Note

Controlling attention through action: Observing actions primes action-related stimulus dimensions

Sabrina Fagioli [a,*], Fabio Ferlazzo [a], Bernhard Hommel [b]

[a] *University of Rome "La Sapienza", Department of Psychology, Cognitive Ergonomics Laboratory, Via dei Marsi, 78, 00185 Rome, Italy*
[b] *Leiden University, Institute for Psychological Research & Leiden Institute for Brain and Cognition, The Netherlands*

Abstract

Previous findings suggest that planning an action "backward-primes" perceptual dimension related to this action: planning a grasp facilitates the processing of visual size information, while planning a reach facilitates the processing of location information. Here we show that dimensional priming of perception through action occurs even in the absence of active action planning. Subjects watched video clips showing a grasping or reaching action before detecting size- or location-defined deviants in visual stimulus sequences. Size deviants were detected faster after seeing a grasp and location deviants were detected faster after seeing a reach. This supports the assumption that perceptual codes and action plans share a common representational medium, and that "attention to action" controls "attention to stimuli".
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Action observation; Visuo-motor priming; Action planning; Visual attention; Dimensional priming; Mirror system

In order to interact successfully with objects and individuals that crowd our environment, we need to continuously select both the appropriate response to the perceptual events we face and the relevant sensory information for accomplishing that act. However, how the perceptual and action information is coded to be promptly available is still a matter of debate. Recently, the hypothesis suggesting a common-coding system for both action and perception (Hommel, Müsseler, Aschersleben, & Prinz, 2001a) has garnered increasing support. This hypothesis challenges the classic assumption that stimuli and responses are separate and hierarchically distributed entities, and rather asserts they are coded as *event*-related features in a common format (i.e., referring to the distal features of the external event).

Given the functional equivalence of perception- and action-related codes, bi-directional priming effects are predicted between the representations of stimuli and actions as a function of the number of feature codes they share (Müsseler & Hommel, 1997; Prinz, 1990, 1997). In line with these predictions, recent behavioral and neurophysiological evidence shows not only that cognitive processes (underlying perception, attention, and mem-

ory) prime the selection of feature-overlapping actions, but also that planning a task and an action affects perceptual processes. For instance, selecting and processing visual stimuli is modulated by task requirements and the actions those stimuli afford (Fisher & Hoellen, 2004; Hannus, Cornelissen, Lindemann, & Bekkering, 2005; Memelink & Hommel, 2006; Tucker & Ellis, 1998), as well as by the type of action the perceiver is currently performing or observing (Bekkering & Neggers, 2002; Bernstein & Cooper, 1997; Craighero, Bello, Fadiga, & Rizzolatti, 2002; Miall et al., 2006; Vogt, Taylor, & Hopkins, 2003; Wohlschläger, 2000). Along the same lines, experiments using positron emission tomography (PET) have revealed increased activity of cortical areas known for being involved in action generation, when subjects discriminate objects compared to non-objects (Grèzes & Decety, 2002).

Further support for the common-coding hypothesis comes from studies addressing the mirror neuron system. Mirror neurons have been described as a special category of visuo-motor neurons, discharging for both performed or observed actions (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Activity of such neurons has been recorded in the premotor cortex of the monkey, but there is increasing evidence for the existence of an analogous system in humans (Iacoboni et al., 1999; see Rizzolatti &

---

* Corresponding author. Tel.: +39 06 49917663; fax: +39 06 4451667.
E-mail address: sabrina.fagioli@uniroma1.it (S. Fagioli).

Craighero, 2004 for a review). Taken together, the neural properties of these cells and the behavioral correlates suggest a dynamic relationship between action and perception, supporting the idea that perceptual and motor features referring to perceived and planned events are coded in the same cognitive format.

To date, most of the few studies on action-perception effects addressed how action modulates the visual discrimination and selection of the action related-features (affordances) *of a particular object*—i.e., of the object the action was targeted at. However, in a recent study, Fagioli, Hommel, and Schubotz (2007) demonstrated that action planning can affect perceptual processing in a more general way. Following up the previous, rather surprising observation in fMRI studies that premotor areas are involved in perceptual tasks (Schubotz & von Cramon, 2003), Fagioli and coworkers found evidence that planning an action does not only prime action-related features belonging to the object of that action but that whole feature dimensions get primed. For instance, preparing for a manual grasp facilitated the visual discrimination of size, while preparing for a reach facilitated the discrimination of spatial location. That is, actions of a particular type seem to "backward prime" feature dimensions that are particularly relevant[1] for these actions, in the sense that the features defined on that dimension specify relevant action parameters. Preparing for a particular type of action thus seems to involve some sort of executive control of visual attention by increasing the weight of action-related perceptual dimensions in stimulus processing (Logan & Gordon, 2001).

Even though the findings of Fagioli et al. (2007) provide evidence that preparing for an action somehow goes along with preparing for the processing of particular stimulus features, these observations do not necessarily require the assumption of the common coding of perceptual and action events. Instead, some separate executive control systems may be responsible for both planning and preparing a particular action and priming the corresponding perceptual dimension(s), much along the lines of Logan and Gordon (2001) ECTVA model of executive control. The aim of the present study therefore was to test whether the backward priming of action-relevant perceptual dimensions can only be observed if people voluntarily prepare for the given action (a situation that is likely to involve possible executive control processes) or whether such priming effects also occur if this action is involuntarily primed (a situation that is unlikely to engage executive control). As mentioned above, premotor action plans are activated through the mere observation of actions, even in the absence of any intention to imitate them (Buccino et al., 2001; Grèzes, Costes, & Decety, 1998). We therefore presented subjects with task-irrelevant film-clips showing a grasping or a reaching action, assuming that this would activate the corresponding action plans in the premotor cortex. If so, and if

activating an action plan is sufficient to backward prime the related perceptual dimensions, this should facilitate the processing of stimulus features falling on these dimensions. That is, we expected that observing a grasping action would facilitate the processing of visual size information, whereas observing a reaching action would facilitate the processing of visual location information.

1. Method

1.1. Participants

Thirteen students (8 males) aged 20–26 voluntarily participated for class credits. Participants were right-handed with normal or corrected to normal vision, and were naive as to the purposes of the experiment.

1.2. Apparatus and stimuli

Participants sat in a dimly lit room facing a 21 in. monitor (Silicon Graphics 550, 800 × 600 pixels, 32 bit color) at a distance of 60 cm. Stimulus presentation and response recording were controlled through the Cogent 2000 toolbox running under Matlab 6.5.

A white asterisk displayed at the geometrical centre of the screen on a black background served as fixation point. Two short 3-s video clips captured through a digital camera (30 frames per second, 24 bit, 320 × 240 pixel) were used as visual primes. The clips showed an actor grasping a cube with her right hand (grasping action) or the same actor reaching-to-touch a dot with her right finger tip (reaching action; see Fig. 1). Both actions were filmed from a frontal view.

A series of seven yellow circles on a black background served as stimuli for the visual discrimination task. On each trial, they were successively displayed at a rate of 600 ms without temporal gaps, along one of the two main diagonal (x/y) axes of the computer screen (at 100/525, 200/450, 300/375, 400/300, 500/225, 600/150, 700/75, and 100/75, 200/150, 300/225, 400/300, 500/375, 600/450, 700/525 pixels, respectively) starting from either the top or the bottom of the screen. The size of the circles alternated from "small" (0.7 cm in diameter, 0.7° of visual angle) to "large" (1.3 cm in diameter, 1.2° of visual angle), or vice versa. Responses were given by pressing a foot pedal placed under the table, on the right side relative to the subject's midline.

1.3. Procedure and design

On each trial, the fixation point appeared for 1 s. It was immediately followed by one of the video clips, which subjects were told to watch passively. Next, one sequence of visual stimuli was presented for the visual discrimination task. On each sequence, seven pictures (yellow circles) were successively displayed along one of the main diagonal axes of the screen. The stimuli alternated their size (small and large) and were presented at the same distance with respect to each other. Both the starting point (top or bottom) and the alternating pattern (small–large) of the stimuli were balanced across trials in order to prevent any bias due to the specific combinations of the stimulus size and stimulus location on the screen (Fig. 2).

Participants were asked to monitor the sequences and to detect any violation of their sequential order. 75% of the sequences contained a size- or location-deviant stimulus, with the size or the spatial location of the fourth, the fifth or the sixth stimulus being repeated rather than followed by the next element of the sequence. Participants were instructed to press the foot pedal as soon as possible whenever they saw the deviant stimulus. No response was required on the regular sequences (25% of the trials). To prevent cognitive strategies as far as possible, size and location deviants occurred randomly intermixed rather than blocked (cf. Fagioli et al., 2007).

Four experimental conditions were produced by combining the observed action (grasping versus reaching) and the stimulus dimension (size versus location) of the deviant sequences. Participants completed 112 trials equally distributed in two blocks of trials.

---

[1] Note that the findings of Fagioli et al. (2007), as well as the present findings, allow conclusions only with respect to the *relative* relevance of perceptual dimensions, and with respect to *relative* facilitation. It makes sense to assume that location information is not entirely irrelevant for grasping and size information is not entirely irrelevant for reaching, so all we can (and want to) say is that for grasping, size is more important than location while the opposite is true for reaching.
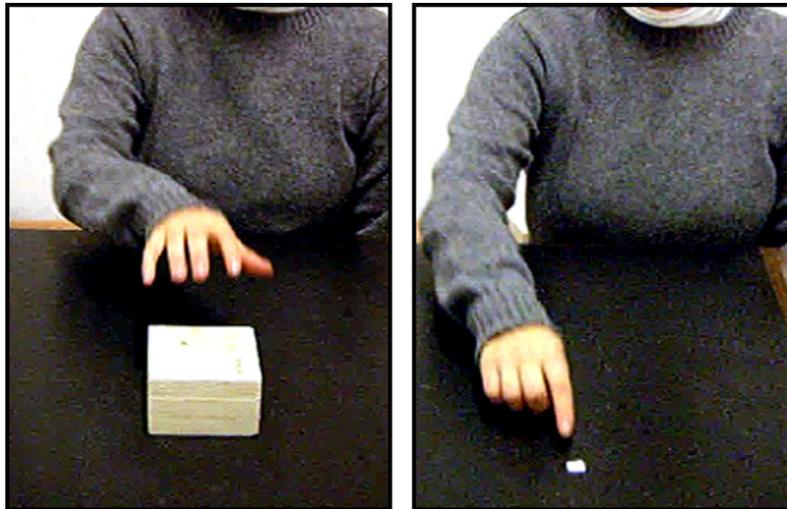
Fig. 1. Frames extracted from the video clips respectively showing the grasping (left frame) and the reaching (right frame) action.

2. Results

Anticipations (reaction times, RTs < 100 ms), missing responses (RTs > 1000 ms) and incorrect responses were considered as errors and excluded from analysis. Mean RTs were computed for each experimental condition and fed into a 2×2 ANOVA, considering action (grasping versus reaching) and stimulus dimension (size versus location) as within-subjects factors. The Mauchley Sphericity test performed on the mean RTs did not show any significant effect, providing evidence that the homoscedasticity assumptions have been met. The analysis revealed a significant two-way interaction ($F(1,12) = 16.53$; $p < .005$), whereas no significant main effect of action and stimulus dimension factors was found ($Fs < 1$). Duncan test confirmed that when participants observed a grasping action they were faster in discriminating size than location deviants (630 versus 650 ms, respectively; $p < .05$; Table 1, upper panel); but when they observed a reaching action they were faster detecting location than size deviants (636 versus 658 ms, respectively; $p < .05$; Table 1, upper panel). An analogous analysis performed on the error rates (incorrect responses, on average 23%) revealed only a main effect of the stimulus dimension factor ($F(1,12) = 5.83$; $p < .05$), with participants being less accurate in discriminating

location than size deviants (5.4% versus 3.1%, respectively). A correlation analysis with mean RTs and mean accuracy did not show any significant correlation ($r = -.53, p > .05$). Hence, there was no sign of a speed-accuracy trade-off.

One might object that our results could be due to a purely perceptual priming effect: location and size discrimination might not have been primed by the actions shown in the video clips

Table 1
The upper panel reports mean reaction times (RTs), standard deviations (S.D.) and error percentages (%) of the main experiment as a function of the observed action and the dimension of the deviant stimulus

Prime	Stimulus dimension		Δ
	Size	Location	
Experiment			
Reaching			
RTs	658	636	21
S.D.	64.19	89.73	
%	11.54	20.88	
Grasping			
RTs	630	650	−20
S.D.	53.41	79.07	
%	10.71	17.58	
Control			
Small object			
RTs	639	606	
S.D.	75.71	96.38	
%	2.92	8.75	
Large object			
RTs	636	618	
S.D.	65.02	105.57	
%	4.17	12.08	

The last column shows the dimensional priming effect (size–location). The lower panel shows reaction times (RTs), standard deviations (S.D.) and error percentages (%) for the control experiment as a function of the observed object and the dimension of the deviant stimulus.
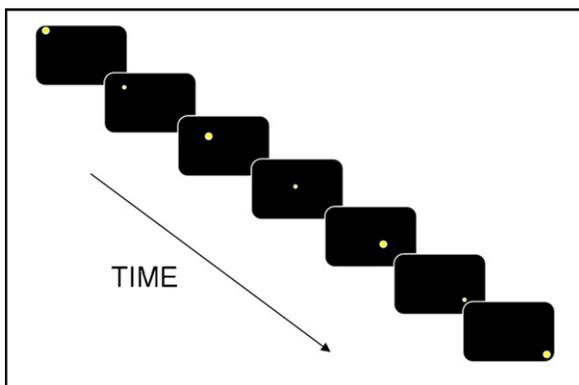

Fig. 2. Example of a regular sequence in the visual discrimination task.

but, rather, by the (small and large) objects these actions aimed at. It is true that perceptual priming is unlikely to account for the whole outcome pattern, as the relationship between object size and actions was random: the reaching video clip (depicting a small object) was randomly followed by both small and large circles during the discrimination task, and the grasping video clip (depicting a large object) was randomly followed by both small and large circles. Nevertheless, to rule out any possible confound we ran a control experiment. It was identical to the first, except that the two video clips were replaced by two still pictures in which only the target object (small or large) – but not the action – was shown. The pictures were extracted as frames from the video clips, and processed in order to erase the actor's arm. The mean reaction times of 12 subjects were analyzed by using the same ANOVA design, with size of object (small versus large) and stimulus dimension (size versus location) as within-subjects factors. This analysis did not show any significant effect or interaction (all $Fs < 1$), ruling out the purely perceptual-priming account. The means and the standard deviations for this control experiment are reported in Table 1 (lower panel).

3. Discussion

Our findings suggest that perceptual feature dimensions can be primed not only through the active preparation of actions that are related to these dimensions (Fagioli et al., 2007) but also as a consequence of merely watching such actions. This provides strong evidence for the idea that activating an action plan is sufficient to backward prime action-related perceptual dimensions. Apparently, then, activating an action representation – be that voluntarily, as in the process of planning an action, or involuntarily, as when observing an action – involves the activation of information about perceptual dimensions that specify parameters of the action. Grasping an object relies on information about the size of the object, so to control the aperture of the hand when approaching it, whereas reaching is mainly driven by location information. Accordingly, biasing attention towards the size dimension when intending to grasp and towards location when intending to reach makes sure that the cognitive system is optimally prepared to process the information needed to control the action.

Even though our findings do not require this assumption, they are consistent with dual route models of action control that distinguish between (1) a (possibly ventral) route that is responsible for the off-line preparation of actions and the recruitment of the feed-forward components of action plans and (2) a (possibly dorsal) feedback-loop that is feeding on-line information into the system to specify the open parameters of the action plan (Glover, 2004; Hommel, Müsseler, Aschersleben, & Prinz, 2001b). According to the logic of this model, it would be the responsibility of the off-line system to specify the relative relevance of the information processed by the on-line loop. In other words, higher-level action planning might control sensorimotor processing through implementing a particular attentional set.

The remaining question is how premotor action control systems could acquire information about suitable perceptual dimensions—how do we get to know that size is important for grasping and location for reaching? Considering the evidence for the common coding of perceptual and action events provides a possible answer to this question. The general idea underlying the common coding principle is that actions are selected and planned not on the basis of proximal motor output but in terms of their (previously acquired and therefore expected) perceptual consequences (Elsner & Hommel, 2001; James, 1890; Lotze, 1852; see Hommel, 1997, 2006). If so, the representation of a grasp is likely to contain information about the hand posture assumed at the end of successful grasps (Rosenbaum, Meulenbroek, Vaughan, & Jansen, 2001); these postures are scaled to object size and, therefore, implicit representations of this feature. Accordingly, selecting a particular grasping action requires particular attention to the size dimension, because this is the most important dimension on which the available grasping representations will differ. If "attention to action" (i.e., response selection) needs to focus on the size dimension anyway, and if perceptual events and action plans are coded in the same representational space, it makes sense to assume that processing perceptual input will be biased in a similar fashion as response related processes—hence, "attention to action" controls "attention to stimuli", without any need for a separate executive controller (as suggested by Logan & Gordon, 2001). The present findings suggest that this is true irrespective of whether the attention to action is induced endogenously (as in the study of Fagioli et al., 2007) or exogenously (as through watching an action).

References

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, *13*, 370–374.

Bernstein, L. J., & Cooper, L. A. (1997). Direction of motion influences perceptual identification of ambiguous figures. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 721–737.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

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

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 229–240.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. *Psychological Research*, *71*, 22–29.

Fisher, M. H., & Hoellen, N. (2004). Space- and object-based attention depend on motor intention. *The Journal of General Psychology*, *13*, 365–377.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Glover, S. (2004). Separate visual representations in the planning and control of action. *Behavioral and Brain Sciences*, *27*, 3–78.

Grèzes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*, 553–582.

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, *40*, 212–222.

Hannus, A., Cornelissen, F. W., Lindemann, O., & Bekkering, H. (2005). Selection-for-action in visual search. *Acta Psychologica*, *118*, 171–191.

Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility. *Psychological Research*, *59*, 248–260.

Hommel, B. (2006). How we do what we want: A neuro-cognitive perspective on human action planning. In R. J. Jorna, W. van Wezel, & A. Meystel (Eds.), *Planning in intelligent systems: Aspects, motivations and methods* (pp. 27–56). New York: John Wiley & Sons.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001a). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001b). Codes and their vicissitudes. *Behavioral and Brain Sciences*, *24*, 910–937.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

James, W. (1890/1981). *The principles of psychology*. Macmillan/Harvard University Press (Original work published 1890).

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*, 393–434.

Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Memelink, J., & Hommel, B. (2006). Tailoring perception and action to the task at hand. *European Journal of Cognitive Psychology*, *18*, 579–592.

Miall, R. C., Stanley, J., Todhunter, S., Levick, C., Lindo, S., & Miall, J. D. (2006). Performing hand actions assists the visual discrimination of similar hand postures. *Neuropsychologia*, *44*, 966–976.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 861–872.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Berlin: Springer Verlag.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neurosciences*, *27*, 169–192.

Rosenbaum, D. A., Meulenbroek, R. G., Vaughan, J., & Jansen, C. (2001). Posture-based motion planning: Applications to grasping. *Psychological Review*, *108*, 709–734.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *Neuroimage, 20*, S120–S131.

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*, 830–846.

Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia*, *41*, 941–951.

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research*, *40*, 925–930.