A person, a dog, and a vase: The effect of avatar type in a perspective-taking task

Marin Dujmović (marin.dujmovic@gmail.com)
Department of Psychology, University of Zadar

Pavle Valerjev (valerjev@unizd.hr)
Department of Psychology, University of Zadar

Abstract
Recent research has shown people process what another person sees even when not explicitly taking the other person perspective. Results show people are slower in their responses when a different number of stimuli can be seen from one's own and the avatar’s (other) perspective. In these incongruent trials, the avatar's perspective interferes with the response about one’s own perspective and vice versa. However, a question remains whether this is purely a consequence of perceptual cueing or if social information contributes to the effect. The goal of our study was to investigate differences in automatic processing depending on avatar type. The experiment was a 2 (perspective) × 2 (congruence) × 2 (avatar type) repeated measures design. Participants were significantly slower for incongruent trials and from the other's perspective. Additionally, perspective by avatar type interaction was significant. Participants were slower from the other's perspective only for human avatars, but not when the avatar was a dog. These results may imply additional time is required to take the perspective of a human being while the perspective of a dog represents a perceptive cue rather than requiring that additional step. We may conclude both perceptual and social information may contribute to perspective taking.

Keywords: mentalizing; perspective taking; dot perspective task; visual attention

Introduction
The question of whether or not we can theorise about what other people see, think and feel is trivial. This ability has been labelled as mentalizing: the ability to represent other people’s mental states and it is based on the concept of theory of mind (Premack & Woodruff, 1978). While this ability seems to be self-evident, the question of whether these processes arise automatically or only through conscious effort is far from trivial. Recent research indicated that processing other people’s perspectives occurs implicitly even if participants are instructed to ignore the other person’s perspective (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010).

The dominant methodological procedure in these types of experiments is the dot perspective task. In this task, participants are asked to take one of two perspectives: their own (self) or the perspective of an observer presented in an image (other). In the congruent situations, the same number of targets can be viewed from both perspectives while in incongruent situations a different number of targets can be viewed from two perspectives. The single trial procedure is as follows: first, the perspective from which the participants are required to observe is cued. Second, the number of stimuli is cued. Finally, participants have to decide whether the cued number of stimuli can be seen from the cued perspective. The usual effect is a slower response in incongruent trials. This effect proves stronger when participants observe from the other person’s perspective when compared to their own, although the interference effect is also present when observing from one’s own perspective (Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2014).

While the existence of the fundamental interference effect has been replicated, there is some debate about whether or not this effect is a consequence of automatic mentalizing or simply a perceptual effect due to attention orientation by perceptual cues (such as the direction of gaze of the avatar in the image). According to Santiesteban et al. (2014), the effect is perceptive in nature. In their experiment avatars were people and arrows. Results showed no difference in the interference effect depending on avatar type. Both a human avatar and an arrow produced the same level of interference. Using a different paradigm Cole and his colleagues (Cole, Smith, & Atkinson, 2015; Cole, Atkinson, Le, & Smith, 2016) reached the same conclusion. They tested the strength of social information by placing barriers between the avatars and the stimuli. There was no effect of the barriers on the interference so they concluded participants did not automatically process the avatar’s mental state. On the other hand, Kragh Nielsen, Slade, Levy, and Holmes (2015) found that the interference effect was lower for arrows when compared to human avatars. They also found a positive correlation between components of the Interpersonal Reactivity Index and interference size when the avatar was a person, but not when it was an arrow. In a modified version of the task, Valerjev & Dujmović (2017) found that there was no effect of avatar skin colour on the size of the interference effects.

Figure 1: Examples of incongruent stimuli for a person (left panel) and a dog (right panel)
The goal of this study was to investigate differences in the level of interference depending on the species of the avatar in the task. If automatic mentalizing does indeed occur it should lead to different sizes of the interference levels depending on avatar type. We expected the interference to be greater when the avatar was a human being compared to when it was a dog.

**Method**

A total of 33 undergraduate psychology students participated in the experiment which was designed and conducted using the E-Prime software package.

The experiment was a 2 (person/dog) × 2 (self/other perspective) × 2 (congruent/incongruent) repeated measures design. Participants could see a person or a dog as the avatar, they could respond from their own or the avatar’s perspective, and the number of stimuli (vases) which could be seen from the two perspectives could be congruent or incongruent. An example of two stimuli depending on avatar type can be seen in Figure 1. The stimuli were designed in the Daz3D software package.

The single trial procedure was a modified version of the dot perspective task. The participants were cued from which perspective they were observing the scene by text (“SELF” or “OTHER”) which was shown for 750 ms prior to the target. The target scenes were then shown and participants had to respond as fast and as accurately as possible by responding to the number of stimuli which could be seen from the cued perspective. In half of the trials the correct response required pressing the numerical key “1” and for the other half the key “2”. The total number of trials was 64, eight per experimental condition and the order of presentation was randomised for each participant.

**Results**

In order to prepare the data for analysis, only correct responses and responses which had response times within the +/-3 standard deviation range were included. Correct responses made up 93.80% of all responses, and 2.68% of the correct responses were excluded due to extreme response times. For each participant mean response times were calculated in each experimental condition and this made up the final data set for analysis.

To determine whether avatar type had an effect on the pattern of response time results a 2 (avatar type) × 2 (perspective) × 2 (congruence) analysis of variance for repeated measures was conducted. The results of this analysis can be seen in Table 1.

**Table 1: ANOVA results for response times**

<table>
<thead>
<tr>
<th>Effect</th>
<th>$F(1, 32)$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruence</td>
<td>15.25**</td>
<td>.32</td>
</tr>
<tr>
<td>Perspective</td>
<td>11.40**</td>
<td>.26</td>
</tr>
<tr>
<td>Avatar</td>
<td>0.09</td>
<td>.00</td>
</tr>
<tr>
<td>Congruence × Perspective</td>
<td>0.06</td>
<td>.01</td>
</tr>
<tr>
<td>Congruence × Avatar</td>
<td>1.42</td>
<td>.04</td>
</tr>
<tr>
<td>Perspective × Avatar</td>
<td>13.04**</td>
<td>.29</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td>0.01</td>
<td>.00</td>
</tr>
</tbody>
</table>

**Figure 2: Response times as a function of avatar type, congruence and perspective**

As expected, participants responded faster in the congruent when compared to incongruent trials and from their own perspective. Tukey HSD post-hoc comparisons showed that participants were significantly slower when responding from the perspective of the avatar compared to their own when the avatar was a person, but this difference was not significant when the avatar was a dog. These results can be seen in Figure 2.
Discussion

The results replicated the consistent findings which show the main effects of congruence and perspective. This, however, was qualified by a significant perspective by avatar type interaction which is the most significant result of the study. Participants were slower from the avatar’s perspective compared to their own only when the avatar was a person. The difference between the two perspectives was not significant when the avatar was a dog. We speculate about two possible explanations for this interaction effect.

First, when participants respond from the person’s perspective, the avatar draws more attention relative to the target stimuli. This additional orienting of attention prolongs response times for human avatars but is not present when the avatar is a dog. This more complex processing may be attributed to mentalizing due to the higher similarity between the participants and a human avatar when compared to a dog.

Second, a dog may serve as a better cue for orienting attention to the target stimuli (due to shape and slant towards the target). This would potentially mean less attention was focused on the avatar when the participants judged from the “other” perspective. As a result, there was no difference between the two perspectives when the avatar was a dog.

Both explanations seem plausible, and they do not necessarily exclude each other. The effect may be due to some social/perceptual interaction.

The main goal of the study was to determine whether the avatar type has an effect on the size of interference effects. The results do not indicate that the difference between congruent and incongruent trials depends on avatar type.

We conclude the debate on automatic mentalizing is still open with the possibly more complex interaction between purely perceptual and social processes impacting the results. Further study and variations of perceptual and social information are needed in order to clarify some of the results presented here.

Acknowledgments

This research was supported by University of Rijeka Research Grant [grant number 13.04.1.3.11].

References


