The Influence of Membership in Outgroups Varying in Competence and Warmth on Observers' Level-2 Visual Perspective Taking

Abstract

Visual perspective taking (VPT), the ability to adopt another person's viewpoint, entails two distinct processes, Level-1 (L1)-VPT and Level-2 (L2-VPT), referring to the ability to perceive whether and how a target sees an object, respectively. Whereas previous efforts investigated the impact of targets' social characteristics on L1-VPT, the present work is the first to do so regarding L2-VPT. Specifically, we investigate the impact of targets' membership in outgroups varying in perceived competence and warmth, the two fundamental dimensions of social perception. Participants in four experiments engaged in a L2-VPT task. Avatars belonged to a low competence low warmth group (LCLW; e.g., the homeless) or to a high competence low warmth group (HCLW; e.g., bankers) in Experiments 1-3, and to a LCLW or high competence high warmth group (HCHW; e.g., female students) in Experiment 4. Participants answered as quickly as possible whether a cued number matched a number present in a scene from either their own or the avatar's perspective. We consistently found support for the presence of both egocentric and altercentric interference, but this was not modulated by group competence and warmth, suggesting that membership in outgroups varying in competence and warmth does not influence L2-VPT. We discuss the findings' implications in light of recent views on VPT.

Keywords: visual perspective taking, social perception, stereotype content model, competence and warmth

Word count: 8293

The Influence of Membership in Outgroups Varying in Competence and Warmth on Observers' Level-2 Visual Perspective Taking

Imagine eating out and noticing a friend with sauce on what is their right cheek for you. To ensure they find the exact spot, you say 'You have sauce on your left cheek'. Your friend grabs a napkin to wipe it off. This scene illustrates one of many situations in which we successfully put ourselves in others' shoes. Perspective taking, the ability to adopt another person's viewpoint, is central to human interaction, with children at the age of 14 months already showing signs of this ability (Moll & Tomasello, 2004; Sodian et al., 2007).

The present work focuses on visual perspective taking (VPT): the inference of a target's visual point of view (Samson et al., 2010; Surtees et al., 2012). We investigate if and how a target's outgroup membership influences observers' ability to infer *how* this target sees an object (i.e., Level-2 rather than Level-1 VPT; Surtees et al., 2012). To do so, we build on the Stereotype Content Model, delineating the way people appraise social groups in their environment (Fiske et al., 2002; for a recent review of social evaluation models, see Abele et al., 2021). As such, we seek to extend knowledge in both literatures by investigating whether membership in outgroups varying in competence and warmth influences observers' L2-VPT.

Visual perspective taking

VPT entails two distinct levels (Flavell et al., 1981; Michelon & Zacks, 2006). The first level (L1-VPT) has been mostly studied using the dot-perspective task (Flavell et al., 1981; Samson et al., 2010). Typically, participants see an avatar facing left or right in a scene, with dots displayed in front of or behind the avatar, or in both locations. Participants' task is to indicate how many dots are visible, either from their own or the avatar's perspective. Answers under the instruction to adopt the avatar's perspective provide a direct measure of VPT, given the clear mission to embrace the other person's point of view. Answers when instructed to take one's own perspective offer an indirect measure of VPT. Indeed, the

intrusion of the avatar's perspective correspond to a spontaneous adoption of the latter's point of view. Response times and errors vary as a function of the adopted perspective and the congruency between what participants and the avatar see. When they both see the same number of dots (i.e., congruent trials), response times and errors are the lowest for both perspectives. In contrast, when they do not see the same number of dots (i.e., incongruent trials), error rates are the highest. Regarding response times, two types of interference emerge on incongruent trials. When participants answer from the avatar's perspective, the difficulty to inhibit their own viewpoint increases response times compared to congruent trials (i.e., *egocentric* interference). This egocentric interference on a measure of VPT suggests that answering from others' perspectives is effortful (Keysar et al., 2003; Qureshi et al., 2010; Samson et al., 2010; Surtees et al., 2012, 2016). Interestingly, when participants answer from their own perspective, response times also increase compared to congruent trials. This so-called *altercentric* interference reveals spontaneous processing of the avatar's perspective, which is inhibited in order to answer from one's own (Samson et al., 2010; Nielsen et al., 2015; Furlanetto et al., 2016; but see Cole et al., 2020).

L1-VPT is influenced by factors related to both participants (e.g., emotions, Bukowski & Samson, 2016; Todd & Simpson, 2016; self-reported empathy, Sulpizio et al., 2015), and avatars (e.g., social relevance, Nielsen et al., 2015; group membership, Simpson & Todd, 2017; Ferguson et al., 2018). However, results regarding the latter are inconsistent. For instance, using the L1-VPT dot-task Simpson and Todd (2017) found egocentric interference for ingroup (vs. outgroup) avatars on incongruent trials, but no difference based on group membership in altercentric interference. Conversely, and also using the dot-task, Ferguson and colleagues (2018) found that for adult participants, the presence of a child (vs. adult) avatar – an outgroup member – reduces altercentric interference for incongruent trials but does not affect egocentric interference. Both studies demonstrate avatars' group memberships

influencing L1-VPT, but differ as to whether this entails egocentric or altercentric interference.

L2-VPT refers to the ability of observers to perceive *how* a target sees an object and is generally investigated using a paradigm developed by Surtees and colleagues (2012). Participants see an avatar facing them and sitting behind a table, and a number either on the wall next to the avatar or on the table. Beforehand, participants are presented with the word 'self' or 'other', indicating whose perspective to adopt, followed by a number. Their task is to decide as quickly as possible if the presented number matches the one seen by the person whose perspective they had to adopt. Importantly, whereas some numbers are non-ambiguous (i.e., they are perceived similarly from either perspective and location; 0 and 8), others are ambiguous (i.e., they are perceived differently depending on location and perspective; 6 and 9). In some trials, visual perspectives of participants and avatars are congruent (e.g., a 9 on the wall is a 9 for both), but in other trials they are incongruent (e.g., a 9 on the table is a 9 for the participant but a 6 for the avatar; see Figure 1).

Error rates increase when the number is ambiguous and displayed on the table. Participants take longer when adopting avatars' perspectives and particularly so on incongruent (vs. congruent) trials, indicating egocentric interference. Like for L1-VPT, the presence of egocentric interference suggests that answering from a target's perspective is effortful. Interestingly, Surtees and colleagues (2012) did not find altercentric interference. However, in a different set of studies using a similar paradigm, Surtees and colleagues (2016) did find such evidence, but only in a 'mixed' condition, where blocks of trials contained both 'self' *and* 'other' trials (vs. a 'blocked' condition, where blocks contained only 'self' *or* 'other' trials). Also, recent work relying on real confederates as targets found altercentric interference in 'blocked' conditions (Elekes et al., 2016, 2017; Freundlieb et al., 2017, 2018). These results suggest that L2-VPT entails a deliberate and cognitively demanding process

when adopting the avatar's point of view, as evidenced by egocentric interference (Surtees et al., 2013, 2016; Janczyk, 2013). At the same time, L2-VPT also reveals the spontaneous adoption of the avatar's point of view, as evidenced by altercentric interference. The lack of altercentric interference in 'blocked' conditions in Surtees and colleagues' (2016) work might be explained by cartoon avatars (vs. real confederates) not sufficing for spontaneous L2-VPT to emerge in 'blocked' conditions. Notwithstanding this, the question remains as to why altercentric interference failed to emerge previously in 'mixed' conditions (i.e., in Surtees et al., 2012).

[Insert Figure 1 here]

Research identified several moderators of L2-VPT linked to participants. For example, time pressure (Todd et al., 2019) and sleep deprivation (Deliens et al., 2018) reduce efficiency. No research to date considered moderators linked to avatars. Moreover, the limited efforts to examine the impact of avatars' group membership on L1-VPT relied exclusively on an ingroup/outgroup distinction. This is particularly unfortunate because work on group perception suggests that not all outgroups are perceived and treated similarly, as documented by abundant work on the Stereotype Content Model and related models (Fiske et al., 2002; Fiske, 2015; Yzerbyt, 2016; see Abele et al., 2021).

The Stereotype Content Model

The Stereotype Content Model (Fiske et al., 2002) delineates the way stereotypes derive from the nature of groups' interdependence as well as from their hierarchical relations. These structural relations trigger perceptions of varying degrees of warmth – providing information on the intent of a group (i.e., cooperation or competition) – and competence – providing information on the ability of a group to carry out its intent. Moreover, the model proposes that these two dimensions of group perception are orthogonal, forming a four-

quadrant space (Cuddy et al., 2007; Cuddy et al., 2009; Lindqvist et al., 2017). Specifically, some groups come across as high on competence but low on warmth (HCLW, e.g., bankers); others are perceived low on competence but high on warmth (LCHW, e.g., disabled people). Still other groups are seen as high on both dimensions (HCHW, e.g., one's ingroup or groups one collaborates with) or low on both (LCLW, e.g., homeless people). These perceptions of groups' competence and warmth trigger different emotions: HCHW elicits pride, HCLW elicits envy, LCHW elicits pity, and LCLW elicits disgust. Moreover, these perceptions are linked to different behavioral intentions, with tendencies to help or harm groups varying as a function of their perceived levels of competence and warmth (Cuddy et al., 2007). Such findings illustrate the necessity to refine a simple *outgroup* distinction by considering outgroups' levels of competence and warmth.

Research reveals that members of low competence low warmth groups do not necessarily come across as truly social agents but are dehumanized instead. For instance, using the well-known trolley dilemma (participants choose between saving several targets on a trolley-track by sacrificing a target on another track), Cikara and colleagues (2010) showed that people more easily sacrificed members of LCLW groups to save people belonging to groups in other quadrants. Similarly, Harris and Fiske (2006, 2007) confirmed previous findings indicating that people explicitly associate LCLW groups with disgust (Fiske et al., 2002), and showed that in people presented with exemplars belonging to these groups, brain areas associated with social cognition (disgust-related areas) are less (more) likely activated. Participants also used fewer 'mental state' verbs when describing a day in the life of LCLW targets compared to targets from other groups (Harris & Fiske, 2009). Moreover, people are more likely to consider LCLW targets as animals (Vaes & Paladino, 2010), judge it morally more acceptable to ostracize LCLW targets (Rudert et al., 2017), and attribute less mental states to LCLW targets (Cameron et al., 2016). Overall, members of LCLW groups tend to be

dehumanized and denied mental states and human-specific abilities. As Harris and Fiske (2011) noted, these results are surprising given that people normally show a spontaneous ability to embrace a target's perspective, even when this target is an animal. Based on the activation of brain regions related to attention and conflict resolution when people face such targets, these authors argue that whilst people perceive these targets' humanity, they actively dehumanize them and pay less attention to such targets.

The current work

We conjectured that a simple *outgroup* label may not reflect the nuances of group perception, nor their implications with respect to such processes as (visual) perspective taking. Accordingly, we aimed to extend previous work by examining for the first time whether people's direct L2-VPT, i.e., when they answer under the instruction to adopt the avatar's perspective, proves sensitive to outgroup's level of competence and warmth. Because direct L2-VPT requires resources (Surtees et al., 2016; Janczyk, 2013), it should reduce people's ability to inhibit prejudice (Crandall & Eshleman, 2003), making the paradigm suitable to investigate the impact of prejudice on VPT. We used the same procedure as Surtees and colleagues (2012), with the important modification of avatars always being outgroup members and the nature of this outgroup in terms of competence and warmth levels varying across trials.

Based on previous work (Fiske et al., 2002; Harris & Fiske, 2006, 2007, 2009, 2011, Cikara et al., 2010; Cameron et al., 2016; Rudert et al., 2017) we hypothesized that LCLW group membership would deteriorate observers' direct L2-VPT, revealing stronger egocentric interference. That is, compared to targets from other quadrants, people should be less efficient to answer from a LCLW target's point of view on incongruent trials which require actively adopting targets' perspectives (Surtees et al., 2013). Indeed, because on congruent trials observers and targets see the same thing, observers may rely solely on their own perspective

to answer, rendering avatars' group irrelevant. To the extent that people actively dehumanize LCLW targets, they can be expected to not only turn their attention away from such targets (Harris & Fiske 2011) but also to attribute them less mental states (Cameron et al., 2016). This in turn would render more difficult inhibiting their own perspective in order to adopt targets' perspectives.

Experiments 1 and 2 investigated the impact of LCLW targets on direct L2-VPT, albeit with different designs. Given the inconclusive results regarding the presence of altercentric interference in the original paradigm (Surtees et al., 2012), we also sought to replicate Surtees and colleagues' original paradigm in Experiment 1. In Experiments 3 and 4, we replaced neutral avatar drawings (see Figure 1) with prototypical pictures of targets to increase ecological validity. In addition, the avatars' social group in the 'self' trials was now conveyed visually, which allowed checking the impact of LCLW targets on indirect L2-VPT, i.e., when participants answer under the instruction to stick to their own perspective. Experiments 1-3 contrasted LCLW targets with targets high on only one of the two dimensions (e.g., HCLW targets). Experiment 4 contrasted LCLW targets with targets from high competence high warmth (HCHW) groups to maximize social distance.

Experiment 1

In addition to investigating the impact of avatars' outgroup memberships on direct L2-VPT, we also investigated previous discrepancies regarding the presence of altercentric interference in similar versions of the L2-VPT task (Surtees et al., 2012; 2016). Accordingly, we created two conditions. In the *no social group* condition, participants took part in a replication of Surtees and colleagues' (2012). In the *social groups* condition, the avatar belonged to a LCLW (i.e., drug addict) or a HCLW group (i.e., politician). We opted against LCHW groups because research shows that these resemble LCLW groups when it comes to indirect measures (Rohmer & Louvet, 2012, 2016). To convey avatars' group membership in

the *social groups* condition, we replaced the 'other' cue by the avatars' group membership (i.e., 'drug addict' or 'politician').

In both conditions, we expected to replicate Surtees and colleagues' (2012) findings of lower efficiency in the incongruent (vs. congruent) trials for the 'other' perspective (i.e., egocentric interference). With recent work supporting the presence of similar differences between (in)congruent trials for the 'self' perspective (i.e., altercentric interference) in conditions akin to this paradigm (Surtees et al., 2016; Elekes et al., 2016, 2017; Freundlieb et al., 2017, 2018), we also expected to find this. In the *social groups* condition, we additionally expected egocentric interference to vary as a function of avatars' social group. Specifically, we expected participants to experience more difficulties taking the perspective of LCLW targets and predicted less efficiency in the other-incongruent trials (i.e., egocentric interference) for the LCLW compared to the HCLW avatar. It was not possible to investigate the impact of group membership on altercentric interference because neither the written cue nor the avatar conveyed any information regarding the avatar's social group in the 'self' trials.

Method

Participants and design. A total of 125 psychology students took part in exchange for partial course credit. The experiment (and all following experiments) took approximatively 25 minutes to complete. Deletion of one duplicated participation resulted in a final sample of 124 (M_{age}=20.2, SD_{age}=2.23, 112 women). Participants were randomly assigned to one of two between-participants conditions (social characteristics: no social group N=66 vs. social groups N=58). The *no social group* condition involved a 2 (number type: non-ambiguous=0 and 8 vs. ambiguous=6 and 9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) repeated measures design. The *social groups* condition involved the same 2 (number type) x 2

(location) design crossed with a *3* (perspective: self vs. HCLW vs. LCLW) repeated measures design.

Power calculations. Because no prior research examined the impact of different social groups on direct L2-VPT, we adopted a conservative approach and ran a power analysis based on a small effect size (Judd et al., 2017). We conducted our power analysis (here and in all following experiments) for a 2 x 2 x 3 repeated measures design using the PANGEA webapp[1], and aimed for 90% power (Lakens, 2021), which indicated that we needed at least 50 participants to detect a small effect (*d*=.2) in the *social groups* condition. To account for potential dropouts and participants not seriously engaging with the task, here and in following experiments we recruited more participants than required.

Procedure. Participants were tested online using the PsyToolkit platform (Stoet, 2010, 2017). After providing informed consent, they read the instructions. To ensure careful reading, each instruction screen had a minimum reading time before participants could go on. Next, the program randomly assigned participants to conditions. In the *no social group condition*, on each trial participants first saw a fixation cross for 750ms. Subsequently, they saw a cue indicating whether they were to adopt their own perspective (if the cue, displayed for 750ms, was 'self') or the avatar's perspective (if the cue was 'other'). Again for 750ms they then saw one of four numbers (0,8,6,9) followed by a black screen for 500ms. Finally, they saw the same scene as in Surtees and colleagues (2012): an avatar sitting at a table and a non-ambiguous or ambiguous number either on the wall or the table (see Figure 1). Participants were to decide as quickly as possible if the number presented earlier matched the number seen by the target whose perspective they had to adopt. They were to press "O" ("N") if this was (was not) the case. The location of the table, the avatar, and the number were counterbalanced (left vs. right of the scene). Participants in the *social groups condition* went

[1]https://jakewestfall.shinyapps.io/pangea; details of all power analyses can be found in respective Supplementary Materials

through the same procedure, but either had to adopt their own perspective if the cue was 'self', or the avatar's if the cue was the avatar's social group (i.e., 'politician' or 'drug addict').

After these instructions, participants could continue after they correctly answered six consecutive practice trials. Participants who made more than 15 mistakes in the overall 30 maximum number of practice trials were briefly interrupted with additional examples to make sure they understood the instructions before moving on to the test trials. After practice trials, participants completed 128 test trials. To ensure that the addition of group labels did not influence whether participants perceived the avatars as social agents or not, they had to indicate their perception of the avatars as social agents (1=*totally disagree*, 5=*totally agree*). Finally, they provided demographic information and were debriefed, compensated, and thanked.

Results

We conducted all analyses using R (R Core Team, 2020) with the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages. All data and R scripts are available on the following link: https://osf.io/7t4zj/?view_only=6f0db97e307e4a14bb31b8a58b026a15. Following Surtees and colleagues (2012), and as in all following experiments, we discarded the trials with non-matching cued and presented numbers, checked and excluded any participant who failed to perform above chance, and excluded trials for which the response time deviated more than 2.5 standard deviations from the mean (less than 1% of the remaining trials were excluded here and in all following experiments).

Perceived socialness of the avatar. The avatar was perceived as a social agent on a similar level in the *no social group* (M=4.29, SD=0.89) and the *social groups* condition (M=4.15, SD=1.06), as shown by Welch's t-test, $t(112.08)=0.75$, $p=.454$ $\eta_p^2<.01$.

Egocentric and altercentric interference. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), here and as in all following experiments we excluded failed trials from analyses. We conducted all analyses on the inverse efficiency score (IES; Simpson & Todd, 2017). This score is obtained by dividing the mean response times by the rate of correct answers, thus considering both speed and accuracy in a single index, with higher (lower) scores indicating lesser (greater) efficiency to answer from the target's perspective. However, across experiments, we also analyzed response times and errors separately (both measures consistently yielded identical results to the IES; see Supplementary Material 1-4).

We submitted the IES to a 2 (social characteristics: no social group vs. social groups) x 2 (number type: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) mixed model analysis with the avatar's social characteristics varying between participants (see Figure 2). The main effects of number type, location, and perspective were significant, confirming that participants were less efficient when the number was ambiguous, $F(1, 7327.57)=1181.30$, $p<.001$, $\eta_p^2=.10$, on the table, $F(1, 7327.48)=419.68$, $p<.001$, $\eta_p^2=.04$, and when they adopted the avatar's perspective, $F(1, 7327.80)=297.04$, $p<.001$, $\eta_p^2=.03$. There was no main effect of the avatar's social characteristics, $F(1, 121.45)=0.28$, $p=.595$, $\eta_p^2<.01$. We observed a significant perspective by avatar's social characteristics interaction, $F(1, 7327.80)=7.15$, $p=.008$, $\eta_p^2<.01$. Participants were less efficient answering from the avatar's perspective in the *no social group* condition, $F(1, 7326.86)=114.81$, $p<.001$, $\eta_p^2=.01$, and even more so in the *social groups* condition, $F(1, 7328.60)=184.08$, $p<.001$, $\eta_p^2=.02$. No other two-way interactions emerged, all Fs<0.87, all ps>.351, all $\eta_p^2<.01$.

The number type × location × perspective interaction was significant, $F(1, 7327.77)=38.03$, $p<.001$, $\eta_p^2<.01$. Follow-up analyses indicated that the number type ×

location interaction was significant for the 'self' trials, $F(1, 7328.01)=55.00$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference based on location, $F(1, 7326.72)=0.83$, $p=.363$, $\eta_p^2<.01$. In contrast, for ambiguous numbers participants proved less efficient when the number was on the table rather than on the wall, $F(1, 7329.08)=125.92$, $p<.001$, $\eta_p^2=.01$. This pattern reveals altercentric interference with a greater difficulty to answer from one's own perspective when what one sees does not correspond with what the avatar sees. Turning to the 'other' trials, the number type × location interaction also proved significant, $F(1, 7327.34)=261.56$, $p<.001$, $\eta_p^2=.02$. Specifically, for non-ambiguous numbers, participants were less efficient when the number was on the table rather than on the wall, $F(1, 7326.71)=8.83$, $p=.003$, $\eta_p^2<.01$. The same effect was even stronger for ambiguous numbers, $F(1, 7328.02)=642.66$, $p<.001$, $\eta_p^2=.06$. This pattern indicates egocentric interference with a greater difficulty to answer from the avatar's perspective when what one sees does not correspond to what the avatar sees.

The avatar's social characteristics × number type × location × perspective interaction was not significant, $F(1, 7327.77)=0.03$, $p=.862$, $\eta_p^2<.01$, suggesting that the observed altercentric and egocentric interferences were not moderated by the avatar's social characteristics.

Impact of the avatar's social group. Our main objective was to investigate whether participants had a harder time adopting the visual perspective of a LCLW compared to a HCLW target. To do so, we only considered data in the *social groups* condition for the trials in which participants answered from the avatar's perspective. We submitted the IES in these trials to a 2 (number type: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: politician vs. drug addict) repeated measures analysis (see Figure 3).

As before, significant main effects of number type and location confirmed that participants were less efficient when the number was ambiguous, $F(1, 1673.06)=423.80$,

$p<.001$, $\eta_p^2=.14$, or on the table, $F(1, 1672.94)=174.59$, $p<.001$, $\eta_p^2=.06$. There was no effect of perspective, $F(1, 1672.03)=0.12$, $p=.734$, $\eta_p^2<.01$.

Not surprisingly, the number type × location interaction was significant, $F(1, 1672.69)=112.42$, $p<.001$, $\eta_p^2=.04$. Specifically, there was no significant difference in efficiency as a function of location for non-ambiguous numbers, $F(1, 1671.79)=3.57$, $p=.059$, $\eta_p^2<.01$, but for ambiguous numbers participants were less efficient when the number was on the table, $F(1, 1673.75)=272.20$, $p<.001$, $\eta_p^2=.10$. No other two-way interactions emerged, all Fs<2.23, all ps>.135, all $\eta_p^2<.01$.

Contrary to our hypothesis, the number type × location × perspective interaction was not significant, $F(1, 1671.91)=0.01$, $p=.910$, $\eta_p^2<.01$.

[Insert Figure 2 here]

[Insert Figure 3 here]

Discussion

As expected, we observed egocentric interference (Surtees et al., 2012, 2016): participants were less efficient answering from the perspective of the avatar on incongruent trials compared to congruent trials. This pattern emerged regardless of whether avatars' social group was cued or not. This dovetails with participants not perceiving a difference in avatars' socialness as a function of whether their group membership was mentioned or not.

More interestingly, and related to our first objective, we also observed altercentric interference (cf. Surtees et al., 2016) using the original design of Surtees and colleagues (2012). That is, participants proved less efficient answering from their own perspective on incongruent compared to congruent trials. Again, this was the case regardless of whether

avatars' social group was cued or not. Moreover, in line with Surtees and colleagues' (2016) findings, the altercentric interference was smaller than the egocentric interference.

The present results, obtained with a high-powered design, first confirm that spontaneous L2-VPT can indeed emerge in mixed blocks (i.e., blocks containing both 'self' and 'other' trials) using cartoon avatars. Regarding our second objective, we did not find stronger egocentric interference for LCLW than HCLW avatars on incongruent trials. Interestingly, we obtained the same results in a replication study of the *social groups* condition (see Supplementary Material 5). Methodological shortcomings detailed below and addressed in Experiment 2 might have contributed to outgroup membership failing to influence egocentric interferences.

Experiment 2

In Experiment 1, half of the trials were 'self' trials, even though our focus was on the difference between the two versions of the 'other' trials. This may have led participants to focus more on the 'self' versus 'other' differentiation, thus minimizing their differentiation between 'politician' versus 'drug addict'. Additionally, the avatar was a member of only one of two social groups. Considering only one social group per quadrant of the Stereotype Content Model (Fiske et al., 2002) may have reduced the range of stereotypes activated in relation to the respective quadrant. Finally, we used social groups as a cue to define avatars' group membership, but participants never had to use this information. Accordingly, they might have paid little attention to avatars' group membership per se, focusing instead on whether the cue was 'self' or 'not self'. Experiment 2 addresses these potential shortcomings (see below). Again, we expected a difference in egocentric interference between the LCLW and HCLW targets. As in Experiment 1, difference in altercentric could not be investigated.

Method

Participants and design. A total of 71 students (M_{age}=20.72, SD_{age}=3.41, 60 women) took part in an experiment involving a 2 (number type: non-ambiguous=0 and 8 vs. ambiguous=6 and 9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) repeated measures design.

Power calculations. Because of the changes in the design, we conducted a new power analysis, which indicated that we needed at least 46 participants to achieve 90% power to detect a small effect (d=.2).

Procedure. The procedure was identical to Experiment 1, except for the following modifications. First, we dropped the no social group condition. Second, we increased the total amount of test trials to 192. Third, because our analyses rely only on 'correct' trials, only a third of these trials were 'incorrect' trials in order to enhance power. Fourth, given that our concern mainly resides in the difference between HCLW and LCLW groups, only a third of the 192 trials were 'self' trials. Fifth, we added two additional social groups in each quadrant, thus considering six social groups ('drug addicts', 'homeless', and 'welfare recipients' for the LCLW quadrant; 'politicians', 'lawyers', and 'managers' for the HCLW quadrant). Finally, immediately following 15 randomly chosen trials, we asked participants to indicate the identity of the avatar's social group to ensure they paid attention to this.

Results

Six participants erred on more than 20% of the random questions regarding the avatar's social group but excluding them from analyses did not change the results. Thus, we decided to keep them.

Inverse Efficiency Score. We submitted the IES to a 2 (number type: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) repeated measures analysis (see Figure

4). Significant effects of number type and location indicated that participants were less efficient when the number was ambiguous, $F(1, 8283.66)=1249.89$, $p<.001$, $\eta_p^2=.09$, or on the table, $F(1, 8283.65)=542.02$, $p<.001$, $\eta_p^2=.04$.

A set of planned contrasts allowed comparing the IES for 'self' versus 'other' (combining HCLW and LCLW avatars), as well as for the HCLW versus LCLW avatars. The 'self' versus 'other' contrast proved significant, $F(1, 8284.30)=192.60$, $p<.001$, $\eta_p^2=.01$. Participants were less efficient for the avatar than for themselves, regardless of avatars' social group. As in Experiment 1, we failed to find significant difference between LCLW and HCLW targets, $F(1, 8284.16)=1.63$, $p=.201$, $\eta_p^2<.01$, and no interaction including this contrast emerged, all Fs<0.14, all ps>.708, all $\eta_p^2<.01$.

The number type × location × 'self' versus 'other' contrast interaction was significant, $F(1, 8283.78)=23.13$, $p<.001$, $\eta_p^2<.01$. Follow-up analyses revealed that the number type × location interaction was significant for 'self' trials, $F(1, 8283.43)=42.19$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference as a function of location, $F(1, 8282.50)<0.01$, $p=.965$, $\eta_p^2<.01$. However, for ambiguous numbers, participants proved less efficient when the number was displayed on the table rather than on the wall, $F(1, 8284.17)=77.27$, $p<.001$, $\eta_p^2<.01$.

Turning to 'other' trials, the number type × location interaction was also significant, $F(1, 8284.16)=442.45$, $p<.001$, $\eta_p^2=.03$. Specifically, for non-ambiguous numbers, participants were less efficient when the number was displayed on the table rather than on the wall, $F(1, 8282.56)=9.54$, $p=.002$, $\eta_p^2<.01$. This difference emerged even more strongly for ambiguous numbers, $F(1, 8285.51)=1017.07$, $p<.001$, $\eta_p^2=.07$.

[Insert Figure 4 here]

Discussion

The results of Experiment 2 replicated the findings of Experiment 1 and of Surtees et al. (2012) regarding egocentric interference. Participants were less efficient when answering from the avatar's perspective for the incongruent trials compared to the congruent trials. Moreover, as in Surtees (2016), we again found altercentric interference in that participants were less efficient answering from their own perspective when the trial was incongruent (vs. congruent). As before, altercentric interference was smaller than egocentric interference. Together, these findings confirm that the L2-VPT task is a robust task to detect both egocentric and altercentric interference.

Once again, we did not find the predicted effect of avatars' social group on direct L2-VPT, despite addressing a series of limitations inherent to Experiment 1. We found no significant difference between the LCLW and the HCLW avatars, regardless of the scenes' characteristics. Furthermore, we obtained the same results in an additional study when not pressuring participants for time (using a design comprising elements from both Experiments 1 and 2; see Supplementary Materials 6). These results suggest that targets' social group does not influence direct L2-VPT. However, issues relating to the ecological validity of the procedure may also account for the findings. We addressed them in Experiment 3.

Experiment 3

One remarkable feature of previous designs is that they failed to capitalize on the strong relationship between visual cues and group stereotypes (Dotsch et al., 2008; [BLINDED]). Indeed, we used visually neutral cartoon characters instead of real people, but in real life people draw different inferences from visual appearances (e.g., Harris & Fiske, 2006; Todorov et al., 2015). Using the same cartoon character in all trials, with only the verbal group cue varying may have blurred differences between the two groups. In Experiment 3, we aimed to achieve a more ecologically valid design by introducing six

pictures depicting prototypical members of LCLW and HCLW groups (three per quadrant). We decided to use only one group per quadrant (i.e., 'drug addict' and 'banker'; cf. Experiment 1) but to alter pictures within each group, also because Experiment 2 suggested that having multiple groups did not change effects.

More importantly, using pictures rather than neutral avatars allowed for an investigation of the impact of targets' social group on indirect measures of L2-VPT, with avatars' social group in the 'self' trials now conveyed visually. Because indirect measures investigate the spontaneous processing of targets' perspectives, we did not expect a difference of altercentric interference as a function of targets' social group[2].

Regarding egocentric interference, we once again expected lesser efficiency in the incongruent trials for the LCLW target compared to the HCLW target.

Method

Participants and design. A total of 60 participants (M_{age}=27.92, SD_{age}=7.69, 21 women) took part on Prolific Academic[3] in exchange of £3.05. The experiment involved a 2 (number type: non-ambiguous=0 and 8 vs ambiguous=6 and 9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (group quadrant: HCLW vs. LCLW)[4] repeated measures design.

Power calculations. Considering the modifications in the design, a power analysis indicated that we needed at least 23 participants to achieve 90% power to detect a small effect (d=.2).

[2] Based on recent findings showing that spontaneous VPT requires targets having visual access to stimuli (Elekes et al., 2017; Freundlieb et al., 2017, 2018), one could alternatively hypothesize that membership in a dehumanized outgroup would impair observers' attention to the target (Harris & Fiske, 2011), which in turn would impair observers' ability to perceive that targets have visual access to the stimuli. This could result in a smaller altercentric interference for LCLW compared to HCLW targets. Fortunately, our statistical tests set out to test both hypotheses. We would like to thank an anonymous reviewer for suggesting relevant literature pointing to this alternative hypothesis.

[3] https://prolific.co/

[4] Neutral avatars in the previous experiments did not have a social group in the 'self' trials, and we thus relied on a 3-levels perspective factor. Here avatars' social group is marked even in the 'self' condition, and we could thus rely on a fully crossed 4 factorial design.

Procedure. Procedures were as before, with three modifications. First, we used pictures as avatars. Second, only a fourth of the trials were 'self' trials (as opposed to one third in Experiment 2). Finally, participants were randomly asked to indicate targets' social group 30 times (as opposed to 15 times in Experiment 2).

Results

Inverse Efficiency Score. We submitted the IES to a 2 (number type: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCLW vs. LCLW) repeated measures analysis (see Figure 5). Significant effects of type of number, location, and perspective indicated that, regardless of avatars' social group participants were less efficient when the number was ambiguous, $F(1, 7299.1)=636.41$, $p<.001$, $\eta_p^2=.06$, on the table, $F(1, 7299.14)=329.69$, $p<.001$, $\eta_p^2=.03$, and when answering from avatars' perspectives, $F(1, 7299.20)=44.27$, $p<.001$, $\eta_p^2<.01$. Contrary to expectations, but in line with our earlier results, there was no main effect of quadrant group, $F(1, 7299.10)=0.29$, $p=.59$, $\eta_p^2<.01$, and no interactions including this factor, all Fs<2.7, all ps>.096, all $\eta_p^2<.01$.

The number type × location × perspective interaction was significant, $F(1, 7299.11)=20.81$, $p<.001$, $\eta_p^2<.01$. Follow-up analyses revealed that the number type × location interaction was significant for 'self' trials, $F(1, 7299.13)=31.55$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, no significant difference in efficiency as a function of location emerged, $F(1, 7299.10)=2.60$, $p=.107$, $\eta_p^2<.01$. However, for ambiguous numbers participants were less efficient when the number was on the table, $F(1, 7299.20)=90.35$, $p<.001$, $\eta_p^2<.01$.

Turning to the 'other' trials, the number type × location interaction was also significant, $F(1, 7299.14)=379.37$, $p<.001$, $\eta_p^2=.03$. Specifically, for non-ambiguous numbers, participants proved less efficient when the number was on the table, $F(1,$

7299.02)=6.93, p=.009, η_p^2<.01. This pattern emerged even more strongly for ambiguous numbers, F(1, 7299.22)=897.45, p<.001, η_p^2=.08.

[Insert Figure 5 here]

Discussion

We replicated our previous results, as well as those of Surtees and colleagues (2016), regarding the presence of both altercentric and egocentric interference. Once again, altercentric interference was less marked than egocentric interference.

Using pictures conveying avatars' group membership instead of neutral avatars allowed us to investigate the potential impact of targets' social group membership on indirect L2-VPT. In line with our expectations and with previous results regarding spontaneous VPT (Surtees et al., 2016; Elekes et al., 2016, 2017; Freundlieb et al., 2017, 2018) we did not observe a significant difference in altercentric interference between LCLW and HCLW targets for the 'self' trials. Contrary to our hypothesis, but replicating our previous results, egocentric interferences between LCLW and HCLW targets on direct measures did not differ.

As explained above, we chose groups from the HCLW quadrant as a comparison because they constitute a less biased and more straightforward point of comparison with LCLW quadrant groups. However, it may be that the social distance between these groups is not large enough to generate differences in L2-VPT. Indeed, while being different on the competence dimension, both these targets come across as lacking warmth (Fiske et al., 2002; Cuddy et al., 2007). Consequently, and to maximize the distance between the target groups, Experiment 4 relied on HCHW group members as comparison.

Experiment 4

To maximize the distance between the targets taken into consideration, we replaced HCLW targets with HCHW targets differing on both fundamental dimensions. This additionally allowed testing for an alternative hypothesis grounded in the *social projection* literature. Because of the active dehumanization associated with LCLW targets (Harris & Fiske, 2006, 2007, 2011), we expected people to direct their attention away from such targets and in turn experience more difficulty to inhibit their own perspective and to take on that of LCLW compared to HCHW targets. However, from a social projection perspective the opposite might prove true. Social projection refers to the process of believing that other people think, feel, and perceive similarly to ourselves (Krueger, 2007) and is especially strong for ingroup members (Robbins & Krueger, 2005). Consequently, if people believe that the more a target resembles them, the more this target will perceive things like them, then they should prove less efficient to take on the perspective of a socially close versus distant target when both perspectives diverge. In other words, from a social projection perspective inhibiting one's point of view for a socially close target (a HCHW (in)group member) could prove especially difficult. At the same time, it should be easier for people to inhibit their perspective for a socially distant target (a LCLW group member) because people should project less (Robbins & Krueger, 2005).

In summary, if inhibiting one's perspective is harder for a dehumanized target, we would expect stronger egocentric interference as a function of targets' social group in the incongruent trials, in the form of less efficiency for LCLW compared to HCHW targets. However, if inhibiting one's perspective is harder for a target perceived as socially closer to the self, we would expect the reverse pattern, with people being less efficient for HCHW compared to LCLW targets. Such results would be in accordance with Simpson and Todd's (2017) findings that social projection impairs observers' direct L1-VPT. As in the previous

studies, we did not expect a difference in altercentric interference as a function of targets' social group membership[5].

Method

Participants and design. A total of 34 psychology students (M_{age}=20.85, SD_{age}=2.56, 29 women[6]) took part in the study in exchange for partial course credit. The experiment involved a 2 (number type: non-ambiguous=0 and 8 vs. ambiguous=6 and 9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) 2 (group quadrant: HCHW vs. LCLW) repeated measures design.

Power calculations. A power analysis indicated that we needed at least 23 participants to achieve 90% power to detect a small effect (d=.2).

Procedure. The procedures were identical to Experiment 3, with two modifications. First, because most participants were female students, we replaced the HCLW group with a HCHW group ('female students'; the LCLW group remained 'homeless') to maximize the perceived social closeness and, in turn, social projection (Krueger, 2007). Second, we relied on two pictures per quadrant avatar instead of three.

Results

Inverse Efficiency Score. We submitted the IES to a 2 (number type: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCHW vs. LCLW) repeated measures analysis (see Figure 6). Significant effect of number type, location, and perspective again indicated that, regardless of avatars' social group, participants were less efficient when the number was ambiguous, F(1, 4100.09)=388.79, p<.001, η_p^2=.06, on the table, F(1, 4100.06)=168.65, p<.001, η_p^2=.03, and when answering from avatars' perspective, F(1, 4100.08)=15.08, p<.001, η_p^2<.01. Again, and

[5] See Footnote 3

[6] The HCHW group avatars depicted female students; results do not differ when excluding men from the analyses.

contrary to our expectations, there was no main effect of quadrant group, $F(1, 4100.05)<0.01$, $p=.997$, $\eta_p^2<.01$, and no interaction including this factor, all Fs<2.97, all ps>.085, all $\eta_p^2<.01$.

The number type × location × perspective interaction was significant, $F(1, 4100.13)=9.03$, $p=.003$, $\eta_p^2<.01$. Follow-up analyses revealed that the number type × location interaction was significant for 'self' trials, $F(1, 4100.12)=22.89$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers there was no significant difference in efficiency as a function of location, $F(1, 4100.04)=0.36$, $p=.546$, $\eta_p^2<.01$. However, for ambiguous numbers, participants proved less efficient when the number was on the table, $F(1, 4100.17)=54.18$, $p<.001$, $\eta_p^2<.01$.

Turning to the 'other' trials, the number type × location interaction was also significant, $F(1, 4100.08)=215.38$, $p<.001$, $\eta_p^2=.04$. Specifically, for non-ambiguous numbers, there was no significant difference in efficiency as a function of location, $F(1, 4100.06)=1.68$, $p=.195$, $\eta_p^2<.01$. However, for ambiguous numbers, participants were less efficient when numbers were on the table, $F(1, 4100.09)=479.54$, $p<.001$, $\eta_p^2=.08$.

[Insert Figure 6 here]

Discussion

Experiment 4 sought to maximize differences between social targets by comparing HCHW (in)groups and LCLW outgroups. This modification allowed us to investigate both our original hypothesis (i.e., a higher egocentric interference for LCLW vs other outgroup members), as well as an alternative social projection hypothesis according to which people should prove less efficient when adopting the perspective of socially close targets (Robbins & Krueger, 2005; Krueger, 2007).

Once again, we replicated our and Surtees and colleagues' (2016) results regarding the presence of both altercentric and egocentric interference: participants had greater difficulties answering on incongruent trials compared to congruent trials, both when answering from their own (indicating altercentric interference) and from targets' perspectives (indicating egocentric interference). Again, altercentric interference was less marked than egocentric interference.

Contrary to our expectations, but in line with our previous results, there was no significant difference in egocentric interference between LCLW and HCHW targets. We thus found no support for our initial hypothesis, nor for an alternative social projection hypothesis. Once again, we failed to observe a significant difference in altercentric interference between LCLW and HCHW targets.

General Discussion

The present experiments explored the impact of others' membership in outgroups varying in competence and warmth on L2-VPT. In doing so, we went beyond a simple *outgroup* label (Simpson & Todd, 2017, Ferguson et al., 2018) by capitalizing on the Stereotype Content Model and relying on work suggesting that groups lacking both competence and warmth (LCLW) are less readily considered as social beings and more likely to be actively dehumanized (Harris & Fiske, 2006, 2007, 2009, 2011; Cikara et al., 2010; Cameron et al., 2016; Rudert et al., 2017). This active dehumanization should entail people attributing less mental states to and directing their attention away from such targets (Harris & Fiske, 2011), and in turn reduce people's inclination to take the perspective especially of members of such groups. Accordingly, we hypothesized that people would show larger egocentric interference in L2-VPT on trials in which their own perspective and that of LCLW group members diverged (i.e., incongruent trials) – compared to other social groups (HCLW groups in Experiments 1–3; HCHW (in)groups in Experiment 4).

As in past research (Surtees et al., 2012, 2016), we consistently observed egocentric interference such that participants proved less efficient when answering from targets' perspectives on incongruent compared to congruent trials. More interestingly, we also consistently observed altercentric interference in that participants were less efficient when answering from their own perspective on incongruent compared to congruent trials. In light of the systematic presence of altercentric interference across our highly powered experiments, including a direct replication of Surtees and colleagues (2012), its absence in the original work (Surtees et al., 2012) seems likely due to some interfering factor (see Todd et al., 2019). Overall, our findings support the claim that people spontaneously consider targets' perspectives, even when it does not prove relevant for the task or when the target is an avatar (Surtees et al., 2016; Elekes et al., 2016, 2017; Freundlieb et al., 2017, 2018).

At the same time, we failed to find support for our hypothesis of a moderating influence of groups' level of competence and warmth on direct L2-VPT. Participants did not prove less efficient to answer while taking the visual perspective of LCLW compared to HCLW or HCHW targets. In Experiment 4, we also tested an alternative hypothesis informed by research on social projection (Krueger, 2007). Indeed, being socially especially close to a group (a HCHW group) might also impair observers' VPT ability because they project their own perspective to close others (Clement & Krueger, 2002). The data failed to provide support for this alternative conjecture. Furthermore, we did not find evidence for a moderating influence of groups' level of competence and warmth on indirect L2-VPT. Overall, the present findings suggest that targets' outgroup characteristics considered in the present work (i.e., competence and warmth) do not influence observers' L2-VPT efficiency. This is surprising in light of past work showing that mere outgroup membership negatively influences the less resource-demanding L1-VPT (Simpson & Todd, 2017; Ferguson et al., 2018).

An explanation for the absence of modulation of egocentric interference by targets' group characteristics can be found in Westra's (2017) suggestion that L2-VPT might only require relying on general knowledge but not actively adopting targets' perspectives: that a six is a nine when viewed from a 180° angle may be common knowledge for adult participants. But with no need to actually adopt targets' perspectives, there is no room for targets' social characteristics intervening. One way future research could circumvent this might be by relying on triple digit numbers or new perspective-dependent stimuli as well as broadening the range of possible answers beyond a simple *yes* vs *no*, thus increasing task difficulty. This should leave even less resources for participants to suppress their prejudice (Crandall & Eshleman, 2003) and render relying on general knowledge insufficient.

Regarding the absence of modulation of altercentric interference by targets' group characteristics, the submentalizing account (Heyes, 2014; Santiesteban et al., 2014) might offer an explanation. According to this account, spontaneous perspective taking is not triggered by social targets but is merely the unintentional processing of attentional cues, with participants not distinguishing between avatars but simply perceiving their directional attention (e.g., gaze or posture). Having said this, evidence for the submentalizing account is inconclusive, with some work showing that avatars' characteristics moderate the spontaneity of L1-VPT (Nielsen et al., 2015; Furlanetto et al., 2016).

Notwithstanding our contribution, some shortcomings remain. Firstly, we considered only a small number of groups, and some participants may have been confronted with groups they would not dehumanize, thus precluding the emergence of stronger interference. Future work could measure what specific LCLW groups participants would feel especially negative about, and idiosyncratically present such groups. This would ensure that each participant would actually perceive the avatars stemming from different social groups as members of LCLW groups. Moreover, the inclusion of more groups on such an idiosyncratic basis would

ensure that the influence of each quadrant of the SCM (Fiske et al., 2002) is tested based on several groups. Such an approach has obvious benefits. Indeed, more accurately representing the number of groups occupying the social space (Cuddy et al., 2009, Binggeli et al., 2014) broadens the range of activated stereotypes and allows for generalization.

Secondly, although we relied on pictures in our later experiments, one might question their ecological validity: picture of drug-addicts might not activate stereotypes the same way as seeing them in the streets. Although the literature on social perception shows that merely showing words or images associated with specific outgroups is enough to trigger different treatments, such as reduced empathy (Xu et al., 2009; Cikara et al., 2014), performing a VPT task with a real person as target differs from a similar task with avatars (Elekes et al., 2016, 2017). In light of the challenge of using real LCLW targets, using more immersive methodologies such as virtual reality could overcome such difficulties while improving ecological validity (Herrera et al., 2018).

Regarding a potential impact of social projection on VPT, the mere presence of socially close (in)group members in the L2-VPT task did not seem sufficient for group membership to impact observers' VPT, contrary to what has been observed regarding L1-VPT (Simpson & Todd 2017; Ferguson et al., 2018). The absence of effect on our direct measure of L2-VPT is rather surprising for two reasons. First, direct L2-VPT is more cognitively demanding than direct L1-VPT (Surtees et al., 2013; 2016) and should thus leave less room to inhibit prejudice (Crandall & Eshleman, 2003). Second, we contrasted targets that are at a great social distance from one another (Cuddy et al., 2009). However, it could be that joint group membership perhaps was not salient or meaningful enough for participants in our design, which would have reduced or precluded the emergence of social projection (Robbins & Krueger, 2005). Still, such an interpretation seems unlikely as participants took part in our study as part of a methodology course and in the faculty of psychology lab, which

together should have rendered salient a *student* identity. Research interested in testing a social projection account (Clement & Krueger, 2002; Krueger, 2007) could ensure ingroup membership salience as in previous work on L1-VPT (Simpson & Todd, 2017), or by manipulating mortality salience (Castano et al., 2002) or entitativity (Hogg et al., 2007; Castano et al., 2003).

By examining the impact of outgroups' level of warmth and competence on L2-VPT, the current work contributes to the literature on visual perspective taking (Surtees et al., 2012, 2016) as well as work on the Stereotype Content Model (Fiske et al., 2002; Cuddy et al., 2007, 2009, Abele et al., 2021). Although previous research showed LCLW group members not being perceived as truly social agents, and that these damaging perceptions engender differential treatments (Harris & Fiske, 2006, 2007, 2009, 2011; Cikara et al., 2010), we did not find evidence of this being the case for neither direct nor indirect L2-VPT. Clearly, more work is required on both L1- and L2-VPT to understand if and under which conditions targets' membership in social outgroups – whether dehumanized or not – interferes with adopting others' visual perspective.

References

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.

Binggeli, S., Krings, F., & Sczesny, S. (2014). Stereotype Content Associated with Immigrant Groups in Switzerland. *Swiss Journal of Psychology*, *73*(3), 123–133. https://doi.org/10.1024/1421-0185/a000133

Bukowski, H., & Samson, D. (2016). Can emotions influence level-1 visual perspective taking? *Cognitive Neuroscience*, *7*(1–4), 182–191. https://doi.org/10.1080/17588928.2015.1043879

Cameron, C. D., Harris, L. T., & Payne, B. K. (2016). The Emotional Cost of Humanity: Anticipated Exhaustion Motivates Dehumanization of Stigmatized Targets. *Social Psychological and Personality Science*, *7*(2), 105–112. https://doi.org/10.1177/1948550615604453

Castano, E., Yzerbyt, V., Paladino, M.-P., & Sacchi, S. (2002). I Belong, therefore, I Exist: Ingroup Identification, Ingroup Entitativity, and Ingroup Bias. *Personality and Social Psychology Bulletin*, *28*(2), 135–143. https://doi.org/10.1177/0146167202282001

Castano, E., Yzerbyt, V., & Bourguignon, D. (2003). We are one and I like it: The impact of ingroup entitativity on ingroup identification. *European Journal of Social Psychology*, *33*(6), 735–754. https://doi.org/10.1002/ejsp.175

Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the trolley track: Neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience*, *5*(4), 404–413. https://doi.org/10.1093/scan/nsq011

Cikara, M., Bruneau, E., Van Bavel, J. J., & Saxe, R. (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of experimental social psychology*, *55*, 110-125. https://doi.org/10.1016/j.jesp.2014.06.007

Clement, R. W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology*, *38*(3), 219-231. https://doi.org/10.1006/jesp.2001.1503

Cole, G. G., Millett, A. C., Samuel, S., & Eacott, M. J. (2020). Perspective-taking: In search of a theory. *Vision*, *4*(2), 30. https://doi.org/10.3390/vision4020030

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. Psychological bulletin, 129(3), 414. https://doi.org/10.1037/0033-2909.129.3.414

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648. https://doi.org/10.1037/0022-3514.92.4.631

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., Bond, M. H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T. T., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodríguez-Bailón, R., Morales, E., Moya, M., … Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, *48*(1), 1–33. https://doi.org/10.1348/014466608X314935

Deliens, G., Bukowski, H., Slama, H., Surtees, A., Cleeremans, A., Samson, D., & Peigneux, P. (2018). The impact of sleep deprivation on visual perspective taking. *Journal of Sleep Research*, *27*(2), 175–183. https://doi.org/10.1111/jsr.12595

Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, *19*(10), 978–980. https://doi.org/10.1111/j.1467-9280.2008.02186.x

Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93-103. https://doi.org/10.1016/j.concog.2016.02.010

Elekes, F., Varga, M., & Király, I. (2017). Level-2 perspectives computed quickly and spontaneously: Evidence from eight-to 9.5-year-old children. *British Journal of Developmental Psychology, 35*(4), 609-622. https://doi.org/10.1111/bjdp.12201

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2018). Age of avatar modulates the altercentric bias in a visual perspective-taking task: ERP and behavioral evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(6), 1298–1319. https://doi.org/10.3758/s13415-018-0641-1

Fiske, S. T. (2015). Intergroup biases: A focus on stereotype content. *Current Opinion in Behavioral Sciences*, *3*, 45–50. https://doi.org/10.1016/j.cobeha.2015.01.010

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology*, *17*(1), 99–103. https://doi.org/10.1037/0012-1649.17.1.99

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*(6), 1065. https://doi.org/10.1037/xhp0000379

Freundlieb, M., Kovács, Á. M., & Sebanz, N. (2018). Reading your mind while you are reading—evidence for spontaneous visuospatial perspective taking during a semantic categorization task. *Psychological science*, *29*(4), 614-622. https://doi.org/10.1177/0956797617740973

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 158–163. https://doi.org/10.1037/xhp0000138

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low. *Psychological Science*, *17*(10), 847–853. https://doi.org/10.1111/j.1467-9280.2006.01793.x

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*(1), 45–51. https://doi.org/10.1093/scan/nsl037

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*(1), 192–231. https://doi.org/10.1080/10463280902954988

Harris, L. T., & Fiske, S. T. (2011). Dehumanized perception: A psychological means to facilitate atrocities, torture, and genocide? *Zeitschrift Für Psychologie*, *219*(3), 175–181. https://doi.org/10.1027/2151-2604/a000065

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, *13*(10), e0204494. https://doi.org/10.1371/journal.pone.0204494

Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131-143. https://doi.org/10.1177/1745691613518076

Hogg, M. A., Sherman, D. K., Dierselhuis, J., Maitner, A. T., & Moffitt, G. (2007). Uncertainty, entitativity, and group identification. *Journal of Experimental Social Psychology*, *43*(1), 135–142. https://doi.org/10.1016/j.jesp.2005.12.008

Janczyk, M. (2013). Level 2 perspective taking entails two processes: Evidence from PRP experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1878–1887. https://doi.org/10.1037/a0033336

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*, 601–625. https://doi.org/10.1146/annurev-psych-122414-033702

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25-41. https://doi.org/10.1016/S0010-0277(03)00064-7

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology*, *18*(1), 1–35. https://doi.org/10.1080/10463280701284645

Kuznetsova, A., Brockhoff, P. B., Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13.

Lakens, D. (2021, January 4). Sample Size Justification. https://doi.org/10.31234/osf.io/9d3yf

Lindqvist, A., Björklund, F., & Bäckström, M. (2017). The perception of the poor: Capturing stereotype content with different measures. *Nordic Psychology*, *69*(4), 231–247. https://doi.org/10.1080/19012276.2016.1270774

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics*, *68*(2), 327–337. https://doi.org/10.3758/BF03193680

Moll, H., & Tomasello, M. (2004). 12-and 18-month-old infants follow gaze to spaces behind barriers. Developmental science, 7(1), F1-F9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. (2015). Inclined to see it your way: Do altercentric intrusion effects in visual perspective taking reflect an intrinsically social process? *Quarterly Journal of Experimental Psychology*, *68*(10), 1931–1951. https://doi.org/10.1080/17470218.2015.1023206

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236. https://doi.org/10.1016/j.cognition.2010.08.003

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

Rohmer, O., & Louvet, E. (2012). Implicit measures of the stereotype content associated with disability. *British Journal of Social Psychology*, *51*(4), 732–740. https://doi.org/10.1111/j.2044-8309.2011.02087.x

Rohmer, O., & Louvet, E. (2016). On dit qu'on les apprécie mais sommes-nous prêts à les côtoyer ? Le cas particulier des personnes en situation de handicap. *Psihologia Sociala*, *38*, 17–34.

Rudert, S. C., Reutner, L., Greifeneder, R., & Walker, M. (2017). Faced with exclusion: Perceived facial warmth and competence influence moral judgments of social exclusion. *Journal of Experimental Social Psychology*, *68*, 101–112. https://doi.org/10.1016/j.jesp.2016.06.005

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 929. https://doi.org/10.1037/a0035175

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, *10*(2), 199–204. https://doi.org/10.1111/j.1467-7687.2007.00580.x

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sulpizio, V., Committeri, G., Metta, E., Lambrey, S., Berthoz, A., & Galati, G. (2015). Visuospatial transformations and personality: Evidence of a relationship between visuospatial perspective taking and self-reported emotional empathy. *Experimental Brain Research*, *233*(7), 2091–2102. https://doi.org/10.1007/s00221-015-4280-2

Surtees, A., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of Level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*(1), 75–86. https://doi.org/10.1111/j.2044-835X.2011.02063.x

Surtees, A., Apperly, I., & Samson, D. (2013). The use of embodied self-rotation for visual and spatial perspective-taking. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00698

Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97–105. https://doi.org/10.1016/j.cognition.2015.12.010

Todd, A. R., & Simpson, A. J. (2016). Anxiety impairs spontaneous perspective calculation: Evidence from a level-1 visual perspective-taking task. *Cognition*, *156*, 88–94. https://doi.org/10.1016/j.cognition.2016.08.004

Todd, A. R., Simpson, A. J., & Cameron, C. D. (2019). Time pressure disrupts level-2, but not level-1, visual perspective calculation: A process-dissociation analysis. *Cognition*, *189*, 41–54. https://doi.org/10.1016/j.cognition.2019.03.002

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. Annual Review of Psychology, 66, 519–45. https://doi.org/10.1146/annurev-psych-113011-143831

Vaes, J., & Paladino, M. P. (2010). The uniquely human content of stereotypes. *Group Processes & Intergroup Relations*, *13*(1), 23–39. https://doi.org/10.1177/1368430209347331

Westra, E. (2017). Spontaneous mindreading: A problem for the two-systems account. *Synthese*, *194*(11), 4559-4581. https://doi.org/10.1007/s11229-016-1159-0

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *Journal of Neuroscience*, *29*(26), 8525–8529. https://doi.org/10.1523/JNEUROSCI.2418-09.2009

Yzerbyt, V. (2016). Intergroup stereotyping. *Current Opinion in Psychology*, *11*, 90–95. https://doi.org/10.1016/j.copsyc.2016.06.009

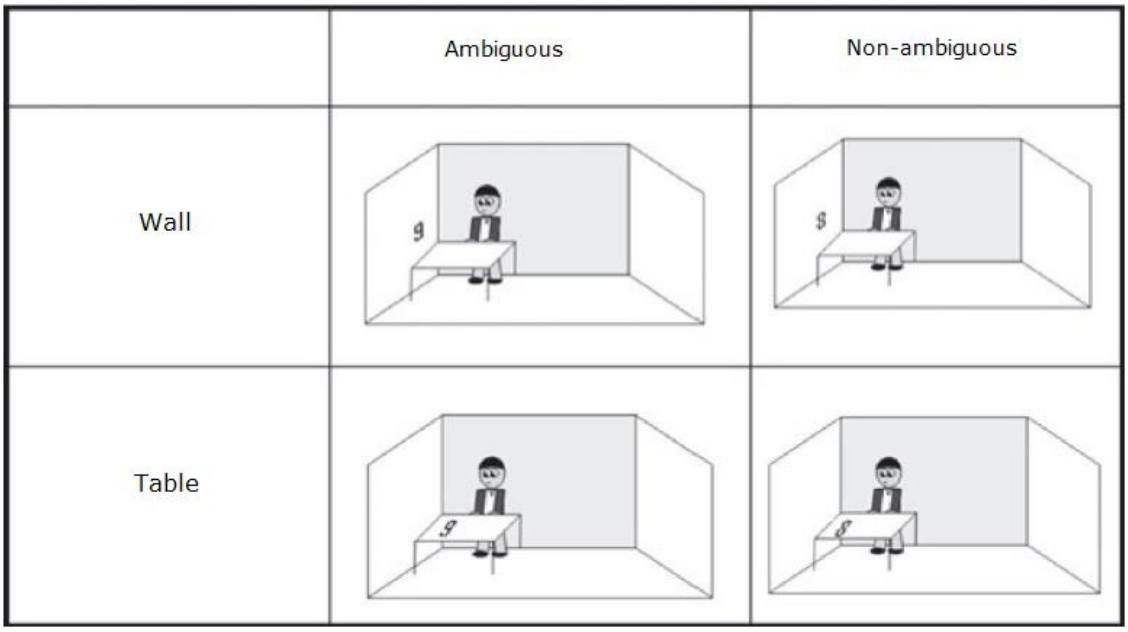

Figure 1. Visual description of the L2-VPT task. The lower-left scene illustrates an incongruent trial (where the number is perceived differently from the participant's or the avatar's perspective), while the lower-right, top-right and top-left scenes illustrate congruent trials. Adapted from "Direct and indirect measures of Level-2 perspective-taking in children and adults", by Surtees, A., Butterfill, S. A., & Apperly, I. A., 2012, *British Journal of Developmental Psychology*, *30*(1), 75–86. Copyright 2012 by John Wiley and Sons.

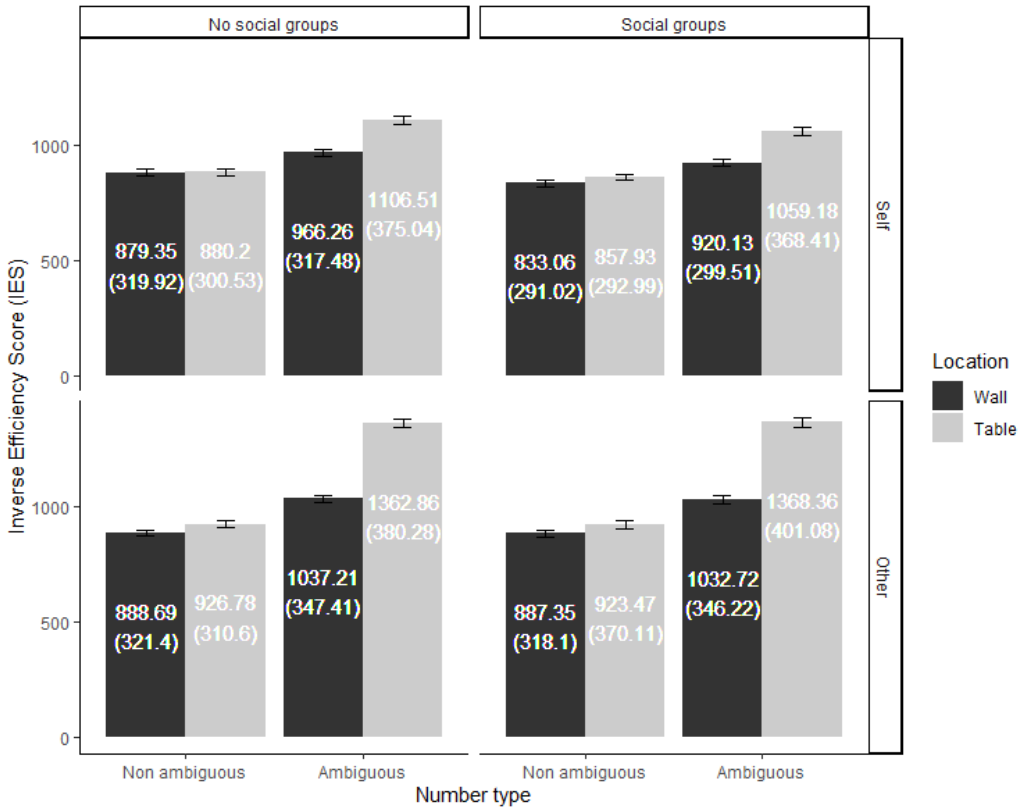

Figure 2. Graph of the inverse efficiency score (IES) as a function of the avatar's social characteristics, number type, location, and perspective. Values are means (standard deviations). Bars represent standard errors.

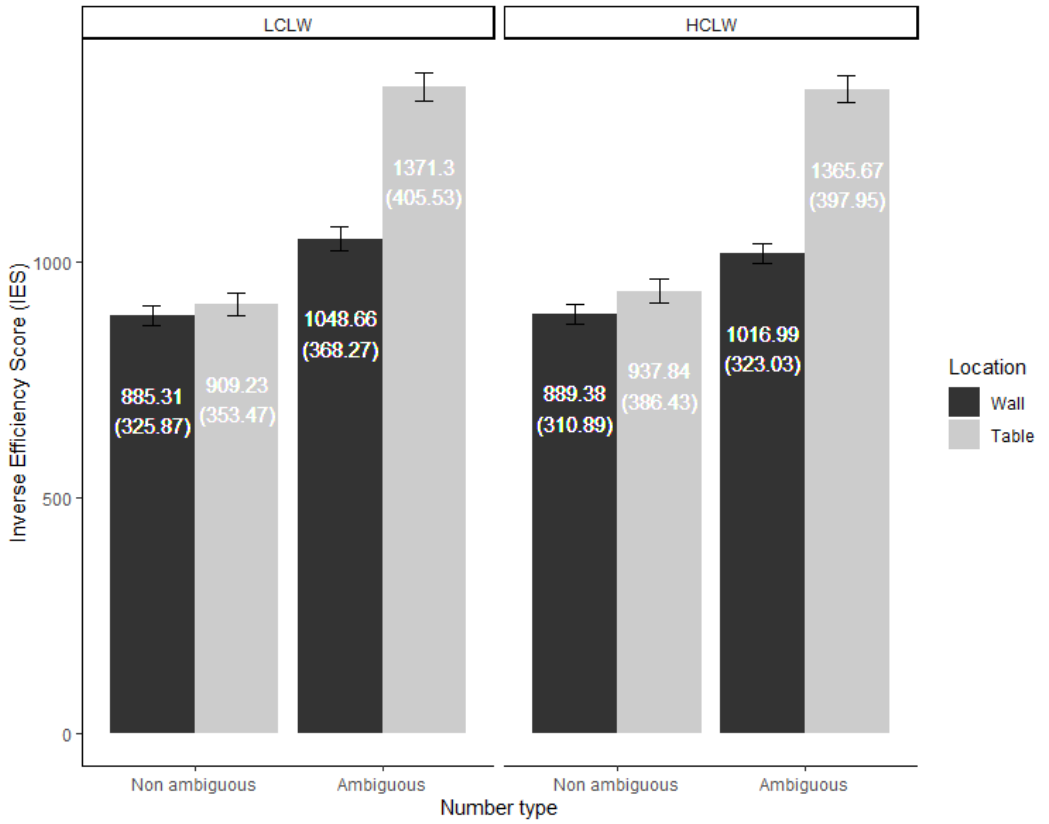

Figure 3. Graph of the inverse efficiency score (IES) as a function of number type, location, and perspective. Values are means (standard deviations). Bars represent standard errors. LCLW=low competence low warmth; HCLW=high competence low warmth.

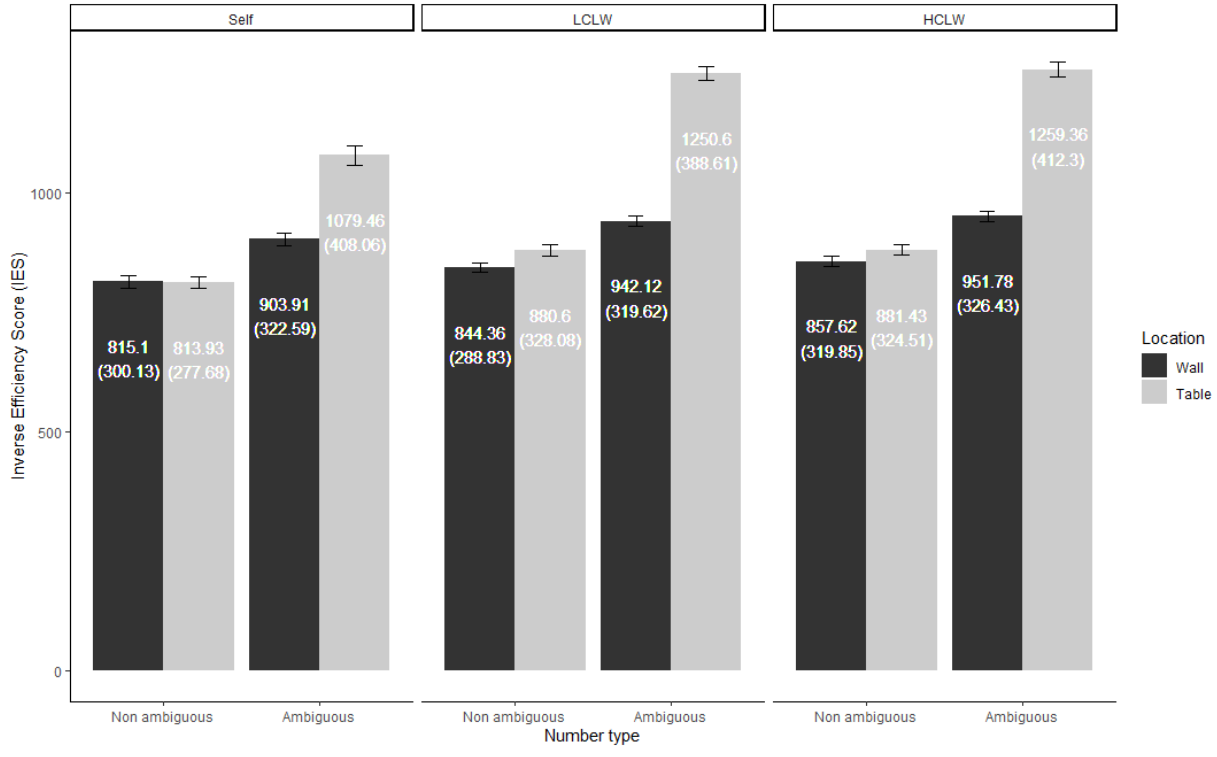

Figure 4. Graph of the inverse efficiency score (IES) as a function of number type, location, and perspective. Values are means (standard deviations). Bars represent standard errors. LCLW=low competence low warmth; HCLW=high competence low warmth.

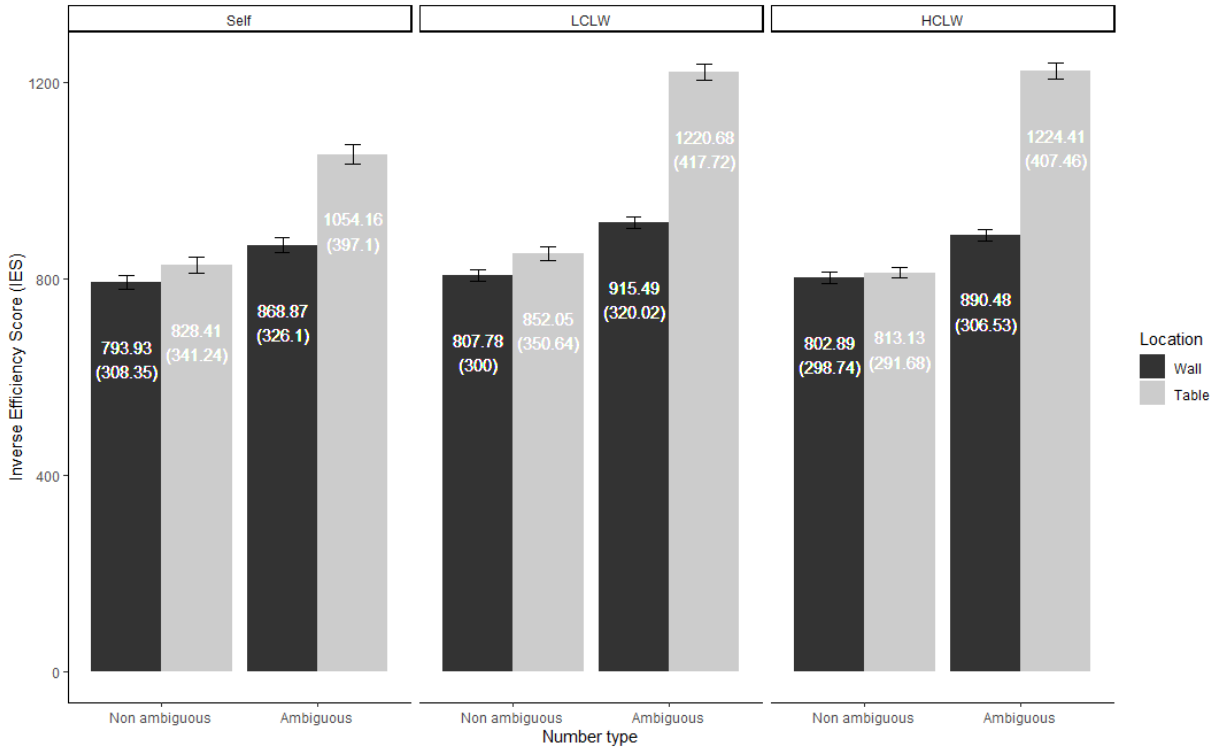

Figure 5. Graph of the inverse efficiency score (IES) as a function of number type, location, and perspective. Values are means (standard deviations). Bars represent standard errors. LCLW=low competence low warmth; HCLW=high competence low warmth.

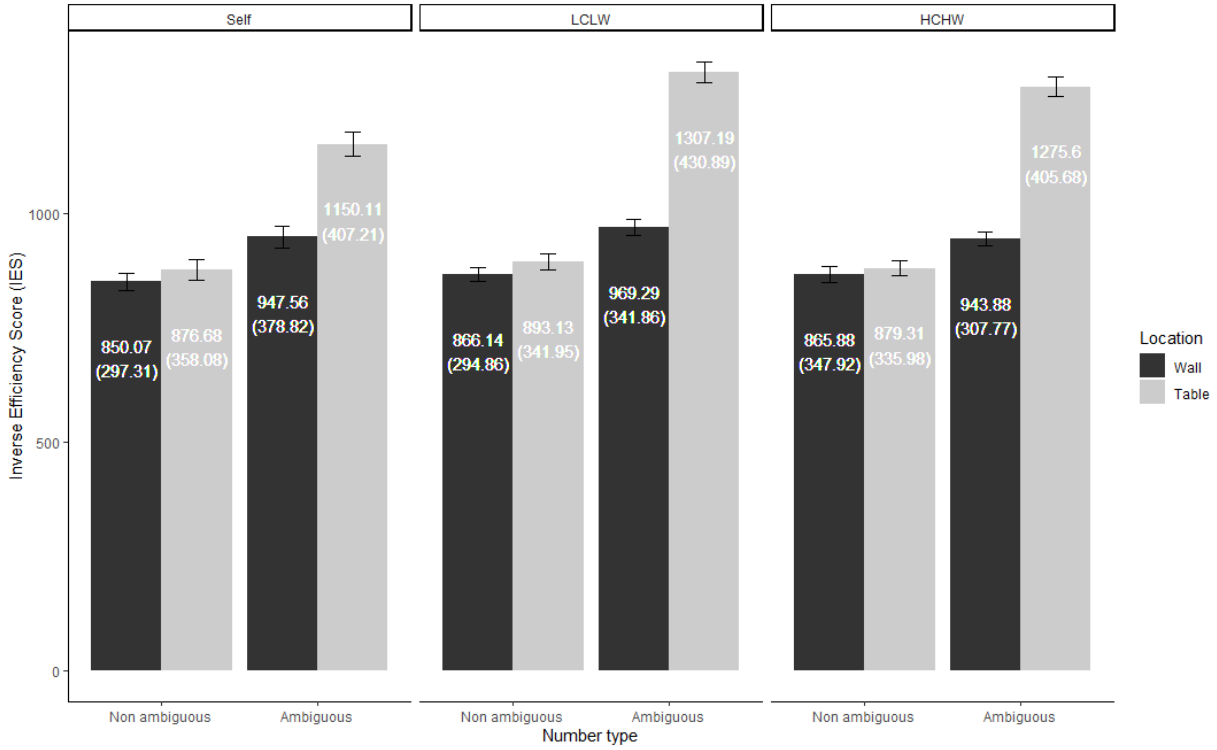

Figure 6. Graph of the inverse efficiency score (IES) as a function of number type, location, and perspective. Values are means (standard deviations). Bars represent standard errors. LCLW=low competence low warmth; HCHW=high competence high warmth.

Experiment 1 – Additional analyses

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) interaction, with participants (I) as random factor and eight replicates. Var(error) was set at 0.2 and var(I*S*L*P) was set at 0.04.

Altercentric interference – response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded the failed trials from the analysis. To investigate the presence of an altercentric interference, we submitted the response time to a 2 (avatar's social characteristics: no social group vs. social groups) x 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) mixed model analysis with the avatar's social characteristics varying between participants, the three remaining factors varying within participants, and participants as random factor (for a visual representation of the data, see Figure 1).

There were significant main effects of type of number, location, and perspective, confirming that participants took longer to answer when the number was ambiguous ($M_{ambiguous}$=1051.86, $SD_{ambiguous}$=362.81, $M_{non\text{-}ambiguous}$=845.82, $SD_{non\text{-}ambiguous}$=296.67), F(1, 7297.51)=1222.05, p<.001, η_p^2=.10, when the number was on the table (M_{table}=1006.36, SD_{table}=374.93, M_{wall}=888.61, SD_{wall}=305.61), F(1, 7297.29)=441.42, p<.001, η_p^2=.04, and when they answered from the perspective of the avatar (M_{other}=994.80, SD_{other}=366.62, M_{self}=896.91, SD_{self}=316.85), F(1, 7297.80)=285.91, p<.001, η_p^2=.03. There was no main effect of the avatar's social characteristics, F(1, 121.32)=1.16, p=.283, η_p^2<.01. We observed a significant perspective × avatar's social characteristics interaction, F(1, 7297.80)=6.70, p=.001, η_p^2<.01, with participants taking more time to answer from the avatar's perspective in the *no social group* condition (M_{other}=1004.89, SD_{other}=360.72, M_{self}=922.95, SD_{self}=326.06), F(1, 7296.76)=111.07, p<.001, η_p^2=.01. This effect emerged even more strongly in the *social groups* condition (M_{other}=983.15, SD_{other}=373.08, M_{self}=866.61, SD_{self}=303.08), F(1,

7298.69)=176.50, $p<.001$, $\eta_p^2=.02$. No further two-way interactions emerged, all Fs<0.42, all ps>.519, all $\eta_p^2<.01$.

The type of number × location × perspective was significant, $F(1, 7297.70)=38.43$, $p<.001$, $\eta_p^2<.01$. Follow-up analyses indicated that the type of number × location interaction was significant for the 'self' trials, $F(1, 7297.92)=59.33$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference based on location ($M_{table}=832.39$, $SD_{table}=280.02$, $M_{wall}=821.92$, $SD_{wall}=293.34$), $F(1, 7296.62)=0.69$, $p=.407$, $\eta_p^2<.01$. In contrast, for ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=1044.04$, $SD_{table}=353.48$, $M_{wall}=903.39$, $SD_{wall}=292.17$), $F(1, 7298.92)=133.18$, $p<.001$, $\eta_p^2=.01$. This pattern indicates the presence of altercentric interference, that is, a greater difficulty to answer from our own point of view when what one sees does not correspond to what the avatar sees (i.e., on incongruent trials).

Turning to the 'other' trials, the type of number × location interaction was also significant, $F(1, 7297.31)=271.54$, $p<.001$, $\eta_p^2=.02$. Specifically, for non-ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=883.47$, $SD_{table}=316.41$, $M_{wall}=845.14$, $SD_{wall}=292.30$), $F(1, 7296.61)=9.93$, $p=.002$, $\eta_p^2<.01$. The same effect was even stronger for ambiguous numbers ($M_{table}=1303.96$, $SD_{table}=362.94$, $M_{wall}=984.31$, $SD_{wall}=318.28$), $F(1, 7297.99)=671.84$, $p<.001$, $\eta_p^2=.06$. This pattern indicates the presence of egocentric interference, that is, a greater difficulty to answer from the avatar's point of view when what one sees does not correspond to what the avatar sees (i.e., on incongruent trials).

The avatar's social characteristics × type of number × location × perspective interaction was not significant, $F(1, 7297.70)=0.01$, $p=.904$, $\eta_p^2<.01$, indicating that the observed altercentric and egocentric interferences were not modulated by the avatar's social characteristics.

Altercentric interference - errors. We submitted the errors (with 0 indicating a failed trial and 1 a successful trial) to a 2 (avatar's social characteristics: no social group vs. social groups) x 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) mixed model analysis with the avatar's social characteristics varying between participants, the three remaining factors varying within participants, and participants as random factor (for a visual representation of the data, see Figure 2). We found a significant main effect of type of number, location and perspective, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.53, z=-4.36, p<.001, when the number was on the table, *OR*=0.26, z=-6.22, p<.001, or when answering from their own perspective, *OR*=1.46, z=2.59, p=.009. There was no main effect of the avatar's social characteristics, z=-0.23, p=.818.

The type of number × location × avatar's social characteristics interaction was significant, z=-3.43, p<.001. Follow-up analyses indicated that the type of number × location interaction was not significant in the *no social group* condition, z=-1.17, p=.241. However, this interaction was significant in the *social groups* condition, z=-5.63, p<.001. Specifically, for non-ambiguous numbers, there was no significant difference based on location, z=-0.06, p=.949. However, participants were less likely to answer correctly for ambiguous numbers they appeared on the table, *OR*=0.08, z=-8.05, p<.001. No further three- of four-way interaction emerged, all zs<|1.274|, ps>.202.

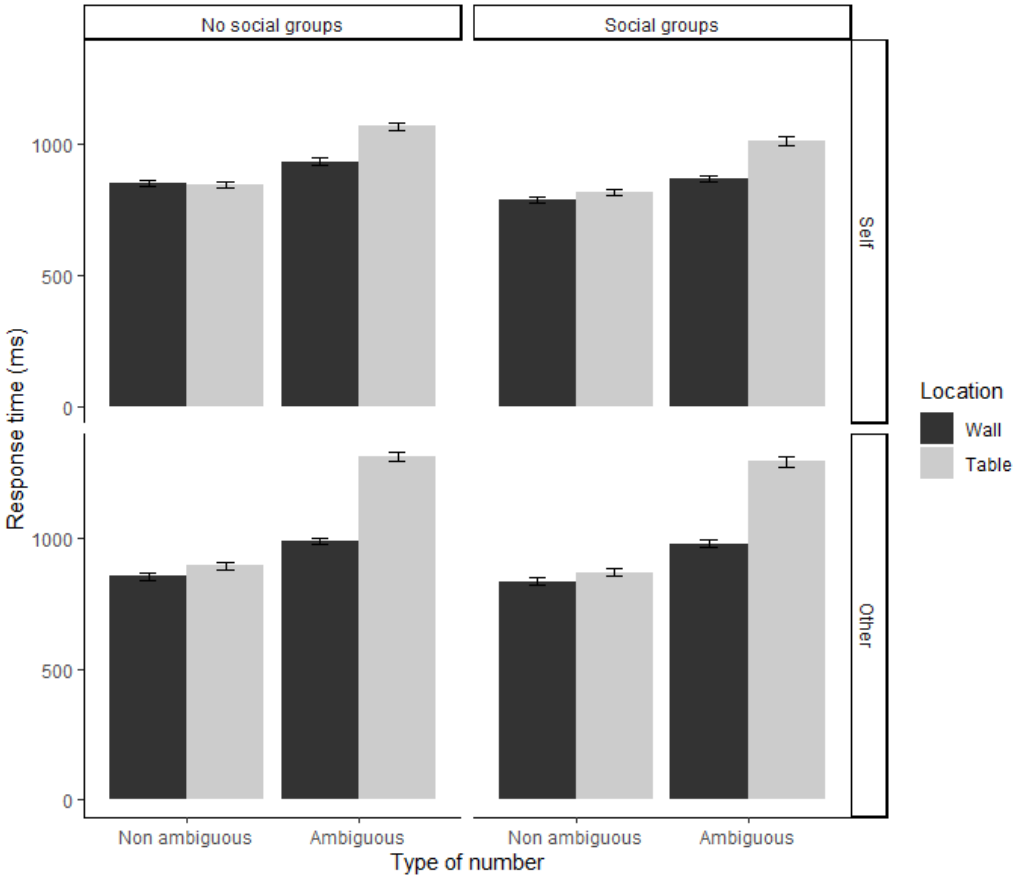

Figure 1. Graph of response time as a function of the avatar's social characteristics, type of number, location, and perspective. Bars represent the standard error of the mean.

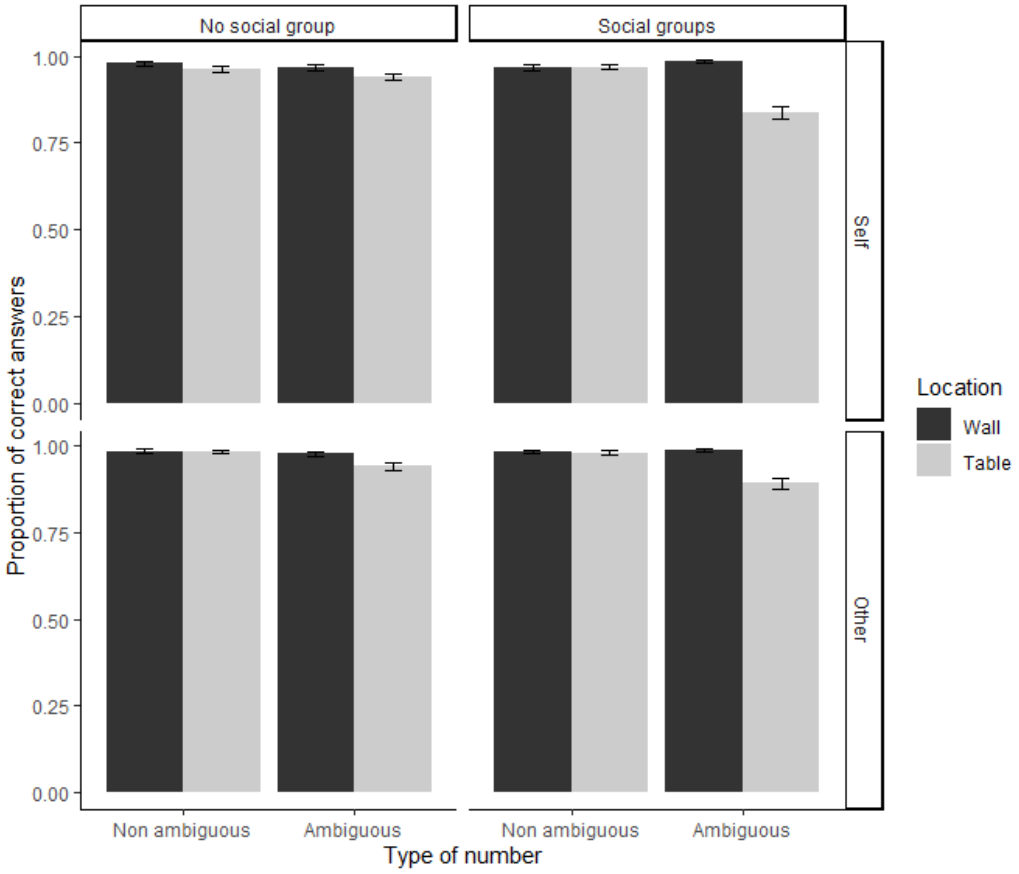

Figure 2. Graph of proportion of correct answers as a function of the avatar's social characteristics, type of number, location, and perspective. Bars represent the standard error of the mean.

Impact of the avatar's social group – response time. Our main objective was to investigate whether participants demonstrate greater difficulty to adopt the visual perspective of a LCLW target compared to a HCLW target. To do so, we only considered data in the *social groups* condition for the trials in which participants had to answer from the avatar's perspective. We submitted the response time in the *social groups* condition to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: politician vs. drug addict) mixed model analysis with the three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 3).

As before, significant main effects of type of number and location confirmed that participants took longer to answer when the number was ambiguous ($M_{ambiguous}$=1123.35, $SD_{ambiguous}$=381.80, $M_{non\text{-}ambiguous}$=852.67, $SD_{non\text{-}ambiguous}$=312.83), F(1, 1661.86)=452.76, p<.001, η_p^2=.15, and when the number was on the table (M_{table}=1064.86, SD_{table}=410.31, M_{wall}=907.46, SD_{wall}=316.86), F(1, 1661.90)=177.90, p<.001, η_p^2=.06. There was no effect of perspective, F(1, 1660.95)=0.30, p=.585, , η_p^2<.01.

Not surprisingly, the type of number × location was significant, F(1, 1661.72)=112.49, p<.001, η_p^2=.04. Specifically, for non-ambiguous numbers, participants took marginally more time when the number was on the table (M_{table}=870.64, SD_{table}=338.32, M_{wall}=834.63, SD_{wall}=284.17), F(1, 1660.72)=3.91, p=.048, η_p^2<.01. This effect emerged even stronger for ambiguous numbers (M_{table}=1292.64, SD_{table}=368.30, M_{wall}=979.65, SD_{wall}=331.07), F(1, 11662.80)=274.78, p<.001, η_p^2=.10. Contrary to our hypothesis, the type of number × location × perspective interaction was not significant, F(1, 1660.77)<0.01, p=.955, η_p^2<.01. No further two-way interactions emerged, all Fs<1.45, all ps>.229, all η_p^2<.01.

Impact of the avatar's social group – errors. We submitted the errors (with 0 indicating a failed trial and 1 a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: politician vs. drug addict) mixed model analysis with the three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 4). We found a significant main effect of type of number and location, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.46, z=-2.31, p=.021, and when the number was on the table, *OR*=0.31, z=-3.54, p<.001. There was no main effect of perspective, z=-0.22, p=.825.

The type of number × location was significant, z=-3.19, p=.001. Specifically, for non-ambiguous numbers, there was no effect of location, z=-0.23, p=.816. However, for

ambiguous numbers, participants were less likely to answer correctly when the number was on the table, *OR*=0.11, *z*=-5.04, *p*<.001. No other interaction proved significant, all *z*s<|0.56|, *p*s>.578.

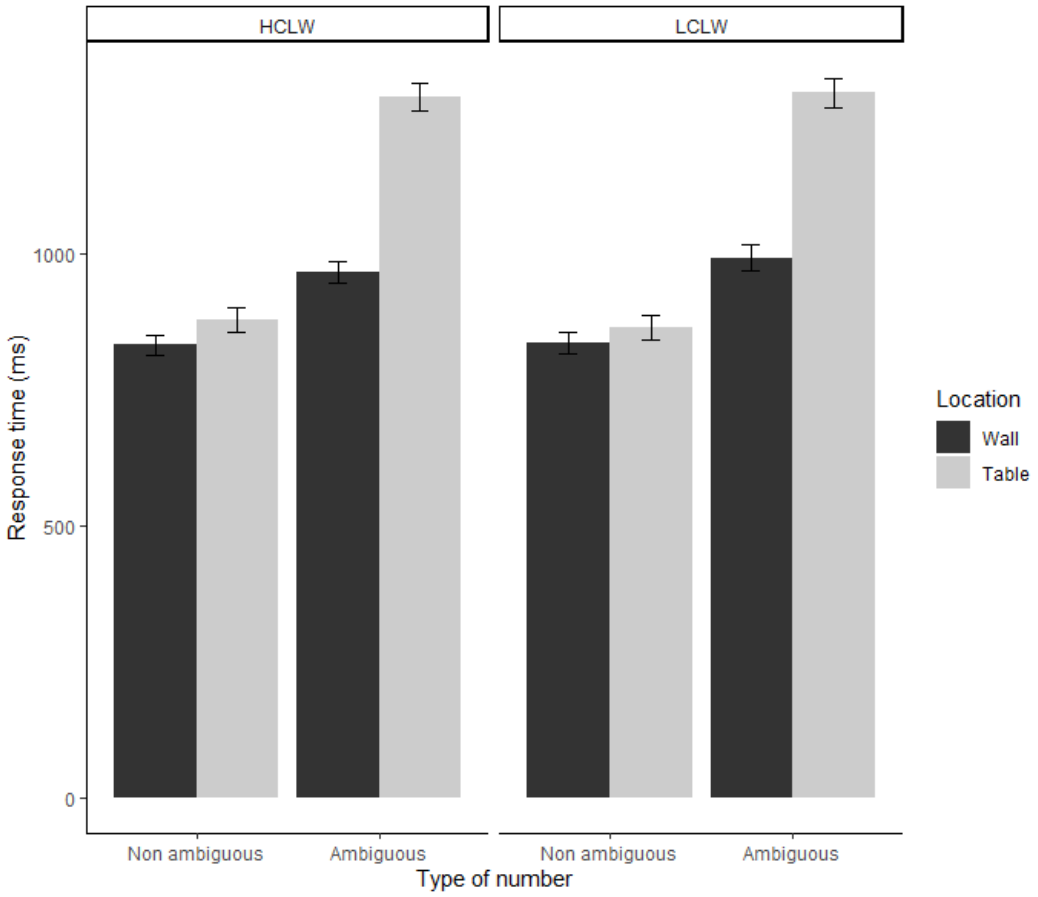

Figure 3. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

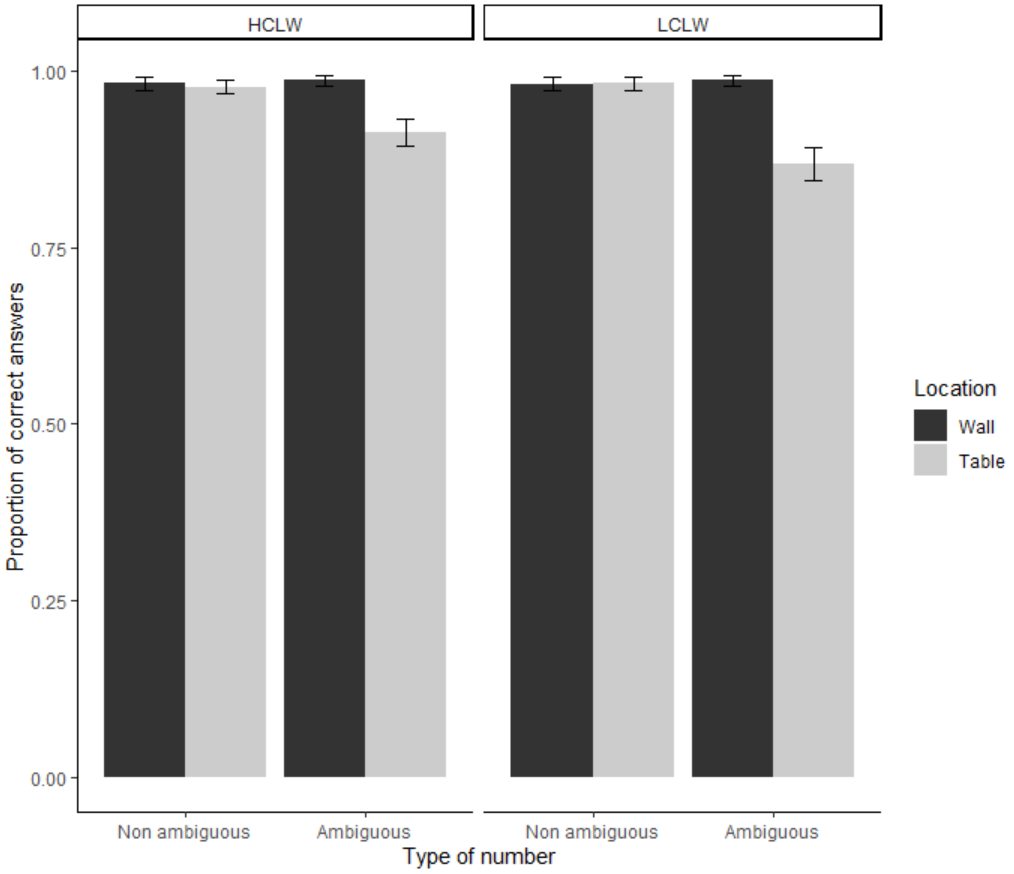

Figure 4. Graph of proportion of correct as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Experiment 2 – Additional analyses

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) interaction, with participants (I) as random factor and 12 replicates. Planned contrasts were used to compare the two avatars.Var(error) was set at 0.2 and var(I*S*L*P) was set at 0.04.

Response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the response time in milliseconds to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 5). Significant effects of type number and of location indicated that participants took more time answering when the number was ambiguous ($M_{ambiguous}$=997.18, $SD_{ambiguous}$=347.98, $M_{non\text{-}ambiguous}$=799.03, $SD_{non\text{-}ambiguous}$=266.07), F(1, 8243.66)=1295.94, p<.001, η_p^2=.10, or on the table (M_{table}=958.49, SD_{table}=356.39, M_{wall}=833.14, SD_{wall}=276.38), F(1, 8243.66)=555.16, p<.001, η_p^2=.05.

A set of contrasts allowed us to compare the response time for ‘self’ versus ‘other’ (both HCLW and LCLW), as well for the HCLW vs. LCLW avatars. The ‘self’ versus ‘other’ contrast proved significant, F(1, 8244.70)=190.68, p<.001, η_p^2=.02, indicating that participants took more time answering for the avatar than for themselves, regardless of the avatar’s social group (M_{other}=912.28, SD_{other}=329.98, M_{self}=833.59, SD_{self}=293.29). There was no significant difference between the LCLW and the HCLW targets, F(1, 8244.49)=1.66, p=.197, η_p^2<.01. No interaction including this factor emerged, all Fs<0.46, all ps>.498, all η_p^2<.01.

The type of number × location × ‘self’ versus ‘other’ contrast was significant, F(1, 8243.88)=25.51, p<.001, η_p^2<.01. Follow-up analyses revealed that the number × location

interaction was significant for the 'self' trials, $F(1, 8243.33)=42.48$, $p<.001$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference in response times as a function of location ($M_{table}=762.18$, $SD_{table}=234.15$, $M_{wall}=763.23$, $SD_{wall}=256.92$), $F(1, 8241.79)<0.01$, $p=.958$, $\eta_p^2<.01$. However, for ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=1008.44$, $SD_{table}=347.09$, $M_{wall}=845.83$, $SD_{wall}=281.40$), $F(1, 8244.53)=77.47$, $p<.001$, $\eta_p^2<.01$.

Turning to the 'other' trials, the type of number × location interaction was also significant, $F(1, 8244.48)=463.20$, $p<.001$, $\eta_p^2=.04$. Specifically, for non-ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=824.60$ms, $SD_{table}=277.41$ms, $M_{wall}=797.93$ms, $SD_{wall}=264.92$ms), $F(1, 8241.92)=8.96$, $p=.003$, $\eta_p^2<.01$. This difference emerged even more strongly for ambiguous numbers ($M_{table}=1187.81$ms, $SD_{table}=359.17$ms, $M_{wall}=888.06$ms, $SD_{wall}=281.70$ms), $F(1, 8246.62)=1050.35$, $p<.001$, $\eta_p^2=.09$.

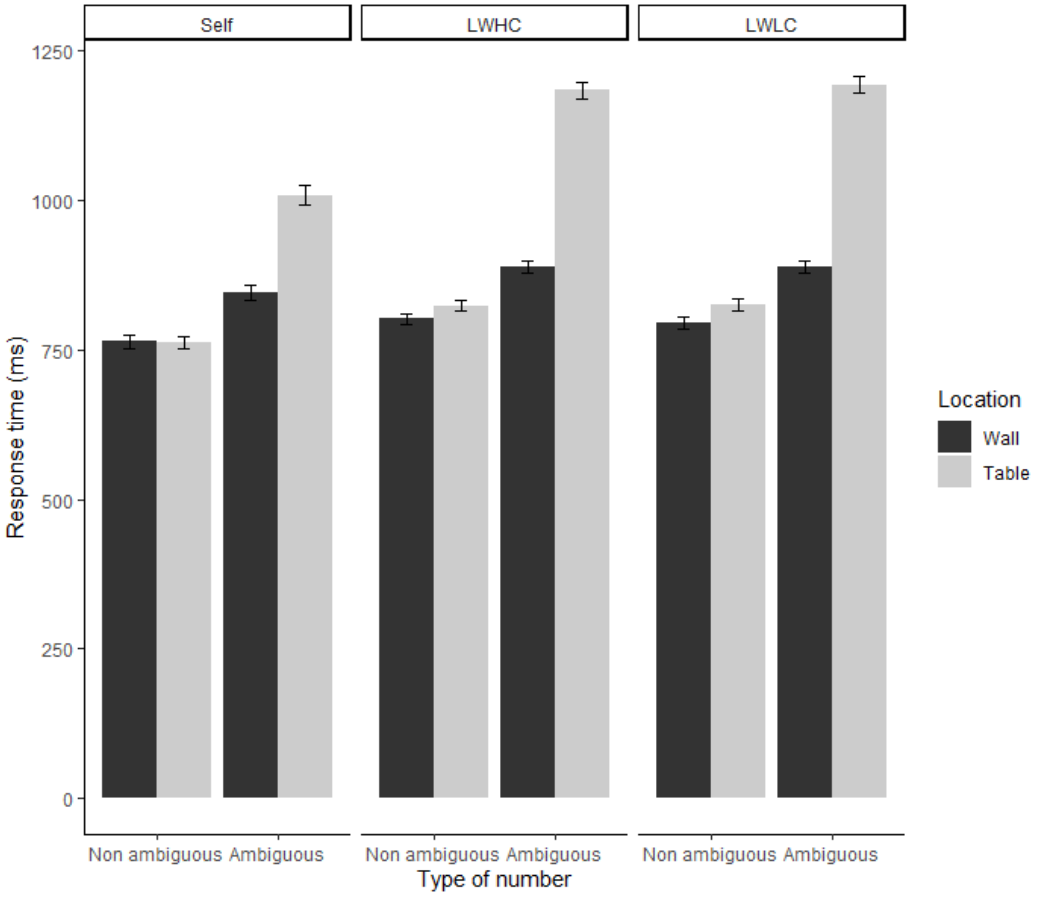

Figure 5. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Errors. We submitted the errors (with 0 indicating a failed trial and 1 indicating a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) logistic mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 6). We found a significant main effect of type of number and location, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.25, z=-9.04, p<.001, or when the number was on the table, *OR*=0.26, z=-8.77, p<.001. Turning to the effect of perspective, the 'self' versus 'other' contrast was not significant, z=-0.34, p=.730. The 'HCLW' versus 'LCLW' contrast proved significant, indicating that participants were

less likely to answer correctly for LCLW groups than for the HCLW groups, *OR*=0.55, *z*=-3.65, *p*<.001. No interaction including this factor emerged, all *z*s<|1.223|, *p*s>.22

The type of number × location interaction was significant, *z*=-8.82, *p*<.001. Specifically, for non-ambiguous numbers, participants were no less likely to answer correctly when the number appeared on the table, *z*=0.03, *p*=.978. However, for ambiguous numbers, they were less likely to answer correctly when they appeared on the table, *OR*=0.07, *z*=-14.91, *p*<.001. The type of number × 'self' versus 'other' contrast interaction was also significant, *z*=2.49, *p*=.028. Specifically, when answering from their perspective, participants were less likely to answer correctly for ambiguous than for non-ambiguous numbers, *OR*=0.14 *z*=-6.19, *p*<.001. Interestingly, this effect was less marked when participants had to answer from the avatar's perspective, *OR*=0.33, *z*=-6.64, *p*<.001. Finally, the location × 'self *vs*. other' contrast interaction also came out significant, *z*=2.19, *p*=.028. That is, when answering from their perspective, participants were less likely to answer correctly when the number was on the table, *OR*=.15 *z*=-5.84, *p*<.001. Again, this effect was less marked when participants had to answer from the avatar's perspective, *OR*=0.34, *z*=-6.61, *p*<.001. No other interaction proved significant, all *z*s<|1.22|, *p*s>.221.

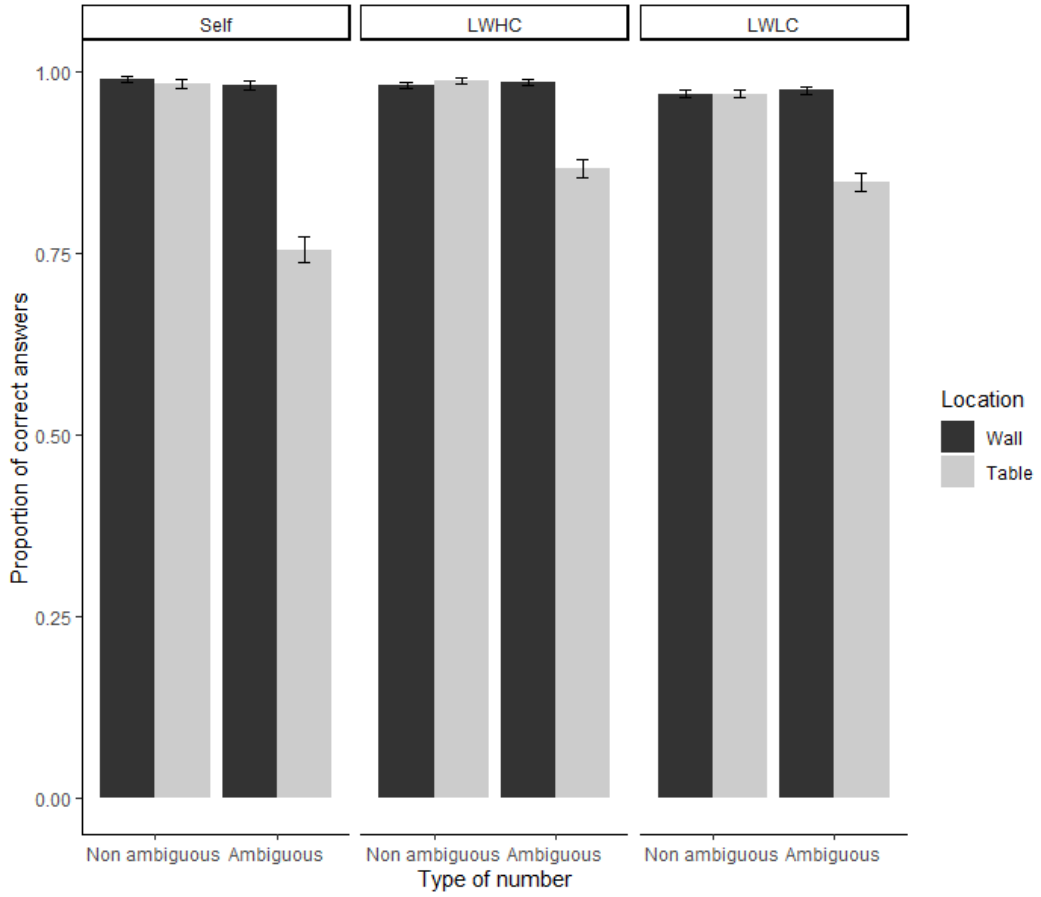

Figure 6. Graph of proportion of correct answers as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Experiment 3 – Additional analyses

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) x quadrant group (Q) interaction, with participants (I) as random factor and 12 replicates. Planned contrasts were used to compare the two avatars.Var(error) was set at 0.111 and var(I*S*L*P*Q) was set at 0.019.

Response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the response time to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCLW vs. LCLW) mixed model analysis with all four factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 7). Significant effects of type number, location and perspective indicated that participants took more time answering when the number was ambiguous ($M_{ambiguous}$=1004.57, $SD_{ambiguous}$=380.25, $M_{non\text{-}ambiguous}$=795.58, $SD_{non\text{-}ambiguous}$=306.72), F(1, 7295.12)=628.81, p<.001, η_p^2=.06, on the table (M_{table}=973.75, SD_{table}=398.13, M_{wall}=825.52, SD_{wall}=302.23), F(1, 7295.14)=329.89, p<.001, η_p^2=.03, and when answering from the avatar's perspective, regardless of the avatar's social group (M_{other}=911.76, SD_{other}=365.16, M_{self}=857.22, SD_{self}=341.82), F(1, 7295.21)=47.37, p<.001, η_p^2<.01. There was no main effect of quadrant group, F(1, 7295.11)=0.26, p=.610, η_p^2<.01, and no interaction including this factor emerged, all Fs<2.81, all ps>.094, all η_p^2<.01.

The type of number × location × perspective interaction was significant, F(1, 7295.12)=20.99, p<.001, η_p^2<.01. Follow-up analyses revealed that the number × location interaction was significant for 'self' trials, F(1, 7295.14)=31.33, p<.001, η_p^2<.01. Specifically, for non-ambiguous numbers, there was no significant difference in response time as a function of location (M_{table}=806.81, SD_{table}=331.80, M_{wall}=773.40, SD_{wall}=299.94), F(1, 7295.11)=2.57, p=.109, η_p^2<.01. However, for ambiguous numbers, participants took longer

to answer when the number was displayed on the table rather than on the wall (M_{table}=1023.16, SD_{table}=376.48, M_{wall}=842.02, SD_{wall}=306.75), F(1, 7295.20)=89.60, p<.001, η_p^2<.01.

Turning to the 'other' trials, the type of number × location interaction was also significant, F(1, 7295.14)=379.73, p<.001, η_p^2=.03. Specifically, for non-ambiguous numbers, participants took more time to answer when the number was displayed on the table rather than on the wall (M_{table}=810.77, SD_{table}=315.06, M_{wall}=783.94, SD_{wall}=290.83), F(1, 7295.03)=7.16, p=.007, η_p^2<.01. This pattern emerged even more strongly for ambiguous numbers (M_{table}=1189.14, SD_{table}=398.25, M_{wall}=878.59, SD_{wall}=303.85), F(1, 7295.23)=900.46, p<.001, η_p^2=.08.

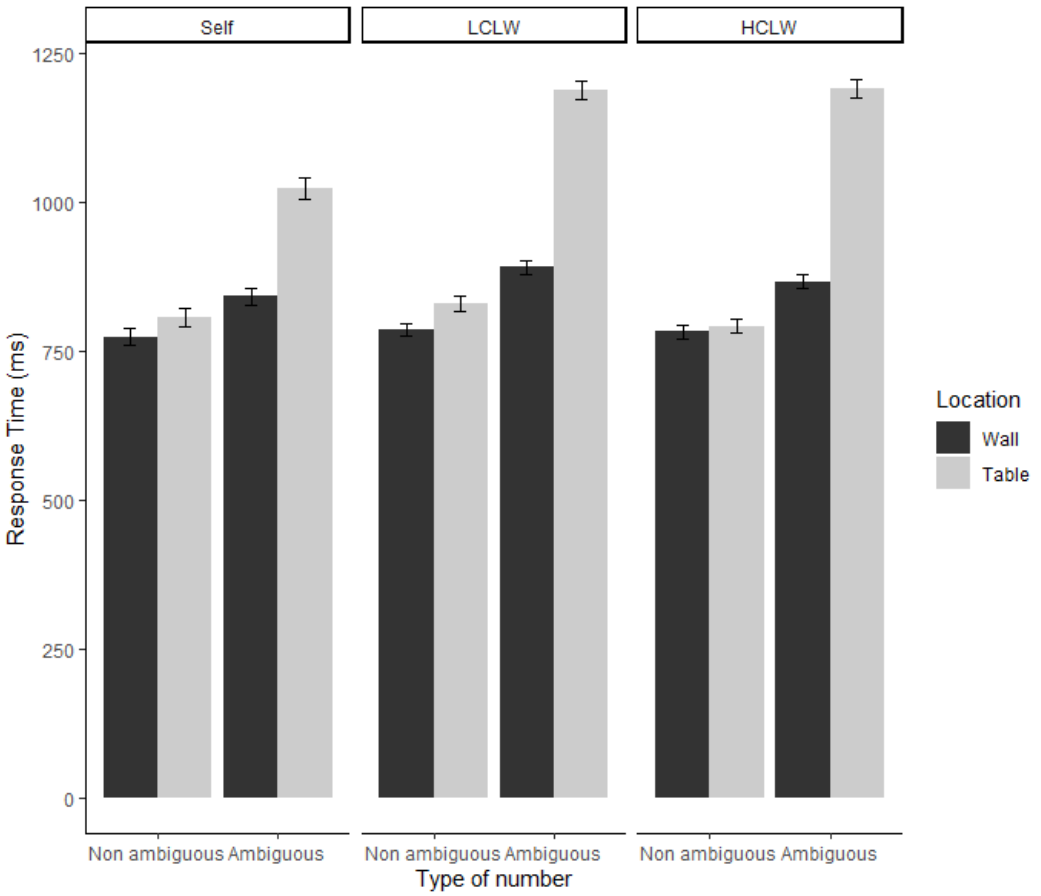

Figure 7. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Errors. We submitted the errors (with 0 indicating a failed trial and 1 indicating a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCLW vs. LCLW) mixed model analysis with all four factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 8). We found a significant main effect of type of number, location and perspective, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.54, z=-3.20, p=.001, when the number was on the table, *OR*=0.47, z=-4.01, p<.001, and when answering from their own perspective, *OR*=1.80, z=3.11, p=.002. There was no significant main effect of quadrant group, z=-1.74, p=.08.

The type of number × location × quadrant group interaction was significant, z=2.06, p=.04. Follow-up analyses revealed that the number × location interaction was significant for HCLW targets, z=-4.83, p<.001. Specifically, for non-ambiguous numbers, there was no significant difference in errors as a function of location z=1.43, p=.150. However, for ambiguous numbers, participants were less likely to answer correctly when the number was displayed on the table rather than on the wall *OR*=0.11, z=-5.69, p<.001.

Turning to the LCLW targets, the type of number × location interaction was also significant z=-2.50, p=.012. Specifically, for non-ambiguous numbers, there was no significant difference in errors as a function of location, z=-0.67, p=.713. However, for ambiguous numbers, participants were less likely to answer correctly when the number was displayed on the table rather than on the wall *OR*=0.25, z=-4.15, p<.001. No other interaction proved significant, all zs<|1.76|, ps>.08.

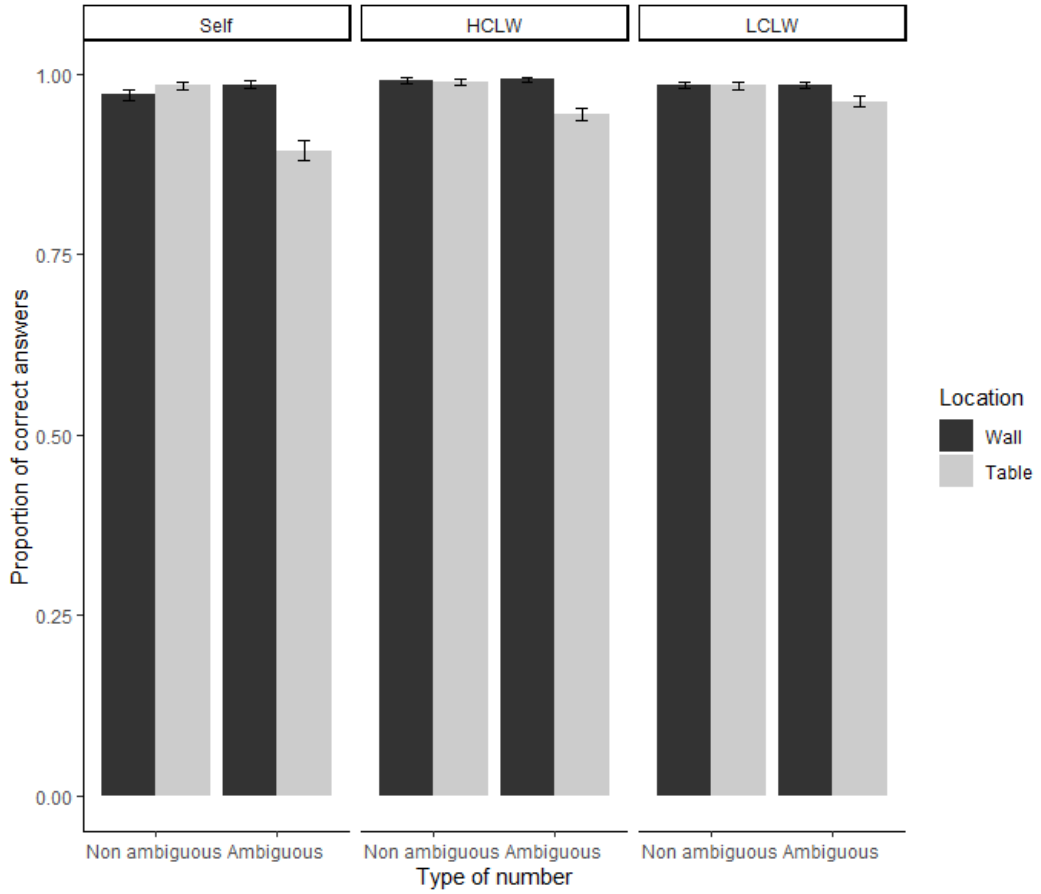

Figure 8. Graph of proportion of correct answers as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Experiment 4 – Additional analyses

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) x quadrant group (Q) interaction, with participants (I) as random factor and 12 replicates. Planned contrasts were used to compare the two avatars.Var(error) was set at 0.111 and var(I*S*L*P*Q) was set at 0.019.

Response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the response time to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCHW vs. LCLW) mixed model analysis with all four factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 9). Significant effects of type number, location and perspective indicated that participants took more time answering when the number was ambiguous ($M_{ambiguous}$=1067.21, $SD_{ambiguous}$=391.33, $M_{non\text{-}ambiguous}$=849.49, $SD_{non\text{-}ambiguous}$=318.41), F(1, 4089.10)=390.31, p<.001, η_p^2=.06, on the table (M_{table}=1032.28, SD_{table}=407.66, M_{wall}=883.80, SD_{wall}=318.44), F(1, 4089.08)=173.80, p<.001, η_p^2=.03, and when answering from the avatar's perspective, regardless of the avatar's social group (M_{other}=966.95, SD_{other}=374.27, M_{self}=925.73, SD_{self}=365.19), F(1, 4089.10)=14.96, p<.001, η_p^2<.01. There was no main effect of quadrant group, F(1, 4089.05)=0.02, p=.875, η_p^2<.01, and no interaction including this factor emerged, all Fs<3.30, all ps>.07, all η_p^2<.01.

The type of number × location × perspective interaction was significant, F(1, 4089.14)=7.22, p=.007, η_p^2<.01. Follow-up analyses revealed that the number × location interaction was significant for 'self' trials, F(1, 4089.13)=25.83, p<.001, η_p^2<.01. Specifically, for non-ambiguous numbers, there was no significant difference in response time as a function of location (M_{table}=854.33, SD_{table}=348.93, M_{wall}=828.66, SD_{wall}=291.07), F(1, 4089.05)=0.36, p=.546, η_p^2<.01. However, for ambiguous numbers, participants took longer

to answer when the number was displayed on the table rather than on the wall (M_{table}=1121.80, SD_{table}=395.48, M_{wall}=913.59, SD_{wall}=351.65), F(1, 4089.18)=60.49, p<.001, η_p^2=.01.

Turning to the 'other' trials, the type of number × location interaction was also significant, F(1, 4089.09)=211.47, p<.001, η_p^2<.04. Specifically, for non-ambiguous numbers, there was no significant difference in response time as a function of location (M_{table}=861.52, SD_{table}=325.16, M_{wall}=842.65, SD_{wall}=309.50), F(1, 4089.07)=1.67, p=.196, η_p^2<.01. However, for ambiguous numbers, participants took longer to answer when the number was displayed on the table rather than on the wall (M_{table}=1249.78, SD_{table}=397.67, M_{wall}=932.58, SD_{wall}=316.34), F(1, 4089.11)=470.31, p<.001, η_p^2=.08.

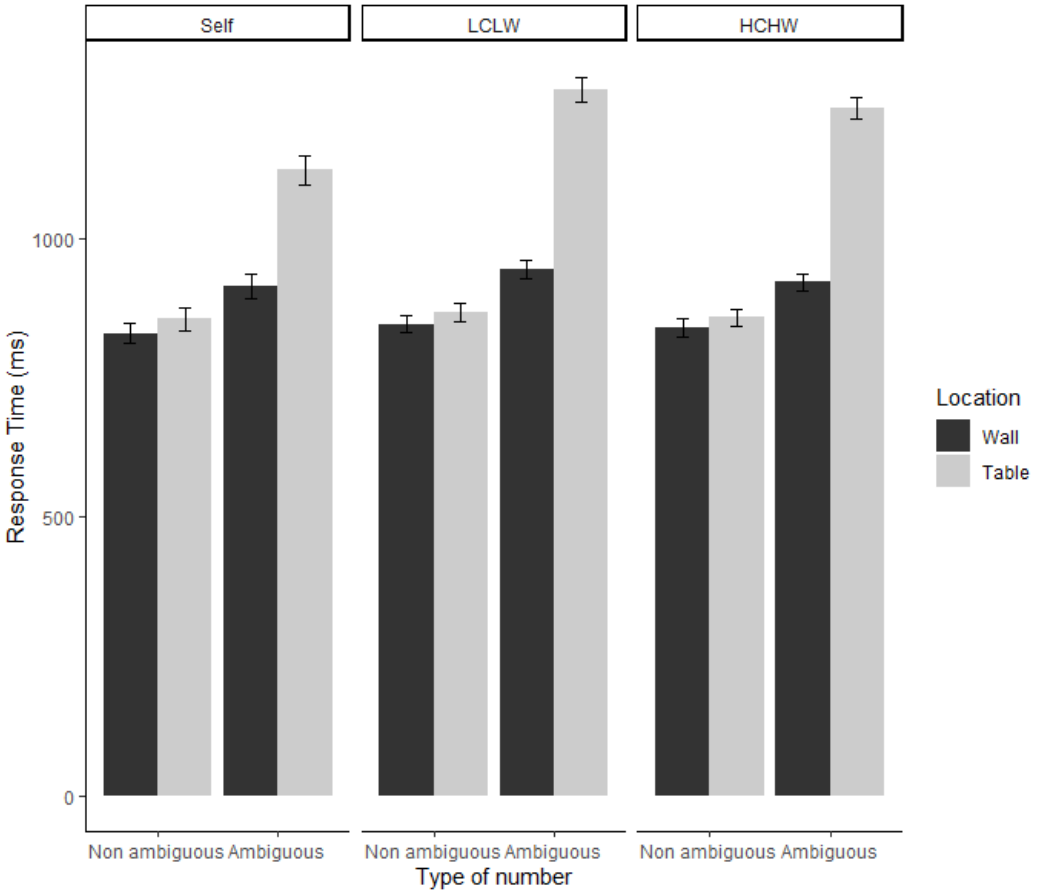

Figure 9. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCHW=high competence high warmth.

Errors. We submitted the errors (with 0 indicating a failed trial and 1 indicating a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 2 (perspective: self vs. other) x 2 (quadrant group: HCLW vs. LCLW) mixed model analysis with all four factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 10). We found no significant main effect, all *z*s<|1.91|, *p*s>.056. The type of number × location interaction was significant, *z*=-4.33, *p*<.001. Specifically, for non-ambiguous numbers, there was no significant difference in errors as a function of location, *z*=1.67, *p*=.094. However, for ambiguous numbers, participants were less likely to answer correctly when the number was displayed on the table rather than on the wall *OR*=0.18, *z*=-4.51, *p*<.001. No other interaction proved significant, all *z*s<|1.25|, *p*s>.211.

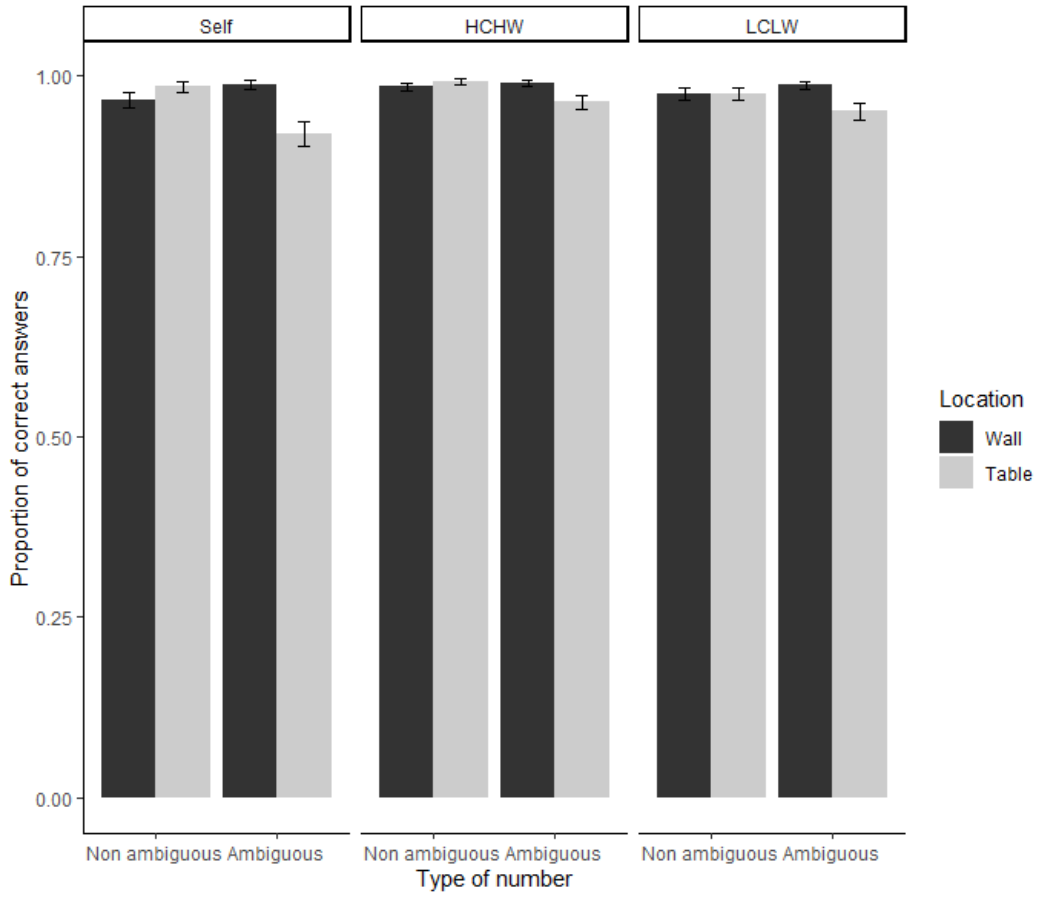

Figure 10. Graph of proportion of correct answers as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCHW=high competence high warmth.

Additional Experiment 1

This experiment consists of an exact replication of the *social groups* condition of our first experiment.

Method

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) interaction, with participants (I) as random factor and eight replicates. Planned contrasts were used to compare the two avatars.Var(error) was set at 0.2 and var(I*S*L*P) was set at 0.04.

Participants and design. A total of 51 Belgian students (M_{age}=20.67, SD_{age}=4.41, 47 women) took part in the 25 min. study in exchange for partial course credits. The experiment involved a 2 (type of number: non-ambiguous=0 and 8 vs. ambiguous=6 and 9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatar vs. LCLW quadrant avatars) within-participant repeated measures design.

Power calculations. We conducted our power analysis for a 2 x 2 x 3 repeated measures design using the PANGEA webapp[7], and aimed for 90% power. This analysis indicated that we needed at least 50 participants to achieve 90% power to detect a small effect (d=.2).

Procedure. The procedure was identical to that of Experiment 1, with the exception that we dropped the no social group condition.

Results

Analyses were performed using R (R Core Team, 2020), with the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages. Again, and following Surtees et al.

[7] https://jakewestfall.shinyapps.io/

(2012), trials for which participants had to press the N (incorrect) button were not considered, leaving out 64 of the 128 trials for each participant. We checked if any participant failed to perform above chance, which was not the case. Furthermore, and also as in Surtees and colleagues (2012), we excluded trials for which the response time was 2.5 SD away from the mean response time (0.88% of the remaining 64 trials per participant).

Inverse Efficiency Score. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the IES to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 1). Significant effects of type number and of location indicated that participants had more difficulties answering when the number was ambiguous ($M_{ambiguous}$=1199.67, $SD_{ambiguous}$=483.02, $M_{non\text{-}ambiguous}$=1016.65, $SD_{non\text{-}ambiguous}$=445.09), F(1, 2828.63)=280.17, p<.001, η_p^2=.05, or on the table (M_{table}=1159.06, SD_{table}=489.61, M_{wall}=1052.99, SD_{wall}=450.23), F(1, 2828.79)=135.64, p<.001, η_p^2=.02.

A set of contrasts allowed us to compare the IES for ‘self’ versus ‘other’ (both HCLW and LCLW), as well for the HCLW vs. LCLW avatars. The ‘self’ versus ‘other’ contrast proved significant, F(1, 2828.07)=163.95, p<.001, η_p^2=.03, indicating that participants had more difficulties answering for the avatar than for themselves, regardless of the avatar’s social group (M_{other}=1185.21, SD_{other}=500.81, M_{self}=1023.03, SD_{self}=427.55). There was no significant difference between the LCLW and the HCLW targets, F(1, 2828.76)=0.82, p=.366, η_p^2<.01. No interaction including this factor emerged, all Fs<2.83, all ps>.092, all η_p^2<.01.

The type of number × location × ‘self’ versus ‘other’ contrast interaction was significant, F(1, 2828.74)=17.84, p<.001, η_p^2<.01. Follow-up analyses revealed that the

number × location interaction was significant for ‘self’ trials, $F(1, 2828.72)=7.95$, $p=.005$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference in difficulty as a function of location ($M_{table}=970.90$, $SD_{table}=412.89$, $M_{wall}=943.38$, $SD_{wall}=365.07$), $F(1, 2828.28)=0.35$, $p=.556$, $\eta_p^2<.01$. However, for ambiguous numbers, participants had more difficulties answering when the number was displayed on the table than on the wall ($M_{table}=1151.32$, $SD_{table}=484.84$, $M_{wall}=1048.44$, $SD_{wall}=423.51$), $F(1, 2828.96)=19.77$, $p<.001$, $\eta_p^2=.01$.

Turning to ‘other’ trials, the type of number × location interaction was also significant, $F(1, 2828.75)=76.78$, $p<.001$, $\eta_p^2=.01$. Specifically, for non-ambiguous numbers, there was no significant difference in difficulty as a function of location ($M_{table}=1085.95$, $SD_{table}=459.11$, $M_{wall}=1066.72$, $SD_{wall}=513.85$), $F(1, 2828.47)=3.15$, $p=.076$, $\eta_p^2<.01$. However, for ambiguous numbers, participants had more difficulties answering when the number was displayed on the table than on the wall ($M_{table}=1484.88$, $SD_{table}=458.01$, $M_{wall}=1155.41$, $SD_{wall}=461.21$), $F(1, 2828.35)=187.74$, $p<.001$, $\eta_p^2=.03$.

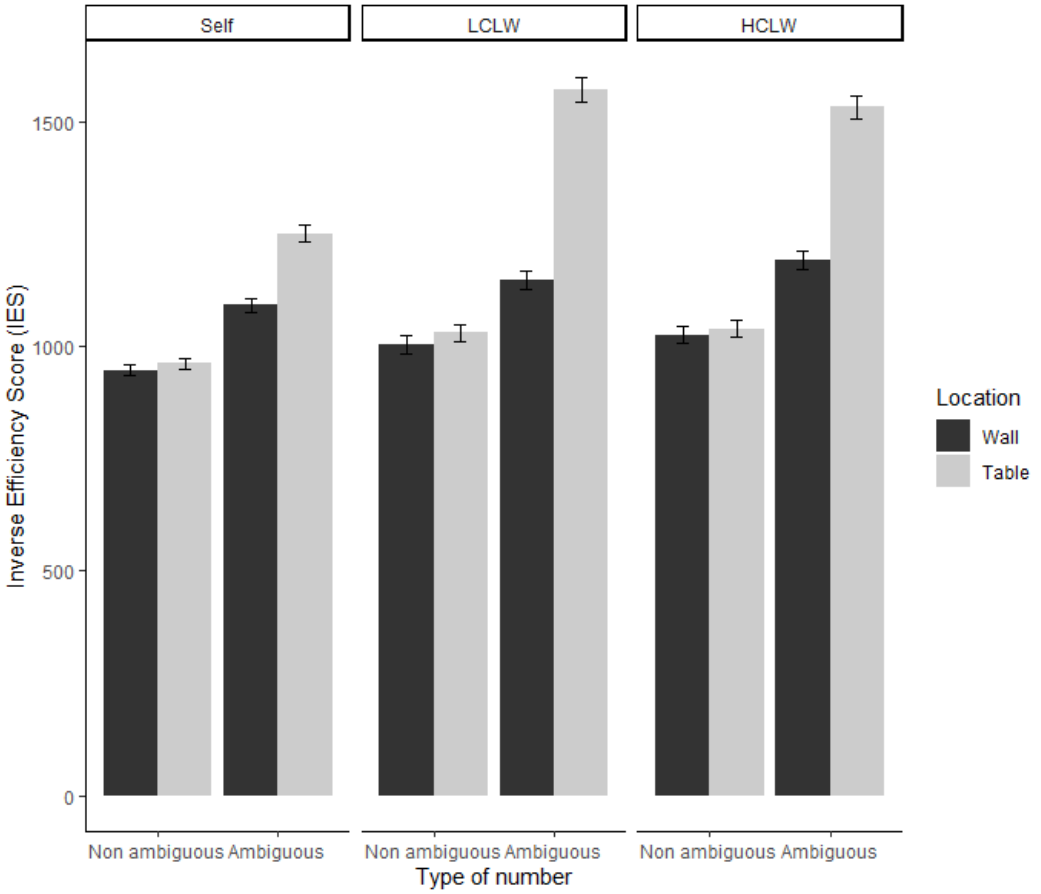

Figure 1. Graph of inverse efficiency score as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the response time in milliseconds to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 2). Significant effects of type number and of location indicated that participants took more time answering when the number was ambiguous ($M_{ambiguous}$=1086.98, $SD_{ambiguous}$=400.79, $M_{non\text{-}ambiguous}$=905.35, $SD_{non\text{-}ambiguous}$=323.24), F(1, 2806.91)=366.29, p<.001, η_p^2=.08, or on the table (M_{table}=1045.14, SD_{table}=402.15, M_{wall}=942.99, SD_{wall}=337.82), F(1, 2807.20)=151.02, p<.001, η_p^2=.03.

A set of contrasts allowed us to compare the response time for 'self' versus 'other' (both HCLW and LCLW), as well for the HCLW vs. LCLW avatars. The 'self' versus 'other' contrast proved significant, $F(1, 2807.93)=220.32$, $p<.001$, $\eta_p^2=.05$, indicating that participants took more time answering for the avatar than for themselves, regardless of the avatar's social group ($M_{other}=1068.28$, $SD_{other}=405.42$, $M_{self}=915.68$, $SD_{self}=321.01$). There was no significant difference between the LCLW and the HCLW targets, $F(1, 2807.19)=1.93$, $p=.165$, $\eta_p^2<.01$.

The location × HCLW versus LCLW avatars contrast interaction was significant, $F(1, 2805.99)=3.93$, $p=.048$, $\eta_p^2<.01$. Specifically, for the HCLW avatar, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=1158.16$, $SD_{table}=446.20$, $M_{wall}=1019.43$, $SD_{wall}=364.27$), $F(1, 2807.16)=51.70$, $p<.001$, $\eta_p^2=.01$. This difference emerged even more strongly for the LCLW avatar ($M_{table}=1152.87$, $SD_{table}=426.12$, $M_{wall}=957.68$, $SD_{wall}=348.55$), $F(1, 2806.63)=96.09$, $p<.001$, $\eta_p^2=.02$. No further interaction including this factor emerged, all Fs<0.57, all ps>.450, all $\eta_p^2<.01$.

The type of number × location × 'self' versus 'other' contrast was significant, $F(1, 2807.18)=28.82$, $p<.001$, $\eta_p^2<.01$. Follow-up analyses revealed that the number × location interaction was significant for the 'self' trials, $F(1, 2807.14)=5.72$, $p=.017$, $\eta_p^2<.01$. Specifically, for non-ambiguous numbers, there was no significant difference in response times as a function of location ($M_{table}=867.19$, $SD_{table}=285.13$, $M_{wall}=853.30$, $SD_{wall}=290.33$), $F(1, 2806.25)=0.56$, $p=.456$, $\eta_p^2<.01$. However, for ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall ($M_{table}=1022.01$, $SD_{table}=363.98$, $M_{wall}=940.51$, $SD_{wall}=323.39$), $F(1, 2807.67)=15.92$, $p<.001$, $\eta_p^2=.01$.

Turning to the 'other' trials, the type of number × location interaction was also significant, $F(1, 2807.27)=99.12$, $p<.001$, $\eta_p^2=.02$. Specifically, for non-ambiguous numbers,

there was no significant difference in response times as a function of location (M_{table}=964.66, SD_{table}=357.81, M_{wall}=937.50, SD_{wall}=342.14), F(1, 2806.47)=2.26, p=.133, η_p^2<.01. However, for ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall (M_{table}=1383.55, SD_{table}=411.57, M_{wall}=1042.56, SD_{wall}=366.07), F(1, 2808.46)=228.23, p<.001, η_p^2=.05.

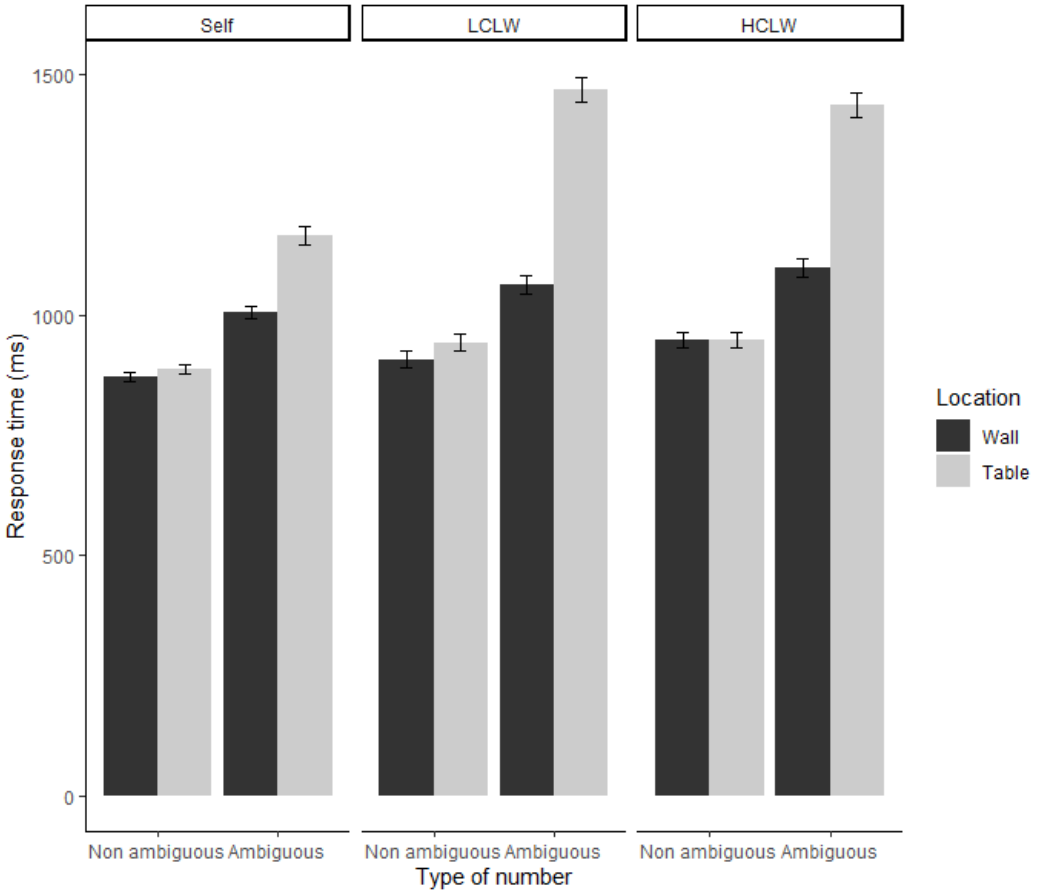

Figure 2. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Errors. We submitted the errors (with 0 indicating a failed trial and 1 indicating a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) logistic mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 3). We found

a significant main effect of type of number and location, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.40, *z*=-5.61, *p*<.001, or when the number was on the table, *OR*=0.38, *z*=-5.99, *p*<.001. Turning to the effect of perspective, the 'self' versus 'other' contrast was not significant, *z*=-0.31, *p*=.753. The HCLW versus LCLW contrast proved significant, indicating that participants were less likely to answer correctly for the LCLW than for the HCLW avatar, *OR*=0.60, *z*=-2.36, *p*=.018. No interaction including this factor emerged, all *z*s<|1.16|, *p*s>.246

The type of number × location interaction was significant, *z*=-4.11, *p*<.001. Specifically, for non-ambiguous numbers, participants were no less likely to answer correctly when the number appeared on the table, *z*=-1.23, *p*=.217. However, for ambiguous numbers, they were less likely to answer correctly when they appeared on the table, *OR*=0.20, *z*=-7.97, *p*<.001. No other interaction proved significant, all *z*s<|1.40|, *p*s>.161.

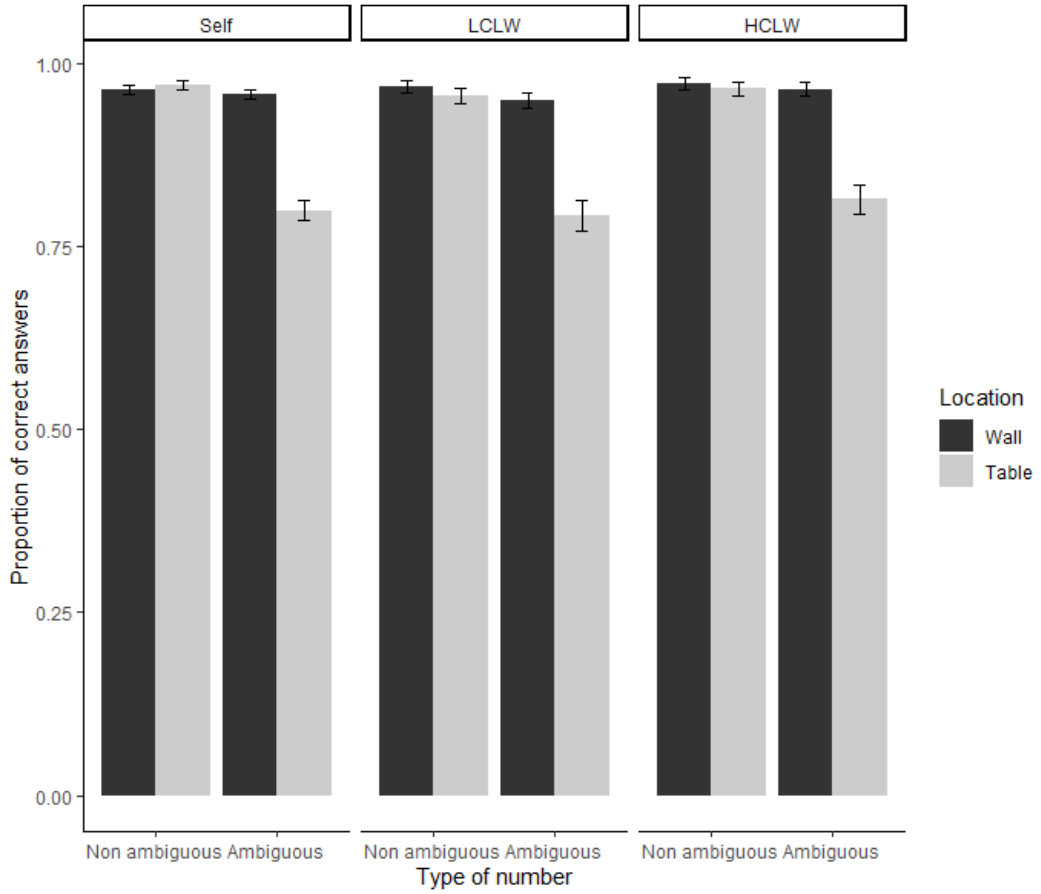

Figure 3. Graph of proportion of correct answers as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Additional Experiment 2

This experiment consists of a replication of the *social groups* condition of our first experiment, without time pressure.

Method

Power calculations. We computed our power analysis to detect the number type (S) x location (L) x perspective (P) interaction, with participants (I) as random factor and eight replicates. Planned contrasts were used to compare the two avatars. Var(error) was set at 0.2 and var(I*S*L*P) was set at 0.04.

Participants and design A total of 101 French-speaking participants (M_{age}=29.49, SD_{age}=9.79, 39 women) took part in the 25 min. study on Prolific Academic[8] (prolific.co) in exchange of a financial compensation of £2.30. The experiment involved a 2 (type of number: non-ambiguous=0 and 8 vs. ambiguous=6 and 9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatar vs. LCLW quadrant avatars) within-participant repeated measures design.

Power calculations. We conducted our power analysis for a 2 x 2 x 3 repeated measures design using the PANGEA webapp, and aimed for 90% power. This analysis indicated that we needed at least 50 participants to achieve 90% power to detect a small effect (d=.2).

Procedure. The procedure was identical to that of Additional Experiment 1, with a few modifications from taken Experiment 2. First, we raised the number of trials to 192. Second, we randomly asked participants to indicate the social group of the avatar in the previous trial. This was asked fifteen times throughout the task. Finally, there was no time pressure. That is, we asked participants to answer as correctly as possible, and asked them to take their time to ensure that they produced a correct answer.

[8] https://prolific.co/

Results

Analyses were performed using R (R Core Team, 2020), with the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages. Again, and following Surtees et al. (2012), trials for which participants had to press the N (incorrect) button were not considered, leaving out 96 of the 192 trials for each participant. We checked if any participant failed to perform above chance, which was not the case. Furthermore, and also as in Surtees and colleagues (2012), we excluded trials for which the response time was 2.5 SD away from the mean response time (0.97% of the remaining 96 trials per participant).

Inverse Efficiency Score. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the IES to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 1). Significant effects of type number and of location indicated that participants had more difficulties answering when the number was ambiguous ($M_{ambiguous}$=1442.07, $SD_{ambiguous}$=877.53, $M_{non\text{-}ambiguous}$=1166.88, $SD_{non\text{-}ambiguous}$=848.17), F(1, 9277.62)=299.09, p<.001, η_p^2=.02, or on the table (M_{table}=1414.77, SD_{table}=921.51, M_{wall}=1194.22, SD_{wall}=806.73), F(1, 9277.44)=199.91, p<.001, η_p^2=.02.

A set of contrasts allowed us to compare the IES for ‘self’ versus ‘other’ (both HCLW and LCLW), as well for the HCLW vs. LCLW avatars. The ‘self’ versus ‘other’ contrast proved significant, F(1, 9277.66)=54.96, p<.001, η_p^2<.01, indicating that participants had more difficulties answering for the avatar than for themselves, regardless of the avatar’s social group (M_{other}=1343.42, SD_{other}=863.41, M_{self}=1220.26, SD_{self}=888.54). There was no significant difference between the LCLW and the HCLW targets, F(1, 9277.05)=0.38,

p=.536, η_p^2<.01. No interaction including this factor emerged, all Fs<0.36, all ps>.546, all η_p^2<.01.

The type of number × location × 'self' versus 'other' contrast interaction was significant, F(1, 9277.49)=6.08, p=.014, η_p^2<.01. Follow-up analyses revealed that the number × location interaction was significant for 'self' trials, F(1, 9277.64)=13.81, p<.001, η_p^2<.01. Specifically, for non-ambiguous numbers, there was no significant difference in difficulty as a function of location (M_{table}=1133.02, SD_{table}=833.32, M_{wall}=1142.50, SD_{wall}=935.36), F(1, 9277.11)=0.05, p=.817, η_p^2<.01. However, for ambiguous numbers, participants had more difficulties answering when the number was displayed on the table than on the wall (M_{table}=1411.64, SD_{table}=849.03, M_{wall}=1210.45, SD_{wall}=904.13), F(1, 9278.22)=24.69, p<.001, η_p^2=.01.

Turning to 'other' trials, the type of number × location interaction was also significant, F(1, 9277.35)=92.09, p<.001, η_p^2=.01. Specifically, for non-ambiguous numbers, participants had more difficulties answering when the number was displayed on the table than on the wall (M_{table}=1232.32, SD_{table}=947.34, M_{wall}=1130.70, SD_{wall}=686.97), F(1, 9277.37)=14.43, p<.001, η_p^2<.01. This effect emerged even more strongly for ambiguous numbers (M_{table}=1750.99, SD_{table}=877.21, M_{wall}=1274.89, SD_{wall}=790.03), F(1, 9277.26)=298.87, p<.001, η_p^2=.02.

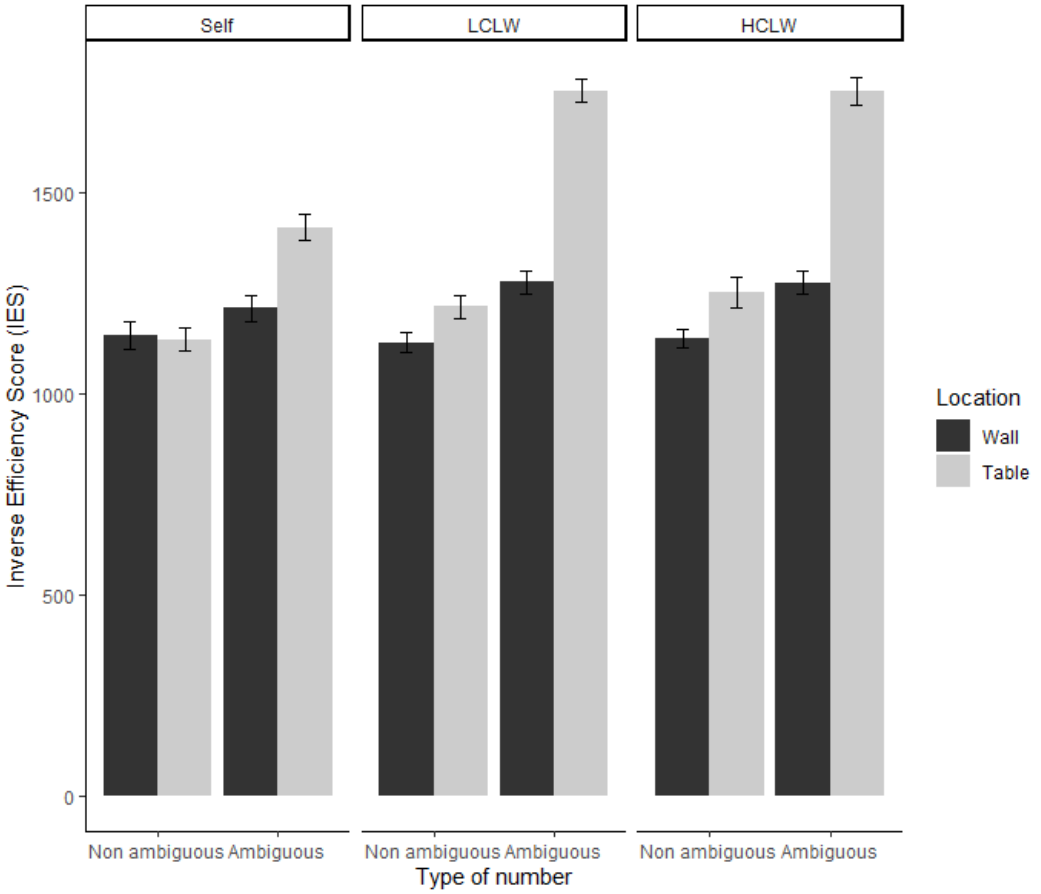

Figure 1. Graph of inverse efficiency score as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Response time. Following previous guidelines (Surtees et al., 2012; Simpson & Todd, 2017), we excluded failed trials from analyses. We submitted the response time in milliseconds to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 2). Significant effects of type number and of location indicated that participants took more time answering when the number was ambiguous ($M_{ambiguous}$=1398.60, $SD_{ambiguous}$=828.16, $M_{non\text{-}ambiguous}$=1130.75, $SD_{non\text{-}ambiguous}$=786.81), $F(1, 9277.74)=327.99$, $p<.001$, $\eta_p^2=.03$, or on the table

(M_{table}=1371.91, SD_{table}=860.55, M_{wall}=1157.47, SD_{wall}=760.88), F(1, 9277.56)=216.80, p<.001, η_p^2=.02.

A set of contrasts allowed us to compare the response time for 'self' versus 'other' (both HCLW and LCLW), as well for the HCLW vs. LCLW avatars. The 'self' versus 'other' contrast proved significant, F(1, 9277.78)=61.37, p<.001, η_p^2<.01, indicating that participants took more time answering for the avatar than for themselves, regardless of the avatar's social group (M_{other}=1302.97, SD_{other}=814.86, M_{self}=1181.91, SD_{self}=819.96). There was no significant difference between the LCLW and the HCLW targets, F(1, 9277.19)=0.24, p=.625, η_p^2<.01. No interaction including this factor emerged, all Fs<0.42, all ps>.515, all η_p^2<.01.

The type of number × location × 'self' versus 'other' contrast was significant, F(1, 9277.62)=6.36, p=.012, η_p^2<.01. Follow-up analyses revealed that the number × location interaction was significant for the 'self' trials, F(1, 9277.76)=15.95, p<.001, η_p^2<.01. Specifically, for non-ambiguous numbers, there was no significant difference in response times as a function of location (M_{table}=1095.95, SD_{table}=743.51, M_{wall}=1105.57, SD_{wall}=859.10), F(1, 9277.24)=0.07, p=.799, η_p^2<.01. However, for ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall (M_{table}=1372.76, SD_{table}=829.31, M_{wall}=1170.28, SD_{wall}=817.37), F(1, 9278.32)=28.45, p<.001, η_p^2<.01.

Turning to the 'other' trials, the type of number × location interaction was also significant, F(1, 9277.47)=101.86, p<.001, η_p^2=.02. Specifically, for non-ambiguous numbers, participants took longer to answer when the number was displayed on the table than on the wall (M_{table}=1193.44, SD_{table}=869.16, M_{wall}=1098.17, SD_{wall}=672.51), F(1, 9277.50)=14.69, p<.001, η_p^2<.01. This effect emerged even more strongly for ambiguous numbers

(M_{table}=1699.15, SD_{table}=817.94, M_{wall}=1235.91, SD_{wall}=756.21), F(1, 9277.70)=314.91, p<.001, η_p^2=.03.

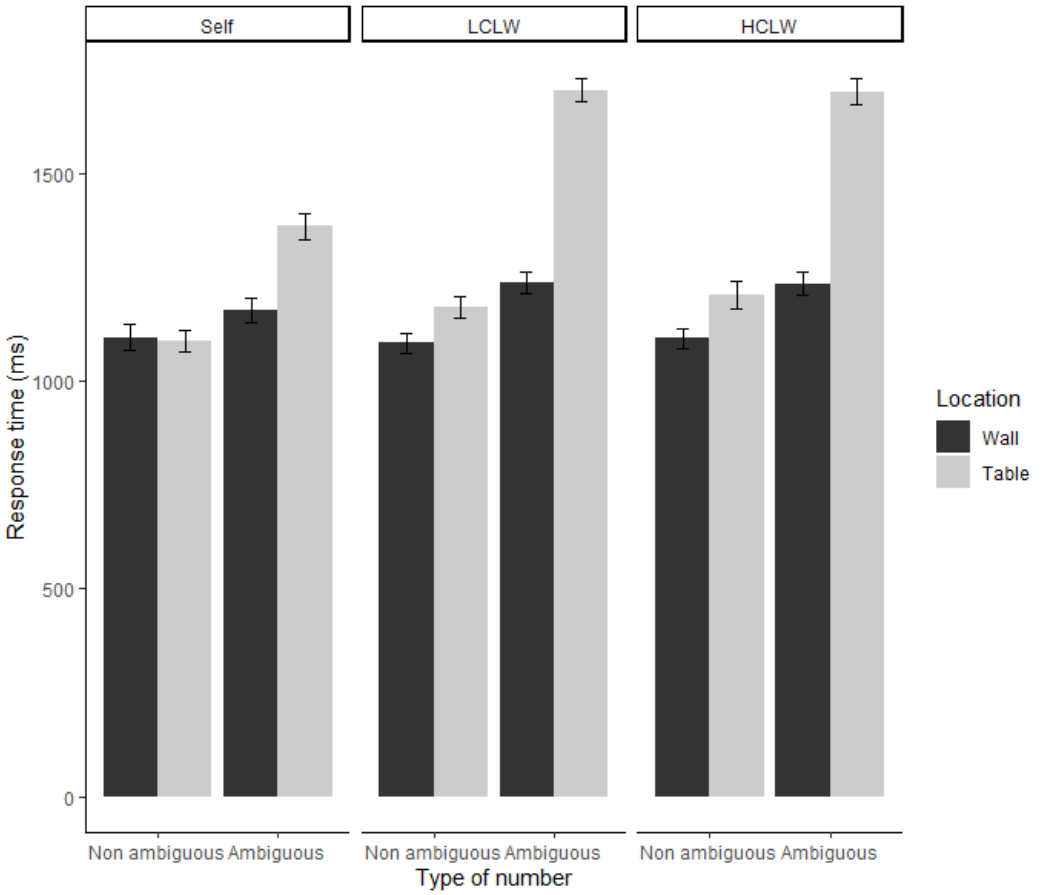

Figure 2. Graph of response time as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.

Errors. We submitted the errors (with 0 indicating a failed trial and 1 indicating a successful trial) to a 2 (type of number: non-ambiguous=0,8 vs. ambiguous=6,9) x 2 (location: wall vs. table) x 3 (perspective: self vs. HCLW quadrant avatars vs. LCLW quadrant avatars) logistic mixed model analysis with all three factors varying within participants and participants as random factor (for a visual representation of the data, see Figure 3). We found a significant main effect of type of number and location, indicating that participants were less likely to answer correctly when the number was ambiguous, *Odds Ratio* (*OR*)=0.59, z=-3.44, p<.001, or when the number was on the table, *OR*=0.38, z=-6.43, p<.001. Turning to the

effect of perspective, the ‘self’ versus ‘other’ contrast was significant, indicating that participants were less likely to answer correctly when answering from their own perspective, *OR*=1.46, *z*=2.50, *p*=.012. The HCLW versus LCLW contrast did not prove significant, indicating that participants were no less likely to answer correctly for the LCLW than for the HCLW avatar, *z*=-0.45, *p*=.651. No interaction including this factor emerged, all *z*s<|1.08|, *p*s>.281

The type of number × location interaction was significant, *z*=-1.66, *p*<.001. Specifically, for non-ambiguous numbers, participants were no less likely to answer correctly when the number appeared on the table, *z*=-0.64, *p*=.521. However, for ambiguous numbers, they were less likely to answer correctly when they appeared on the table, *OR*=0.16, *z*=-8.69, *p*<.001. No other interaction proved significant, all *z*s<|1.51|, *p*s>.132.

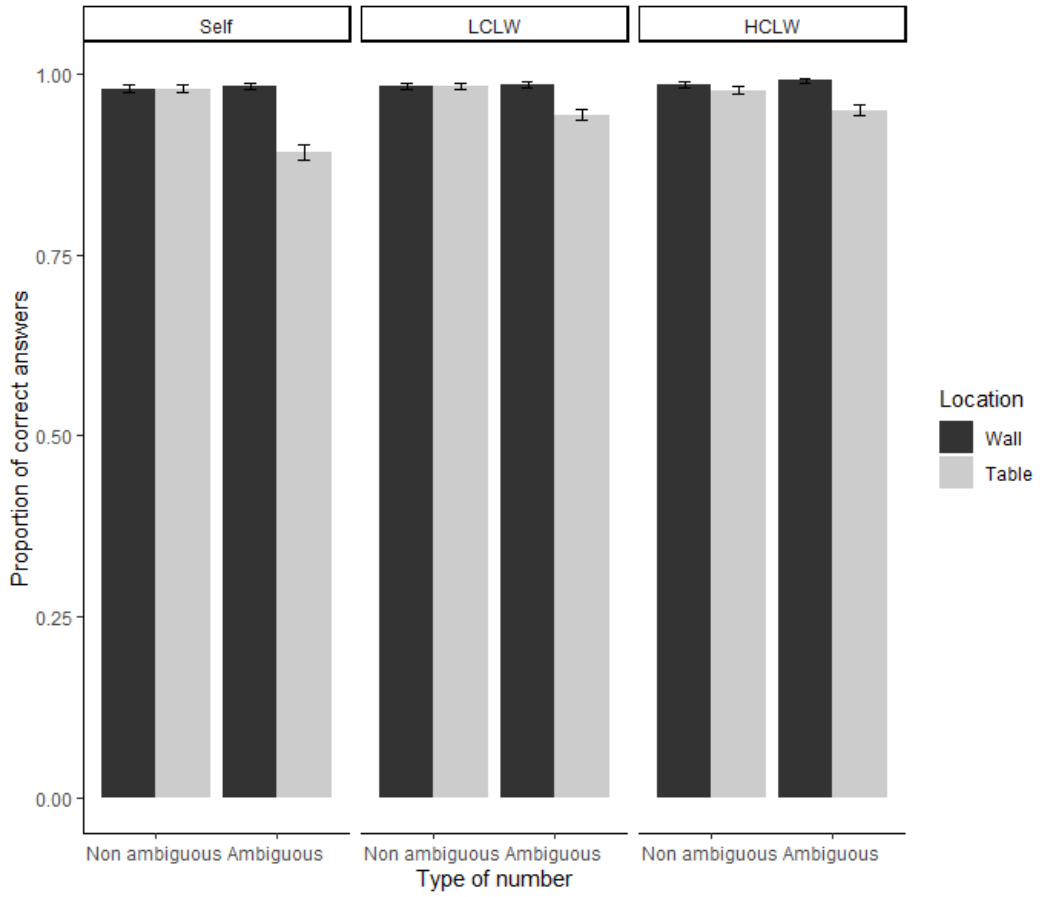

Figure 3. Graph of proportion of correct answers as a function of type of number, location, and perspective. Bars represent the standard error of the mean. LCLW=low competence low warmth; HCLW=high competence low warmth.