Cross-Modal Interactions between Olfaction and Vision When Grasping

Umberto Castiello¹,2, Gesualdo M. Zucco¹, Valentina Parma¹, Caterina Ansuini¹ and Roberto Tirindelli³

¹Department of General Psychology, University of Padua, Via Venezia 8, 35131 Padua, Italy, ²Department of Psychology, Royal Holloway, University of London, London, UK and ³Department of Neuroscience, University of Parma, Parma, Italy

Correspondence to be sent to: Umberto Castiello, Department of General Psychology, University of Padua, Via Venezia 8, 35131 Padua, Italy. e-mail: umberto.castiello@unipd.it

Abstract

This study used kinematics to investigate the integration between vision and olfaction during grasping movements. Participants were requested to smell an odorant and then grasp an object presented in central vision. The results indicate that if the target was small (e.g., a strawberry), the time and amplitude of maximum hand aperture were later and greater, respectively, when the odor evoked a larger object (e.g., an orange) than when the odor evoked an object of a similar size as the target or no odor was presented. Conversely, the time and amplitude of maximum hand aperture were earlier and reduced, respectively, when the target was large (e.g., a peach) and the odor evoked a smaller sized object (e.g., an almond) than when the odor evoked an object of a similar size as the target or no odor was presented. Taken together, these results support the evidence of cross-modal links between olfaction and vision and extend this notion to goal-directed actions.

Key words: cross-modal interactions, kinematics, olfaction, reach-to-grasp, vision

Introduction

Many events in everyday life are registered by the sense organs of more than one modality. Consequently, the coordination and integration of information derived from different sensory systems are essential for providing a unified perception of our environment to control actions within it. The investigation of multisensory coding during natural actions is still in its infancy, but recent studies seem to suggest that cross-modal links in motor control are potentially numerous and substantial (Gentilucci et al., 1998; Patchay et al., 2003, 2005). In these experiments, participants reached and grasped with one hand a visual target of different sizes while holding another unseen object (distractor) of different size in the other hand. Results indicated that proprioceptively guided manipulation with one hand influenced finger shaping of visuomotor grasping of the other hand when the two objects differed in size. Maximum grip aperture (the maximum distance between the tips of the index finger and thumb) of a visually guided reach to grasp was proportional to the diameter of a nonvisible distractor object manipulated proprioceptively with the other hand. This effect was the first demonstration of a cross-modal interference effect linked specifically to interaction with objects.

The aforementioned interference effects were partly explained in terms of action-based attentional mechanisms that may serve to select the target from competing distractors (Tipper et al., 1998; Castiello, 1999). In these terms, target and distractor both evoke grasping representations that interact in a mutually suppressive or competitive way. Interference is thus the result of the competition between the target and the potential distractors’ action representation in whatever modality they are presented.

To date, “action” cross-modal interference research has focused on links between vision and proprioception. No studies have examined the nature of such cross-modal interference effects involving other sensory modalities such as olfaction (for a review, see Doty, 2001). Indeed, a large body of literature suggests that odors can influence behavioral performance. These influences modulate simple reaction times to auditory and visual stimuli (Zucco and Tressoldi, 1989; Millot et al., 2002), voice pitch (Millot et al., 2002), airflow motor control (Bensafi et al., 2003), emotional behavior (Herz and Cupchik, 1995), memory (Zucco, 2003), and attention (Spence et al., 2001).

Furthermore, previous evidence has revealed that olfactory information interacts with visual information, contributing
to our rich multisensory experience in the environment (for review, see Verhagen and Engelen, 2006). For example, in one study (Gottfried and Dolan, 2003), subjects participated in an olfactory detection task whereby odors and pictures were delivered separately or together. The odor–picture combinations could be semantically congruent or incongruent. The task for the subjects was to decide whether they smelled an odor or not. Results showed that detection, intensity, and ratings were enhanced when an odor was presented together with a congruent picture in comparison to a non-matching picture (Gottfried and Dolan, 2003), which suggests the presence of a mutual interaction between visual and olfactory perceptions.

Therefore, in the present study, we investigated the nature of multisensory interaction between olfactory cues and vision during the performance of a grasping action toward a visual stimulus. Specifically, we administered an odor pertaining to an object requiring a hand aperture that could be similar in size to or different in size from that required by the visual target. On the basis of previous accounts of the cross-modal interference effects of grasping, we predict that if the internal representation of the object evoked by the odor is associated with a type of hand aperture that differs from that associated with the target, then interference effects should be evident on grasping kinematics. As hypothesized, we found that the size of the object evoked by the odorant influenced the subsequent grasping kinematics.

Materials and methods

Subjects

Sixteen subjects (12 females and 4 males, aged 20–29 years, mean age 24 years) took part in the experiment. All subjects were right handed and reported normal olfactory function, no history of olfactory dysfunction, and normal or corrected-to-normal vision in a confidential report. All subjects were naive as to the purpose of the experiment. All subjects gave informed consent to participate in the study. The experimental procedures were approved by the Institutional Review Board at the University of Padova and were in accordance with the Declaration of Helsinki. The experimental session lasted approximately 30 min.

Stimuli and design

The target stimuli consisted of six plastic objects grouped on the basis of the maximum hand aperture they naturally require (if not grasped by the stem): large (apple, orange, and peach; see Figure 1A, top panel) and small (almond, garlic, and strawberry; see Figure 1A, bottom panel). Plastic objects were used so as to maintain their visual attributes and size similar throughout the entire period of experimentation. The size of the stimuli within each of the two “maximum hand aperture” categories was broadly similar.

Six olfactory stimuli corresponding to the visual stimuli described above were administered. The stimuli were extracted by the Sniffin sticks olfactory kit test (a validated and commonly utilized test), based on penlike odor-dispersing devices. The test allows the assessment of odor detection, identification, and discrimination (Kobal et al., 1996). Each odorant was kept by the experimenter at approximately 2 cm from both nostrils for all subjects. The stimuli were delivered in a well-ventilated room. All subjects underwent an odor familiarization test that was administered before the experimental session.

The visual/olfactory stimuli combinations (see Table 1) produced six experimental conditions: 1) congruent small condition, in which the odor may evoke the representation of an object requiring a type of small hand aperture similar to that for the target; 2) congruent large condition, in which the odor evoked the representation of an object requiring a type of large hand aperture similar to that for the target; 3) incongruent small condition, in which the odor evoked the representation of an object requiring a type of hand aperture smaller than that required by the target; 4) incongruent large condition, in which the odor evoked the representation of an object requiring a type of hand aperture larger than that required by the target; 5) small control condition, in which the subjects grasped the small target but no odor was presented; and 6) large control condition, in which the subjects grasped the large target but no odor was presented.

Apparatus and procedure

Hand kinematics were measured by means of resistive sensors embedded in a glove (CyberGlove, Virtual Technologies, Palo Alto, CA). In particular, we analyzed the data collected from the sensors located on the distal phalanx of the index finger, the first carpometacarpal joint and on the distal phalanx of the thumb. The sensor’s linearity was 0.62% of maximum nonlinearity over the full range of hand motion. The sensor’s resolution was 0.5° that remains constant over the entire range of joint motion. Sampling occurred at 12-ms intervals.

At the start of the procedure, participants had the right hand (wearing the CyberGlove) in a pronated position with the palm pressing a switch (Figure 1B). The sequence of events was as follows: 1) vision was occluded before the target was positioned on the working surface; 2) an auditory tone was presented, the experimenter presented the odorant, and the subjects were instructed to smell it; 3) as soon as a 2-s time period elapsed, a subsequent tone indicated to the experimenter to remove the odorant; 4) after a variable interval (500–1000 ms), an auditory tone, which was distinguishable from the previous tones, was presented; and 5) upon hearing this auditory tone, participants were instructed to start the reaching movement toward the target, grasp it, and lift it up. Vision was controlled using spectacles fitted with liquid crystal lenses (Translucent Technologies Inc., Toronto,
Ontario, Canada) that rendered the target stimulus visually accessible by changing from opaque to clear. Vision was restored upon the release of the starting switch. The target stimulus rested on a second pressure switch positioned at the center of the working surface, that is, at 33 cm from the starting position at the subject’s midline (Figure 1B). Participants were instructed to reach at a natural speed. Each subject performed a total of 48 trials (eight trials for each experimental condition). The order of presentation of the experimental conditions was counterbalanced within and across subjects. Subjects were given 10-min practice blocks in which to familiarize themselves with the tasks. Sufficient time (20 s) was allowed between trials to recover from any odor adaptation.

**Dependent measures and data analysis**

Kinematic analysis was confined to the time and amplitude of maximum hand aperture. As previously demonstrated, these dependent measures are most effective in revealing the effects of task-irrelevant information on the unfolding of grasping movements (for a review, see Castiello, 1999). The maximum hand aperture was studied by analyzing the angular distance formed by the sensors described above. The time at which maximum hand aperture occurred was calculated as a percentage of the total movement duration. Movement duration was calculated as the time from the release of the starting switch to the time at which the
target was lifted from the switch located underneath the object. Movement duration was normalized and included in the multivariate analysis of variance (MANOVA) to investigate hand aperture at different epochs during the reach across experimental conditions. Another reason for normalization was that kinematic differences may be better understood when the occurrence of kinematic events is expressed in terms relative to the overall reach duration. The time and amplitude of maximum hand aperture were analyzed using a 2 (target size: large vs. small) x 3 (condition: congruent vs. incongruent vs. control) x 10 (normalized time: from 10% to 100% of the reach) within-subject MANOVA. Main effects were used to explore the means of interest. Bonferroni corrections (alpha level: $P < 0.05$) were applied.

**Results**

The MANOVA was significant for the main factor target size [$F(1,15) = 8.939$, mean square error (MSE) = 173.955, $P < 0.001$], the interaction between target size and time [$F(18,270) = 12.698$, MSE = 90.529], the interaction between target size and condition [$F(2,14) = 12.031$, MSE = 106.399, $P < 0.0001$], and the three-way interaction target size by time by condition [$F(18,270) = 14.804$, MSE = 458.769, $P = 0.0001$].

**The effect of object size**

Maximum hand aperture was greater for larger than for smaller targets [main effect of target size; $50^\circ$ (SE = 2.0) vs. $43^\circ$ (SE = 2.1), respectively; $P < 0.01$], and the maximum hand aperture was reached earlier for small rather than for large targets [interaction target size by time; $51\%$ (SE = 2.1) vs. $65\%$ (SE = 3.2) of reaching duration, respectively; $P < 0.01$].

**Table 1** Visual target–odor combinations for the congruent and the incongruent experimental conditions

<table>
<thead>
<tr>
<th>Congruent conditions</th>
<th>Incongruent conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>Almond–garlic</td>
<td>Apple–orange</td>
</tr>
<tr>
<td>Garlic–strawberry</td>
<td>Orange–peach</td>
</tr>
<tr>
<td>Almond–strawberry</td>
<td>Apple–peach</td>
</tr>
<tr>
<td>Small/large</td>
<td>Large/small</td>
</tr>
<tr>
<td>Apple–garlic</td>
<td></td>
</tr>
<tr>
<td>Strawberry–apple</td>
<td>Apple–garlic</td>
</tr>
<tr>
<td>Strawberry–orange</td>
<td>Orange–garlic</td>
</tr>
<tr>
<td>Strawberry–peach</td>
<td>Orange–strawberry</td>
</tr>
<tr>
<td>Garlic–apple</td>
<td>Peach–almond</td>
</tr>
<tr>
<td>Garlic–orange</td>
<td>Peach–garlic</td>
</tr>
<tr>
<td>Garlic–peach</td>
<td>Peach–strawberry</td>
</tr>
</tbody>
</table>

**Incongruent large condition**

When participants grasped a small target, preceded by an odor evoking an object requiring an incongruent large grasp, the amplitude of maximum hand aperture was larger than when the same target was grasped preceded by an odor evoking a congruent grasp or no odor was presented (interaction target size by condition; congruent condition = $43^\circ$, SE = 2.8; control condition = $42^\circ$, SE = 3.2; incongruent condition = $50^\circ$, SE = 2.2; $P$ values $< 0.001$; see Figure 2). As represented in Figure 3a, if the target required a small grasp and the odor evoked the representation of an object requiring a large grasp, maximum hand aperture for the small target was reached later (as classically happens for a large target grasped in isolation) than when the small target was grasped in congruent or control conditions (congruent condition = $61\%$, SE = 3.0; control condition = $61\%$, SE = 2.1; incongruent condition = $67\%$, SE = 2.4; $P$ values $< 0.01$). For the congruent and control conditions, maximum hand aperture occurred at a similar time ($P > 0.05$).

**Incongruent small condition**

When participants grasped a large target preceded by an odor evoking an object requiring an incongruent small grasp, the amplitude of maximum hand aperture was smaller than when the same target was grasped preceded by an odor evoking a congruent grasp or when no odor was presented (interaction target size by condition; congruent condition = $50^\circ$, SE = 2.2; control condition = $49^\circ$, SE = 2.3; incongruent condition = $42^\circ$, SE = 2.7; $P$ values $< 0.001$; see Figure 2). It is worth noting that although there was a reduction in hand aperture for the incongruent small condition, the angle measured at the joints of interest still allows for the “large” object to be grasped appropriately. As shown in Figure 3b (interaction target size by time by condition), when the target...
required a large grasp and the odor evoked the representation of an object requiring a small grasp, maximum hand aperture for the large target was reached earlier (as classically happens for a small target grasped in isolation) than when the same target was grasped in the congruent and control conditions (congruent condition = 67%, SE = 2.4; control condition = 68%, SE = 2.1; incongruent condition = 61%, SE = 2.2; P values < 0.01). For the congruent and control conditions, maximum hand aperture occurred at a similar time (P > 0.05).

Discussion

We set out to investigate the integration between vision and olfaction during a reach-to-grasp movement. In accordance with classical descriptions for this movement, the results for the control condition indicated that maximum hand aperture was correlated with object size. That is, maximal hand aperture was smaller for small than for large targets (see Castiello, 2005, for a review) and vice versa. Noticeably, for the incongruent conditions, information gained from the olfactory stimulus appeared to leak in and influence the target action pathways. The hand aperture for the action toward the target assumed the features of the hand aperture related to the object evoked by the olfactory stimulus.

The present results are compatible with those obtained in similar experiments in which visual and proprioceptive information from the distractor affected selectively the grasp component of a reach-to-grasp movement (Castiello, 1999; Patchay et al., 2003, 2005). When the distractor was smaller or bigger than the target, the amplitude of maximum grip aperture was, respectively, smaller or bigger than in the conditions where the size of the distractor was neither taken into account nor altered.

We suggest that the present cross-modal interference effects may indicate that the type of object representation evoked on the basis of the olfactory information contains information about the action that the object requires. This action competes with the action programmed for the target object and, when incompatible, generates the reported interference effect. The objects evoked by the olfactory stimuli utilized in the present study required similar or different types of grasp than the target object (i.e., small hand aperture = precision grip, large aperture = whole-hand grasp). Thus, when the object evoked by the olfactory stimulus requires a type of grasp different from that required by the target, parallel computations for different types of grasp, one for the target and one for the attended distractor, may be at the origin of the changes found for the kinematics of the action directed to the target. This view is supported by neurophysiological and behavioral evidence. In the first case, different types of grasp such as precision grip and whole-hand grasp (which correspond to the small and large hand aperture pattern adopted by our subjects, respectively) are subserved by different neural populations (Rizzolatti et al., 1988; Sakata et al., 1995). In the second case, the kinematics differ for precise and whole-hand types of grasp (Gentilucci et al., 1991; Castiello, 1996).

In light of this evidence, it can be suggested that conflicts may emerge when the distractor and target objects require different prehensile patterns in order to be grasped or manipulated. Neuronal populations, kinematic planning, and

![Graphical representation for the interaction type of target by condition by time. Panel (A) represents how the time of maximum hand aperture (filled arrow) is delayed for an action toward a small target when olfactory information evokes an object requiring an incongruent large grasp. Panel (B) represents how the time of maximum hand aperture (filled arrow) is anticipated for an action toward a large target when olfactory information evokes an object requiring an incongruent small grasp. The time course of maximum hand aperture is expressed in terms relative to the overall reach duration (%). Please note that the final hand aperture is similar for the large and small objects, given that we are measuring an angle that in the final hand position changes very little depending on the size of the target. The graphs represent hand aperture averaged across trials and subjects for each experimental condition.](http://chemse.oxfordjournals.org/)

Figure 3
functional properties for the irrelevant object evoked by the olfactory stimulus are alerted and interfere with neuronal populations, kinematic planning, and functional properties activated and executed for the target object. In other words, the objects evoked by the olfactory stimuli automatically activate their motor responses without the participant’s intention to act. This may suggest that the representation of olfactory stimuli is already present in memory and implicitly and automatically encoded in terms of the action it evokes. This is in line with previous evidence suggesting that both storage of and access to olfactory stimuli are automatic and implicit (e.g., Zucco, 2003). Several studies suggest that olfactory memory represents a unique and separate memory system when compared to other sensory modalities (e.g., Richardson and Zucco, 1989; Engen, 1991; Schab, 1991; Herz and Engen, 1996; Zucco, 2003). For instance, memory is rarely affected by retention intervals, incidental or intentional learning, retroactive interference as well as interference tasks, or the adoption of strategies as to improve learning (see Engen and Ross, 1973; Lawless and Engen, 1977; Goldman and Seamon, 1992; Bromley and Doty, 1996). These anomalies seem to suggest that olfaction is a distinctive sensory system that is not able to give rise to conscious representations of odors (Zucco, 2003).

All in all, the present results complement the literature on grasping interference effects by demonstrating that the mechanisms of selection for the control of an overt hand action take into account not only irrelevant visual (e.g., Castiello, 1999) and proprioceptive (e.g., Patchay et al., 2005) information but also olfactory information. However, before this conclusion can be accepted, there are some issues that must be addressed. First, another possible explanation for our results, and one that we cannot exclude, is that our participants were able to accurately label the presented odors. In these circumstances, verbal conscious representations of the distractor would have been responsible for the reported “motor” effects. This possibility is, however, unlikely given that we know from previous literature that naming an odor is an extremely hard cognitive task. For example, when asked to do so, people are usually affected by the so-called “tip of the tongue” state (see Lawless and Engen, 1977). Rather than naming the correct name of the odorants, people make use of autobiographical life episodes or verbal interjections or name hedonically the odorants (Richardson and Zucco, 1989; Schab, 1991; Chobor, 1992; Herz and Engen, 1996; Larsson, 2002; Zucco, 2003).

Second, it cannot be excluded that it is purely the visual, mental image of the odorant’s object that interferes with grasp kinematics. In this respect, the claim that we can experience olfactory imagery is itself a matter of some controversy. Several authors claim that olfactory imagery does not occur (e.g., Engen, 1991; Herz, 2000; Zucco, 2003), whereas others argue that it does (e.g., Cain and Algom, 1997; Elmes, 1998). Although it is outside the scope of the present study to solve this matter, a reconciling view that may fit with the present results has been as recently suggested. Stevenson and Case (2005) propose that olfactory imagery can occur if language is bypassed. In such circumstances, the capacity to form olfactory images is likely to be as good as the capacity to form images using any other modality (Stevenson and Case, 2005). By following this line of reasoning, it can be proposed that the type of representation built on olfactory information is task dependent. If the task is motoric in nature, then the type of representation built on olfactory information might not need to pass through a semantic route involving language but through a pragmatic route that is concerned with how to act upon the represented object, that is, a motor olfactory image might be created.

Third, comparative literature may argue against the idea that olfactory stimuli may affect grasping movements. For example, Metz and Whishaw (2000) found that rats maintain the same grip size regardless of the size of a strongly scented object for which they are reaching. However, as the authors themselves suggest, there might be differences in terms of skilled movement organization between rats and primates, the former being highly constrained indicating an innate organization, perhaps at the level of the motor cortex, and the latter being under the control of phylogenetically younger parietal areas that may control higher levels of action behavior as those tackled in the present study. In this connection, neurophysiological evidence in macaques indicates that different groups of cells in the anterior–lateral bank of the intraparietal sulcus (area AIP) respond to different types of grasp such as precision grip and whole-hand grasp depending on object properties (e.g., size). Similarly, visually guided grasping neuroimaging studies in humans (e.g., Frey et al., 2005) have identified the anterior intraparietal sulcus as the functional homologue of macaque area AIP. It is thus a reasonable assumption that in the present study the conflict between relevant- and irrelevant-task information is resolved within regions of the parietal cortex that form a major component of the “dorsal” visual stream that is thought to be fundamentally involved in selective attention (Culham and Kanwisher, 2001), the control of action (Goodale and Milner, 1992), and the integration of object perceptual and motor properties (Sakata et al., 1995). In order to test this further, we think it would be interesting to perform neuroimaging studies looking at the brain activation in areas concerning olfaction and control of hand configurations. This question is currently being addressed in our laboratory.

In conclusion, the results of the present study confirm and extend previous evidence of cross-modal interactions between olfaction and vision. Crucially, the present study is the first to reveal the effects of olfactory information on the process of selection for the control of a goal-directed action. The suggestion is that task-irrelevant olfactory stimuli can trigger motor-related object representations that compete with the motor-related representation for the task-relevant (the to be grasped) object.