the appropriate action with the object if it was a target-object and to withhold from responding if it was a no-go object. If the depicted object was a target-object, subjects released the starting button, grasped the target-object in the appropriate way, brought the target-object to the designated end location and put the object back on the table. The next trial was initiated when the subject returned to the starting button. If the depicted object was a no-go object no response was required and the next trial was initiated after 2–3 s. During the training block, each object was presented 12 times, resulting in 12 actions with each of the target-objects (12 times action towards the nose; 12 times action towards the ear) and 24 no-go trials, resulting in a total number of 48 trials per block.

2.6. Test phase

The timeline for a single trial in the test phase is depicted in Fig. 2. Each trial started with a fixation cross for 500 ms, followed by a picture of a single object for 1000 ms. Next a fixation cross appeared for 1000 ms, followed by a picture representing an actor using the object. Subjects were required to respond to the second picture, by indicating whether the object

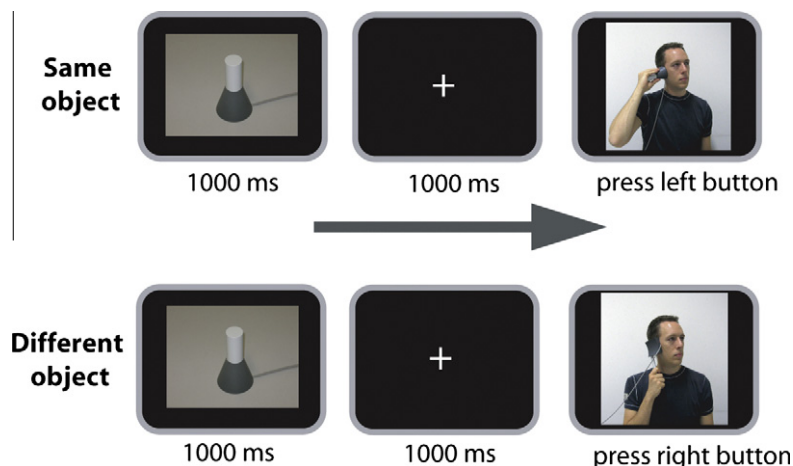


Fig. 2. Experimental procedure during test blocks. Subjects conducted an object recognition task, indicating whether an actor was using the same object as the single object presented in the first picture. If the object in the two subsequent pictures was the same, subjects responded by pressing the left button (upper part of figure; target trials) and if a different object was presented in the second picture subjects pressed the right button (lower part of figure; catch trials). For each object, four different action pictures were created in which the correctness of the end location and the grip applied to the object were independently manipulated.

in the action picture was the same object as presented in the first picture or not (object recognition task). If the objects in both pictures were the same, subjects responded by pressing the left button whereas if the objects differed subjects responded by pressing the right button. In this way we ensured that participants always responded to target trials with the hand that was unrelated to the object's use and that the response to target trials was always made with the same hand. Reaction times were measured relative to onset of the second stimulus. In 67% of all trials both pictures represented the same object (target trials) and in 33% of all trials a different object was presented in the second picture (catch trials). Accordingly, target trials consisted of two subsequent pictures, the first showing an object only and a second showing an object used by an actor, in which the same object was presented. For each of the four objects, four different pictures composed of the factor-level combinations for End location (correct, incorrect) and Grip (correct, incorrect) could be presented, which were repeated six times each, resulting in 96 target trials. Forty-eight catch-trials were included as filler items in which a different object was presented in the second picture, resulting in a total number of 144 trials during each test block. The first eight trials were used as practice trials in which the subject familiarized with the task. Training and test blocks were presented in an alternating fashion and each block was repeated three times to allow testing for the effects of training over time. Analysis focused on the test blocks in which eventual effects of training should become apparent in a facilitation of reaction times for observation of trained objects being used in a correct fashion. Please note that for the analysis of reaction times we only focused on target trials in which the two subsequent pictures represented the same object. In filler trials a different object was presented in the second picture and therefore it was not possible to investigate the effects of a previously activated action representation on subsequent object recognition. Finally, to control for the possible confound that the effects observed could be attributed to differences in stimulus salience, we also analyzed reaction times to non-trained objects. Because the trained object sets were counterbalanced between participants, overall the reaction times to trained and untrained objects were obtained in response to exactly the same pictures. Accordingly, any difference in the reaction time-effects between trained and untrained objects should be attributable to the training of the subject with two of the four objects. Although the subject did not know about the correct goal and grip for the untrained objects, we analyzed reaction times to untrained objects using a 2 (End location: correct vs. incorrect) \times 2 (Grip: correct vs. incorrect) repeated measures ANOVA, similar to the reaction times to trained objects.

2.7. Results

2.7.1. Training blocks

During the training blocks subjects grasped the wrong object in less than 1% of all trials. Incorrect end-postures (moving the object to an incorrect position near the face) were detected in 8.3% of all trials. The incorrect end-postures likely reflect that when subjects grasped the ear-object they tended to slightly tilt their head and accordingly the ear-object was detected in the area that was previously defined as the nose area. Error messages on the screen encouraged subjects not to move their head during the transportation phase. Reaction and movement times are represented in Table 1 and differences between blocks were analyzed using separate *t*-tests.

For reaction times, a significant difference was only found between the first block (672 ms) and the third block (634 ms), $t(15) = 2.2$, $p < .05$. For object grasping times, marginally significant differences were found between the first block (838 ms) and the second block (700 ms), $t(15) = 1.9$, $p = .07$ and between the first block (838 ms) and the third block (695 ms), $t(15) = 2.1$, $p = .057$. For end posture times, no significant differences were found between blocks (p 's $> .16$).

2.7.2. Test blocks

Subjects incorrectly responded to action pictures in 4.5% of all trials. No difference was found between errors in response to action pictures representing trained compared to untrained objects ($F < 1$). Trials with incorrect responses or trials that exceeded the subject's mean reaction time by more than two standard deviations were excluded from subsequent analysis.

Reaction times averaged over test blocks for trained objects are represented in Fig. 3 and were separately analyzed using a 2 (End location: correct vs. incorrect) \times 2 (Grip: correct vs. incorrect) repeated measures ANOVA. Analysis of the reaction

Table 1

Reaction times and movement times for the three training blocks in the first and second experiment. Reaction and Movement Times are represented in milliseconds and standard errors are between brackets.

	Block 1	Block 2	Block 3
<i>Experiment 1</i>			
Reaction times	672 (22)	652 (20)	634 (22)
Object grasping times	838 (83)	700 (71)	695 (75)
End posture times	1667 (154)	1604 (128)	1431 (84)
<i>Experiment 2</i>			
Reaction times	701 (43)	675 (42)	644 (38)
Object grasping times	730 (42)	713 (34)	679 (37)
End posture times	1597 (68)	1628 (98)	1503 (82)

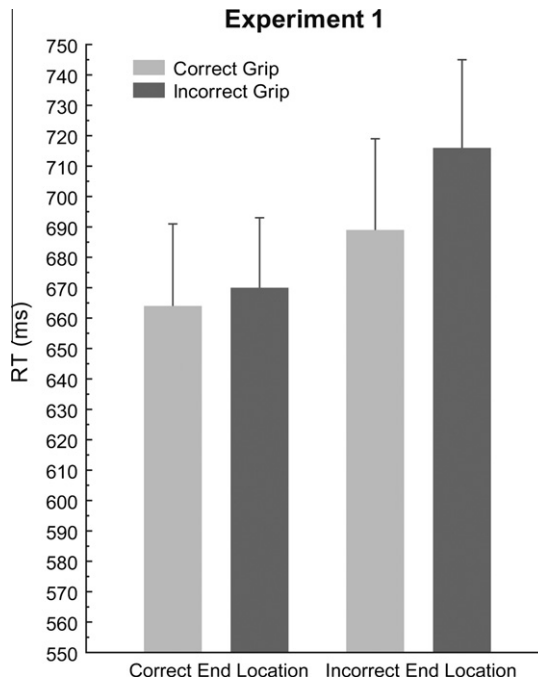


Fig. 3. Reaction times in Experiment 1 for the trained objects for the different stimulus categories. Bars on the left represent pictures in which the trained object was presented at a correct end location, bars on the right indicate objects at an incorrect end location. Light bars represent objects grasped with a correct grip and dark bars represent objects being grasped with an incorrect grip. Error bars represent standard errors.

times to trained objects revealed a significant main-effect of End location, $F(1, 15) = 10.7$, $p < .01$, $\eta^2 = .42$, reflecting faster reaction times to pictures representing a target-object at a correct (667 ms) compared to an incorrect end location (703 ms). In addition, a main effect of Grip was found, $F(1, 15) = 4.8$, $p < .05$, $\eta^2 = .24$ reflecting faster reaction times in response to correct grips (677 ms) compared to incorrect grips (693 ms). No interaction was found between End location and Grip, $F(1, 15) = 1.1$, $p > .30$, $\eta^2 = .07$.

To investigate whether the reaction time data differed between test blocks, Block was included as an additional variable, using a three (Block: 1st, 2nd and 3rd) \times two (End location: correct vs. incorrect) \times two (Grip: correct vs. incorrect) repeated measures ANOVA. In the analysis of the reaction time data, Block did never interact with any of the other variables (F 's < 1) and accordingly, averaging reaction times over blocks seems warranted. Finally, the analysis of reaction times for untrained objects did not reveal significant main-effects or interaction effects (all F 's < 1). Therefore it is unlikely that the reaction time effects observed in response to trained objects can be attributed to differences at a perceptual level.

2.8. Discussion

In the first experiment subjects trained to use novel objects that elicited an action effect (sound or odor) when used in a proper fashion. After training, subjects performed an object recognition task in which the correctness of the action (end location at which the object was presented and the grip applied to the object) was manipulated. A strong interference with reaction times was found when subjects responded to pictures in which the trained object was held at an incorrect end location near the face. Furthermore, reaction times were slower when trained objects were grasped with an incorrect grip. Importantly, reaction time effects cannot be explained on the basis of perceptual differences between the stimulus categories, because no interference effects were observed for the untrained objects. The present findings suggest that learning to use a novel object that elicits a meaningful action effect results in the acquisition of action semantics, representing both the appropriate grip and the correct end location of the object. Interestingly, observing an object at an incorrect end location caused a stronger interference with object identification than observing an object being grasped with an incorrect grip. In addition, reaction time effects of incorrect end locations and grips seemed to be additive, with the strongest interference occurring when both the end location and the grip applied to the object were incorrect. Thereby the first experiment provides support for the idea that the dominance of end locations over grips is a general principle underlying learning the use of novel objects, and is not restricted to the use of familiar objects and tools.

A possible functional mechanism underlying the findings from the first experiment could be the ideomotor principle, according to which actions are represented in terms of the effects they produce (Hommel, 1996; Hommel, Musseler, Ascherleben, & Prinz, 2001; Kunde, Hoffmann, & Zellmann, 2002; Prinz, 1997). On this account, the observation of an object

automatically activates the associated action effect, which – according to the ideomotor principle – should activate the motor program required for using the object (Paulus, Hunnius, Vissers, & Bekkering, in press). The activated motor program in turn facilitates the recognition of objects that are used in a proper fashion (van Elk et al., 2008a). According to this interpretation, the interference effects of incorrect end locations and grips found in the first experiment reflect the incongruence between the displayed action and the acquired action representations that were activated by retrieving the action effect in response to observing the picture of a single object. Following this line of reasoning, an interesting question is how important the action effect of an object (in this case a sound- or odor-effect) is in facilitating the acquisition of novel action representations. For example, it could well be that learning to use objects without a clearly defined action effect results in the acquisition of different action representations, in which the relation between the end location and the grip is more arbitrary and gets less well established. To address this question we conducted a second experiment in which subjects trained with objects that did not have a clear action effect (no odor or sound was presented when the object was brought towards the nose or ear).

3. Experiment 2

The same experimental setup and procedure as in the first experiment were used. Subjects learned to perform actions with novel objects and subsequently they performed an object recognition task, in which the correctness of the action context (end location at which the object was presented and grip applied to the object) was manipulated. In contrast to the first experiment, the objects that the subject used did not elicit an action effect. Subjects were trained to always bring one object towards their nose and one other object always towards their ear. In the second experiment 16 right-handed subjects participated (four males, mean age = 23.3 years) who had not participated in the previous experiment.

3.1. Results

3.1.1. Training blocks

During the training blocks subjects grasped the wrong object in less than 1% of all trials. Incorrect end-postures (moving the object to an incorrect position near the face) were detected in 4.0% of all trials. Reaction and movement times are represented in Table 1 and differences between blocks were analyzed using separate *t*-tests.

For reaction times, significant differences were found between the first block (701 ms) and the second block (675 ms), $t(15) = 2.3$, $p < .05$ and between the first block (701 ms) and the third block (644 ms), $t(15) = 5.1$, $p < .001$. A marginally significant difference was found for reaction times between the second block (675 ms) and the third block (644 ms), $t(15) = 2.1$,

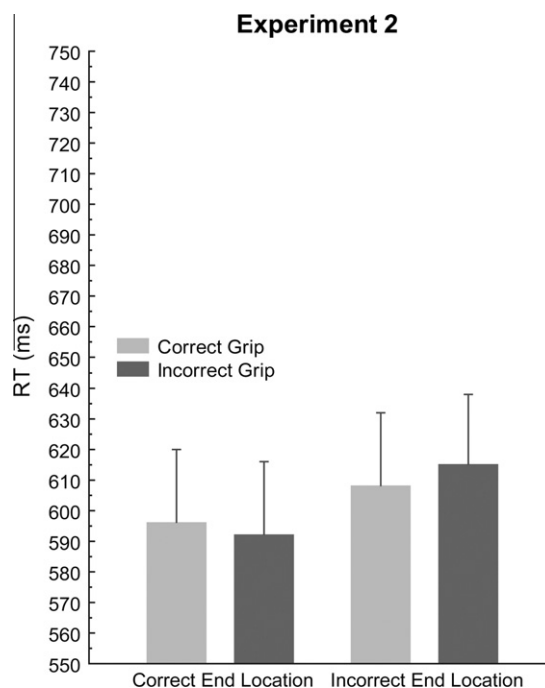


Fig. 4. Reaction times in Experiment 2 for the trained objects for the different stimulus categories. Bars on the left represent pictures in which the trained object was presented at a correct end location, bars on the right indicate objects at an incorrect end location. Light bars represent objects grasped with a correct grip and dark bars represent objects being grasped with an incorrect grip. Error bars represent standard errors.

$p = .053$. For object grasping times marginally significant differences were found between the first block (730 ms) and the third block (679 ms), $t(15) = 2.0$, $p = .06$ and between the second block (713 ms) and the third block (679 ms) $t(15) = 2.0$, $p = .06$. No significant differences were found between blocks in end posture times (p 's $> .11$). Data from the training blocks of the second experiment confirms findings from the first study, by showing the largest decrease of reaction times and object grasping times between the first and the third block.

3.1.2. Test blocks

In the second experiment subjects incorrectly responded to action pictures in 8.7% of all trials. Trials with incorrect responses or trials that exceeded the subject's mean reaction time by more than two standard deviations were excluded from subsequent analysis. Reaction times for trained and untrained objects in the second experiment are represented in Fig. 4 and were analyzed separately using a two (End location: correct vs. incorrect) \times two (Grip: correct vs. incorrect) repeated measures ANOVA.

Analysis of the reaction times to trained objects revealed a significant main-effect of End location, $F(1, 15) = 11.5$, $p < .005$, $\eta^2 = .43$, reflecting faster reaction times to pictures representing a trained object at a correct (594 ms) compared to an incorrect end location (611 ms). No main-effect of Grip was found and no interaction was found between End location and Grip (F 's < 1).

To investigate whether the reaction time data differed between test blocks, Block was included as an additional variable, using a three (Block: 1st, 2nd and 3rd) \times two (End location: correct vs. incorrect) \times two (Grip: correct vs. incorrect) repeated measures ANOVA. In the analysis of the reaction time data, Block did never interact with any of the other variables (F 's < 1) and accordingly, averaging reaction times over blocks seems warranted. To control for the possible confound that reaction time effects to target objects actually reflect differences in saliency between the stimulus categories, we also analyzed reaction times to untrained objects using a two (End location: correct vs. incorrect) \times two (Grip: correct vs. incorrect) repeated measures ANOVA. No main-effect of End location or Grip was found for the untrained objects and no interaction was found between End location and Grip (p 's $> .14$).

3.2. Discussion

In the second experiment subjects learned to use novel objects that did not elicit an action effect when brought towards the nose or ear. In the subsequent object recognition task subjects responded slower to pictures in which a trained object was presented at an incorrect end location, thereby replicating the findings from the first experiment. In contrast to experiment 1 no main effect of Grip was found. The relevance of the present findings and theoretical implications will be discussed in the final section.

4. General discussion

The present study investigated whether the dominance of goals over grips, observed in previous studies for the use of well-known objects, applies to the acquisition of novel action representations as well. Previous studies have shown that action knowledge about objects is automatically activated when the object is perceived (Bub et al., 2008) and facilitates the subsequent recognition of objects (van Elk et al., 2008a). Accordingly we expected that the training with novel objects should result in the acquisition of action representations, which in turn should facilitate the recognition of objects being used in a proper fashion.

In the first experiment it was found that subjects who learned to use novel objects that elicited an action effect responded slower to pictures in which the object was used in an incorrect fashion. More precisely, when presented with a picture of the target-object during the test phase, object identification was found delayed when the object was presented at an incorrect end location or when the object was grasped with an incorrect grip. Interference effects of end location- and grip-violations during the recognition of objects can be regarded as a mismatch between the acquired action representation of the object and the displayed action (van Elk et al., 2008a). Therefore, the interference effects of end location- and grip-violations likely reflect that subjects who learned to use a novel object with a clear action effect acquired a strong action representation about the correct use of an object. In the second experiment we assessed the relative importance of action effects for the acquisition of novel action representations. In line with the first experiment, it was found that after training subjects responded slower if the object was presented at an incorrect compared to a correct end location. Thereby the second experiment underlines the importance of knowledge about end locations for the acquisition of novel action representations, even if the objects do not have functional significance.

Interestingly and in contrast to the first experiment, in the second experiment observing the object being grasped with an incorrect grip did not interfere with object identification. It could be that in the first experiment the grip applied to the object was more important, as the object only elicited an action effect, when grasped in an appropriate way. Therefore, the training with functional objects may have resulted in the acquisition of action semantics, specifying both the end location and the grip required for grasping the object. Previous studies have used the term 'action semantics' to refer to the action representations that enable us to use objects in a meaningful way (Hodges, Spatt, & Patterson, 1999; Noppeney, 2008; Springer & Prinz, 2010; van Elk et al., 2008b, 2009; Vingerhoets, Vandekerckhove, Honore, Vandemaele, & Achten, 2010). In these cases,

the term 'semantics' was used because it captures the fact that this action knowledge does not consist of arbitrary sensorimotor associations, but is related to the meaningful use of objects. In other words: although at a neural level action knowledge may be correctly described in terms of specific sensorimotor associations, at a behavioral level such a description does not capture the fact that these associations are behaviorally meaningful. For instance, we bring cups to our mouths in order to drink and as a consequence the association between cup and mouth is not arbitrary but reflects a behaviorally meaningful action. Accordingly, with respect to the present study it seems warranted to suggest that in the first experiment subjects acquired novel action semantics, as the learned actions were behaviorally meaningful. In contrast, in the second experiment the relation between the way in which the object was grasped and the end location was more arbitrary and accordingly grip-related information was less well established than in the first experiment.

At least two possible functional mechanisms could underlie the influence of action effects on the acquisition of novel action representations, namely the rationality principle and the acquisition of modality-specific representations. First, it could be that by assessing the efficiency of an action in achieving a particular action effect people pay special attention to the correct grip. Several studies have shown that people assess the rationality of the specific grip chosen to perform an action with respect to the given end location, which is known as the efficiency principle (Csibra & Gergely, 2007). Accordingly, by estimating the efficiency of a particular grip in relation to the object's function subjects may incorporate the specific grip information into their representation of the object.

Secondly, it could be that the functional significance of an object (i.e. the object can be used for hearing or smelling) resulted in a deeper processing of the object properties through additional modality-specific representations, which are represented in a distributed neural architecture (Allport, 1985; Martin, 2007). In the language domain classical studies concerning the learning of words have shown that the processing of meaningful information facilitates the retrieval of this information (Craik & Lockhart, 1972; Craik & Tulving, 1975). Although speculative, one intriguing possibility is that training with objects that elicit an action effect equips the object with functional significance and results in a richer semantic representation, including both auditory and olfactory features (cf. Allport, 1985; Martin, 2007). These modality-specific representations may in turn facilitate the retrieval of action representations during object recognition (cf. Siakaluk, Pexman, Aguilera, Owen, & Sears, 2008). Although at present the precise neural mechanisms underlying the acquisition of novel action representations remain unclear this suggestion opens interesting possibilities for future research.

An intriguing question is how crucial motor experience is for the acquisition of novel action representations. Although the present study did not directly address this issue, it must be pointed out that subjects did not show the same advantage for correct compared to incorrect end locations for untrained objects that were only perceptually presented during the training phase. Furthermore, in a recent study the importance of motor experience for the acquisition of novel object knowledge was directly demonstrated (Paulus, Lindemann, & Bekkering, 2009). When subjects performed a concurrent motor task during the learning of novel object-end location associations, no subsequent advantage was observed when objects were presented at a correct compared to an incorrect end location. Finally, several studies indicate that only the active but not the passive use of a tool results in the tool being incorporated in the body schema, thereby further underlining the central importance of action experience for novel tool use (Maravita & Iriki, 2004; Maravita, Spence, Kennett, & Driver, 2002).

At a perceptual level, the present study shows that previous action experience with an object facilitates object recognition, when the object is presented at the expected end location. Related to this issue, a recurring discussion regarding the hierarchical organization of the motor system concerns the question whether at a perceptual level the dominance of end locations over grips does not merely reflect an advantage of global over local visual processing. Information about the end location is spatially defined, whereas grip-related information is more locally defined. Thus, information regarding the grip can only be inferred after attention has been directed to the relevant end location. In addition, visual information regarding the grip may be more difficult to perceptually discern and remember than information about the end location. Although we acknowledge that the dominance of end locations over grips can be described at different levels, we would like to point out that an attentional or memory explanation cannot fully account for the findings in this and other studies. First, in the present study participants were primed with a picture of a single object and accordingly were able to retrieve information about both the end location and the grip applied to the object, well before the presentation of the action snapshot. Second, even though the grip applied to the object was made perceptually more salient (e.g. applying different colors to different graspable parts of the object), still information about the end location had a stronger effect on performance than grip-related information. Finally, in a recent study on action planning, using a simple reach and grasp task, a similar dominance of end locations over grips was observed as in the present study (van Schie, van Elk, Bekkering, *in preparation*), thereby providing further support for the notion that the dominance of goal- or end locations over grips reflects a fundamental principle of the human motor system, rather than a by-product of attentional processing.

A related issue when it comes to the hierarchical view of the motor system is whether the observed dominance of end locations over grips truly reflects a hierarchical relation or is a mere necessity following from the temporal relation between grips and end locations during action planning and execution. By definition end locations and grips are reciprocally related: once an end location is selected the required grip is automatically determined and vice versa: once the grip is selected the possible range of end locations is restricted. This reciprocity between end locations and grips was exploited in a recent study on action planning (van Schie et al., *in preparation*). In this experiment we employed a simple reach- and grasp-paradigm, whereby the end location and the grip of the movement concurred (e.g. movement towards the left object always entailed a power grip; movement towards the right object always entailed a precision grip). Participants were instructed to perform the movement either based on grip-related information (power vs. precision grip) or on information about the end location

(grasp the left vs. the right object). It was found that action planning based on information about the end location was faster, probably because the grip required for grasping the object could be determined on the fly. In contrast, planning actions based on information about the grip was slower, because this entailed an additional selection process (i.e. determining which object should be grasped). This experiment shows a more efficient process of action planning if the grip is determined by the end location than if end location is determined by the grip. Typically, in our everyday actions the grip is indeed determined by the end location (i.e. grasping a wine bottle to place it in a wine cooler requires a different grip than grasping a wine bottle to pour a glass of wine). Whether this relationship is described as a 'natural dependency' or a 'hierarchical relationship' is a matter of theoretical debate, but at least it shows that the default mode of action planning is the case where end locations determine the selection of grips.

Previous studies on the acquisition of novel object semantics have typically investigated how at a perceptual level the categorization of novel objects is facilitated by perceptual features of the objects (e.g. Desmarais, Dixon, & Roy, 2007; Vuong & Tarr, 2006) or by expertise visual training with the objects (Behrmann, Marotta, Gauthier, Tarr, & McKeef, 2005; Tarr & Cheng, 2003). However, relatively little is known about the effects of direct action experience on object recognition. Interestingly, one recent study suggests that pairing novel objects with specific actions (e.g. pull, slide or twist) may facilitate the ease with which these objects can be identified (Desmarais et al., 2007). This effect is likely mediated by the acquisition of motor representations that become associated with the object and that in turn facilitate object recognition. Preliminary evidence for the involvement of motor representations in the acquisition of novel objects was found in a study by Kiefer, Sim, Liebich, Hauk, and Tanaka (2007). In the training phase subjects made either a pointing movement towards a new 2D object or made an action pantomime regarding a detail feature of the object that afforded a specific action (e.g. pantomime grasping the object at the handle). After several weeks of training subjects conducted a categorization task while their EEG was recorded. Interestingly, an early activation in frontal motor regions and a later activation in occipito-parietal visual-motor areas was found for subjects in the pantomime compared to the pointing condition, suggesting that the pantomime training established visuo-motor representations regarding the object's use whereas simple pointing did not. Further evidence for the role of visuo-motor representations in the acquisition of novel functional object representations was found in a recent study in which subjects verbally acquired functional knowledge about novel objects (Paulus et al., 2009). Only when subjects performed a concurrent motor task during the training phase, but not when performing another attention demanding task, subjects performed worse on the subsequent recognition of an actor using objects in the correct fashion. Together these studies suggest that learning to use novel objects relies on the acquisition of motor representations reflecting the end location related to the object's typical use.

5. Conclusion

The aim of the present study was to investigate if the dominance of end locations over grips, previously observed for actions with well-known objects, applies to the learning of novel actions as well. It was found that after training with novel objects, subjects acquired an action representation specifying the end location towards which the objects were directed. Thereby the present study suggests that the hierarchical view of the motor system reflects a general principle, underlying the learning of novel object use as well.

Acknowledgments

This work was supported by the Netherlands Organization for Scientific Research (Grant 453-05-001 awarded to HB) and by the Marie Curie Intra European Fellowship within the Seventh European Community Framework Program (IEF Grant 252713 to MVE).

References

- Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 32–60). Edinburgh: Churchill Livingstone.
- Barrett, T. M., Davis, E. F., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology*, *43*(2), 352–368.
- Behrmann, M., Marotta, J., Gauthier, I., Tarr, M. J., & McKeef, T. J. (2005). Behavioral change and its neural correlates in visual agnosia after expertise training. *Journal of Cognitive Neuroscience*, *17*(4), 554–568.
- Bourgeois, K. S., Khawar, A. W., Neal, S. A., & Lockman, J. J. (2005). Infant manual exploration of objects, surfaces, and their interrelations. *Infancy*, *8*(3), 233–252.
- 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.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294.
- Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*(1), 60–78.
- Desmarais, G., Dixon, M. J., & Roy, E. A. (2007). A role for action knowledge in visual object identification. *Memory & Cognition*, *35*(7), 1712–1723.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, *308*(5722), 662–667.
- Grafton, S. T., & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Sciences*, *26*(4), 590–616.
- Hodges, J. R., Spatt, J., & Patterson, K. (1999). "What" and "how": Evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proceedings of the National Academy of Science of the United States of America*, *96*(16), 9444–9448.

- Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, 59(3), 176–186.
- Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5), 849–878. discussion 878–937.
- Kiefer, M., Sim, E. J., Liebich, S., Hauk, O., & Tanaka, J. (2007). Experience-dependent plasticity of conceptual representations in human sensory-motor areas. *Journal of Cognitive Neuroscience*, 19(3), 525–542.
- Kunde, W., Hoffmann, J., & Zellmann, P. (2002). The impact of anticipated action effects on action planning. *Acta Psychologica (Amst)*, 109(2), 137–155.
- Lindemann, O., Stenneken, P., van Schie, H. T., & Bekkering, H. (2006). Semantic activation in action planning. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 633–643.
- Majdandzic, J., Grol, M. J., van Schie, H. T., Verhagen, L., Toni, I., & Bekkering, H. (2007). The role of immediate and final goals in action planning: An fMRI study. *Neuroimage*, 37(2), 589–598.
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, 8(2), 79–86.
- Maravita, A., Spence, C., Kennett, S., & Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition*, 83(2), B25–34.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology*, 35(4), 1091–1101.
- Noppeney, U. (2008). The neural systems of tool and action semantics: A perspective from functional imaging. *Journal of Physiology Paris*, 102(1–3), 40–49.
- Paulus, M., Hunnius, S., Vissers, M., Bekkering, H. (in press). Bridging the gap between the other and me: The functional role of motor resonance and action effects in infants' imitation. *Developmental Science*. doi:10.1111/j.1467-7687.2011.01040.x.
- Paulus, M., Lindemann, O., & Bekkering, H. (2009). Motor simulation in verbal knowledge acquisition. *Quarterly Journal of Experimental Psychology (Colchester)*, 1, 8.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9(2), 129–154.
- Rosenbaum Vaughan, J., Barnes, H. J., & Jorgensen, M. J. (1992). Time course of movement planning: Selection of handgrips for object manipulation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 18(5), 1058–1073.
- Siakaluk, P. D., Pexman, P. M., Aguilera, L., Owen, W. J., & Sears, C. R. (2008). Evidence for the activation of sensorimotor information during visual word recognition: The body-object interaction effect. *Cognition*, 106(1), 433–443.
- Springer, A., & Prinz, W. (2010). Action semantics modulate action prediction. *Quarterly Journal of Experimental Psychology (Colchester)*, 63(11), 2141–2158.
- Tarr, M. J., & Cheng, Y. D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences*, 7(1), 23–30.
- van Elk, M., van Schie, H. T., & Bekkering, H. (2008a). Conceptual knowledge for understanding other's actions is organized primarily around action goals. *Experimental Brain Research*, 189(1), 99–107.
- van Elk, M., van Schie, H. T., & Bekkering, H. (2008b). Semantics in action: An electrophysiological study on the use of semantic knowledge for action. *Journal of Physiology Paris*, 102(1–3), 95–100.
- van Elk, M., van Schie, H. T., & Bekkering, H. (2009). Action semantic knowledge about objects is supported by functional motor activation. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4), 1118–1128.
- van Schie, H. T., & Bekkering, H. (2007). Neural mechanisms underlying immediate and final action goals in object use reflected by slow wave brain potentials. *Brain Research*, 1148, 183–197.
- van Schie, van Elk & Bekkering (in prep). What makes a goal a goal? Dominance of goals over grips reflect the importance of spatial features in action planning.
- Vingerhoets, G., Vandekerckhove, E., Honore, P., Vandemaële, P., & Achten, E. (2010). Neural correlates of pantomiming familiar and unfamiliar tools: Action semantics versus mechanical problem solving? *Human Brain Mapping*, 32(6), 905–918.
- Vuong, Q. C., & Tarr, M. J. (2006). Structural similarity and spatiotemporal noise effects on learning dynamic novel objects. *Perception*, 35(4), 497–510.