

Reading Your Mind While You Are Reading—Evidence for Spontaneous Visuospatial Perspective Taking During a Semantic Categorization Task

Psychological Science
2018, Vol. 29(4) 614–622
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617740973
www.psychologicalscience.org/PS


Martin Freundlieb, Ágnes M. Kovács, and Natalie Sebanz

Department of Cognitive Science, Central European University

Abstract

Recent studies have demonstrated people's propensity to adopt others' visuospatial perspectives (VSPs) in a shared physical context. The present study investigated whether spontaneous VSP taking occurs in mental space where another person's perspective matters for mental activities rather than physical actions. Participants sat at a 90° angle to a confederate and performed a semantic categorization task on written words. From the participants' point of view, words were always displayed vertically, while for the confederate, these words appeared either the right way up or upside down, depending on the confederate's sitting position. Participants took longer to categorize words that were upside down for the confederate, suggesting that they adopted the confederate's VSP without being prompted to do so. Importantly, the effect disappeared if the other's visual access was impeded by opaque goggles. This demonstrates that human adults show a spontaneous sensitivity to others' VSP in the context of mental activities, such as joint reading.

Keywords

visuospatial perspective taking, mental space, word recognition

Received 5/4/17; Revision accepted 10/16/17

From passing the basketball to a fellow teammate to handing over a knife at the dinner table, being able to adopt other people's visuospatial perspectives (VSPs) is pivotal for successfully engaging in a large variety of social interactions. Recent research has provided evidence that people adopt others' VSP spontaneously, computing the relative location of objects from another's orientation without being prompted to do so. We seem to be equipped with mechanisms allowing us to spontaneously take into account not only *whether* somebody else can or cannot see a certain object (visual perspective taking, or Level 1 perspective taking; see Flavell, Everett, Croft, & Flavell, 1981) but also *how* objects look from another's point of view (VSP taking, or Level 2 perspective taking; cf. Flavell et al., 1981).

For example, when being asked to give verbal descriptions of the spatial relations among an array of objects, observers spontaneously adopted the VSP of another person facing them (Lozano, Hard, & Tversky, 2007; Tversky & Hard, 2009; cf. Cavallo, Ansuini,

Capozzi, Tversky, & Becchio, 2016). Furthermore, when participants were asked to indicate the spatial locations of stimuli arranged vertically in front of them with left and right responses, and a task partner was sitting at a 90° angle next to them, they spontaneously adopted the other person's spatial reference frame, processing the stimuli in terms of the other's left and right (Freundlieb, Kovács, & Sebanz, 2016; Freundlieb, Sebanz, & Kovács, 2017).

Nearly all of the evidence for spontaneous VSP taking comes from tasks where an observed agent or task partner could physically act on objects. In these tasks, the physical location of an object (say, an apple; Cavallo et al., 2016) varied along a spatial dimension (e.g., appeared to the right vs. to the left of somebody else),

Corresponding Author:

Martin Freundlieb, Department of Cognitive Science, Central European University, Oktober 6 utca 7, 1051, Budapest, Hungary
E-mail: Freundlieb_Martin@phd.ceu.edu

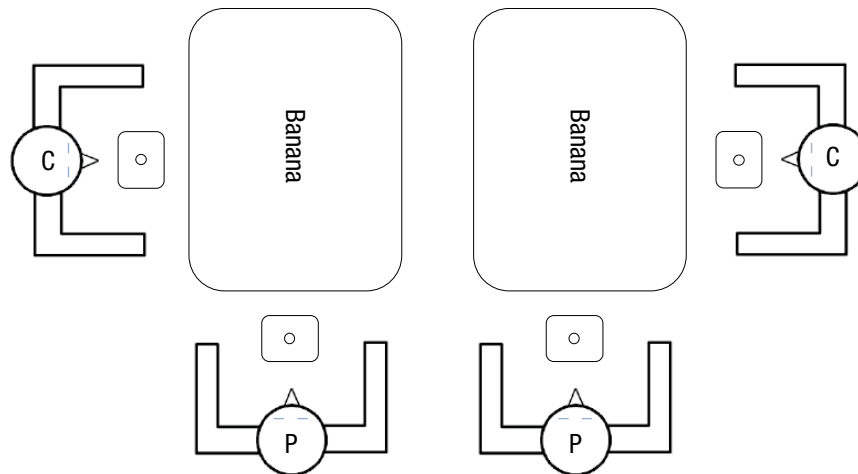


Fig. 1. Setup for Experiment 1. The participant (P) sat at a 90° angle to the confederate (C). An example of a trial in the congruent condition is shown on the left, and an example of a trial in the incongruent condition is shown on the right.

and participants' left versus right responses reflected how they, or the other person, would physically interact with the object. This raises the question of whether spontaneous VSP taking has a signature limit to the effect that it operates only in a shared physical realm, where different perspectives imply different actions on objects, or whether VSP taking extends to mental space¹ in which spatial relations matter for cognitive processes rather than for physical actions. For example, when the newspaper is oriented at a right angle from you at the breakfast table, will it be easier for you to read its headlines if the paper happens to be aligned with your partner's perspective? Reading is a prototypical case of a mental activity where objects (words) are manipulated by the mind rather than by our hands. Yet the sharing of semantic information may involve more sophisticated mechanisms that do not manifest in spontaneous effects on VSP taking. In particular, processing semantic information through reading is a cultural skill acquired late in our evolutionary history—it is therefore important to determine whether differences in VSPs can lead to spontaneous perspective taking in this domain.

Some first evidence for VSP taking in mental space comes from a task involving numerical cognition (Surtees, Apperly, & Samson, 2016; see also Elekes, Varga, & Király, 2016). In a joint numerical judgment task, participants were slower to indicate whether a number was smaller or larger than 5 when the numerical value of a digit was different for a task partner sitting opposite (e.g., on trials where they saw a “6” while their partner saw a “9”). This indicates that participants also computed the symbol from the other's viewpoint. However, importantly, as participants were asked to respond to smaller numbers with a left response and

to larger numbers with a right response, one could still argue that these results are based on the spatial-numerical associations of response codes (cf. Dehaene, Bossini, & Giraux, 1993) and, thus, still strongly relate to the action space.

To investigate whether spontaneous VSP taking occurs not only in physical but also in mental space, we developed a novel task in which participants were required to read words in order to perform a semantic categorization task. Across three experiments, we asked whether participants are faster in processing words when they are oriented such that they can be easily read by another individual, compared with an orientation that is the same from the participants' point of view but difficult to read from another's perspective.

Experiment 1

Participants sat at a 90° angle to a confederate and performed a semantic categorization task with words being displayed on a horizontally mounted computer screen (see Fig. 1). The stimuli always belonged to one of two categories, and participants were instructed to respond only to stimuli from their own category and not to respond to stimuli from the other category. All word stimuli were displayed in the same orientation (90° angle clockwise²) from the participants' perspective. The confederate sat to the right or to the left side of the participant. From the confederate's perspective, words thus either appeared the right way up (*congruent condition*), when he sat to the participant's left, or upside down (*incongruent condition*) when he sat to the participant's right. If participants spontaneously adopted the confederate's VSP, then it should be easier

for them to read words that are the right way up for the confederate and harder to read words that are upside down for the confederate, resulting in a congruency effect.

Method

Participants. We based our sample size on a previously published study testing VSP taking in a paired samples design (Freundlieb et al., 2016). Prior to data collection, we decided to test 16 participants and set the significance level (α) to .05. Sixteen participants (mean age = 20.06 years; 12 women; 15 right-handed) signed up for this study and received gift vouchers for their participation. All were naive to the purpose of the study, reported normal or corrected-to-normal vision, signed informed consent prior to the experiment, and were debriefed after the experiment. All 16 participants met the inclusion criterion of having more than 90% successful trials in each experimental condition.

Stimuli and apparatus. The stimuli consisted of single nouns in Hungarian (subtending between 3.3° and 7.3° of visual angle, depending on the length of the word). Each word item belonged to one of two categories: (a) animals or (b) fruits and vegetables. Each of the two categories contained 32 items that had been used in a prior study on social memory and were controlled for frequency (see Elekes et al., 2016; Table S1 in the Supplemental Material available online presents lists of all word items). In order to rule out carryover effects between the two experimental conditions, we randomly split the 32 items from each of the two categories in halves, resulting in four sublists (Animals 1, Animals 2, Fruits and Vegetables 1, Fruits and Vegetables 2), each containing 16 items. This way, participants responded to a unique list of word items in each of the two conditions. During the trials, single word items were always presented in the same orientation (90° clockwise from the participants' perspective) and at the same central position on a horizontally arranged 27-in. iMac (Mid 2011; see Fig. 1). The monitor was mounted at a height of about 63 cm from the floor. Responses were given on two button boxes (ioLab Response Box, GitHub, San Francisco, CA), which both the participant and the confederate placed on their lap. The button boxes were partially covered with a piece of carton so that only the button used to respond (i.e., the most central button) was visible.

Design and procedure. Viewing distance to the screen was approximately 70 cm, both for the participant and for the confederate, who was oriented at a 90° angle to the participant. A young adult male acted as the confederate. During the instruction phase, the experimenter

assigned both the participant and the confederate to one of the two categories (animals vs. fruits and vegetables) and asked them to respond with a button press only to word items from their own category and not to respond to word items from the other's category (*go/no-go task*). Each trial started with the presentation of a fixation cross (subtending 1.31° of visual angle, presented in the center of the screen) for 350 ms. Subsequently, the screen turned blank for 100 ms, after which a word item was shown for 1,200 ms. The word items were randomly chosen from two sublists (e.g., Animals 1 and Fruits and Vegetables 1 in the first condition and then Animals 2 and Fruits and Vegetables 2 in the second condition), and each sublist was—consecutively—repeated four times per condition, with items presented in a random order. Participants performed two conditions (congruent and incongruent), each containing 128 trials (2 categories \times 16 items per sublist \times 4 repetitions of each sublist). They were asked to respond as fast and as accurately as possible and not to tilt their heads during the experiment.

To establish different congruency relations, we varied the sitting position of the confederate. While the participant always sat at the narrow end of the rectangular screen, the confederate switched between the two long ends during the experiment. In the congruent condition, the confederate sat to the participant's left, so that words were oriented toward him, or the right way up. In contrast, in the incongruent condition, the confederate sat to the participant's right, so that words appeared upside down (see Fig. 1). Before each condition, eight practice trials familiarized the participants with the task. These were later excluded from the statistical analysis.

The order of conditions (congruent vs. incongruent), the assigned category (animals vs. fruits and vegetables), as well as the starting sublist (Animals 1 vs. Animals 2 vs. and Fruits and Vegetables 1 vs. Fruits and Vegetables 2, respectively) was counterbalanced across participants.

Data analysis. Data were collected for participants only. Errors (i.e., missed button presses during participants' own trials or button presses during the confederate's trials) and reaction times (RTs) more than 2 standard deviations from each participant's condition mean were excluded from the RT analysis. Both the two condition means for correct-response RTs and errors for each participant were subjected to separate two-tailed, paired-samples *t* tests.

Results

For the RT analysis, 0.34% of the trials were removed as errors, and 4.49% were removed for being more than

2 standard deviations away from each participant's condition mean, leaving 95.17% of the raw data as correct-response trials. Generally, the removal of these outliers did not result in changes of the significance patterns observed in this experiment. Comparing the number of errors in the congruent and incongruent conditions showed that participants made significantly more errors in the congruent condition ($M = 1.56$, $SD = 1.15$) compared with the incongruent condition ($M = 0.56$, $SD = 0.81$), $t(15) = 4.14$, $p = .001$, $d = 1.0$.³

The RT analysis revealed that participants were significantly faster in the congruent ($M = 578.0$, $SE = 15.29$) than in the incongruent ($M = 618.4$, $SE = 19.47$) condition, $t(15) = 4.49$, $p < .001$, two-tailed, $d = 1.13$ (see Fig. 2). A post hoc power analysis (using G* Power; see Faul, Erdfelder, Lang, & Buchner, 2007) revealed that, given the mean of difference $M_z = 40.33$, the standard deviation of difference $SD_z = 35.00$, and the effect size of $d = 1.13$, we achieved a power of $1 - \beta = 0.98$. In order to test whether the specific category assigned to the participants (animals vs. fruits and vegetables), the order of conditions (starting with the congruent vs. the incongruent condition), or the order of the sublists (Animals 1, Animals 2, Fruits and Vegetables 1, Fruits and Vegetables 2) influenced the results, we conducted a repeated measures analysis of variance (ANOVA) with congruency as a within-subjects factor and category, order of condition, and order of sublists as between-subjects factors. The results yielded only a main effect of congruency, $F(1, 11) = 21.11$, $p = .001$, $\eta_p^2 = .66$, but no effect of category, $F(1, 11) < 1.0$, $p > .250$, $\eta_p^2 < .01$; order of conditions, $F(1, 11) < 1.0$, $p > .250$, $\eta_p^2 = .06$; and order of sublists, $F(1, 11) < 1.0$, $p > .250$, $\eta_p^2 = .09$, or any two-way interactions between congruency and the between-subjects factors, all F s < 1.8 , p s $> .211$, $\eta_p^2 < .24$.

These results suggest that participants spontaneously adopted the other person's VSP when performing a semantic categorization task together. This facilitated the processing of words oriented such that the confederate could easily read them and/or impaired processing of words that were oriented upside down from the confederate's perspective. An open question is whether the active engagement of the confederate was necessary for triggering spontaneous VSP taking in mental space.

Experiment 2

Experiment 2 investigated whether the mere presence of another person with a diverging VSP is sufficient for participants to spontaneously adopt the other's VSP. In one block, the confederate performed the same task as in Experiment 1, while in the other block he was instructed to just watch the stimuli on the screen.

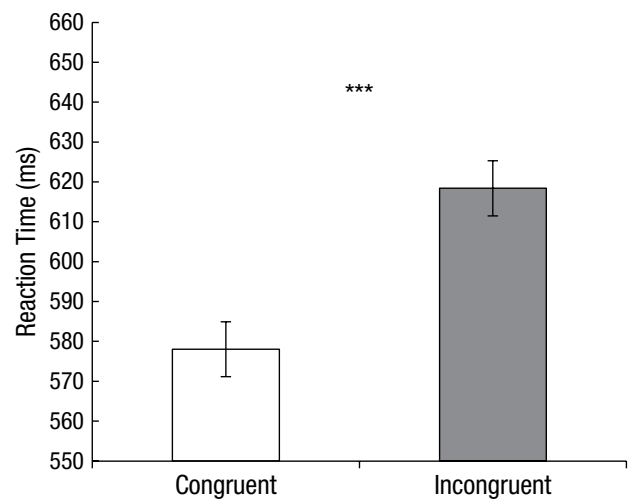


Fig. 2. Mean reaction time in the congruent and incongruent conditions in Experiment 1. Error bars display within-subjects confidence intervals according to Loftus and Masson (1994). The asterisks indicate a significant difference between conditions ($p < .001$).

Because reading is a mental activity that does not necessarily manifest in physical actions, a passive individual can still engage in it. Therefore, we predicted that the presence of a passive confederate with a divergent VSP would be sufficient for participants to adopt his VSP, leading to a congruency effect.

Method

Participants. Prior to data collection, we decided to obtain data from 32 participants and set the significance level (α) to .05. Changing the experimental paradigm to a 2×2 factorial design led us to double our initial sample size. This sample size is identical to that utilized in a previously published study on VSP taking that used a similar factorial design (Surtees et al., 2016). Thirty-three participants (mean age = 21.68 years; 20 women; 29 right-handed) signed up for this study and received gift vouchers for their participation. One participant with severely reduced vision forgot to bring his glasses and was therefore excluded from the analysis. Each of the 32 participants (mean age = 21.66 years; 20 women; 28 right-handed) was naive to the purpose of the study, reported normal or corrected-to-normal vision, signed informed consent prior to the experiment, and was debriefed after the experiment.

Stimuli and apparatus. The stimuli and the apparatus were identical to those in Experiment 1.

Procedure. Participants performed both of the two conditions (congruent vs. incongruent) in two different blocks (other-active vs. other-passive). Each condition contained 128 trials, resulting in a total of 512 trials in Experiment 2.

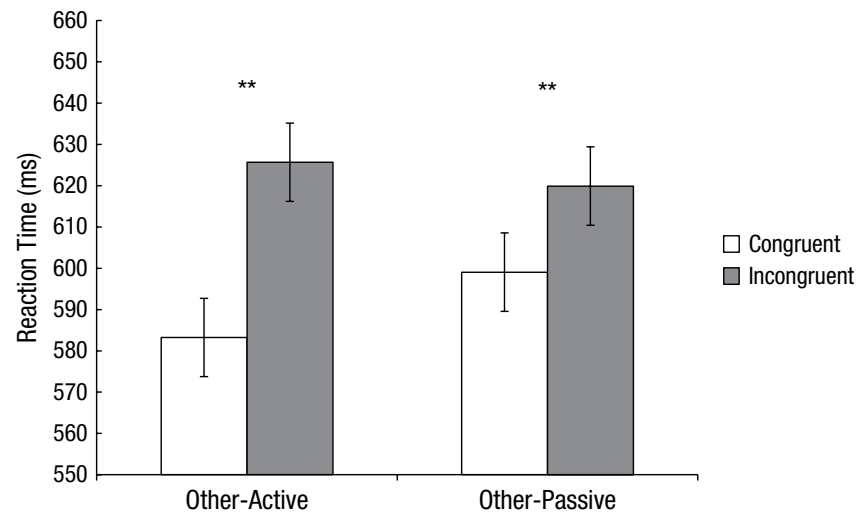


Fig. 3. Mean reaction time in the other-active and other-passive blocks in Experiment 2. Error bars display within-subjects confidence intervals according to Loftus and Masson (1994). The asterisks indicate significant differences between conditions ($p < .01$).

The participants' task was the same as in Experiment 1. While the other-active block was an exact replication of Experiment 1, in the other-passive block, the confederate was instructed not to respond to his category but instead just to watch the stimuli on the screen. Before each condition, eight practice trials familiarized the participants with the task. These were later excluded from the analysis.

The order of congruency (congruent vs. incongruent), the order of the blocks (other-active vs. other-passive first), the order of the starting sublist (Animals 1 vs. Animals 2 vs. Fruits and Vegetables 1 vs. Fruits and Vegetables 2), and the assigned category (animals vs. fruits and vegetables) was counterbalanced across participants.

Data analysis. Errors (i.e., missed button presses during participants' own trials or button presses during the confederate's trials) and RTs more than 2 standard deviations from each participant's condition mean were excluded from the RT analysis. Both of the two condition means for correct-response RTs for each participant as well as his or her errors were subjected to separate two-way, repeated measures ANOVAs with the factors congruency (congruent vs. incongruent) and activity other (other-active vs. other-passive).

Results

We removed 0.31% of the trials as errors and 4.34% as outliers, leaving 95.35% of the raw data as correct-response trials. The removal of these outliers did not result in changes of the significance patterns observed

in this experiment. The error analysis did not reveal any statistically significant results for congruency, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 < .01$; activity other, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 = .01$; or their interaction, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 < .01$.

The RT analysis revealed a significant main effect of congruency, $F(1, 31) = 40.68$, $p < .001$, $\eta_p^2 = .57$, with RTs being generally faster during the congruent than during the incongruent condition (see Fig. 3). This effect was moderated by a significant interaction between congruency and activity other, $F(1, 31) = 5.56$, $p = .025$, $\eta_p^2 = .15$. The difference score between congruent and incongruent trials was significantly higher in the other-active condition ($M = 42.42$, $SE = 7.43$), compared with the other-passive condition ($M = 20.85$, $SE = 5.99$), $t(31) = 2.36$, $p = .025$, two-tailed, $d = 0.41$. In addition, post hoc t tests revealed a significant congruency effect between congruent and incongruent trials both in the other-active condition, $t(31) = 5.71$, $p < .001$, $d = 1.0$, as well as the other-passive condition, $t(31) = 3.48$, $p = .002$, two-tailed, $d = 0.61$. Finally, there was no main effect of activity other, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 = .01$.

We conducted further analyses to check whether the results of Experiments 1 and 2 were influenced by demand characteristics. Given the within-subjects design used in both experiments, the difference between congruent and incongruent trials could in principle have been obvious to participants and guided their performance. However, during the debriefing, none of the participants reported anything that indicated their being aware of our hypotheses. As one block in Experiment 2 (other-active) was a direct replication of Experiment

1, we pooled the performance of those participants from Experiments 1 and 2 who started with the congruent condition first ($n = 16$) and compared it with the performance of those participants from Experiments 1 and 2 who started with the incongruent condition first ($n = 16$). We then conducted an independent-samples t test ($N = 32$) between the two groups to investigate whether the previously found effects can be established in a between-subjects analysis. The results showed a significant difference in RTs between the congruent-first ($M = 573$, $SE = 12.1$) and the incongruent-first ($M = 645.54$, $SE = 23.3$) group, $t(30) = 2.73$, $p = .01$, two-tailed, $d = 0.97$. This speaks against the possibility that the findings from Experiments 1 and 2 were based on demand characteristics.

In summary, the results of Experiment 2 replicate Experiment 1 and indicate that participants were sensitive to the fact that in the absence of any overt responses, the confederate could still read the stimuli. While this demonstrates that VSP taking in mental space can occur in the absence of direct evidence of another's engagement, we also found that the effect of VSP taking was larger when the confederate actively performed a task that required reading the words. A related question raised by the results of Experiment 2 is whether the congruency effect observed in the passive condition might simply be due to the bodily orientation of the observer rather than to him reading the words.

Experiment 3

Another agent's facial features (such as his forehead, eyes, and nose) facing toward objects may automatically trigger a shift of attention regardless of whether that agent has visual access to the stimuli or not (cf. Heyes, 2014). Alternatively, the other's visual access may be crucial for spontaneous VSP taking to occur. To address this question, we manipulated the confederate's ability to see the stimuli in Experiment 3. If participants' responses systematically changed in accordance with the confederate seeing or not seeing the stimuli, this would support the claim that ascribing visual access to another agent is a precondition for adopting his or her VSP and processing the stimuli as if seen from that person's point of view.

Method

Participants. Thirty-two participants (mean age = 21.94 years; 20 women; 29 right-handed) signed up for this experiment and received gift vouchers for their participation. All were naive to the purpose of the experiment, reported normal or corrected-to-normal vision, signed informed consent prior to the experiment, and were debriefed after

the experiment. All 32 participants met the inclusion criterion of having more than 90% successful trials in each experimental condition.

Stimuli and apparatus. The stimuli were identical to those in Experiments 1 and 2. The only difference in the apparatus was that the confederate wore a pair of lift-front goggles (Lux Optical, Worldwide Euro Protection, Luxembourg) throughout the experiment. These goggles had small shutters that could be either lifted up (in which case, one had unhindered vision through transparent Plexiglas) or flapped down (in which case, black tape on the shutters blocked vision).

Procedure. Participants performed the two conditions (congruent vs. incongruent) in two different blocks (blindfolded vs. seeing). Each condition contained 128 trials, resulting in a total of 512 trials, as in Experiment 2. The participants' task was identical to that in Experiments 1 and 2. While the other-seeing block replicated the other-passive block in Experiment 2 (the only exception being that the confederate wore the transparent goggles), in the other-blindfolded block, the confederate was instructed to flap the shutters of his goggles down and wait until the end of the block. Before each condition, eight practice trials familiarized the participants with the task. These were later excluded from the analysis.

The order of congruency (congruent vs. incongruent), the order of the blocks (blindfolded vs. seeing first), the order of the starting sublist (Animals 1 vs. Animals 2 vs. Fruits and Vegetables 1 vs. Fruits and Vegetables 2), and the assigned category (animals vs. fruits and vegetables) were counterbalanced across participants.

Data analysis. Errors (i.e., missed button presses during participants' own trials or button presses during the confederate's trials) and RTs more than 2 standard deviations from each participant's condition mean were excluded from the RT analysis. Both the two condition means for correct-response RTs for each participant as well as his or her errors were subjected to two-way, repeated measures ANOVAs with the factors congruency (congruent vs. incongruent) and vision other (other-seeing vs. other-blindfolded).

Results

We removed 0.56% of the trials as errors and 4.65% as outliers, leaving 94.79% of the raw data as correct-response trials. The removal of these outliers did not result in changes of the significance patterns observed in this experiment. The error analysis revealed a significant main effect of vision other, $F(1, 31) = 5.18$, $p = .030$, $\eta_p^2 = .14$, showing that participants made more errors when the confederate was blindfolded ($M = 1.25$,

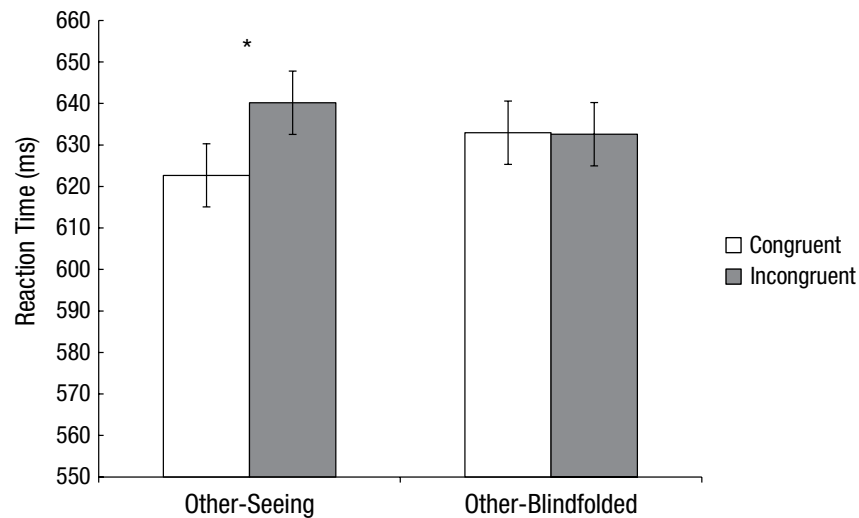


Fig. 4. Mean reaction time in the other-seeing and other-blindfolded blocks in Experiment 3. Error bars display within-subjects confidence intervals according to Loftus and Masson (1994). The asterisk indicates a significant difference between conditions ($p < .05$).

$SE = 0.21$) than when he had visual access to the stimuli ($M = 0.83$, $SE = 0.16$). Neither the main effect of congruency, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 = .03$, nor the interaction between the two factors, $F(1, 31) = 1.56$, $p = .221$, $\eta_p^2 = .05$, was significant.

The RT analysis revealed a significant interaction between congruency and vision other, $F(1, 31) = 5.88$, $p = .021$, $\eta_p^2 = .16$. In post hoc analyses, pairwise comparisons showed a significant difference in RTs between the congruent ($M = 622.68$, $SD = 68.81$) and the incongruent ($M = 640.18$, $SD = 73.02$) blocks only in the other-seeing condition, $t(31) = 2.22$, $p = .034$, two-tailed, $d = 0.39$, but not in the other-blindfolded condition, $t(31) < 1.0$, $p > .250$, two-tailed, $d < 0.01$ (see Fig. 4). The other pairwise comparisons did not reach significance (all $ps > .218$, all $ds < 0.2$). Furthermore, there was neither a significant main effect of congruency, $F(1, 31) = 1.72$, $p = .200$, $\eta_p^2 = .05$, nor of vision other, $F(1, 31) < 1.0$, $p > .250$, $\eta_p^2 < .01$. This indicates that the confederate's ability to see the stimuli was necessary for participants to adopt his VSP, while cues about the orientation of the confederate relative to the stimuli (e.g., the direction of his body, forehead, nose) were not sufficient for triggering spontaneous VSP taking.

General Discussion

The aim of this study was to investigate whether spontaneous VSP taking occurs in mental space, where spatial relations matter for mental rather than physical actions. Across three experiments, we found that participants reliably adopted the VSP of a confederate in

the context of a semantic categorization task that involved reading words. Specifically, we found that participants were faster to categorize words that were oriented in an upright way from the point of view of a confederate, compared with words oriented upside down from the confederate's point of view. This shows that VSP taking is not restricted to situations involving physical interactions with objects or action planning (cf. Creem-Regehr, Gagnon, Geuss, & Stefanucci, 2013; Freundlieb et al., 2016; Freundlieb et al., 2017) because it also extends to the mental space.

Our findings advance the understanding of perspective taking in two ways. First, they underline the pervasiveness of the phenomenon of VSP taking in human interactions—operating not only in physical space, where it potentially facilitates motor coordination, but also in mental space, where it might facilitate interpersonal communication. The propensity to adopt other people's VSP therefore helps overcome differences not only in the ways in which our bodies relate to the world, but also in the way our minds process information, helping to create shared meaning. Second, our results indicate continuity of the mechanisms underlying spontaneous VSP taking, allowing for processing of objects that are manipulated mentally from another's perspective in much the same way as objects that are manipulated physically. This, in turn, can inform the long-standing debate on the origins of mental perspective taking (or theory of mind) and its relation to spatial perspective taking (Butterfill & Apperly, 2013).

Furthermore, our results provide evidence for spontaneous VSP taking in mental space. The orientation of

the words and the sitting position of the confederate were completely irrelevant for the participants' task, and participants were never prompted to adopt the other's perspective. This extends earlier studies on VSP taking in physical space, where participants were asked to provide responses about spatial arrangements of objects from a particular perspective (Cavallo et al., 2016; Furlanetto, Becchio, Samson, & Apperly, 2016).

A further important finding of our study is that participants spontaneously adopted somebody else's VSP even if that other person was not explicitly assigned a task and was just passively observing the stimuli (Experiment 2). Previous studies have shown evidence for spontaneous VSP taking only if the task was performed within a cooperative context (Surtees et al., 2016) or if the responses given by the other person indicated his or her constant engagement in the shared task (Freundlieb et al., 2016). We believe that this discrepancy between earlier results of VSP taking in physical space and the present results can be explained by differences in the tasks involved. For example, as the rules of the task used by Freundlieb and colleagues (2016) were completely arbitrary, it was only through the confederate's overt responses that participants could verify the confederate's participation in the task. In contrast, participants in the current task could still assume that the passive confederate processed the stimuli in a meaningful way—as written words automatically trigger reading and recognition processes (especially given that the confederate was instructed to watch the stimuli; cf. Strijkers, Bertrand, & Grainger, 2015; Stroop, 1935). This “passive” participation (likely involving reading) might have been sufficient for participants to perceive the task as being interactive (or as a “team context”; see Surtees et al., 2016) and, hence, to spontaneously adopt the confederate's VSP. Finally, the results of Experiment 2 suggest that although the activity of the other was not a necessary factor for the congruency effect to occur, it seemed to have further increased VSP-taking effects.

But what exactly led to the observed effects in the current study in the first place? It could be claimed that the congruency effects found in Experiments 1 and 2 were simply based on a domain-general mechanism picking up directional cues (such as somebody else's body orientation) and thereby redirecting participants' attention (cf. Heyes, 2014). Importantly, such a mechanism would elicit the same effect regardless of whether the agent exhibiting such directional cues had visual access to the stimuli or not. However, in Experiment 3, we replicated the results obtained in Experiment 2 and showed that spontaneous VSP taking disappeared if the confederate's visual access to the stimuli was blocked. Importantly, this corroborates that participants were taking into account how the confederate saw the stimuli

and rules out the possibility that directional cues about the confederate's facial features (such as the orientation of his body, forehead, and nose) were sufficient for triggering spontaneous VSP-taking effects.

We think that the mechanism underlying the effects reported here involves a modulation of the processes involved in word reading that was prompted by the spontaneous adoption of the other person's VSP. Words are not processed and transformed as integral units over the entire range of orientations (Koriat & Norman, 1985). Instead, Koriat and Norman (1985) proposed that when stimuli are close to the upright canonical orientation ($\pm 60^\circ$), word recognition relies on whole-word units, whereas at more extreme orientations (beyond 120° deviations), it appears to be based on sequential letter identification. They suggest that for intermediate orientations (60° – 120° , which coincides with the orientation used in our study) “word recognition may rely on units larger than single letters” (Koriat and Norman, 1985, p. 507). In our study, adopting the other person's VSP might have led participants to process the words more holistically in the congruent condition and in a letter-by-letter fashion in the incongruent condition, creating the observed differences in reaction times. Future experiments will have to determine and disentangle the effects of VSP taking on semantic, orthographic, and lower visual levels of word processing.

Action Editor

Edward S. Awh served as action editor for this article.

Author Contributions

M. Freundlieb, Á. M. Kovács, and N. Sebanz developed the study concept. All authors contributed to the study design. Testing and data collection were performed by M. Freundlieb. M. Freundlieb analyzed the data and drafted the manuscript. N. Sebanz and Á. M. Kovács provided critical revisions. All authors approved the final version of the manuscript for submission. Á. M. Kovács and N. Sebanz contributed equally to this study.

Acknowledgments

We thank Stephen Butterfill and Fruzsina Elekes for useful discussions and Matteo Frisoni and Anna Maria Lisincki for their help running this study.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by the European Research Council, partly under the European Union's Seventh Framework Programme (FP7/2007-2013) Grants 284236-REPCOLLAB and 616072-JAXPERTISE.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797617740973>

Notes

1. Note that we use the term *mental space* to highlight the contrast to the physical interaction space. It therefore differs from the notion of mental space as defined by Fauconnier (1994).
2. As the general reading direction is from the top of a page to the bottom of a page, we thought it would be more natural for participants to read top to bottom (or clockwise) compared with bottom to top (or anticlockwise). Therefore, we chose to use stimuli that were rotated in a clockwise manner. However, Yu, Park, Gerold, and Legge (2010) showed no difference in performance when participants read words that were rotated clockwise versus counterclockwise.
3. The fact that there was a significant effect of congruency in the error analysis could indicate that there was a speed/accuracy trade-off. However, this pattern was not observed in subsequent experiments. As participants generally committed very few errors (1.2% and 0.4% in the congruent and incongruent condition, respectively), the speed/accuracy trade-off in this experiment is unlikely to account for the robust effect in RTs.

References

- Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language, 28*, 606–637.
- Cavallo, A., Ansuini, C., Capozzi, F., Tversky, B., & Becchio, C. (2016). When far becomes near: Perspective taking induces social remapping of spatial relations. *Psychological Science, 28*, 69–79.
- Creem-Regehr, S. H., Gagnon, K. T., Geuss, M. N., & Stefanucci, J. K. (2013). Relating spatial perspective taking to the perception of other's affordances: Providing a foundation for predicting the future behavior of others. *Frontiers in Human Neuroscience, 7*, Article 596. doi:10.3389/fnhum.2013.00596
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.
- Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous Level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93–103.
- Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge, England: Cambridge University Press.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*, 99–103.
- Freundlieb, M., Kovács, Á. M., & Sebanz, N. (2016). When do humans spontaneously adopt another's visuospatial perspective? *Journal of Experimental Psychology: Human Perception and Performance, 42*, 401–412. doi:10.1037/xhp0000153
- Freundlieb, M., Sebanz, N., & Kovács, Á. M. (2017). Out of your sight, out of my mind: Knowledge about another person's visual access modulates spontaneous visuospatial perspective-taking. *Journal of Experimental Psychology: Human Perception and Performance, 43*, 1065–1072.
- Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in Level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 158–163. doi:10.1037/xhp0000138
- Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science, 9*, 131–143.
- Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 490–508.
- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.
- Lozano, S. C., Hard, B. M., & Tversky, B. (2007). Putting action in perspective. *Cognition, 103*, 480–490.
- Strijkers, K., Bertrand, D., & Grainger, J. (2015). Seeing the same words differently: The time course of automaticity and top-down intention in reading. *Journal of Cognitive Neuroscience, 27*, 1542–1551.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.
- Surtees, A., Apperly, I. A., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition, 150*, 43–52.
- Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition, 110*, 124–129.
- Yu, D., Park, H., Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text. *Journal of Vision, 10*(2), Article 21. doi:10.1167/10.2.21