Title
Partner effects and individual differences on perspective taking

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Loy, Jia
Demberg, Vera

Publication Date
2022

Peer reviewed

Partner effects and individual differences on perspective taking

Jia E. Loy (jialoy@coli.uni-saarland.de) and Vera Demberg (vera@coli.uni-saarland.de)
Department of Computer Science and Department of Language Science and Technology, Saarland University

Abstract

Spatial perspective taking, in which people mentally adopt another person's view of the world, is a crucial component of everyday communication. We investigate spatial perspective taking in listeners interpreting ambiguous instructions from a partner, looking at how this behaviour varies with a human vs. computer partner (Exp. 1 and 2), and with individual differences in social and cognitive abilities (Exp. 3). Listeners' perspective taking tendencies vary with their individual differences in spatial orientation ability, with more othercentricism associated with better spatial orientation. In addition, partner identity influences perspective taking; however, in contrast to previous work, we find higher levels of egocentricism with a computer than a human partner. Our results highlight the importance of taking into account both external factors and individual differences in understanding spatial perspective taking.

Keywords: perspective taking; individual differences; audience design

Introduction

When communicating about the world, people often have to describe spatial locations to an interlocutor (e.g., when giving directions). If interlocutors' viewpoints differ, speakers can opt to formulate the description from their own (an egocentric) or their partner's (an othercentric) perspective. Considerable work on perspective choice and interaction has documented behaviour in spatial description tasks. Studies show that speakers exhibit adaptive tendencies in perspective taking based on various cognitive and social factors; for instance, speakers tend to be more egocentric when adopting a partner's perspective is computationally difficult (Galati & Avraamides, 2015), and conversely more othercentric when they perceive their partner's capabilities to be limited in some way (Schober, 1993; Shelton & McNamara, 2004). These findings are consistent with the assumption that perspective taking involves cognitive effort (cf. Horton & Keysar, 1996), but that speakers may be willing to invest in this if pragmatically motivated, such as to ensure mutual understanding (cf. Clark & Wilkes-Gibbs, 1986).

However, egocentric and othercentric descriptions are not the only linguistic choices; speakers may produce ambiguous spatial descriptions (e.g., "it's on the left"), leaving the burden of disambiguation to the listener. Thus, an equally important question is how and when listeners adapt to a speaker's spatial perspective in comprehension.

Duran, Dale, and Kreuz (2011) investigated spatial perspective taking in listeners using a computerised task in which participants followed instructions from a virtual speaker who referenced one of two identical objects on a tabletop with potentially ambiguous spatial terms *front*, *back*, *left*, or *right* (e.g., "give me the folder on the left"). The authors manipulated whether the speaker and participant had the same or different perspectives of the table. Thus, when perspectives differed, participants could respond egocentrically or othercentrically. Results showed that spatial perspective taking performance decreased with increasing misalignment between the participant and speaker's perspectives, supporting the view that perspective taking is cognitively demanding.

Duran et al. (2011) additionally classified participants as "egocentric", "othercentric", or "mixed" based on their dominant mode of response. The authors observed that mixed responders were in the minority, with most falling on extreme ends of a bimodal distribution with a roughly even split. A subsequent experiment showed that altering listeners' expectations about their partner's perspective taking ability by telling them that their partner was a real human increased the proportion of egocentric responders, consistent with the view that spatial perspective choice is part of a collaborative effort sensitive to one's beliefs about an interlocutor. Notably, the emergence of distinct response groups may also reflect participant tendencies based on individual differences, which may modulate one's willingness to take perspective.

Although Duran et al. did not investigate individual differences in their study, a growing body of research highlights a relationship between various cognitive and social abilities and spatial perspective taking. One line of research proposes that spatial perspective taking is an embodied cognitive process involving a mental rotation of one's self into a target orientation (Hegarty & Waller, 2004; Kessler & Thomson, 2010). To test this, Hegarty and colleagues developed the Object Perspective Test (OPT; Kozhevnikov & Hegarty, 2001; Hegarty & Waller, 2004), a test of spatial orientation in which participants have to take on an imagined perspective within an array of objects and indicate a direction to a target object. Confirmatory factor analyses showed that this test loaded on the same factor as other tests of spatial orientation ability (e.g., the Money Road Map Test; Schultz, 1991), supporting the view that spatial perspective taking draws on the ability to make egocentric spatial transformations.

Another cognitive ability often implicated is inhibitory control, based on the rationale that taking another perspective requires ignoring a dominant egocentric perspective. Frick and Baumeler (2017) found a relationship between spatial

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

perspective taking and inhibitory control in children, even after controlling for other individual differences such as age and IQ. Studies on (non-spatial) perspective taking also report a role of this ability (Brown-Schmidt, 2009; Wardlow, 2013), suggesting that the cognitive challenge in perspective taking lies in the need to suppress one's egocentric perspective.

Finally, a recent line of work highlights a link between spatial and social skills, in particular social and communication ability as measured by the Autism Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). This research is motivated by the argument that taking another's perspective is inherently social in nature (cf. Tversky & Hard, 2009); correspondingly, results tend to show that lower social skills are associated with poorer perspective taking performance (Kessler & Wang, 2012; Xiao, Xu, Sui, & Zhou, 2021; Job, Kirsch, Inard, Arnold, & Auvray, 2021). Individuals with Autism Spectrum Disorders (ASDs) have also been shown to perform more poorly on perspective taking tasks, a finding that has been attributed to social impairments such as difficulty with theory of mind (Hamilton, Brindley, & Frith, 2009).

While previous studies support a relationship between spatial perspective taking and various individual differences, many of these have so far not been investigated together. The majority of studies also focus on speakers' perspective choice; less is known about how such differences may modulate this behaviour in comprehension. Here, we investigate spatial perspective taking in listeners, with the aim of exploring whether othercentric perspective tendencies are linked to an individual's social and cognitive abilities. We employ a collaborative task in which participants carry out a partner's instructions, which can be interpreted egocentrically or othercentrically. Experiments 1 and 2 are conceptual replications of Duran et al.'s comparison of spatial perspective taking with a human vs. computer partner. Surprisingly, we find the opposite result in that listeners are more egocentric with a computer than a human partner in both experiments. In Experiment 3, we explore the role of individual differences in spatial perspective taking with a computer partner. We find that othercentric tendencies are associated with spatial orientation ability, but not with social skills or inhibitory control.

Experiment 1

Participants respond to pre-recorded instructions from a partner requesting for objects on a tabletop in a web-based virtual environment. We varied the location of the partner's avatar such that the participant and partner are either side-by-side or across from each other. In the latter configuration, requests for an object on the left/right are spatially ambiguous and can be interpreted egocentrically or othercentrically. We focused on left/right references since larger effects were reported for these in Duran et al. (2011). In addition, we varied partner identity (computer vs. human) between-subjects in attempt to replicate Duran et al.'s (Exp 1 vs. 3) partner effects.[1]

[1]Preregistration details can be found at `https://osf.io.cz42t`

Methods

Participants 524 participants were recruited on Amazon Mechanical Turk (AMT). We excluded data from participants who: (a) were (self-reported) nonnative speakers of English (4), (b) failed to meet accuracy criteria (<80% accuracy on unambiguous trials; 71), or (c) indicated suspicion about the authenticity of their partner or the interaction (175). Thus, the final dataset consisted of 274 participants: 134 and 140 in the human and computer partner conditions respectively.

Materials, design and procedure The task was described as an activity in which two users carry out a joint task in a virtual workspace. The participant's goal was to move objects in the workspace following their partner's instructions. Displays showed the objects arranged in pre-determined locations (top, left, bottom, or right) on a tabletop, and two avatars representing the participant and their partner. The participant's avatar was always at the bottom of the table; we manipulated whether the partner's avatar was next to the participant (same perspective) or across from the participant (different perspective; see Figure 1).

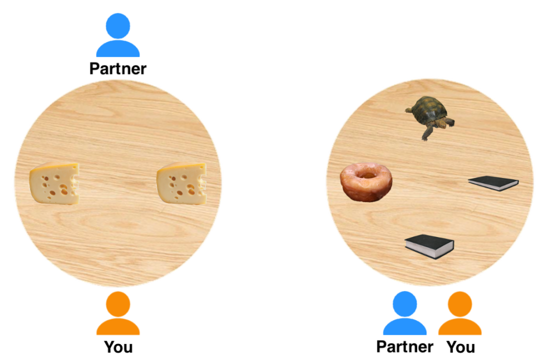

Figure 1: Example of a critical trial in the different perspective condition (left) and a filler trial (right).

On each trial, participants heard a pre-recorded utterance from their partner requesting an object, which they had to click and drag to them. The experiment had 12 critical trials (six same and six different perspective) and 36 filler trials. Critical utterances were of the form "Give me the <*object*> on the left/right"). The accompanying display showed two identical objects in the left and right positions on the table. Thus, on different perspective trials the critical utterance was always spatially ambiguous.

Filler trials were included to reduce the salience of critical displays. These varied the number of objects shown (two/three/four), and included different contrasts relevant for referent identification: colour (e.g., red/green apple; $n = 8$); size (e.g., long/short ruler; $n = 8$); location (left/right, with the addition of distractor objects; $n = 8$); and no contrast (identifiable by the bare noun; $n = 12$). Half of the filler trials used the same perspective configuration and the other half used the different perspective configuration.

We manipulated partner identity as either a human (another worker on AMT) or a simulated computer. Participants were told the task took place in real-time; in fact partner utterances were pre-recorded and we simulated live interaction with variable utterance onset latencies and waiting times (e.g., a delay ostensibly for a human partner to sign up). Human utterances were recorded by a female native speaker of North American English; computer utterances were synthesised with the Apple Macintosh built-in text-to-speech function ("Agnes" voice). After the task, participants completed a post-test questionnaire with questions aimed at verifying whether those in the human partner condition suspected the authenticity of the interaction (as in Duran et al., 2011).

Results

The experiment yielded two dependent measures: the object selected by participants on each trial and their response times. Following Duran et al. (2011), we categorised participants as "egocentric" or "othercentric" based on their dominant pattern of object selection on different perspective critical trials (>70% of trials in the relevant response category); participants who were neither were categorised as "mixed".

We conducted two planned statistical analyses. The first examined perspective taking based on participants' object selection, using logistic mixed effects regression to analyse the outcome of whether or not participants chose the object from an egocentric perspective on each trial. The second examined processing cost based on trial response times (from utterance offset), using linear mixed effects regression to analyse log-transformed response times. Models included perspective (same vs. different) and partner (human vs. computer) as fixed effects (sum-coded), and participant and item random intercepts and by-participant random slopes for perspective.

Table 1: Breakdown of responder types in each experiment.

Experiment	Partner	Egocentric	Othercentric	Mixed
1	Human	108 (81%)	17 (13%)	9 (6%)
	Computer	125 (89%)	13 (9%)	2 (1%)
2	Human	9 (8%)	95 (83%)	10 (9%)
	Computer	39 (31%)	81 (64%)	7 (6%)
3	Computer	68 (38%)	48 (27%)	64 (36%)

Distribution of responders Table 1 shows the breakdown of egocentric, othercentric, and mixed responders. Following Duran et al., we compared the rate of egocentric and othercentric responders across partner conditions. For egocentric responders, there was a marginally significant 8% decrease in the human compared to computer partner condition, $\chi^2(1) = 3.40, p = .06$. For othercentric responders, there was a non-significant 5% increase in the human compared to computer partner condition, $\chi^2 = 0.80, p = .3$.

Perspective taking Figure 2 shows the percentage of trials on which participants chose the object from an egocentric perspective. The model showed main effects of perspective, with participants less likely to respond egocentrically on different perspective trials, $\beta = -2.43$, $SE = 1.21$, $p = .04$; and of partner, with participants more likely respond egocentrically with a computer partner, $\beta = 1.69$, $SE = 0.70$, $p = .02$.

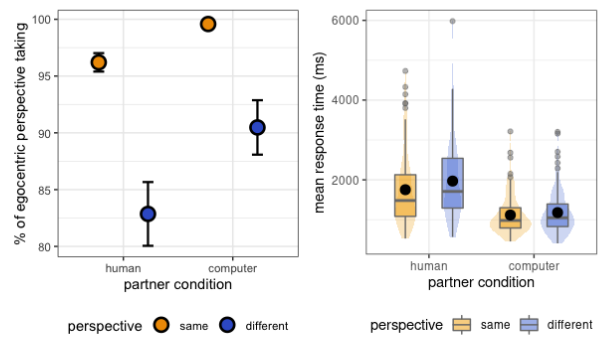

Figure 2: Experiment 1 results. Left: Percentage of trials on which participants chose the object from an egocentric perspective. Error bars represent ±1 standard error of by-participant means. Right: Raw mean trial response times.

Processing cost Figure 2 shows participants' mean response times in the experiment. The model showed main effects of perspective, with longer response times observed in the different perspective condition, $\beta = 0.06$, $SE = 0.02$, $t = 2.81$; and of partner, with shorter response times observed in the computer partner condition, $\beta = -0.45$, $SE = 0.06$, $t = -7.97$. To investigate whether longer response times with the human partner were due to a larger proportion of othercentric responders, we ran the same analysis including responder group (egocentric vs. othercentric; sum-coded) as a predictor. This model revealed a perspective by responder group interaction, $\beta = 0.21$, $SE = 0.06$, $t = 3.17$. Separate models confirmed that othercentric responders were significantly slower on different compared to same perspective trials, $\beta = 0.26$, $SE = 0.07$, $t = 3.25$ while this difference was not significant in egocentric responders, $t = 1.28$.

Discussion

The main finding from Experiment 1 was that partner identity influenced listeners' spatial perspective taking behaviour. This is in line with previous work which shows that perspective taking tendencies are sensitive to social cues about one's partner (e.g., Schober, 2009; Duran et al., 2011). However, in contrast to Duran et al., we observed higher rates of egocentricism with a computer compared to a human partner. We return to this point in the general discussion.

We found longer response times following requests from the human partner, driven by higher rates of othercentric responding in that condition. This suggests that there are greater processing costs involved in taking another's perspective, and is in line with Duran et al. (2011), who observed larger effects of perspective taking on response times in othercentric compared to egocentric responders.

Finally, we note that our rates of egocentricism were high (>80%) in both partner conditions. While some researchers argue that this perspective is default and othercentricism is cognitively costly (Epley, Keysar, Van Boven, & Gilovich, 2004), our distribution is at odds with Duran et al. (2011), who found a more even split with a similar task. This discrepancy may be attributed to their inclusion of front/back spatial references, whereas we exclusively used left/right references, which are known to be harder to produce and comprehend (Franklin & Tversky, 1990). Thus, we repeated the experiment with front/back references in attempt to elicit a more even distribution of egocentricism and othercentricism.

Experiment 2

We replaced left/right references with front/back ones.[2]

Methods

Participants 508 participants were recruited on AMT. We excluded data from participants who were nonnative speakers of English (7), who failed to meet accuracy criteria (146), or who indicated suspicion about their partner or the interaction (114). The final dataset consisted of 241 participants: 114 and 127 in the human and computer partner conditions.

Materials, design, and procedure These were identical to Experiment 1 other than the replacement of critical utterances with front/back references. The objects appeared in the top and bottom positions on the table on these trials.

Results

Analysis procedures were the same as in Experiment 1.

Distribution of responders Table 1 shows the breakdown of responders groups. For egocentric responders, there was a 23% decrease in the human compared to computer partner condition, $\chi^2(1) = 18.20, p < .001$. For othercentric responders, there was a 19% increase in the human compared to computer partner condition, $\chi^2 = 10.69, p = .001$.

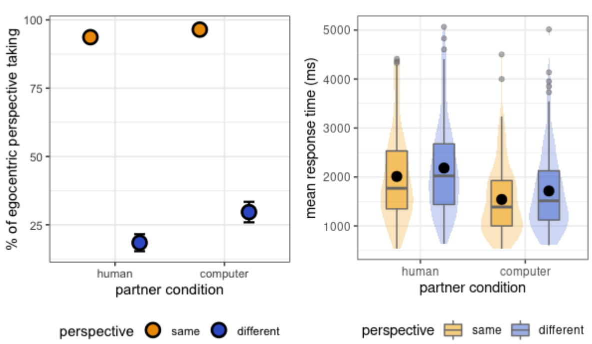

Figure 3: Experiment 2 results. Left: Percentage of trials on which participants chose the object from an egocentric perspective. Error bars represent ± 1 standard error of by-participant means. Right: Raw mean trial response times.

Perspective taking Figure 3 shows the percentage of trials on which participants responded egocentrically. The model showed main effects of perspective, with participants less likely to respond egocentrically on different perspective trials, $\beta = -13.75$, $SE = 1.21$, $p < .001$; and of partner, with participants more likely to respond egocentrically with a computer partner, $\beta = 1.10$, $SE = 0.54$, $p = .04$.

Processing cost Figure 3 shows participants' mean response times in the experiment. The model showed main effects of perspective, with longer response times in the different perspective condition, $\beta = 0.08$, $SE = 0.02$, $t = 3.85$; and of partner, with shorter response times with a computer partner, $\beta = -0.24$, $SE = 0.05$, $t = -4.55$. An additional analysis including responder group (egocentric vs. othercentric) as a predictor showed no effect of responder type nor its interaction with any of the other predictors (all $|t| < 1.1$).

Discussion

Our partner effects here corroborate those of Experiment 1: Listeners were more likely to interpret ambiguous spatial references egocentrically with a computer than a human. These results confirm that spatial reasoning is a flexible process that adapts dynamically to social cues such as the identity of one's partner. As in Experiment 1, we saw longer response times on different perspective trials, reflecting a cognitive cost in taking another perspective. Unlike Experiment 1 though, there was no difference between egocentric and othercentric responders, with both groups taking longer on different perspective trials. This may reflect differences between left/right and front/back perspective taking, the latter being computationally simpler and thus eliciting a smaller cost. Studies on spatial cognition report a similar disadvantage for left/right discrimination compared to other body-oriented dimensions such as front/back or near/far (Newcombe & Huttenlocher, 1992; Farrell, 1979), suggesting that spatial perspective taking is not equal across dimensions. In line with this, we saw a higher proportion of othercentricism in Experiment 2 (>70%) than Experiment 1. However, the similar response times across responder groups may suggest that front/back perspective taking was too trivial to evoke a clear distinction between egocentric and othercentric behaviour. Thus, in Experiment 3 we included references in both spatial dimensions.

Experiment 3

The main task was similar to Experiments 1 and 2, but including left/right and front/back spatial references. As the focus was on the role of individual differences in perspective taking, we omitted the partner manipulation, hence all participants interacted with a computer partner.[3] Participants completed the main task followed by four individual differences tests in the order: Autism Quotient (AQ), Object Perspective Test (OPT), Stroop task, directional discrimination test.[4]

Methods

Participants 291 participants were recruited on AMT. We excluded participants who were nonnative speakers of English (2) or did not meet accuracy criteria on the main task

[2]Preregistration details at `https://osf.io/s3tqu`

[3]The choice of a computer rather than human partner was logistical, as it eliminated the need to exclude participants who did not believe they were interacting with a real person.

[4]Preregistration details at `https://osf.io/76xdz`

(36). An additional 73 participants failed to complete the individual differences battery or to meet criteria on the Stroop task (>90% accuracy), leaving 180 participants in the final analyses.

Main task We made the following changes to the design:

(a) The number of critical trials increased to 16 (eight same and eight different perspective, with four left/right and four front/back utterances in each condition).

(b) Location contrast fillers were omitted, leaving colour contrast, size contrast, and no contrast filler displays (12 each); the total number of filler trials remained the same.

(c) All participants interacted with a computer partner.

Autism Quotient We measured social skills using the AQ, a 50-item self-administered questionnaire designed to assess autism-like traits (Baron-Cohen et al., 2001). For each participant, we derived a score based on questions from the social and communication subscales (cf. Shelton, Clements-Stephens, Lam, Pak, & Murray, 2012), using Austin's (2005) scoring strategy. Higher scores reflect poorer social skills.

Object Perspective Test We measured spatial orientation ability using a web-based version of the OPT. On each trial participants saw an array of seven objects, had to imagine themselves at one object and facing another, and indicate the direction to a third object. A deviation score for each participant was derived by taking the mean angle deviation between the participant's response and the correct response across trials. Higher scores reflect poorer spatial orientation ability.

Stroop task We measured inhibitory control using the colour-word Stroop task (Stroop, 1935). On each trial participants pressed a key (r, b, g, y, or p) corresponding to the colour of a word (red, blue, green, yellow, or purple). A participant's Stroop effect is derived from taking the difference between their mean response time on incongruent trials (colour and text of the word mismatch) and congruent trials (colour and text of the word match). Higher Stroop scores indicate poorer inhibitory control.

Directional discrimination task We implemented a novel task designed to test participants' differential ability to respond to front/back vs. left/right spatial references. On each trial, participants had to click as quickly as possible on one of four identical moles following a recorded instruction. A score for each participant was derived from the difference between their mean response time on front/back and left/right trials. However due to data loss from a number of participants misinterpreting the instructions as well as issues writing data to the server, we omitted this task from subsequent analyses.[5]

Results

We first analysed data from the main task for the two outcome variables: object selected and response times. Models included a single predictor, perspective (same vs. different; sum-coded). To explore the mediating role of individual differences in spatial perspective taking, we constructed a full model for each of the two outcome variables, with perspective and its interaction with each of the three measures (AQ score, OPT deviation score, and Stroop difference score) as predictors. Measures were entered as scaled and centred continuous variables. We subsequently constructed a final model by eliminating predictors that were not significant in the full model. All models included participant and item random intercepts and by-participant random slopes for perspective.

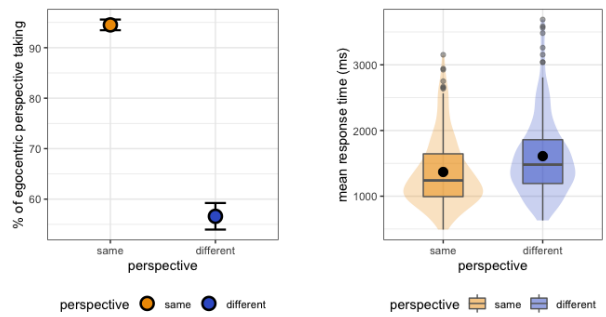

Figure 4: Experiment 3 results. Left: Percentage of trials on which participants chose the object from an egocentric perspective. Error bars represent ±1 standard error of by-participant means. Right: Raw mean trial response times.

Main task: perspective taking and processing cost Table 1 shows the breakdown of responder groups in the experiment. Figure 4 shows the percentage of trials on which participants responded egocentrically and their mean response times. Participants were less likely to respond egocentrically on different perspective trials, $\beta = -7.86$, $SE = 0.90$, $p < .001$; and took longer to respond on different perspective trials, $\beta = 0.15$, $SE = 0.02$, $t = 6.40$. An analysis including responder group (egocentric vs. othercentric) showed a perspective by responder group interaction, $\beta = 0.13$, $SE = 0.05$, $t = 2.51$, reflecting a larger slowdown on different perspective trials in othercentric compared to egocentric responders.

Contribution of individual differences The full model for perspective taking including all three individual differences measures showed an effect of perspective, with participants less likely to respond egocentrically on different perspective trials, $\beta = -7.61$, $SE = 0.97$, $p < .001$; and an interaction with OPT deviation score, $\beta = 3.50$, $SE = 0.98$, $p < .001$, driven by higher deviation scores being associated with more egocentric perspective taking on different perspective trials, $\beta = 1.73$, $SE = 0.51$, $p < .001$; no corresponding relationship was observed on same perspective trials ($p > .1$). This interaction is illustrated in Figure 5. Neither AQ score nor Stroop difference score modulated perspective taking tendencies (all $p > .6$). The final model including only perspective and its interaction with OPT score showed an effect of perspective, $\beta = -7.57$, $SE = 0.97$, $p < .001$ and a significant interaction, $\beta = 3.65$, $SE = 0.97$, $p < .001$. This model had an R^2 of 0.89, and was significantly better than a model including only perspective, $\chi^2(1) = 14.08, p < .001$.

[5]Analyses on the subset of usable participants from this task do not change the significance of the individual differences results.

The full model for response times showed an effect of perspective, with longer response times on different perspective trials, $\beta = 0.15$, $SE = 0.02$, $t = 6.24$; and a marginal interaction with OPT, $\beta = -0.11$, $SE = 0.07$, $t = -1.64$, driven by higher deviation scores being associated with longer response times on same perspective trials, $\beta = 0.16$, $SE = 0.07$, $t = 2.24$; no corresponding relationship was observed on different perspective trials ($|t| < 0.9$).[6] This interaction is illustrated in Figure 5. Neither AQ score nor Stroop difference score modulated participants' response times (all $|t| < 0.7$). The final model including only perspective and its interaction with OPT score showed an effect of perspective, $\beta = 0.15$, $SE = 0.02$, $t = 6.44$ and a marginal interaction, $\beta = -0.11$, $SE = 0.06$, $t = -1.71$. This model had an R^2 of 0.25, and was a marginally better fit than a model including only perspective, $\chi^2(1) = 2.90, p = .09$.

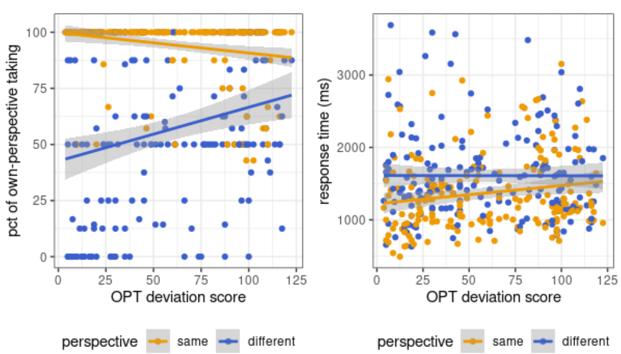

Figure 5: Relationship between OPT deviation score and egocentric perspective taking (left), and mean response times in the main task (right). Dots represent individual participants.

Discussion

The main finding from Experiment 3 revealed the role of individual differences on spatial perspective taking. Listeners with poorer spatial orientation ability, as shown by higher OPT scores, are less inclined to interpret ambiguous spatial references from their partner's perspective. This result highlights the cognitively challenging aspect of spatial perspective taking, even in a simplistic comprehension task. Another approach, for instance, could have been for othercentric listeners to adopt a simple heuristic of choosing the "opposite" object from their egocentric perspective on different perspective trials; however the finding that perspective taking is related to spatial orientation ability suggests that listeners did not do this, but instead invested in the mental operation of orienting themselves with their partner's perspective.

General Discussion

In this paper, we investigated spatial perspective taking in listeners in a task of simulated interaction. We found that othercentric perspective taking incurs a cognitive cost, with listeners' willingness to invest in this cost varying with (a) their perception of their partner's identity (Exp. 1 and 2), and (b) their spatial orientation ability (Exp. 3).

Surprisingly, our partner effects were in the opposite direction to Duran et al.'s (2011), who found higher rates of egocentricism with a human than a computer partner. It is possible that methodological differences between the studies contributed to the disparity. For instance, Duran et al. used a simpler design and the same stimuli throughout, whereas we included a number of varied fillers to distract listeners from the perspective manipulation. This may have increased their overall egocentric tendencies by reducing their attention towards the aspect of spatial ambiguity; however we see no reason for such an effect to differ across partner conditions. Another explanation, however, relates to a shift in expectations towards computers as communicative partners. Duran et al. attribute their results to listeners inferring about their partner's ability to collaborate and shifting the burden of perspective taking onto a partner who is human (cf. Tenbrink, Fischer, & Moratz, 2002). However, more recent studies have found greater egocentric tendencies in speakers interacting with robots compared to humans (e.g. Carlson, Skubic, Miller, Huo, & Alexenko, 2014). Together, these findings may reflect a shift in our perceptions about computers' capabilities in response to recent developments in Artificial Intelligence (Williams, Park, & Breazeal, 2019).

Beyond the external cue of partner identity, listeners' perspective taking tendencies varied with their own spatial ability. This result has implications for the question of the mechanism underlying spatial perspective taking. While we do not provide a direct test of relevant mechanism(s), we note that our individual difference measures target largely distinct processes that have separately been linked to perspective taking. Our results are consistent with the theory that spatial perspective taking relies on an embodied cognitive process of mental self-rotation (Kessler & Thomson, 2010). Perhaps surprisingly, we did not find a role of social skills, which has been implicated in several studies. This may be due to partner-specific attributions in human-computer interaction. Xiao et al. (2021) found that speakers' othercentric tendencies correlated with their social skills when addressing humans but not robots, suggesting that speakers only regard humans as social partners. A similar argument may apply to listeners responding to a computer; it is possible we would find a mediating role of social skills in the same task with a human partner.

Together, our results reveal the complex nature of spatial perspective taking, with listeners' tendencies depending on both external (e.g., partner-specific information) and internal (e.g., individual differences) factors. While our study falls short of investigating these factors in combination, we highlight the importance of a multidimensional approach to the study of perspective taking. Considering the interplay of multiple sources of information would contribute to a better understanding of the phenomenon.

[6] We do not have an explanation for the effect of OPT in same perspective trials. We speculate that some othercentric responders may have still been considering their partner's perspective even when perspectives aligned (cf. Duran et al., 2011). However, as the interaction is highly marginal we refrain from over-interpreting it.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 programme (ERC Starting Grant "Individualized Interaction in Discourse", grant agreement No. 948878).

References

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the autism spectrum quotient (aq). *Personality and Individual Differences*, *38*(2), 451–460.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (aq): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*(1), 5–17.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*(5), 893–900.

Carlson, L., Skubic, M., Miller, J., Huo, Z., & Alexenko, T. (2014). Strategies for human-driven robot comprehension of spatial descriptions by older adults in a robot fetch task. *Topics in Cognitive Science*, *6*(3), 513–533.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, *121*(1), 22–40.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327.

Farrell, W. S. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(1), 42.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental PGychology: general*, *119*(1), 63.

Frick, A., & Baumeler, D. (2017). The relation between spatial perspective taking and inhibitory control in 6-year-old children. *Psychological research*, *81*(4), 730–739.

Galati, A., & Avraamides, M. N. (2015). Social and representational cues jointly influence spatial perspective-taking. *Cognitive Science*, *39*(4), 739–765.

Hamilton, A. F. d. C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, *113*(1), 37–44.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117.

Job, X., Kirsch, L., Inard, S., Arnold, G., & Auvray, M. (2021). Spatial perspective taking is related to social intelligence and attachment style. *Personality and Individual Differences*, *168*, 109726.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition*, *114*(1), 72–88.

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: differences predicted by sex and social skills score. *Spatial Cognition & Computation*, *12*(2-3), 133–158.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745–756.

Newcombe, N., & Huttenlocher, J. (1992). Children's early ability to solve perspective-taking problems. *Developmental psychology*, *28*(4), 635–643.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*(1), 1–24.

Schober, M. F. (2009). Spatial dialogue between partners with mismatched abilities. *Spatial Language and Dialogue*, *1*, 23–39.

Schultz, K. (1991). The contribution of solution strategy to spatial performance. *Canadian Journal of Psychology*, *45*(4), 474–491.

Shelton, A. L., Clements-Stephens, A. M., Lam, W. Y., Pak, D. M., & Murray, A. J. (2012). Should social savvy equal good spatial skills? the interaction of social skills with spatial perspective taking. *Journal of Experimental Psychology: General*, *141*(2), 199.

Shelton, A. L., & McNamara, T. P. (2004). Spatial memory and perspective taking. *Memory & Cognition*, *32*(3), 416–426.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Tenbrink, T., Fischer, K., & Moratz, R. (2002). Spatial strategies in human-robot communication. *KI*, *16*(4), 19–23.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124–129.

Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, *20*(4), 766–772.

Williams, R., Park, H. W., & Breazeal, C. (2019). A is for artificial intelligence: the impact of artificial intelligence activities on young children's perceptions of robots. In *Proceedings of the 2019 chi conference* (pp. 1–11).

Xiao, C., Xu, L., Sui, Y., & Zhou, R. (2021). Do people regard robots as human-like social partners? Evidence from perspective-taking in spatial descriptions. *Frontiers in Psychology*, 4039.