Cognition 176 (2018) 65–73

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Mental space maps into the future

Anna Belardinelli[a], Johannes Lohmann[a,*], Alessandro Farnè[b,c], Martin V. Butz[a]

[a] *Cognitive Modeling, Department of Computer Science and Department of Psychology, Faculty of Science, University of Tübingen, Germany*
[b] *Integrative Multisensory Perception Action & Cognition Team – ImpAct, Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, Lyon, France*
[c] *Hospices Civils de Lyon, Neuro-immersion, Lyon, France*

ARTICLE INFO

Keywords:
Event segmentation theory
Theory of event coding
Anticipatory behavioral control
Peripersonal space
Cross-modal congruency paradigm

ABSTRACT

It has been suggested that our mind anticipates the future to act in a goal-directed, event-oriented manner. Here we asked whether peripersonal hand space, that is, the space surrounding one's hands, is dynamically and adaptively mapped into the future while planning and executing a goal-directed object manipulation. We thus combined the crossmodal congruency paradigm (CCP), which has been used to study selective interactions between vision and touch within peripersonal space, with an object manipulation task. We expected crossmodal interactions in anticipation of the upcoming, currently planned object grasp, which varied trial-by-trial depending on the object's orientation. Our results confirm that visual distractors close to the future finger positions selectively influence vibrotactile perceptions. Moreover, vibrotactile stimulation influences gaze behavior in the light of the anticipated grasp. Both influences become apparent partly even before the hand starts to move, soon after visual target object onset. These results thus support theories of event encodings and anticipatory behavior, showing that peripersonal hand space is flexibly remapped onto a future, currently actively inferred hand position.

1. Introduction

Over the last two decades, various theories on "anticipatory behavior" suggest that our mind predicts immediate but also more distant future consequences of self-executed actions and perceived events and thus acts in an anticipatory, goal-directed manner (Friston, 2009; Friston, Daunizeau, & Kiebel, 2009; Hoffmann, 1993, 2003; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1990). These theories imply that encodings of future states are activated before actual goal-directed motion takes place and that goal-directed, active inference processes focus the mind on those aspects that are believed to be critical for achieving a particular goal, such as a successful object grasp.

A recent theory extension (Butz, 2016) complements the event segmentation theory (EST) proposed in cognitive psychology (Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks & Tversky, 2001). EST suggests that our mind segments perceptions into events and event boundaries. For example, drinking from a bottle can be segmented into reaching, grasping, transporting, and actual drinking. Combined with anticipatory behavior, the extended theory suggests that our mind focuses on desired future event boundaries and subsequent events. For example, our eyes fixate an object when preparing an object grasp in such a way that the intended grasp type and the subsequent object manipulation can be inferred (Belardinelli, Stepper, & Butz, 2016).

Furthermore, it has been proposed that event boundaries will typically involve spatial predictive encodings, which characterize when an event boundary is likely to occur (Butz, 2016). Thus, anticipatory spatial remappings towards upcoming event boundaries can be expected to be present even before the actual goal-directed action towards that event boundary unfolds. Interestingly, event-oriented segmentations are well-suited for enabling hierarchical planning and motor control (Botvinick, Niv, & Barto, 2009; Botvinick & Weinstein, 2014; Wolpert, Diedrichsen, & Flanagan, 2011) and they offer an explanation of how event encodings (Hommel et al., 2001) may develop.

In this work, we asked whether the peripersonal space (PPS) surrounding one's hands is mapped into the future onto the next anticipated event boundary, that is, an object grasp. Behavioral neuroscience has indicated that PPS encodes the space surrounding bodily surfaces, such as the hand or the face, regardless of where the surface is currently positioned, integrating all relevant multisensory information available (di Pellegrino, Làdavas, & Farnè, 1997; Fogassi et al., 1996; Holmes & Spence, 2004; Làdavas, Zeloni, & Farnè, 1998). Moreover, it has been proposed that PPS encodings exist for enabling effective spatial interactions (Graziano & Cooke, 2006). While typically being anchored to a body part, it was shown that PPS partially remaps around a tool upon

---

* Corresponding author at: Cognitive Modeling, Department of Computer Science, Sand 14, 72076 Tübingen, Germany.
E-mail addresses: belardinelli@informatik.uni-tuebingen.de (A. Belardinelli), johannes.lohmann@uni-tuebingen.de (J. Lohmann), alessandro.farne@inserm.fr (A. Farnè), martin.butz@uni-tuebingen.de (M.V. Butz).

https://doi.org/10.1016/j.cognition.2018.03.007
Received 16 June 2017; Received in revised form 4 March 2018; Accepted 6 March 2018
0010-0277/

tool usage (Holmes, 2012).

One means to explore multimodal sensory interactions in PPS is the crossmodal congruency paradigm (CCP; see e.g. Spence, Pavani, Maravita, & Holmes, 2004), where task-irrelevant visual stimuli interfere with tactile perceptions if the former are presented close to the tactually stimulated body part. Typically, participants are slower and less accurate in identifying whether the thumb or the index finger was stimulated if concurrently a LED is flashed at the other finger location (incongruent condition), whereas a flash at the location of the stimulated finger prompts a faster response (congruent condition). In the case of a concurrent grasping task, vision-to-touch interference can actually occur at a distance – before the hand gets close to the target object – when visual stimuli are presented on the object, even at the time the hand just starts to move (Brozzoli, Cardinali, Pavani, & Farnè, 2010; Brozzoli, Pavani, Urquizar, Cardinali, & Farnè, 2009). These results emphasize the importance of PPS encodings for executing manipulative interactions (Brozzoli, Ehrsson, & Farnè, 2014; Brozzoli, Makin, Cardinali, Holmes, & Farnè, 2012). In previous studies, however, the required (pinch) grasp was instructed and thus predictable. Moreover, the object was always visible to the participants throughout the trial, and thus also already before the go signal for the motor task. As a result, the congruent configurations were fixed and the role and adaptivity of the spatial mapping remained elusive.

We thus asked if PPS is adaptively remapped into the future in anticipation of the next event boundary and subsequent event, that is, the intended grasp followed by an object manipulation. We expected to observe an anticipatory crossmodal congruency effect (aCCE), in which visual distractors near the future finger position at the to-be grasped object should affect responses to tactile vibrations on the fingers depending on the planned grasp type. Moreover, we reasoned that an eye gaze preference towards that object side where the stimulated finger would be placed should be detectable. While effects of touch on visual perception have been reported for static and moving visuo-tactile stimuli (Driver & Spence, 1998; Gray & Tan, 2002), to the best of our knowledge, no haptics-related oculomotor effects have been reported in anticipation of an upcoming hand-object interaction.

To investigate the hypothesized aCCE, we combined CCP with an object manipulation task: participants had to reach and virtually grasp a bottle on a touchscreen, displace it slightly to the right, and put it back down in an upright orientation. Additionally, participants had to verbally report the finger (thumb or index) on which a vibrotactile stimulation was felt. Concurrently with the vibration, sometimes a visual distractor was presented on one side of the bottle about where the thumb or index finger would be placed.

The critical manipulation was the variation of the bottle's orientation at visual onset – either upright or upside-down – calling for different hand-grasp orientations – either overhand or underhand – in anticipation of the intended manipulation (Herbort & Butz, 2011, 2012; Rosenbaum, Chapman, Weigelt, Weiss, & van der Wel, 2012; Rosenbaum, Marchak, Barnes, Siotta, & Jorgensen, 1990). Critically, the two grasp types yield different crossmodal congruencies, seeing that the index finger (thumb) will be placed on the right (left) or left (right) side of the bottle, depending on whether the bottle is presented upright or upside down (see Fig. 1). In this case, hence, congruency is flexibly defined according to the unfolding motor planning, which can be decided only once the target object is shown (i.e. upon visual onset). Furthermore, to investigate the dynamics of such an anticipatory PPS remapping, the visuo-tactile stimulation was applied at different points in time around movement onset and two complementary experiments were conducted. Taken together the results show that PPS is indeed mapped highly adaptively and purposefully into the future supporting the upcoming and unfolding motor behavior.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Sixteen participants ($M = 24.9$, $SD = 4.9$ years, 8 female) took part in the experiment. Three participants (2 female) were excluded because of bad quality eye-tracking data. The sample size was determined comparable to that of previous studies in the literature on CCE (e.g., Brozzoli et al., 2010). All participants were right-handed, had corrected-to-normal vision, and reported normal tactile sensitivity. They were compensated for their participation with money or course credit and signed the informed consent form. The experiment was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki.

2.1.2. Apparatus

The visual target stimulus used in both experiments was a picture (320 × 960 pixel, 7.1° × 22.1°) of a 0.5 L plastic bottle full of water on a white background. Stimulus presentation was done on a 1680 × 1050 pixel (37.1° × 24.2°) touchscreen with a refresh rate of 60 Hz.

A red dot (10 pixel, 0.23° radius) was used as visual distractor and presented (in the corresponding conditions) for 200 ms either on the right side or the left side 80 pixel (1.84°) away from the center and at middle height on the bottle. The simultaneous tactile stimulation was delivered to the acting hand by means of LilyPad Arduino vibration motors (2.0 cm diameter, 0.8 cm thickness), applying a vibration for 200 ms to either the thumb or the index fingertip. The motors were controlled via an Arduino Uno microcontroller (Arduino S.R.L.) running custom C software.

Eye movements were collected by means of a binocular remote eyetracker (EyeFollower, LC Technologies), working at 120 Hz and with an accuracy $< 0.4°$ even under head movements. Each participant was calibrated with a 9-point calibration procedure.

A keyboard was placed between the participant and the monitor to record reaction times. Participants had to hold down the space bar to start the next trial, releasing it when initiating the reach towards the bottle. Motion times were measured as the time between space bar release and first touch on the touchscreen. Additionally, the touchscreen information was used to confirm the grasp type.

Verbal responses were collected via a headset microphone using a custom C# program, based on the Microsoft Speech API 5.4. The API provides a time-stamp at the beginning of each utterance.

The whole experiment was implemented in Matlab R2013a, using the Psychophysics Toolbox extension (Brainard, 1997).

2.1.3. Design and procedure

Participants sat at a table in front of the apparatus and at about 70 cm from the screen. The experiment consisted in a dual task paradigm. As a first task, participants were requested to fixate a fixation cross on the left of the screen 420 pixel (9.3°) away from the center of the screen and to keep their right hand on the spacebar until the stimulus appeared in the center of the screen. Upon stimulus presentation they had to plan the grasp and displacement of the bottle on the screen to a target location on the right side of the screen, (420 pixel, 9.3° away from the center), denoted by a flat gray ellipse. The grasp was detected by the first touch on the screen and determined the disappearance of the bottle picture. A second touch on the right half of the screen was considered as the placing of the bottle on the target location (and determined its reappearance there). The return of the hand to the spacebar signaled the end of the trial (see Fig. 2 for an exemplary trial time course). The bottle could be presented either upright or upside down (factor orientation). Participants were instructed to grasp the bottle and place it back in an upright orientation, so to enforce a supine (underhand) grasp when the bottle was presented upside down.