



Visual perspective taking for avatars in a Simon task

Christian Böffel¹ · Jochen Müsseler¹

Published online: 5 September 2018
© The Psychonomic Society, Inc. 2018

Abstract

In modern digital applications, users often interact with virtual representations of themselves or others, called *avatars*. We examined how these avatars and their perspectives influence stimulus–response compatibility in a Simon task. Participants responded to light/dark blue stimuli with left/right key presses in the presence of a task-irrelevant avatar. Changes in stimulus–response compatibility were used to quantify changes in the mental representation of the task and perspective taking toward this avatar. Experiments 1 and 2 showed that perspective taking for an avatar occurred in orthogonal stimulus–response mappings, causing a compatibility effect from the avatar’s point of view. In the following two experiments we introduced a larger variety of angular disparities between the participant and avatar. In Experiment 3, the Simon effect with lateralized stimulus positions remained largely unaffected by the avatar, pointing toward an absence of perspective taking. In Experiment 4, after avatar hand movements were added in order to strengthen the participants’ sense of agency over the avatar, a spatial compatibility effect from the avatar’s perspective was observed again, and hints of the selective use of perspective taking on a trial-by-trial basis were found. Overall, the results indicate that users can incorporate the perspective of an avatar into their mental representation of a situation, even when this perspective is unnecessary to complete a task, but that certain contextual requirements have to be met.

Keywords Avatar · Orthogonal compatibility · Perspective taking · Simon effect · Stimulus–response compatibility · Action effects

The ever-changing world of modern technology continually imposes new tasks on users. Avatars are often used to represent users in virtual environments, and the rising trend toward use of virtual reality can lead to diffusion between the real and virtual worlds, as users put themselves in virtual characters’ shoes and see the world through their eyes. Flavell, Green, Flavell, Watson, and Campione (1986) referred to such phenomena as *level 2 perspective taking*. Originally the target of this change in perspective was another person, but the same concept can be applied to interactions with avatars. The difference between one’s own and the avatar’s point of view can be quantified by measuring the angular disparity between the two perspectives. Generally, larger angular disparities between the person and the target are associated with higher perspective-taking costs, such as increased reaction times (for an overview, see Avraamides, Hatzipanayioti, & Galati,

2015). Seeing the world from a different point of view includes knowledge of how a scene would look from this new perspective. This phenomenon is often characterized as being more complex than the mere judgment about someone else’s ability to see a certain object from their point of view, a process that is called *level 1 perspective taking* (for a comparison of the two concepts, see Michelon & Zacks, 2006). At first glance, perspective taking has a lot in common with the mental rotation of objects, and it is sometimes referred to as *mental self-rotation* (Hegarty & Waller, 2004). However, there are some pronounced differences between the two concepts. When mentally rotating an object, the time needed to perform the rotation is linearly dependent on the angular disparity (Cooper & Shepard, 1973; Janczyk, 2013; Shepard & Metzler, 1971). In contrast, perspective taking is often characterized by a discontinuous relationship between angles and reaction times, with an increase in reaction times starting between rotations of 60° and 90°, and little to no increase for smaller angles (Janczyk, 2013). Janczyk provided substantial evidence that two distinct processes are involved in perspective taking. The first one is used to bridge small gaps in perspective (up to 60°) and likely is effortless, in the sense that no central capacity is involved. The second process, however,

✉ Christian Böffel
boeffel@psych.rwth-aachen.de

¹ Institute of Psychology, RWTH Aachen University,
52066 Aachen, Germany

does require central capacity and is needed when larger disparities must be overcome.

Freundlieb, Kovács, and Sebanz (2016) demonstrated that the adoption of someone else's perspective can also occur spontaneously in social situations toward a confederate. Although perspective taking was not required to solve the task used in their study, Freundlieb et al. (2016) still found evidence for it. In a later experiment, they demonstrated that this process is modulated by the targeted person's visual access and is only present when the confederate can actually see the stimuli (Freundlieb, Sebanz, & Kovács, 2017). On the basis of these results, it seems that level 1 perspective taking could be a precursor for level 2 perspective taking.

Our aim was to examine the degree to which visual perspective taking toward a nonhuman target, such as an avatar, can be invoked in similar situations, when it is not necessary to complete the given task. In comparison to Freundlieb et al. (2016), who used a real person as a confederate, these avatar tasks can be viewed as less social. The avatar is merely a rough sketch of a person, and not an actual human being. Depending on the cover story used and the exact properties of the situation, an avatar is often more like a tool and less like a social partner. On the basis of the spontaneous nature of perspective taking, we expected it to be at least partly driven by bottom-up processes, and therefore predicted that perspective taking could be invoked as a result of situational features, even when the target was an avatar and not a person. Or, to put it differently: In the correct circumstances, a person has no choice other than to adopt the perspective of an avatar to some degree, even when this process is not useful, or is potentially detrimental, to the task at hand.

To quantify these changes in perspective, we incorporated avatars as a target for perspective taking into a Simon task. It is important to note that these avatars were irrelevant to the task and could be seen as distractors. In the following paragraphs, we will elaborate on the concept of stimulus–response compatibility in general and the Simon effect as a special case of spatial compatibility, and show how these concepts are useful to quantify perspective taking.

Stimulus–response compatibility and the Simon effect

When performing a task, certain mappings of stimuli to responses result in faster reaction times and fewer errors than do other mappings. Fitts and Deininger (1954) were the first to examine these stimulus–response (SR) compatibility effects in a systematic manner. They found that mappings in which the stimuli and responses were spatially similar improved performance relative to random mappings. The paradigms that are used to examine such compatibility effects can be very simple and often only feature two different stimulus and response

locations. When two different stimulus properties (such as two different colors, or pitches of tones) indicate either ipsi- or contralateral responses, an ipsilateral, and therefore spatially corresponding, SR ensemble leads to better performance than does a contralateral, and therefore noncorresponding, one. In this case, the spatially corresponding conditions can be identified as compatible, whereas the noncorresponding conditions are incompatible. Spatial correspondence is one of the strongest factors that causes compatibility effects.

The examination of SR compatibility has long tradition, and a wide variety of tasks have been used to study this phenomenon. Some of the more prominent experiments were conducted by J. R. Simon and colleagues: Simon and Small (1969) used a mapping of high-pitched and low-pitched monaural tones to left and right key presses, and Simon and Rudell (1967) used the spoken words “left” and “right” in a similar fashion. The stimuli were randomly presented to the participants' left and right ears. Although the stimulus location was task-irrelevant, Simon and Small found a spatial congruency effect between ear and response location (for a similar effect using visual stimuli, see Brebner, Shephard, & Cairney, 1972). Such a compatibility effect caused by the irrelevant position of a stimulus is called a *Simon effect* (Hedge & Marsh, 1975; for an overview, see Hommel, 2011).

Probably the best known theoretical framework for these SR compatibility effects is the *dimensional overlap* model, proposed by Kornblum, Hasbroucq, and Osman (1990). The authors suggested that SR ensembles are compatible when the dimensional overlap between the stimuli and responses is sufficiently large. These overlaps often occur in the spatial dimension—for instance, on the horizontal, left–right dimension. Importantly, the overlapping dimensions do not have to be relevant to the task itself, allowing the model to explain the Simon effect. The only necessary condition for compatibility effects is that the dimensional overlap be sufficiently large (Kornblum et al., 1990). The model proposes that when a dimensional overlap is present, stimulus features cause the activation of corresponding responses via a direct and automatic route. A stimulus on the right, for example, automatically activates a right key press when the overlapping dimension is the horizontal position. A second and indirect route is based on the mapping, often defined by an instruction. When both routes call for the same response, the ensemble is compatible, and trials in which both routes lead to different responses are incompatible. Showing some similarities to the overlapping dimensions in Kornblum's model, other theoretical frameworks have focused on the idea of common coding. The *theory of event coding* (Hommel, 2015; Hommel, Müsseler, Aschersleben, & Prinz, 2001), for example, argues that perception and action share the same basic representational units. This theory views perception and action planning as being closely related, although functionally different. As a

consequence, there should be no fundamental difference between stimulus and response codes, so that perception and action can influence each other. In a Simon task, the stimuli and responses share task-irrelevant features, such as spatial information. For example, when participants should answer a green stimulus with a left key press—regardless of stimulus position—and the stimulus happens to appear on the left, the stimulus position will lead to the abstract feature code “LEFT.” This feature code now facilitates motor responses that use the same feature code, in this case a left key press. It is important to note that responding to the same stimulus by saying “left” would also create a response that shares the same abstract “LEFT” feature code, even though the response is not bound to a physical location in space. Both responses could therefore recruit the same feature code, even when the modality is different. The more feature codes that stimuli and responses have in common, usually, the higher is their resulting degree of compatibility. Another framework for SR compatibility is the *response discrimination account* (Ansorge & Wühr, 2004). In a series of experiments, Ansorge and Wühr provided compelling evidence in favor of the idea that compatibility effects only arise when the responses are shown on an axis that is also used to discriminate between responses. In the majority of classical compatibility experiments, this is a given. Often horizontal stimulus sets are combined with a horizontal response set, so that the stimuli are indeed presented on a discriminating axis. There are, however, some surprising compatibility effects that seem to defy most of the traditional explanations for SR compatibility.

Orthogonal compatibility

In a compatibility experiment, Bauer and Miller (1982) assigned a horizontal stimulus set to a vertical response set and observed a peculiar SR compatibility effect. The right–up, left–down mapping resulted in faster reaction times when the responses were performed with the right hand. In contrast, a left–up, right–down advantage was observed when participants responded using their left hand. As a consequence, these mappings seemed to vary in compatibility: For the right-hand condition, the combinations of down–left and up–right were compatible, whereas the down–right and up–left combinations were incompatible. For the left-hand condition, the compatibility associations were reversed.

Cho and Proctor (2005) observed a similar up–right mapping advantage by assigning stimuli that were positioned above or below a fixation point to left or right responses. This up–right/down–left advantage is usually smaller and more volatile than classic compatibility effects and is known as the *orthogonal compatibility effect*. Nishimura and Yokosawa (2006) reported a 12-ms advantage in favor of

orthogonally compatible conditions, whereas Cho, Proctor, and Yamaguchi (2008) weren’t able to observe this mapping advantage in a similar setup. Overall the up–right/down–left advantage in orthogonal mappings appears to be a lot smaller than the regular Simon effect, and therefore is harder to measure.

Although small, the orthogonal compatibility effect is remarkable because there is no obvious relation between the spatial properties of the stimuli and responses—no overlapping dimensions. Weeks and Proctor (1990) proposed a possible explanation for this effect by pointing out that asymmetries in the coding process could explain this effect. They argued that assigning positive and negative polarities to stimulus dimensions introduces compatibility on an abstract level. A positive polarization can be attributed to a higher salience of the top position, as compared to the lower salience (and therefore negative polarization) of the lower position on the vertical dimension. Overall, right responses seem to be more salient than left responses for right-handed people. However, a response in the left dimension—for example, after the response panel was shifted to the left—can cause higher salience for left than for right responses. This explanation is known as the *salient-feature hypothesis* (Cho & Proctor, 2003; Weeks, Proctor, & Beyak, 1995), and it manages to close the gap in the dimensional overlap model by reintroducing an overlap on the abstract dimension of polarity. This polarity of the spatial dimension, rather than the spatial dimension itself, is now the relevant dimension that causes the observed compatibility effects. Other researchers have also been interested in the topic of orthogonal compatibility and have put forward other hypotheses for its origin. Lippa and Adam (2001), for example, argued that end-state comfort might be an important factor in orthogonal mappings, and that participants prefer mappings that lead to the more comfortable end state. They demonstrated in their first experiment that the often found up–right/down–left advantage is reversed when the response panel is shifted to the left. This observation is still in line with the salient-feature hypothesis, because the left shift of the response panel could have caused an increased salience for the left position. But in their following experiment, Lippa and Adam managed to find conditions in which a response panel on the right was also related to an up–left advantage, a finding that is irreconcilable with the salient-feature hypothesis.

Although compatibility effects are sometimes seen as problematic when examining the underlying mechanisms of perspective taking (May & Wendt, 2013), we will make the case that they can be a valuable tool when determining whether perspective taking occurs. Orthogonal compatibility effects without an eccentricity manipulation are generally smaller than compatibility effects with lateralized stimulus and

response positions, so we decided that an orthogonal setup would be more suitable to reveal compatibility changes based on perspective taking with an avatar, because the orthogonal compatibility is easier to influence and overcome than is its lateralized counterpart. We will now lay out how a change of SR compatibility as a result of the avatar presentation can be an indicator of a change in the mental representation of a task. We believe these compatibility changes can therefore be useful as an objective performance measure to quantify perspective taking.

Avatars, perspective taking, and compatibility

We aimed to investigate whether the presence of an avatar and its orientation can change the mental representation of a task as a result of visual perspective taking. An example of such an avatar is shown in Fig. 1. When we take a look at both scenarios in the figure, we can see that there are two conflicting ways to represent the location of the blue disc: On the one hand, we can code it from our own point of view, and in this case the disc would be on the left in the first panel and on the right in the second one. Coding the stimulus from our own perspective is usually the default assumption (Gardner & Potts, 2011; Taylor, Flynn, Edmonds, & Gardner, 2016). On the other hand, we could represent the stimulus position from the avatar's perspective. As a result, in both cases the disc would be represented as on the left from the avatar's viewpoint. In the left panel, both representations lead to the same outcome, and the stimulus would be represented as on the left.

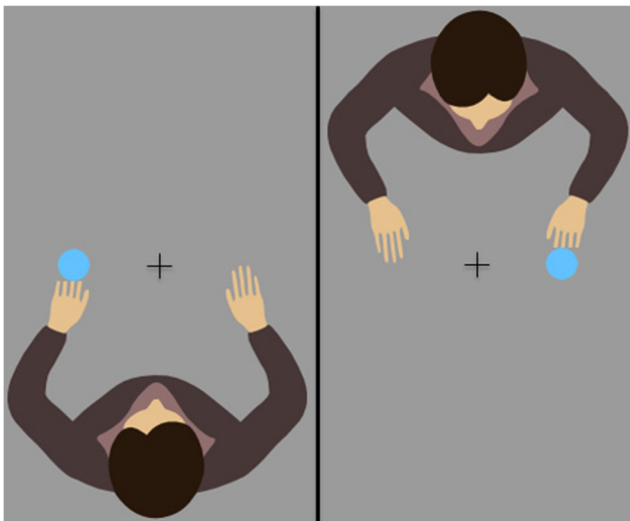


Fig. 1 Avatars in egocentric (0° rotation; left) and rotated (180° rotation; right) perspectives. In both cases the blue disc is on the avatar's left, but from the perspective of the viewer, the disc changes position from left to right.

In the right panel we have two conflicting representations. If we assume that perspective taking is successful, the representation from the avatar's point of view should overwrite the egocentric representation. This switch in coding should be measurable with the help of SR compatibility effects, in that perspective taking should cause the 180° -rotated condition on the right, in which the stimulus is on the avatar's left, to activate a left response. This condition should therefore be compatible with a left key press. Without perspective taking, we would expect that the stimulus on the right would activate a right response. This right key press should be compatible because it is spatially corresponding to the stimulus from our own point of view. To demonstrate these effects, we aimed to influence the typically observed effects in Simon tasks by introducing into the paradigm avatars with which the participants could identify. These avatars could also be seen as distractors that would carry spatial information in a similar way to how the target stimuli do. However, there is a crucial difference between avatars and targets in a Simon task: Both the avatars and stimuli carry spatial information that is irrelevant to the task, but the targets themselves still contain task-relevant features, whereas the avatars do not. It might therefore be easier to ignore the avatar position than the stimulus position. The avatars also might resemble an accessory stimulus (for an example of the use of accessory stimuli, see Nishimura & Yokosawa, 2010), but in contrast to typical accessory stimuli, the avatars would remain visible throughout the experiment, and not suddenly appear before stimulus presentation.

The issue of SR compatibility in the context of visual perspective taking has been raised before by other authors. May and Wendt (2013), for example, pointed out that in the paradigms that are typically used to observe perspective taking, the participant is often required to judge on which side of a body a certain object is presented. These tasks therefore often contain only ipsilateral responses, and therefore cannot accurately determine the influence of spatial compatibility on these measurements. The combination of a compatibility task and a perspective-taking task seems to be a logical step to bridge this gap, and so allow us to observe how SR compatibility is affected by perspective taking. If perspective taking with an avatar changes the spatial compatibility relations, this would be observable as a dependency of the spatial compatibility on the position of the avatar. Such a result could therefore be interpreted as evidence that the mental representation of the task had changed.

The goals of this study were to further our understanding of visual perspective taking and to demonstrate the usefulness of the Simon effect as a tool to quantify visual perspective taking. We also wanted to show that perspective taking can occur with regard to targets that are task-irrelevant and that do not engage in actions on their own.

Experiment 1

In Experiment 1, we examined whether the orthogonal Simon effect could be influenced by perspective taking toward an avatar that was present during the task. The avatar itself was placed to the left or right of a vertically arranged stimulus set (Fig. 2), manipulated by rotating the avatar 90° clockwise (referred to as “ 90° ”) or 90° counterclockwise (referred to as “ -90° ”), respectively, from the participant’s position. When the avatar is displayed on the left-hand side of a vertical set of stimuli (90° rotation) facing the central fixation point, stimuli presented above the fixation point can be coded as “left” from the avatar’s point of view. This causes a spatial correspondence to a key press on the participant’s left. This particular scenario can be seen in light of the response discrimination account (Ansorge & Wühr, 2004), and a compatibility effect relative to the avatar’s point of view would indicate that the stimuli are presented on a response-discriminating axis.

We expected a change in the frame of reference in which the stimulus position was coded relative to the avatar’s midline instead of the person’s own. As a result, the responses and stimuli would be coded on the same dimension and could therefore share spatial features. If this were the case, the result would be a dimensional overlap that would cause a spatial compatibility effect similar to the effect in a classic setup with left and right stimuli and responses. To put this in terms of the theory of event coding: Both the stimulus and response would now relate to the same, abstract feature code. We therefore believed that this avatar-related modification of the measured compatibility effects would allow inferences about the nature of the spatial codes formed and would therefore provide an objective performance measure to quantify the degree of perspective taking. Overall, a match between response and stimulus codes of some kind would be the most straightforward explanation if we were to observe such an avatar-based compatibility effect.

Additionally, we instructed half of the participants to imagine controlling the avatar’s hands and to adopt its point of

view while performing the task (steer group), in order to examine whether the effect would be subject to top-down modulation. To accomplish this, the participants had to perform a mental self-rotation of 90° either clockwise or counterclockwise (i.e., -90°), to mentally align themselves with the left or right avatar position, respectively. When the avatar was presented on the left side, its left hand was pointing toward the upper stimulus position and its right hand toward the lower, resulting in an association that would oppose the usually advantageous and compatible up–right combination. If the avatar were on the right side, the resulting association would correspond to the established up–right, down–left advantage. The other half of the participants were instructed to ignore the avatar while performing the same task (ignore group). Both groups differed only in terms of the instructions.

Franz, Sebastian, Hust, and Norris (2008) found evidence that perspective taking is independent of central processing, and is therefore not subject to a central processing bottleneck when combined with tasks that require the allocation of central processing resources, such as the retrieval of SR mappings. However, this is apparently only the case in tasks with rotations up to 60° . Since the rotations in the present experiment were 90° and -90° , they were most likely associated with cognitive effort, and therefore subject to capacity limitations (Janczyk, 2013). As a consequence, perspective taking should be related to higher overall reaction times when compared to ignoring the avatar, because processing of the SR mapping and perspective taking would both demand central capacity. This would make perspective taking an overall uneconomical strategy that should be avoided in order to maximize performance. However, this conclusion would only be applicable if perspective taking is voluntary and subject to a sufficient degree of top-down modulation. We expected top-down modulation to be relevant and to observe a higher amount of perspective taking in the steer group than in the ignore group, on the basis of observations in an experiment that had used a comparable manipulation with different

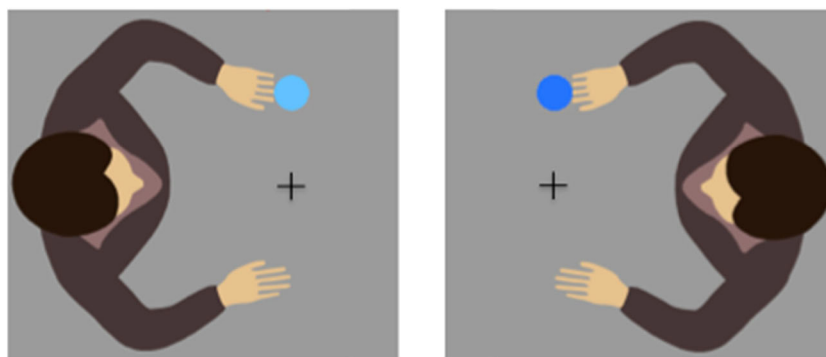


Fig. 2 Examples of the conditions in Experiment 1. (Left) 90° rotation, here with a light blue stimulus on top. (Right) -90° rotation, here with a dark blue stimulus on top.

instructions in non-Simon perspective-taking tasks (Böffel & Müsseler 2018, Müsseler, Ruland, & Böffel 2018).

Hypotheses

We believed that the typical up–right/down–left advantage of orthogonal SR ensembles would be overwritten by a Simon effect defined by the avatar’s point of view: When the avatar was presented on the right, we expected an advantage for up–right/down–left pairs, but when the avatar was on the left, an advantage of up–left/down–right ensembles was predicted, as a result of perspective taking. Because orthogonal compatibility is defined by the typical up–right/down–left advantage, our predicted effect would manifest itself in the form of an interaction between the factors orthogonal compatibility and avatar rotation. We further predicted that this interaction would be influenced by the instructions used, with larger effects with the *steer* than with the *ignore* instructions.

Method

Participants and sensitivity In total, 24 students (22 female, two male) from RWTH Aachen University, with a mean age of $M = 24.2$ ($SD = 8.4$), participated in this experiment for course credit or a monetary compensation of €5. All participants of the present and the following experiments had normal or corrected-to-normal vision and gave informed consent to the terms of data collection, use, and storage, in accordance with the Declaration of Helsinki (World Medical Association, 2013).

The total sample size allowed for a detection of compatibility effects of about $\eta_p^2 = .26$ overall and about $\eta_p^2 = .45$ within each group individually, with a power of $(1 - \beta) = .80$ according to G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). For a comparison, the effect observed in the perspective-taking task of Freundlieb et al.’s (2016) experiment had an effect size of $\eta_p^2 = .44$. All effect sizes (η_p^2) reported in this study incorporate correlations between paired measures (Lakens, 2013).

Apparatus and stimuli Matlab and the Psychophysics Toolbox extension, version 3.0 (Brainard, 1997; Pelli, 1997), were used to control the experiment. The stimuli were presented on a 22-in. CRT monitor (Iiyama Visionmaster Pro 514, with a resolution of $1,024 \times 768$ pixels at 100 Hz). The participants were seated approximately 70 cm in front of the monitor and responded with their left and right index fingers on left and right response keys, each in a distance of 5 cm from the participant’s midline.

Dark blue (RGB 36 115 254) and light blue circles (RGB 98 193 254), each with a diameter of 50 pixels (1.79°), were

used as the targets, presented at a distance of 45 pixels (1.61°) above or below a central fixation cross in front of a gray background (RGB 155 155 155). The avatar covered an area of roughly 240×200 pixels ($8.73^\circ \times 8.56^\circ$) and was positioned in such a way that its hands would point toward the stimulus positions from either the left or the right side facing the fixation cross (Fig. 2). The avatar positions were achieved by rotating the avatar by 90 and -90° . The avatar position was blocked so that it was switched after the first half of the experiment. The mapping of light and dark blue stimuli to left and right responses and the starting positions of the avatar were counterbalanced between participants.

Procedure The experiment was conducted in a dimly lit room. After being informed of the terms of data collection and storage, half of the participants were instructed that the avatar was irrelevant to the task and were asked to ignore it (*ignore* group), and the other half were told to adopt the avatar’s point of view and imagine steering the avatar’s hands (*steer* group). Regardless of the instructions used, the avatar did not show any visual action effect after a key press. As a consequence, the differences between the groups were based solely on the two different instructions.

Each participant performed two blocks, with a right (-90°) or a left (90°) avatar position, respectively. Each of these two main blocks started with 20 practice trials and was followed by eight subblocks with five repetitions of each condition. The order of trials was randomized within each subblock, and the order of the initial avatar position and the mapping of light and dark blue stimuli to left and right responses was counterbalanced between participants. Each condition was repeated 40 times over the course of the experiment, resulting in a total of 320 trials per participant. The participants needed between 20 and 30 min to complete the experiment.

Each individual trial started with the presentation of the fixation cross and the avatar, which remained visible throughout the experiment. After a delay of 750 ms, targets were presented above or below the fixation cross. If a response was incorrect, slower than 1,000 ms, or faster than 100 ms, it was labeled as an error and followed by a feedback tone. The waiting period between the response and the beginning of the next trial was 1,500 ms, which increased by an additional 1,500 ms after an error had occurred.

Design We labeled the up–right and down–left SR ensembles as compatible, and the down–right/up–left pairings as incompatible, in reference to the orthogonal Simon effect. This resulted in a $2 \times 2 \times 2$ design with the within-subjects factors orthogonal compatibility (incompatible vs. compatible) and avatar rotation (90° vs. -90°), and the between-subjects factor instructions (*steer* vs. *ignore*).

Results and discussion

Reaction times (RTs) longer than 1,000 ms or shorter than 100 ms were regarded as errors and were removed from the analyses. A total of 20 trials (0.26%) were excluded in this way, along with 151 false responses (1.97%), for a total of 171 errors (2.23%). The mean RTs and percentage errors were analyzed separately using $2 \times 2 \times 2$ mixed design analyses of variance (ANOVAs) with repeated measurements on two factors.

Reaction times The instructions factor did not cause a significant main effect or interaction with the other factors (all p s > .25). A significant interaction between the position of the avatar and orthogonal compatibility was observed, with $F(1, 22) = 14.81$, $p < .001$, $\eta_p^2 = .40$. The up–right advantage was 16 ms when combined with the -90° avatar rotation, and -15 ms with the 90° avatar rotation, reversing the orthogonal compatibility effect when the avatar was presented to the left of the stimulus set (Fig. 3). Post-hoc two-tailed t tests showed that both effects differed significantly from zero, with $t(23) = -2.55$, $p = .018$, and $t(23) = 3.49$, $p = .002$, respectively. No other significant RT effects were observed.

Percentage errors A significant interaction between the factors avatar position and orthogonal compatibility was observed, with $F(1, 22) = 26.65$, $p < .001$, $\eta_p^2 = .55$. The up–right advantage amounted to 1.4% for an avatar position on the right (-90°), and to -1.2% for an avatar on the left (90°), again reversing the orthogonal compatibility effect in the left avatar position. Two-tailed post-hoc t tests showed that both effects differed significantly from zero, with $t(23) = 3.76$, $p = .001$, and $t(23) = -2.60$, $p = .016$, respectively. Other significant effects were not observed.

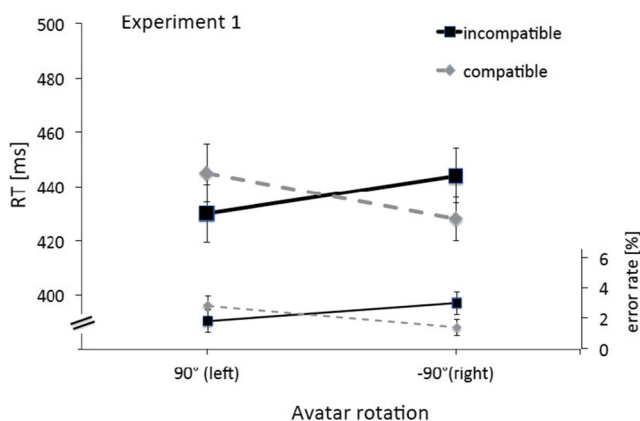


Fig. 3 Mean reaction times (RTs; thick lines) and error rates (thin lines) as a function of orthogonal compatibility and avatar rotation. Error bars represent 95% within-subjects confidence intervals (Morey, 2008).

No significant differences between the steer and ignore instructions were observed in terms of either mean RTs or percentage errors. The further lack of a significant interaction between the factors instructions, orthogonal compatibility, and avatar rotation was contrary to our initial hypothesis of significant top-down modulation of perspective taking by instructions. We therefore cannot reject the null hypothesis with sufficient confidence. This shows that the top-down modulation of the effect was most likely unsuccessful, and might therefore underline the importance of stimulus-driven, bottom-up processing for perspective taking. However, these results cannot completely rule out an influence of the instructions. If we were to assume that the effect size of the three-way interaction on the population level was close to an observed $\eta_p^2 = .05$, the experiment would lack the power to reliably detect an effect of this magnitude. Because the effect was most likely very small at best, however, we will treat both instructions as largely interchangeable. Overall, this is a surprising finding and could point to the conclusion that the participants in the ignore group were in fact unable to ignore the avatar. One alternative view would be that the steer group also ignored the avatar, contrary to the instructions, but we believe that the observed influence of the avatar rotation supports the first explanation. If both groups had been able to ignore the avatar, its rotation should not have mattered for SR compatibility. There is another explanation, however: It is possible that the participants in both groups knowingly attended to the avatar, and that one group ignored the ignore instructions. This could have been the result of a resource-saving strategy in the ignore group. If perspective taking is the natural way to interact with these scenarios and has to be suppressed, actively ignoring the avatar might potentially be associated with cognitive control.

Although ignoring the avatar could help eliminate the additional costs of conditions that were noncorresponding from the avatar's point of view, it would also eliminate the potential benefits in conditions that spatially corresponded with the perspective taking. If these effects are symmetrical, then there would be no overall advantage of the ignore strategy, so using additional resources to pursue it would be inefficient. Overall, we think that if the participants were able to ignore the avatar effectively, we would have been unable to measure any influence of the avatar and its rotation on orthogonal compatibility.

Furthermore, we managed to eliminate the up–right/down–left advantage expected in orthogonal mappings. In the conflict of orthogonal and avatar-induced compatibility, the avatar compatibility came out on top and effectively overwrote the orthogonal compatibility effect. This is especially remarkable because not only was the avatar's position irrelevant to the task, but so was the avatar itself. This points us to the conclusion that perspective taking did take place, even though it was unnecessary. Combined with the results regarding the

instructions, it seems very likely that perspective taking with avatars may be an involuntary and spontaneous process. Because of its requirement for central capacity, it could potentially reduce the performance in the primary task.

However, if we look back at the theories regarding orthogonal compatibility, there was an alternative explanation to consider: The effects could also be explained using the salient-feature hypothesis (Weeks & Proctor, 1990; Weeks et al., 1995). The avatar position could have increased the salience of the corresponding pole of the spatial axis and caused a shift in polarity. When the avatar was presented on the left side, it would make the left location more salient and associate it with positive polarity. If this were the case, the shift in polarity could lead to the up–left/down–right mapping being compatible in terms of polarity. In the second experiment, we aimed to address this possibility and try to differentiate between perspective taking and salient-feature coding as two possible explanations of the effects observed in Experiment 1.

Experiment 2

In this experiment we wanted to show that the increased salience of the location where the avatar was presented was insufficient to explain the reversal of orthogonal compatibility observed in Experiment 1. Weeks and Proctor (1990) argued that orthogonal compatibility is the result of the corresponding polarity between up–right- and down–left-oriented SR ensembles. The presentation of a distractor such as the avatar would probably increase the polarity of the hemifield it was presented in, resulting in a positive polarity of the “left” position when the avatar was also presented on the left side. To contrast this idea with the perspective-taking hypothesis, we modified the first experiment by replacing the avatar with geometrical figures that should have also increased the salience of the location they were presented at, but without inducing perspective taking. If the mere presence of a stimulus on the left side of the targets causes a shift of the left side’s polarity toward positive values, the same effect should be achievable by presenting easily noticeable, and therefore salient, geometrical figures. We used two different alternative distractors to replace the avatar in the first two blocks of the experiment: a black disc and an arch (Fig. 4). We expected the avatar effect to be more closely related to perspective taking than to a shift in polarity, and to observe a differential effect of the distractor types.

Hypotheses

In the conditions that used the avatar as a distractor, we predicted a replication of the first experiment’s results. In contrast, we expected to observe a significantly reduced influence of the other distractors as compared to the avatar. This

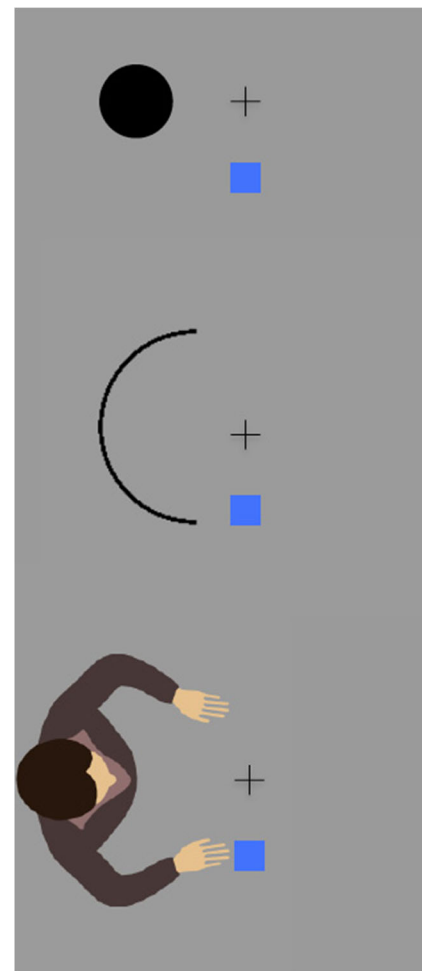


Fig. 4 Schematic examples of the conditions in Experiment 2: (Top) Distractor disc. (Middle) Distractor arch. (Bottom) Distractor avatar. Only the 90°-rotated conditions with dark blue targets are shown.

prediction should be observable in the form of a significant interaction between the factors distractor type, orthogonal compatibility, and distractor rotation.

Method

Participants and sensitivity Sixteen naive students (14 female, two male) from the same pool as in Experiment 1, with a mean age of $M = 22.4$ ($SD = 4.4$), participated in the experiment. Participation was compensated with course credit. On the basis of the observations in Experiment 1, we assumed that the population effect size of the interaction between avatar rotation and orthogonal compatibility is close to the measured effect size of $\eta_p^2 = .40$. According to G*Power (Faul et al., 2007), a sample size of 16 participants would suffice to measure effects of this magnitude. This sample size would yield a power of $(1 - \beta) = .84$, assuming a true effect size of $\eta_p^2 = .40$.

Apparatus, stimuli, and procedure The experimental setup used was similar to that of Experiment 1 and was based on

the same hardware and software used for stimulus presentation. We also used the same target positions and distractor (avatar) rotations described above. We used two different colored squares instead of discs as the target stimuli, with a length and width of 24 pixels (0.86°) and the same light and dark blue color as in Experiment 1. This change was made in order to use a round distractor that was perceptually similar to the avatar's head without introducing an overlap between distractor and target in the dimension of shape.

The experiment consisted of three blocks in which different distractors were presented on the left or right side of the fixation cross (Fig. 4). The distractor location was changed after half the trials in each block, and half of the participants started with a left and the other half with a right distractor position. In the first block, this distractor was a black disc (50 pixels, 1.79°), presented at a 125-pixel (4.46°) distance from the fixation cross in the middle of the screen. In the second block, a black arch replaced the circle, with a gestalt similar to the avatar's arm span, and in the third block the same avatar was used as in Experiment 1. In the third block we therefore aimed to replicate the first experiment's effects, while contrasting them with a possible salience—or polarity-based—effect. All participants performed all three blocks in the same order (disc–arch–avatar). We did not balance the order of distractors between participants, because we assumed that early presentation of the avatar would lead to a different interpretation of the distractors in the following blocks. This could potentially lead the participant to interpret the curved line as arms or the circle as a head, effectively rendering our manipulation inconsequential.

In contrast to Experiment 1, we instructed all participants to ignore the displayed distractors. This made the same instructions applicable to all distractor types, which was more useful than using a steer instruction that could differentially have affected the interpretation of the distractors. No other remarks regarding the shape or meaning of the distractors were given.

Design We continued to label the up–right and down–left SR ensembles as compatible and the down–right/up–left ensembles as incompatible. The result was a $2 \times 2 \times 3$ design with three within-subjects factors: orthogonal compatibility (incompatible vs. compatible), distractor rotation (90° vs. -90°), and distractor type (disc vs. arch vs. avatar).

Results and discussion

The mean RTs and percentage errors were analyzed separately using repeated measure ANOVAs. Applying the same RT criterion as in Experiment 1, a total of 82 trials (0.7%) were excluded, along with 293 false responses (2.3%), for a total of 375 errors (3.0%).

Reaction times A significant interaction of the factors orthogonal compatibility and distractor rotation was observed, with $F(1, 15) = 8.24$, $p = .012$, $\eta_p^2 = .35$. A rotation of -90° (distractor on the right) was associated with an orthogonal compatibility effect of 6 ms, and a rotation of 90° (distractor on the left) produced an orthogonal compatibility effect of -5 ms. This interaction was significantly modulated by the factor distractor type. The observed influences of distractor rotation and orthogonal compatibility, already found in Experiment 1, were only present in the block that used the avatar (Fig. 5), resulting in a three way interaction, with $F(2, 30) = 4.43$, $p = .021$, $\eta_p^2 = .23$. The numerically largest compatibility effects were observed when the avatar was used as a distractor: 14-ms advantage of incompatible conditions in the -90° rotation, and 14-ms advantage of compatible mappings in the 90° rotation, replicating the results of the first experiment.

We further observed a tendency toward a main effect of distractor type, $F(2, 30) = 3.18$, $p = .056$, $\eta_p^2 = .18$, or put differently, an increase of mean RTs over the course of the experiment: The distractor disc was associated with the fastest ($M_{\text{disc}} = 450$ ms), the distractor arch with intermediate ($M_{\text{arch}} = 456$ ms), and the distractor avatar ($M_{\text{avatar}} = 460$ ms) with the slowest mean RTs.

Percentage errors We observed a significant interaction between the factors orthogonal compatibility and distractor rotation, $F(1, 15) = 11.90$, $p = .004$, $\eta_p^2 = .44$, with larger advantages of orthogonally incompatible mappings in the -90° conditions. The results further showed a marginal main effect of distractor type, $F(2, 30) = 2.70$, $p = .083$, $\eta_p^2 = .15$, with increasing error rates over the course of the three blocks.

The results of Experiment 2 support our hypothesis that the observed avatar influence on orthogonal compatibility is a result of perspective taking rather than of a shift in polarity, because the compatibility relations observed in Experiment 1

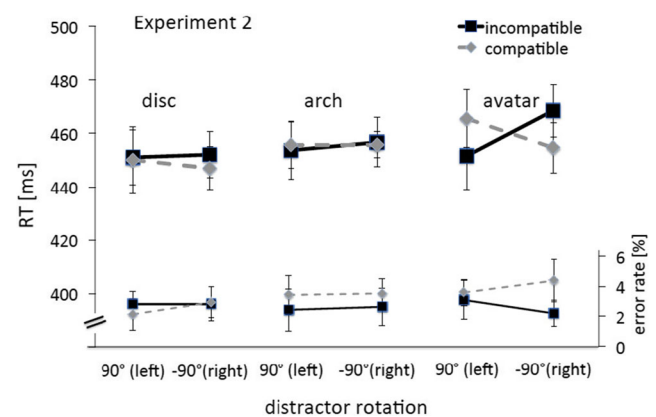


Fig. 5 Mean reaction times (RTs; thick lines) and error rates (thin lines) as a function of orthogonal compatibility, distractor type, and distractor rotation. Error bars represent 95% within-subjects confidence intervals (Morey, 2008).

were only present when the avatar was used as a distractor. The observed main effects of distractor type can most likely be attributed to fatigue over the course of the experiment.

Experiment 3

The previous two experiments showed that participants tend to adopt the perspective of rotated avatars with orthogonal stimulus positions. In the present experiment, we extended the avatar rotations in order to observe perspective taking with larger angular disparities. Because the stimuli were rotated along with the avatar, this change also introduced lateralized stimulus positions. This led to varying degrees of spatial correspondence on the horizontal axis and set the bar higher for an avatar-based compatibility effect. Although the orthogonal Simon effect is usually weak, the standard Simon effect with lateralized mappings is very robust. We wanted to test whether the measured avatar influence can also challenge this strong effect. We expected to measure an influence of the avatar, but predicted that it would be reduced in conditions with further lateralized stimulus positions, because of a greater dimensional overlap between the stimulus and response locations. Larger angular disparities are also associated with increased effort in perspective taking, and this should make perspective taking even less desirable.

To achieve our experimental goals, the avatar and stimulus locations were rotated by 15°, 75°, 105°, and 165° clockwise, from the participant's view (Fig. 6), where the 15° and 165° conditions were associated with a stronger lateralization of the stimulus locations than were the 75° and 105° conditions. We used 15° and 165° instead of 0° and 180° in order to include two conditions that had the same degree of stimulus laterality as each other and were comparable to the spatial relations in a typical Simon task, while avoiding the 0°/180° conditions, which could lead the participant to interpret the scene as being mirrored instead of rotated. This would be problematic, because people usually tend to prefer a mirror explanation over rotation in such scenarios (Sutter & Müsseler, 2010).

Hypotheses

On the basis of the results of the previous experiments, we expected to observe a Simon effect defined by spatial correspondence as seen from the avatar's point of view. This predicted effect would be measurable in the form of an interaction between the factors spatial correspondence and avatar rotation. We expected an advantage of spatially corresponding conditions with rotations of 15° and 75°, but an advantage of spatially noncorresponding conditions with avatar rotations of 105° and 165°, and therefore to find a Simon effect based

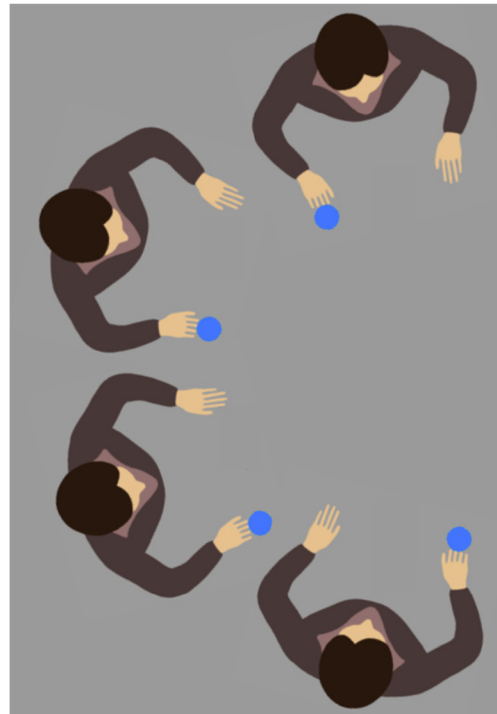


Fig. 6 The four different avatar rotations used in Experiments 3 and 4 (top to bottom: 165°, 105°, 75°, and 15°).

on the spatial correspondence as seen from the *avatar's* point of view, rather than from the participant's own perspective.

Method

Participants and power In total, 16 students (15 female, one male) with a mean age of $M = 24.5$ years ($SD = 9.5$) participated in this experiment for course credit. This sample size would yield a power of $(1 - \beta) = .84$, assuming an effect size of $\eta_p^2 = .40$.

Apparatus, stimuli, and procedure We used the same method as in the previous experiments but the disc-type stimuli from Experiment 1. We changed the degree of rotation of the avatar and the stimuli to 15°, 75°, 105°, and 165° clockwise from the participant's point of view (Fig. 6). We created four different sequences of these rotations that were balanced using the Latin square. Each of the four resulting sequences was completed by four participants, who were all presented with the steer instruction from Experiment 1, in order to make this experiment comparable to a planned follow-up (Exp. 4) that would introduce action effects and include actual steering of the avatar's hands by the participants.

As in the previous experiments, the avatar rotations were presented blockwise. Each block started with 20 practice trials, and each condition was repeated a total of 40 times throughout the experiment, amounting to a total of 640 trials

per participant, excluding practice trials. The participants needed about 45 min to complete the experiment.

Design The factor spatial correspondence was defined from the participant's point of view. The result was a 2×4 design with two within-subjects factors: spatial correspondence (noncorresponding vs. corresponding) and distractor rotation (15° , 75° , 105° , and 165°).

Results and discussion

Overall, 67 trials (0.65%) were marked as errors on the basis of the same RT criterion as in the previous experiments. In addition, 354 trials had incorrect responses (3.46%), for a total of 421 errors (4.11%). RTs and percentage errors were analyzed separately using repeated measures ANOVAs and are shown in Fig. 7. We observed a significant influence of the factor spatial correspondence on the mean RTs, $F(1, 15) = 6.60$, $p < .021$, $\eta_p^2 = .31$, which overall favored spatially corresponding SR ensembles ($M_{\text{noncorr}} = 462$ ms vs. $M_{\text{corr}} = 452$ ms). No other significant effects were observed.

The results indicated a Simon effect aligned with spatial correspondence from the participant's point of view and independent of the avatar's rotation. This was contrary to our prediction, and we therefore had to reject our hypothesis. The most straightforward conclusion is that perspective taking did not take place at higher angular disparities, and that participants simply ignored the avatar. This explanation is supported by the fact that no significant increase in RTs was observed with larger rotations, as would be expected in the case of perspective taking. Since the avatar was irrelevant to the task, it was much easier not to adopt its perspective in order to free central capacity that could be used to complete the response selection stage of the task. In the subsequent experiment, we introduced conditions that should make it

more difficult to ignore the avatar, with the goal of provoking perspective taking in conditions in which it was not observed in this experiment.

Experiment 4

In Experiment 4 we aimed to increase the avatar's influence by providing it with appropriate action effects, in the form of hand movement. The avatar's hand movements corresponded to the participant's response hand (left key presses were followed by movements of the avatar's left hand, and vice versa for right key presses), to strengthen the feeling of control over the avatar, and thereby the sense of agency. We expected this change to force the avatar's movement to be included in the participant's action code as a result of action effect anticipation. With this change, the avatar still remained irrelevant to the task; its hand movements, however, did become a relevant tool for action control—for example, as a means to monitor the response's correctness and the registration of the response by the apparatus. We expected that the irrelevant spatial position of the avatar would become part of the event code, in the same way that irrelevant spatial information related to the target is incorporated into its feature code. To realign the location of effector and its action effect, perspective taking should take place.

Hypotheses

We predicted that the inclusion of action effects would lead to a reappearance of the avatar's influence that had been observed in the first two experiments, which would renew the hypotheses made in Experiment 3.

Method

Participants and power In total, 16 students (13 female, three male) with a mean age of $M = 22.1$ years ($SD = 3.6$) participated in this experiment for course credit. Because the sample size was the same as in Experiment 3, it would also lead to a power of $(1 - \beta) = .84$, assuming an effect size of $\eta_p^2 = .40$.

Stimuli, procedure, and design These were the same as in Experiment 3; only an action effect of the avatar was added, increasing the correspondence of the hand to the participant's response. The hand movement occurred as soon as possible after the corresponding key press (with the next valid frame), and the hand remained in this position until the start of the following trial.

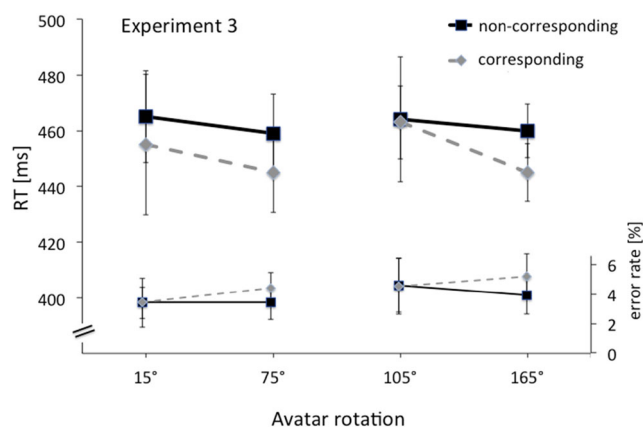


Fig. 7 Mean reaction times (RTs; thick lines) and error rates (thin lines) as a function of spatial SR correspondence and avatar rotation. Error bars represent 95% within-subjects confidence intervals (Morey, 2008).

Results and discussion

Overall, 66 trials (0.64%) were marked as errors, on the basis of the same RT criteria as in the previous experiments. In addition, 347 trials had incorrect responses (3.39%), for a total of 413 errors (4.03%).

Reaction times We found a marginally significant effect of the factor spatial correspondence, with $F(1, 15) = 4.16, p = .059, \eta_p^2 = .22$: Spatially corresponding trials were associated with faster mean RTs overall, as compared to noncorresponding trials ($M_{\text{corr}} = 455$ ms, $M_{\text{noncorr}} = 461$). We further observed a significant interaction between the factors degree of rotation and spatial correspondence, with $F(3, 45) = 26.54, p < .001, \eta_p^2 = .64$. The mean RTs, shown in Fig. 8, display a reversal of the Simon effect for degrees of rotation larger than 90° . Although we did not find a significant main effect of the factor rotation, overall a trend was observed, connecting higher degrees of rotation to slower RTs, $F(3, 45) = 2.26, p = .094, \eta_p^2 = .13$.

Percentage errors We observed a significant interaction between the factors spatial correspondence and rotation, $F(1.88, 28.23) = 6.05, p = .007, \eta_p^2 = .29$ (degrees of freedom were Greenhouse–Geisser adjusted as a result of sphericity violation). The results are in line with the RT data, pointing toward a reversal of the Simon effect in conditions with 105° and 165° avatar rotations.

Combined analysis of Experiments 3 and 4 To estimate the effect size of the action effect manipulation, the mean RTs and percentage errors of the previous two experiments were jointly analyzed in a $4 \times 2 \times 2$ mixed ANOVA, with the within-subjects factors avatar rotation and spatial correspondence, and the between-subjects factor action effect.

We observed a significant three-way interaction of spatial correspondence, rotation, and action effect in mean RTs, $F(3,$

$90) = 11.38, p < .001, \eta_p^2 = .28$, which shows how strongly the interaction between spatial correspondence and avatar rotation was influenced by the presence of action effects. Without action effects, compatibility was based on spatial correspondence from the participant's point of view, but after action effects were added, compatibility sided with spatial correspondence as seen from the avatar's perspective instead. This effect was complemented by an interaction of spatial correspondence, rotation, and action effect in percentage errors, although the effect was smaller, $F(3, 90) = 3.23, p = .043, \eta_p^2 = .10$ (degrees of freedom were Greenhouse–Geisser-adjusted as a result of sphericity violation). No significant effects other than the ones reported in the separate analyses were found.

These results support our hypothesis that action effects did nudge participants into taking the avatar's perspective, which was clearly demonstrated by the reversal of the Simon effect in avatar positions that were rotated by 105° and 165° . Surprisingly, this change in compatibility relationships was not accompanied by a significant overall increase in RTs with increased angular disparity, although a trend was observed ($p = .09$). If we take a look at the RT pattern in Fig. 8, the conditions that are compatible from the avatar's point of view (15° compatible, 75° compatible, 105° incompatible, and 165° incompatible) show a different dependency on rotation than in the conditions that are incompatible from the avatar's viewpoint. An exploratory AVOVA of the avatar-compatible condition RTs with only the factor rotation was performed to investigate this, and the influence of rotation was highly significant, $F(3, 45) = 5.51, p = .003, \eta_p^2 = .27$. The corresponding analysis of the avatar-incompatible conditions did not show such an effect, $F(3, 45) = 0.26, p = .85, \eta_p^2 = .02$. This suggests that the participants might have approached the task differently on a trial-by-trial basis. They may have used perspective taking if it led to a compatible situation, resulting in an increase of the mean RT with rotation, but not if the result was incompatible from the avatar's perspective. In this way, they were able to reap the compatibility benefits without the additional costs of incompatible conditions after perspective taking. Overall, such a combination of strategies should be the most beneficial, but it was unexpected that participants would be able to switch between the two so quickly.

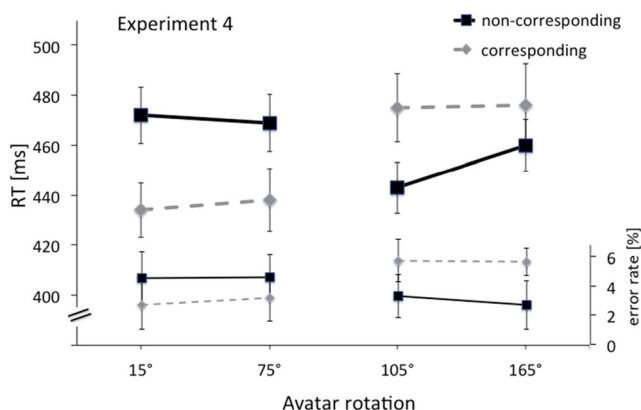


Fig. 8 Mean reaction times (RTs; thick lines) and error rates (thin lines) as a function of spatial SR correspondence and avatar rotation. Error bars represent 95% within-subjects confidence intervals (Morey, 2008).

General discussion

Overall, the results lead us to the conclusion that perspective taking can be invoked toward avatars, even when the avatar is not relevant to the task. However, it also became apparent that this process is influenced by the characteristics of the situation at hand.

The results of Experiments 1 and 2 indicate a general tendency to alter the frame of reference in order to create a dimensional overlap between the stimuli and responses, when

they are orthogonal. Lippa and Adam (2001) attributed the tendency to perform a mental self-rotation in orthogonal mappings to increased comfort with the rotation outcome. It allows the participants to recode left/right–up/down stimulus–response pairs as left/right, compatible/incompatible pairs. Because this representational shift effectively eliminates the up–down axis from the equation, perspective taking might be a valid mechanism to reduce the number of relevant dimensions, which is possibly advantageous for working memory allocation. Our results show that participants prefer the direction of realignment suggested by the avatar over the naturally occurring tendency to map the up and down positions to right and left positions, respectively. These results are in accordance with Freundlieb et al. (2016), who observed a similar compatibility effect defined by the position of a confederate. However, one major difference between Freundlieb et al.'s (2016) study and this one is that the avatars we used in the first experiments did not perform a task themselves. Although perspective taking might be useful in order to understand and interpret a confederate's actions (Tversky & Hard, 2009), this aspect was most likely irrelevant in our experiments, and probably is not a necessary condition for visual perspective taking overall. This also raises some questions about the importance of theory of mind for visual perspective taking.

On the basis of the second experiment, we concluded that a shift in spatial polarity alone, as proposed by Weeks and Proctor (1990), was insufficient to explain the observed compatibility effects in Experiment 1. We currently believe that perspective taking and the resulting different stimulus coding is the most straightforward explanation for the observed effects. Looking back at the major theories of stimulus–response compatibility, the results allow us to make the following observations. On the basis of the dimensional overlap model of Kornblum et al. (1990), the measured compatibility effects point to a dimensional overlap that is caused by the avatar, showing that the avatar is fundamentally different from the other distractors used in the second experiment. It appears the stimuli were coded from the avatar's point of view as either right or left, and therefore had a dimensional overlap with the right or left responses. This dimensional overlap was absent with other distractors, for which therefore no compatibility effects were observed. With regard to the theory of event coding, we think that stimuli and responses activate the same feature codes in conditions that are spatially corresponding from the avatar's point of view. When we look at the response discrimination account of Ansgore and Wühr (2004), we can draw a similar conclusion. In our first two experiments, the account would predict no Simon effect without an alteration of stimulus coding, because the stimuli appear on the vertical axis whereas the responses are made on the horizontal. But with the avatar present, we did indeed find a compatibility effect, suggesting that the stimulus position were part of the axis that discriminated between responses. This

avatar-based compatibility effect can be reconciled with the response discrimination account if we accept that the stimulus positions were represented from the avatar's point of view instead of the person's. The stimuli were represented on the horizontal axis instead, the same axis that discriminated between responses and the avatar again caused a match between stimuli and responses that was not present before. In essence, the first two experiments show that compatibility based on the avatar's point of view is crucial when it is in conflict with the orthogonal Simon effect. Contrary to the labels we used in our experiments, the faster conditions should be labeled as compatible, which should be defined on the basis of the avatar's point of view; the labeling based on the orthogonal Simon effect should be abandoned.

However, the situation is different when the stimulus and response positions were already lateralized (Exps. 3 and 4) and larger angular disparities had to be overcome to realign oneself with the avatar. Here, stimuli and responses contained a dimensional overlap in the dimension laterality even without perspective taking. Although perspective taking is in this case not useful to reduce the number of relevant dimensions and not needed to match the stimulus to the response codes, the participants could still use it for strategic reasons, for example to increase the dimensional overlap further or to produce it in noncorresponding conditions with rotations larger than 90°. The perspective-taking strategy would lead to an overall performance advantage if the advantage in noncorresponding trials outweighed the disadvantage in corresponding ones.

When comparing the mean RTs of Experiments 3 and 4, it is evident that the observed avatar effects are a lot larger in Experiment 4 and most likely are absent in Experiment 3. However, it is possible that the avatar influence was still present in Experiment 3 but simply was not large enough to be measured with the method used, due to a lack of statistical power. Both Experiments 3 and 4 used the steer instructions, which asked the participant to imagine controlling the avatar's hands with the respective keys and to try to take the avatar's point of view. Since the only difference between the two experiments was the introduction of action effects, the differences in observations between the two are likely a result of this manipulation. The manipulation also introduced varying degrees of response–effect correspondence, depending on the avatar rotation. Correspondence between responses and action effects typically leads to compatibility effects that are conceptually similar to those of SR compatibility (Kunde, 2001; Müsseler, Kunde, Gausepohl, & Heuer, 2008; Müsseler & Skottke, 2011). However, only a small and not significant overall main effect of rotation was observed, which led us to the conclusion that response–effect compatibility is most likely insufficient to explain the differences between Experiments 3 and 4. Perspective taking could, however, be a valid strategy to reduce response–effect incompatibilities, because it results in spatially corresponding representations of response and

effect locations. Perspective taking is associated with an increase of RTs with higher angular disparities (Janczyk, 2013; Lippa & Adam, 2001), and this information leads us to a remarkable observation: Only conditions that were spatially corresponding from the avatar's point of view showed this increase in RT with rotation, whereas the avatar-incompatible conditions did not. This might indicate that participants can switch between the avatar perspective and their own very quickly, and that they might prefer the strategy that produces a compatible ensemble over the other. This, however, is difficult to reconcile with the idea of automatic activation of a response based on the stimulus location, because this automatic activation should preempt the selection of a strategy. There are some ways to solve this dilemma. It is possible that the observed compatibility effects are a result of a general advantage of ipsi- over contralateral mappings in working memory rather than automatic activation, or that the processing of stimulus position is updated or recoded after the strategy is selected. This pattern was absent in Experiment 3, and it seems likely that perspective taking only took place in the presence of appropriate action effects, and therefore of perceived control over the avatar. The additional benefit of eliminating response–effect incompatibilities might have been an additional incentive for perspective taking. To follow up on the relationship between the interpretation of avatar movements and perspective taking, we conducted an experiment that demonstrated a possible connection between body ownership and avatar-based compatibility effects (Böffel & Müsseler 2018). It therefore seems plausible that the differences between Experiments 3 and 4 might be related to differences in perceived ownership as well.

Overall, the results point to the conclusion that perspective taking toward avatars can occur in tasks even when it is not mandated by the primary task itself. Our experiments show that contextual features are crucial to whether perspective taking occurs, and they demonstrate that action effects and agency over the avatar are of vital importance to the processes that lead to perspective taking in avatar control and may allow for strategic or selective use of perspective taking.

Author note This study was supported by the Deutsche Forschungsgemeinschaft (Grant DFG MU 1298/11). The authors thank Marina Papke for her assistance recruiting participants and collecting the data. The authors declare that this study was conducted in the absence of any financial relationships that could be viewed as conflicts of interest.

References

- Ansorge, U., & Wühr, P. (2004). A response-discrimination account of the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 365–377. <https://doi.org/10.1037/0096-1523.30.2.365>
- Avraamides, M. N., Hatzipanayioti, A., & Galati, A. (2015). What's so difficult with adopting imagined perspectives? *Cognitive Processing*, 16, 121–124. <https://doi.org/10.1007/s10339-015-0728-3>
- Bauer, D. W., & Miller, J. (1982). Stimulus–response compatibility and the motor system. *Quarterly Journal of Experimental Psychology*, 34A, 367–380. <https://doi.org/10.1080/14640748208400849>
- Böffel, C., & Müsseler, J. (2018). Perceived ownership of avatars influences visual perspective taking. *Frontiers in Psychology*, 9(May), 1–9. <https://doi.org/10.3389/fpsyg.2018.00743>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. <https://doi.org/10.1163/156856897X00357>
- Brebner, J., Shephard, M., & Cairney, P. (1972). Spatial relationships and S–R compatibility. *Acta Psychologica*, 36, 1–15. [https://doi.org/10.1016/0001-6918\(72\)90040-6](https://doi.org/10.1016/0001-6918(72)90040-6)
- Cho, Y. S., & Proctor, R. W. (2003). Stimulus and response representations underlying orthogonal stimulus–response compatibility effects. *Psychonomic Bulletin & Review*, 10, 45–73. <https://doi.org/10.3758/BF03196468>
- Cho, Y. S., & Proctor, R. W. (2005). Representing response position relative to display location: Influence on orthogonal stimulus–response compatibility. *Quarterly Journal of Experimental Psychology*, 58A, 839–864. <https://doi.org/10.1080/02724980443000359>
- Cho, Y. S., Proctor, R. W., & Yamaguchi, M. (2008). Influences of response position and hand posture on the orthogonal Simon effect. *Quarterly Journal of Experimental Psychology*, 61, 1020–1035. <https://doi.org/10.1080/17470210701467979>
- Cooper, L. A., & Shepard, R. N. (1973). The time required to prepare for a rotated stimulus. *Memory & Cognition*, 1, 246–250. <https://doi.org/10.3758/BF03198104>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Fitts, P. M., & Deininger, R. L. (1954). S–R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology*, 48, 483–492. <https://doi.org/10.1037/h0054967>
- Flavell, J. H., Green, F. L., Flavell, E. R., Watson, M. W., & Campione, J. C. (1986). Development of knowledge about the appearance–reality distinction. *Monographs of the Society for Research in Child Development*, 51(1), 1–87. <https://doi.org/10.2307/1165866>
- Franz, E. A., Sebastian, A., Hust, C., & Norris, T. (2008). Viewer perspective affects central bottleneck requirements in spatial translation tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 398–412. <https://doi.org/10.1037/0096-1523.34.2.398>
- 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. <https://doi.org/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. <https://doi.org/10.1037/xhp0000379>
- Gardner, M. R., & Potts, R. (2011). Domain general mechanisms account for imagined transformations of whole body perspective. *Acta Psychologica*, 137, 371–381. <https://doi.org/10.1016/j.actpsy.2011.04.006>
- Hedge, A., & Marsh, N. W. A. (1975). The effect of irrelevant spatial correspondences on two-choice response-time. *Acta Psychologica*, 39, 427–439. [https://doi.org/10.1016/0001-6918\(75\)90041-4](https://doi.org/10.1016/0001-6918(75)90041-4)

- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175–191. <https://doi.org/10.1016/j.intell.2003.12.001>
- Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, *136*, 189–202. <https://doi.org/10.1016/j.actpsy.2010.04.011>
- Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology*, *6*, 1318:1–5. <https://doi.org/10.3389/fpsyg.2015.01318>
- 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, disc. 878–937. <https://doi.org/10.1017/S0140525X01000103>
- Janczyk, M. (2013). Level 2 perspective taking entails two processes: Evidence from PRP experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1878–1887. <https://doi.org/10.1037/a0033336>
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review*, *97*, 253–270. <https://doi.org/10.1037/0033-295X.97.2.253>
- Kunde, W. (2001). Response–effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 387–394. <https://doi.org/10.1037/0096-1523.27.2.387>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t* tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lippa, Y., & Adam, J. J. (2001). An explanation of orthogonal S–R compatibility effects that vary with hand or response position: the end-state comfort hypothesis. *Perception & Psychophysics*, *63*, 156–174. <https://doi.org/10.3758/BF03200510>
- May, M., & Wendt, M. (2013). Visual perspective taking and laterality decisions: Problems and possible solutions. *Frontiers in Human Neuroscience*, *7*, 549:1–7. <https://doi.org/10.3389/fnhum.2013.00549>
- Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics*, *68*, 327–337. <https://doi.org/10.3758/BF03193680>
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>
- Müsseler, J., Kunde, W., Gausepohl, D., & Heuer, H. (2008). Does a tool eliminate spatial compatibility effects? *European Journal of Cognitive Psychology*, *20*, 211–231. <https://doi.org/10.1080/09541440701275815>
- Müsseler, J., Ruhland, L., & Böffel, C. (2018). Reversed effect of spatial compatibility when taking avatar’s perspective. *Quarterly Journal of Experimental Psychology*. Advance online publication. <https://doi.org/10.1177/1747021818799240>
- Müsseler, J., & Skottke, E.-M. (2011). Compatibility relationships with simple lever tools. *Human Factors*, *53*, 383–390. <https://doi.org/10.1177/0018720811408599>
- Nishimura, A., & Yokosawa, K. (2006). Orthogonal stimulus–response compatibility effects emerge even when the stimulus position is task irrelevant. *Quarterly Journal of Experimental Psychology*, *59*, 1021–1032. <https://doi.org/10.1080/17470210500416243>
- Nishimura, A., & Yokosawa, K. (2010). Visual and auditory accessory stimulus offset and the Simon effect. *Attention, Perception, & Psychophysics*, *72*, 1965–1974. <https://doi.org/10.3758/APP.72.7.1965>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:<https://doi.org/10.1163/156856897X00366>
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703. <https://doi.org/10.1126/science.171.3972.701>
- Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304. <https://doi.org/10.1037/h0020586>
- Simon, J. R., & Small, A. M., Jr. (1969). Processing auditory information: Interference from an irrelevant cue. *Journal of Applied Psychology*, *53*, 433–435. <https://doi.org/10.1037/h0028034>
- Sutter, C., & Müsseler, J. (2010). Action control while seeing mirror images of one’s own movements: Effects of perspective on spatial compatibility. *Quarterly Journal of Experimental Psychology*, *63*, 1757–1769. <https://doi.org/10.1080/17470210903511244>
- Taylor, A., Flynn, M., Edmonds, C. J., & Gardner, M. R. (2016). Observed bodies generate object-based spatial codes. *Acta Psychologica*, *169*, 71–78. <https://doi.org/10.1016/j.actpsy.2016.05.009>
- Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*, 124–129. <https://doi.org/10.1016/j.cognition.2008.10.008>
- Weeks, D. J., & Proctor, R. W. (1990). Salient-features coding in the translation between orthogonal stimulus and response dimensions. *Journal of Experimental Psychology: General*, *119*, 355–366. <https://doi.org/10.1037/0096-3445.119.4.355>
- Weeks, D. J., Proctor, R. W., & Beyak, B. (1995). Stimulus–response compatibility for vertically oriented stimuli and horizontally oriented responses: Evidence for spatial coding. *Quarterly Journal of Experimental Psychology*, *48A*, 367–383. <https://doi.org/10.1080/14640749508401395>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Journal of the American Medical Association*, *310*, 2191–2194. <https://doi.org/10.1001/jama.2013.281053>