Hands in Thought and Motion

Johannes Lohmann (johannes.lohmann@uni-tuebingen.de)
Chair of Cognitive Modeling, Sand 14
Tübingen, 72076 Germany

Martin V. Butz (martin.butz@uni-tuebingen.de)
Chair of Cognitive Modeling, Sand 14
Tübingen, 72076 Germany

Abstract

Theories of event-predictive, anticipatory behavior control suggest that complex action planning and control is segmented into sequences of anticipated subgoals and according behavioral events, which accomplish the subgoals. Here we focus on the cognitive dynamics during successive subgoal activations. We combined a virtual object interaction task (prehension and transport of a bottle) with a crossmodal congruency task. Anticipatory crossmodal congruency effects (aCCEs) occur at the goal of the current behavior, before the goal is reached. These aCCEs appear to be stronger during prehension, while visual distractors at the currently irrelevant movement target have no effect. While the results so far provide only partial support for the proposed anticipatory, sequential control process, the paradigm is well-suited to probe the dynamic changes of spatial body representations in object interactions.

Keywords: Event Predictive Cognition; Anticipatory Behavioral Control; Peripersonal Space; Virtual Reality

Introduction

Natural object interactions are carried out in a way to facilitate possible, or planned subsequent actions. For instance, the actual grasp that is used to pick up a bottle has been shown to be modulated by the expected next action (either drinking or handing it over; Belardinelli, Stepper, & Butz, 2016). This predictive mode of acting implies that the final outcome, or the final outcome possibilities, are considered at the beginning of sequential object manipulations (Flanagan, Vetter, Johansson, & Wolpert, 2003; Hoffmann, 2003). According to event-predictive theories of anticipatory motor control, the initiation of goal-directed actions requires the activation of event-predictive structures or *schemata* (EPSs; e.g. Butz, 2016; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Richmond & Zacks, 2017), which consist of event-predictive encodings and event boundary encodings, where the latter mark event beginning, endings, or transitions. These ESPs are assumed to be hierarchically structured. For instance, the overall goal of preparing a cup of tea is composed of various subgoals, or events, like boiling water, fetching tea and possibly milk, a cup, etc. While there is quite some evidence for this event structure in perception and action, it remains largely open how it is realized in a predictive manner. Eye-tracking studies have shown that the eyes are tuned selectively to the next contact point of the index finger in an object interaction task already before the actual action unfolds (Belardinelli et al., 2016). Apparently, movement planning entails a prediction of the final hand posture at the target object.

It has been argued that this prediction is realized by remapping a certain spatial body representation – the so called peripersonal hand space (PPHS) – upon the grasping target. In order to test this assumption, one typical indicator of PPHS, the selective interaction between vision and touch, has been probed with the crossmodal congruency paradigm (Spence, Pavani, Maravita, & Holmes, 2004). In crossmodal congruency tasks, participants are requested to indicate the location of a tactile stimulation. Visual distractors presented close to the stimulated body part interfere with tactile perception. For instance, participants are slower to identify whether thumb or index finger received a tactile stimulation, if a LED is flashed at the non-stimulated finger (incongruent), compared to trials where the distractor is presented at the stimulated finger (congruent). If PPHS is indeed remapped towards the grasping target, one would expect anticipatory crossmodal congruency effects (aCCE) at the target location, even before movement initiation. Support for this notion comes from real world (Brozzoli, Pavani, Urquizar, Cardinali, & Farnè, 2009; Brozzoli, Cardinali, Pavani, & Farnè, 2010), pantomimic (Belardinelli, Lohmann, Farnè, & Butz, 2018), and virtual reality studies (Lohmann, Belardinelli, & Butz, 2019).

The aCCE seems to be a useful tool to investigate the mechanisms of event-predictive, anticipatory motor control, and the properties of ESPs. If the assumed hierarchical structure of ESPs is realized on the behavioral level, one would expect the aCCE only at the currently relevant event boundary, as the PPHS should only be remapped onto the current interaction goal. If action planning is realized by activating the final goal along with all subgoals, one would expect an aCCE to be measurable at the final location of an object interaction even at the beginning of the movement.

We asked participants to grasp and carry a bottle in a virtual reality (VR). The interaction was composed of a sequence of two subgoals: Grasping the bottle at a pick-up location and carrying it to a placement location. At different times before and during either the prehension, or the transport, participants received a tactile stimulation at the thumb or index finger. At the same time, a visual distractor was presented at the currently relevant, that is, the next subgoal, or the currently irrelevant location, that is, the final goal or the previous subgoal. Participants were requested to name the stimulated finger as fast as possible. An aCCE would be indicated by faster responses in cases when the visual stimulus coincides with the future finger position of the stimulated finger. The main investigation focused on the dynamics of the (a)CCE: does it occur only at the currently relevant event boundary, or is it also measurable at previously relevant or prospectively relevant event boundaries?

Method

Participants

In order to determine an appropriate sample size, we conducted a power analysis using our earlier data regarding aCCE (Experiment 1 in Lohmann et al., 2019). For the sought three-way interaction between visual distractor, stimulated finger and bottle orientation, we previously observed an effect sizes of $\eta_p^2 = .67$. Given a power of 0.9 and an alpha level of 0.05, a lower bound for the sample size of 18 was determined. The power analysis was performed by means of the Monte Carlo method. Nineteen participants from the University of Tübingen participated in the experiment (eleven females). Their age ranged from 19 to 28 years (M = 21.5, SD = 2.25). Two participants were left-handed and all participants had normal or corrected-to-normal vision. Participants provided informed consent and received either course credit or a monetary compensation for their participation. The experimental protocol was approved by the board of ethics in psychological research of the University of Tübingen. Three participants (one male, two female, one left-handed) had difficulties with the virtual grasping procedure and could not complete the experiment. The respective data were not considered in the analysis.

Apparatus

To immerse participants in VR, they were equipped with an Oculus Rift© DK2 stereoscopic head-mounted display (Oculus VR LLC, Menlo Park, California). Hand movements were tracked by means of a Leap Motion© near-infrared sensor (Leap Motion Inc, San Francisco, California, SDK version 3.2.1). Positional data regarding the palm, phalanges and fingers were obtained from the Leap Motion© sensor. These data were used to render a hand model in VR. Participants were also equipped with a headset in order to respond verbally to the tactile stimulation. Speech recognition was implemented by means of the Microsoft Speech API 5.4. The whole experiment was implemented with the Unity® engine 2018.2.18 using the C# interface provided by the API. In order to be able to support and observe the participants, the scene was rendered in parallel on the Oculus Rift and a computer screen.

Vibrotactile stimulation was delivered by means of two shaftless vibration motors (10 mm × 3.4 mm) attached to the tip of the thumb and the index finger of the participants. The motors were controlled via an Arduino Uno microcontroller (Arduino S.R.L., Scarmagno, Italy) running custom C software. Control commands were send through an USB connection from the Unity® program. The wiring diagram as well as additional information regarding the components can be found at the first author's webpage. [1]

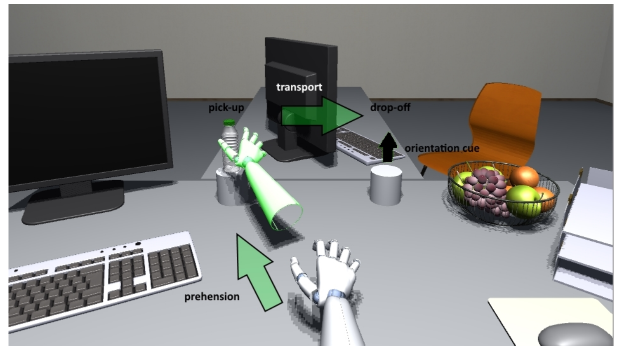

Figure 1: The VR scene and the different parts of the interaction. First, participants had to grasp a bottle from a pick-up location (prehension). Second, they had to put it down at a drop-down location (transport). After grasping the bottle (indicated by the green hand in the image), an orientation cue (the black arrow) appeared at the drop-down location, indicating the requested bottle orientation.

Virtual Reality Setup

The VR setup put participants in an office where they sat in front of a desk (see Fig. 1). The desk surface in VR corresponded with a desk surface in the real world. Two pedestals, one on the left, and one on the right side of the participants, were placed on the table, 45 cm away from the participants initial hand position, and 45 cm away from each other. Participants were requested to grasp a 3D model of a plastic bottle either oriented upright or upside down, which appeared always on the left pedestal. The bottle was 15 cm in height, subtending a visual angle of 6.1° at the left pedestal (the viewing distance was approximately 1.4 m). The right pedestal served as the target location, where participants were requested to place the bottle. After grasping the bottle at the left pedestal, an arrow occurred at the right pedestal, either pointing upward, or downward, indicating the final orientation of the bottle. The arrow was 15 cm in height and subtended a visual angle of 6.1° at the right pedestal. This cue remained visible throughout the rest of the trial. Participants were requested to place the bottle either upright (arrow up) or upside down (arrow down) at the right pedestal.

Instructions and feedback were presented in different textfields, aligned at eye-height. At the beginning of a trial, a fixation cross appeared at the left pedestal. The fixation cross was 10 cm wide and 10 cm high, subtending a visual angle of 4.1°. The visual distractor was realized by means of a red, spherical flash with a diameter of 8 cm (equal to a visual angle of 3.3°) appearing above and slightly to the left or right of the pick-up or the drop-down location (see Fig. 2).

[1] https://uni-tuebingen.de/de/26084

Procedure

At the beginning of the experiment, participants received a verbal instruction regarding the VR equipment. Then they were equipped with vibration motors and familiarized with the tactile stimulation. Participants were then seated comfortably on an arm chair and put on the HMD. Before the actual experiment, participants performed a grasping training and trained the verbal response until they felt comfortable with both tasks. In the grasp training, participants performed the prehension and the transport without receiving a tactile stimulation. In the verbal response training, participants had to put their hand in the starting position. Then the bottle appeared at the initial location, after a variable time interval between 50 ms and 350 ms, a visual distractor appeared either on the left or the right site of the bottle. At the same time a tactile stimulus was delivered to the participants' thumb or index finger, and participants were requested to indicate the stimulated finger verbally by saying "index" or "thumb" (i.e., in German "Zeigefinger" or "Daumen"). In the verbal response training, participants did not perform a grasping movement, but remained with their hand in the starting position.

The actual experiment combined both tasks in a dual-task paradigm. At the beginning of each trial, participants had to move their right hand into a designated starting position, consisting of red, transparent spheres. The spheres indicated the required positions of the fingers and the palm. The spheres turned green when the respective fingers were in position. Furthermore, participants had to maintain a stable looking direction on a fixation cross. Once both requirements were met for 1000 ms, the fixation cross as well as the visible markers of the initial position disappeared and a bottle appeared on the left pedestal. The bottle was either oriented upright, or upside down. Participants were instructed to grasp the bottle with a power grasp, and to transport it to the pedestal to the right. At the beginning of a trial, participants did not know, whether they had to place the bottle in an upright orientation, or upside down. After grasping the bottle at the pick-up location, an arrow appeared at the goal location and remained visible until the end of the trial. If the arrow pointed upwards, participants were requested to put down the bottle in an upright orientation. If the arrow pointed down, participants were requested to put down the bottle upside down. We did not explicitly instruct a supine (underhand) grasp in case of bottles oriented upside down, however, all participants performed this kind of grasp.

Besides the grasp-and-carry task, participants had to discriminate which finger received a vibrotactile stimulation and to report the stimulated finger as fast as possible upon vibration detection. The onset of the tactile stimulation varied from trial to trial. The tactile stimulation could either be delivered during the prehension or the transport, either at the beginning of the respective movement, or after covering half of the distance to the current movement target. A visual distractor appeared at the same time either at the pick-up, or drop-down location. Hence, the distractor could appear either at the currently relevant, or irrelevant movement target. The location of the distractor was above the pedestals, slightly shifted either to the left or the right side, roughly corresponding to the locations of the index finger, or thumb, if a bottle would be grasped or released at the respective location. Depending on the bottle orientation, this was expected to yield different congruent and incongruent conditions with respect to the aCCE (see Fig. 2).

The experiment consisted of 512 trials, presented in a single block. The experiment was self-paced and participants could pause between trials. The whole procedure took between 90 and 120 minutes, including preparation and training.

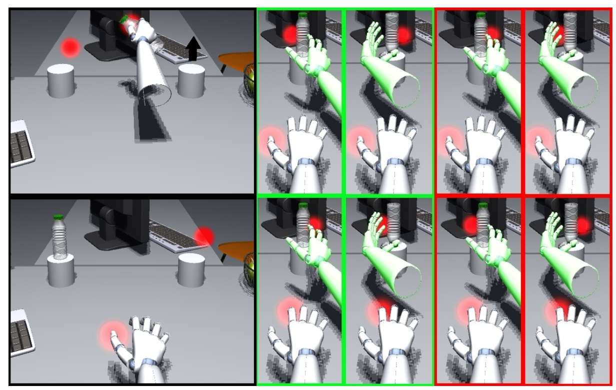

Figure 2: On the right side, the different congruency conditions with respect to the future hand position (transparent green hand), depending on bottle orientation are shown. The stimulated finger is indicated by a red flash (not visible to the participants). Red frames indicate incongruent conditions, congruent conditions are marked with a green frame. In our setup, the conditions shown on the right side correspond to *relevant* stimulation at the *pick-up* location, as the visual distractor occurs at the current movement goal, which is the pick-up location in this case. On the left side, two examples for irrelevant stimulations are shown. The upper left example shows *irrelevant* stimulation at the *pick-up* location during the transport action (SOA2). The lower left example shows *irrelevant* stimulation at the *drop-down* location at movement onset (SOA1). In both cases, the thumb receives the tactile stimulation, hence, these are incongruent conditions, as the target bottle orientation is upright in both cases, and the visual distractor appears at the right side of the respective location, which corresponds to the index finger position when grasping an upright bottle.

Factors, Measures, Data Treatment

We varied seven factors across trials. First, the bottle could be initially oriented upright or upside down (*initial orientation*). Second, the requested final bottle orientation could be upright or upside down (*final orientation*). Third, the visual distractor could appear either on the left or the right side (*distractor*). Fourth, the tactile stimulation could be applied either to the thumb or to the index finger (*stimulation*). Fifth, the visual distractor could either occur at the pick-up, or the drop-down location (*stimulation location*). Sixth, the visual distractor could either occur at the currently relevant movement target, or not (*relevance*). For instance, a visual distractor at the drop-down location (right pedestal), while the participants were reaching for the bottle at the pick-up location (left pedestal), was coded as irrelevant. Seventh, we varied the onset of the tactile stimulation and the visual distractor, relative to the current movement target (*SOA*): At movement onset (SOA1), or after the hand traveled half-way to the current target location (SOA2). We repeated the 2 (initial orientation) × 2 (final orientation) × 2 (distractor) × 2 (stimulation) × 2 (stimulation location) × 2 (relevance) × 2 (SOA) factor combinations 4 times, yielding 512 trials. Due to the small number of repetitions per factor combination, we did not filter out any correct response times, however, in order to avoid outlier effects, all analyses were carried out on median response times, instead of mean response times. The primary dependent measure were the verbal response times for naming the stimulated finger. Data from error trials (wrong verbal response, 2.1% of the trials) were excluded from the response time analyses.

Congruency

To evaluate our hypotheses, aCCEs were of key interest. In previous work (Brozzoli et al., 2009, 2010; Belardinelli et al., 2018; Lohmann et al., 2019), aCCEs have been operationalized in terms of a three-way interaction between the bottle orientation, the visual distractor, and the site of the tactile stimulation (cf. Fig. 2, right side). For instance, in the case of an upright bottle, a tactile stimulation of the index finger along with a visual distractor on the right side of the bottle is considered congruent, since the visual distractor matches the future finger position. Here, we determined congruency with respect to the current movement target, the relevance factor indicates whether the visual distractor appears at the current movement target or not. For instance, an irrelevant stimulation at the pick-up location would mean that the visual distractor is displayed at the pick-up location (left pedestal) while the participants are already moving towards the drop-down location (right pedestal). Congruency would be determined with respect to the currently relevant bottle orientation. In the example from above, this would be the final bottle orientation (see also Fig. 2, left side for more examples). To focus the analysis, we recoded the data accordingly and obtained a congruency factor, combining the visual distractor and tactile stimulus factor. We also report an analysis of the respective individual response time differences (incongruent minus congruent conditions).

Results

Verbal response times from correct trials from the 16 considered participants were analyzed with a 2 (initial orientation) × 2 (final orientation) × 2 (congruency) × 2 (stimulation location) × 2 (relevance) × 2 (SOA) repeated measures ANOVA. Verbal response times differences between incongruent and congruent conditions were further analyzed with a 2 (initial orientation) × 2 (final orientation) × 2 (stimulation location) × 2 (relevance) × 2 (SOA) repeated measures ANOVA. All reported post-hoc comparisons were submitted to a Holm-Bonferroni correction. The analyses were carried out with R (R Core Team, 2016) and the `ez` package (Lawrence, 2015). In case of violations of the assumption of sphericity, p-values were submitted to a Greenhouse-Geisser adjustment. Bayes factors were calculated with the `BayesFactor` package (Morey & Rouder, 2018).

Verbal Response Times

The ANOVA yielded a significant main effect for congruency ($F(1,15) = 4.75$, $p = .046$, $\eta_p^2 = .24$), as well as significant interactions between stimulation location and relevance ($F(1,15) = 65.85$, $p < .001$, $\eta_p^2 = .81$), stimulation location, relevance, and final bottle orientation ($F(1,15) = 5.78$, $p = .030$, $\eta_p^2 = .28$), stimulation location, relevance, and SOA ($F(1,15) = 20.21$, $p < .001$, $\eta_p^2 = .57$), and stimulation location, relevance, and congruency ($F(1,15) = 4.46$, $p = .050$, $\eta_p^2 = .23$; all remaining p's $\geq .058$).

Participants responded faster in case of congruent stimulation ($\text{Mdn}_{congruent}$ = 691 ms vs. $\text{Mdn}_{incongruent}$ = 698 ms). As indicated by the interaction between stimulation location and relevance, responses in the transport part of the movement - relevant stimulation at the drop-down location (Mdn = 653 ms), and irrelevant stimulation at the pick-up location (Mdn = 652 ms) - were faster, than responses to stimulation during prehension - relevant stimulation at the pick-up location (Mdn = 735 ms), and irrelevant stimulation at the drop-down location (Mdn = 738 ms; all respective p's $< .001$). This pattern was modified by two three-way interactions. Regarding the interaction between stimulation location, relevance, and final bottle orientation, responses during transport seemed to be faster if the bottle had to be placed upright (Mdn = 645 ms and Mdn = 646 ms, respectively), compared to cases were it had to be placed upside down (Mdn = 659 ms and Mdn = 660 ms, respectively). However, after adjusting for multiple comparisons, none of the respective differences reached significance. With respect to the interaction between stimulation location, relevance, and SOA condition, responses in both parts of the movement seemed faster in case of stimulation during the movement, compared to stimulation at movement onset. However, the respective differences were only found to be significant in the transport phase (Mdn_{SOA2} = 619 ms and Mdn_{SOA2} = 620 ms vs. Mdn_{SOA1} = 686 ms and Mdn_{SOA1} = 685 ms; all respective ps $< .018$).

The three-way interaction between stimulation location, relevance, and congruency is shown in Fig. 3 (left side). In general, response times during transport were faster than during prehension. For our hypotheses, the most relevant question is whether the respective congruent conditions in the case of relevant stimulation indeed yields the fastest response times. This is true for the prehension, congruent stimulation at the relevant location yielded faster responses than irrele-

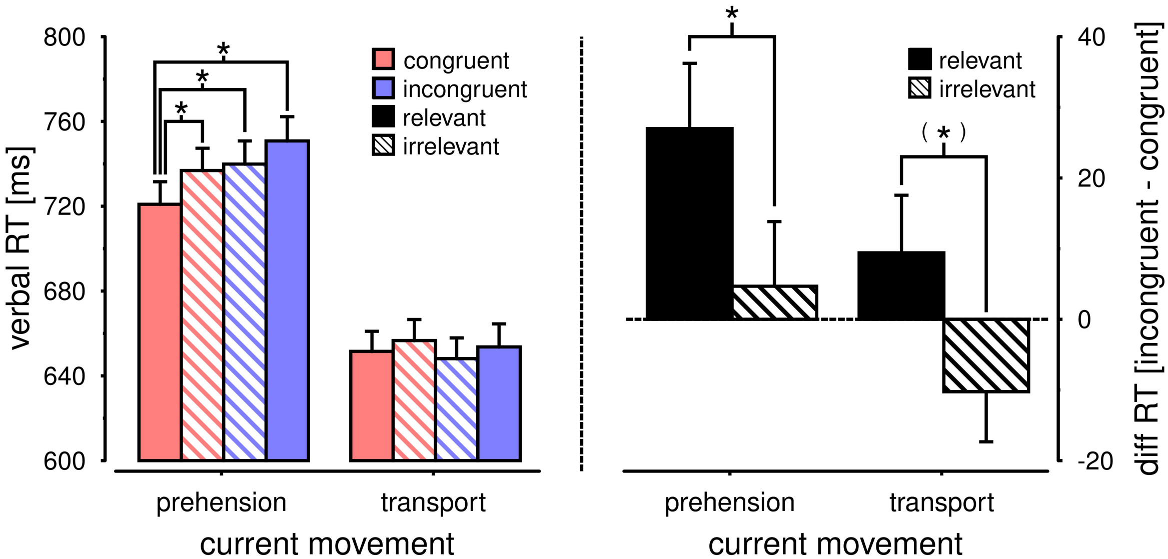

Figure 3: On the left, the stimulation location × relevance × congruency interaction on the verbal response times is shown. RTs are generally faster during transport compared to prehension. Significant modulations due to congruency are only observed during prehension. During transport, congruent as well as incongruent RTs, both for stimulation at the relevant and the irrelevant location are rather similar. On the right side, the stimulation location × relevance interaction on the individual response time differences is shown. During prehension, the congruency effect is stronger at relevant locations. A similar tendency can be observed for the transport. However, the response time difference for relevant locations does not seem to differ significantly from 0. Significant differences are indicated with an asterisk. Asterisks in brackets indicate comparisons that only approached significance. Error bars indicate the standard error of the mean.

vant congruent stimulation at the goal location ($t(15) = 2.91$, $p = .010$), irrelevant incongruent stimulation at the goal location ($t(15) = 3.01$, $p = .009$), and incongruent stimulation at the relevant location ($t(15) = 2.82$, $p = .013$). However, for the transport none of the relevant comparisons reached significance after adjusting for multiple comparisons.

Verbal Response Time Differences

To focus on the effects involving congruency, we conducted a second ANOVA on the individual response time differences between incongruent and congruent conditions. The ANOVA only yielded a significant main effect of relevance ($F(1,15) = 6.18$, $p = .025$, $\eta_p^2 = .29$; all remaining p's $\geq .060$). In case of relevant stimulation, the response time difference was larger, than in case of irrelevant stimulation ($Mdn_{relevant} = 18$ ms vs. $Mdn_{irrelevant} = -3$ ms). For both means, we calculated the Bayes factor for a comparison against 0. For stimulation at relevant locations, the estimated Bayes factor suggested that the data were 3.1 more likely to be larger than 0 than equal to 0. For irrelevant stimulation, this seems unlikely ($BF_{10} = 0.3$). The three-way interaction between stimulation location, relevance, and congruency from the previous analysis, should be reflected by a two-way interaction between stimulation location and relevance in the response time differences. This interaction did not reach significance ($F(1,15) = 4.12$, $p = .060$, $\eta_p^2 = .22$), but since the absence of a congruency effect for the transport movement is not in line with our hypotheses, we obtained the Bayes factors for the respective comparisons against zero and compared the differences by means of t-tests. Indeed the Bayes factors imply, that only stimulation at the pick-up location during prehension yields a congruency effect ($BF_{10} = 4.5$; Mdn = 26 ms). For irrelevant stimulation during prehension ($BF_{10} = 0.3$; Mdn = 4 ms), a congruency effect seems unlikely. Regarding the transport movement, neither stimulation at the irrelevant location ($BF_{10} = 0.4$; Mdn = -10 ms), nor stimulation at the relevant location ($BF_{10} = 0.4$; Mdn = 9 ms) seems to yield a congruency effect. The t-tests showed that the congruency effect in the prehension is significantly larger when the distractor appears at the relevant location, compared to when it appears at the irrelevant location ($t(15) = 2.10$, $p = .026$). For the transport movement, this difference only approached significance ($t(15) = 1.63$, $p = .062$; see also Fig. 3, right side). The case of irrelevant stimulation at movement onset during transport is special, as the visual distractor is shown at the current hand position. To assure that there is no crossmodal congruency effect with respect to the current (instead of the future) bottle orientation, we calculated congruency with respect to the current bottle orientation and tested the respective difference against zero. The according Bayes factor indicated no congruency effect ($BF_{10} = 0.4$; Mdn = -9 ms).

Discussion

Our goal was to investigate the dynamic, anticipatory remapping of PPHS during a sequential object interaction, composed of a prehension and a transport movement. In order to do so, we combined a virtual object interaction with a crossmodal congruency paradigm, to assess anticipatory crossmodal congruency effects (aCCEs). Participants had to grasp and transport virtual bottles while receiving a tactile stimulation on their right thumb or index finger along with a visual stimulation at the currently relevant, or the currently irrelevant movement target. The position of the visual distractor relative to the movement target could either match the future finger location or not. We expected aCCEs only to occur at the currently relevant movement target. That is, during prehension we only expected an aCCE if the visual distractor was shown at the pick-up location, while, during transport, we expected an aCCE only if the visual distractor was presented at the drop-down location. The results confirm the first assumption, which is in line with the typical aCCE reported in previous studies (Belardinelli et al., 2018; Brozzoli et al., 2009, 2010; Lohmann, Belardinelli, & Butz, 2018). However, our results show only a weak tendency for an aCCE during the transport phase. This might be due to the comparatively small sample size. However, other factors of the experiment itself seem to be the reason for this absence of an aCCE during the transport movement. The generally faster responses during transport compared to prehension already indicate that the interaction between the crossmodal congruency task and the motor task is weaker during transport.

Predictability of movement outcomes can affect the movement planning strategy in sequential actions, resulting in higher, or lower, couplings of the partial movements (e.g. Lewkowicz & Delevoye-Turrell, 2019). In order to encourage participants to pay close attention to the respective movement targets, we reduced the overall planning security, by showing the cue for the final bottle orientation only upon grasping the bottle at the pick-up location. Hence, participants had almost no time to plan their movement with respect to the bottle orientation at the movement onset of the transport movement. Accordingly, the aCCE at the first SOA in the transport phase was likely to be smaller than at the first SOA during prehension. For the second SOA, that is, stimulation after covering half of the distance to the current movement target, the aCCE should be similar both during prehension and transport. Descriptively, this is indeed what we observe, but due to the overall smaller effect size of the aCCE during the transport movement, the validity of this trend needs yet to be verified. In order to so, we plan to increase planning certainty by either keeping the final bottle orientation constant, or by showing the orientation cue already at the beginning of the trial.

Even though we found only partial evidence for our hypothesis, of a fast, dynamic remapping of PPHS at event-boundaries, the experimental design seems to be well-suited to investigate the realization of dynamic, event-predictive, anticipatory behavior control on a sensorimotor level. The Bayes factor analysis shows that the aCCE is selective, as it is not observed at task-irrelevant locations. This is especially surprising for stimulations at the pick-up location at movement onset for the transport movement, where the visual distractor appears very close to the hand. Hence, the results imply that the selective remapping of PPHS is a part of the anticipatory behavior control processes that are unfolding during sequential, goal-directed hand movements.

Besides investigating mechanisms of anticipatory behavior control, the aCCE might also be useful to understand the role of peripersonal space in social action understanding. Previous research has shown that peripersonal space can be remapped to include other agents in social situations (Maister, Cardini, Zamariola, Serino, & Tsakiris, 2015). Schaefer, Heinze, and Rotte (2012) proposed that this might be due to peripersonal space being involved in an embodied simulation of other bodies. If this remapping indeed serves action understanding, one could expect an aCCE for the hand of an interaction partner in a joint action scenario. The concept of peripersonal space has also been adapted for human-robot interactions (Roncone, Hoffmann, Pattacini, Fadiga, & Metta, 2016; Nguyen, Hoffmann, Roncone, Pattacini, & Metta, 2018). Here, current research focuses mostly on the defensive purpose of peripersonal space, that is, avoiding or making contact with approaching objects. Understanding the predictive role of peripersonal space in action control might also facilitate action understanding in artificial agents, enabling predictive human-robot interactions.

Acknowledgments

Support and funding for this study was provided by the Leibniz-Wissenschafts Campus Tübingen (WCT) and the DFG research unit *Modal and Amodal Cognition: Functions and Interactions* (FOR 2718).

References

Belardinelli, A., Lohmann, J. Y., Farnè, A., & Butz, M. V. (2018). Mental space maps into the future. *Cognition*, *176*, 65–73.

Belardinelli, A., Stepper, M. Y., & Butz, M. V. (2016). It's in the eyes: Planning precise manual actions before execution. *Journal of Vision*, *16*(1), 18. doi: 10.1167/16.1.18

Brozzoli, C., Cardinali, L., Pavani, F., & Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia*, *48*(3), 796–802.

Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., & Farnè, A. (2009). Grasping actions remap peripersonal space. *Neuroreport*, *20*(10).

Butz, M. V. (2016). Toward a unified sub-symbolic computational theory of cognition. *Frontiers in Psychology*, *7*, doi: 10.3389/fpsyg.2016.00925.

Flanagan, J. R., Vetter, P., Johansson, R. S., & Wolpert, D. M. (2003). Prediction precedes control in motor learning. *Current Biology*, *13*(2), 146–150.

Hoffmann, J. (2003). Anticipatory behavioral control. In *Anticipatory behavior in adaptive learning systems* (pp. 44–65). Springer.

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

Lawrence, M. A. (2015). ez: Easy analysis and visualization of factorial experiments [Computer software manual].

Lewkowicz, D., & Delevoye-Turrell, Y. N. (2019). Predictable real-time constraints reveal anticipatory strategies of coupled planning in a sequential pick and place task. *Quarterly Journal of Experimental Psychology*. doi: 10.1177/1747021819888081

Lohmann, J., Belardinelli, A., & Butz, M. V. (2018, July). Are you sure how to move? Expected uncertainty modulates anticipatory crossmodal interactions. In *Proceedings of the 40th annual conference of the cognitive science society.* Madison, USA.

Lohmann, J., Belardinelli, A., & Butz, M. V. (2019). Hands ahead in mind and motion: Active inference in peripersonal hand space. *Vision*, *3*(2). doi: 10.3390/vision3020015

Maister, L., Cardini, F., Zamariola, G., Serino, A., & Tsakiris, M. (2015). Your place or mine: Shared sensory experiences elicit a remapping of peripersonal space. *Neuropsychologia*, *70*, 455–461.

Morey, R. D., & Rouder, J. N. (2018). Bayesfactor: Computation of bayes factors for common designs [Computer software manual]. (R package version 0.9.12-4.2)

Nguyen, D. H. P., Hoffmann, M., Roncone, A., Pattacini, U., & Metta, G. (2018). Compact real-time avoidance on a humanoid robot for human-robot interaction. In *Proceedings of the 2018 acm/ieee international conference on human-robot interaction* (pp. 416–424).

R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria.

Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences*, *21*(12), 962–980. doi: 10.1016/j.tics.2017.08.005

Roncone, A., Hoffmann, M., Pattacini, U., Fadiga, L., & Metta, G. (2016). Peripersonal space and margin of safety around the body: learning visuo-tactile associations in a humanoid robot with artificial skin. *PloS one*, *11*(10), e0163713. doi: 10.1371/journal.pone.0163713

Schaefer, M., Heinze, H.-J., & Rotte, M. (2012). Close to you: embodied simulation for peripersonal space in primary somatosensory cortex. *PLoS One*, *7*(8), e42308. doi: 10.1371/journal.pone.0042308

Spence, C., Pavani, F., Maravita, A., & Holmes, N. (2004). Multisensory contributions to the 3-d representation of visuotactile peripersonal space in humans: Evidence from the crossmodal congruency task. *Journal of Physiology*, *98*(1), 171–189.