Journal of Experimental Psychology: Human Perception and Performance

The Extent of Default Visual Perspective Taking in Complex Layouts

Lewis J. Baker, Daniel T. Levin, and Megan M. Saylor

Online First Publication, November 9, 2015. http://dx.doi.org/10.1037/xhp0000164

Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/15/$12.00 http://dx.doi.org/10.1037/xhp0000164

The Extent of Default Visual Perspective Taking in Complex Layouts

Lewis J. Baker, Daniel T. Levin, and Megan M. Saylor
Vanderbilt University

Emerging research suggests that visual perspective taking might be based in part on a default, early developing cognitive process. This hypothesis receives support from experiments demonstrating that adults experience interference from task-irrelevant perspectives of depicted agents even when participants are making judgments about their own perspective. However, a number of recent articles conclude that this self-judgment interference effect may be because of simple directional cues alone, and might, therefore, not reflect processes specific to visual perspective taking. In 3 studies, we demonstrate that self-judgment interference is constrained by agents' apparent line-of-sight access to subspaces in realistic rendered scenes. Participants displayed processing costs when their perspective conflicted with that of an avatar, who faced in the direction of all possible targets but could not see some of the targets because of occlusion. This interference effect occurred using 2 different configurations of occluders, and disappeared when windows were added to the occluders, allowing avatars line of sight access to all of the targets visible to the participant. These results demonstrate that default perspective taking is not attributable to directional cues alone but instead reflects a relatively sophisticated calculation of an agent's line of sight.

Keywords: visual perspective taking, theory of mind, subitizing, gaze, line of sight

Supplemental materials: http://dx.doi.org/10.1037/xhp0000164.supp

Successful social interaction often requires putting aside one's own views to take the perspective of another. Workplace gossip, defensive driving, and lively argument all depend on taking perspectives to infer what another person is seeing, thinking, feeling, or desiring. The ability to take perspectives requires representation of another's mental state, discrimination of that mental state as separate from one's own, and the use of knowledge to make judgments based on these inferences. Although perspective taking has been characterized as a cognitively effortful process (e.g., Lin, Keysar, & Epley, 2010), recent findings indicate that observers may represent simple forms of another's visual perspective by default (Kovács, Téglás, & Endress, 2010), and that these representations can quickly and efficiently guide judgments about "belief-like states" (Apperly & Butterfill, 2009; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). These judgments are theorized to operate in parallel with the effortful processes associated with adult theory-of-mind (TOM; Apperly & Butterfill, 2009), yielding efficient responses to basic social stimuli that can be modified with executive control.

In the first demonstration of this effect, Samson et al. (2010) had participants judge the number of objects that they or a cartoon avatar could see (Figure 1a). Participants were slower and less accurate when their perspective differed from the avatar, even when reporting from their own perspective. However, recent debate has questioned the degree to which this self-judgment interference denotes true perspective taking (Heyes, 2014). Subsequent research has demonstrated self-judgment interference for arrows instead of avatars, thus, suggesting that the default system does not reflect implicit perspective taking (Santiesteban, Catmur, Bird, Hopkins, & Heyes, 2014), but rather a simple and well-known form of directional cuing (e.g., Friesen, Ristic, & Kingstone, 2004).

In three experiments, we demonstrate default representation of another's perspective even when perspective could not be predicted from a simple directional cue. We then demonstrate that variations in occluders and targets that affect line of sight modulate the self-judgment interference effect. These findings suggest a system that calculates the line-of-sight between agents and attended objects, and responds to more complex scenes than previously observed.

Lewis J. Baker, Daniel T. Levin, and Megan M. Saylor, Department of Psychology and Human Development, Vanderbilt University.

This material is based upon work supported by the National Science Foundation (NSF) Graduate Research Fellowship 2013139545 to L.J.B. and NSF Award 0826701 to D.T.L. The authors thank Timothy McNamara and Ian Apperly for critical comments to a draft, and Kristopher Preacher for consultation on statistical procedures.

Correspondence concerning this article should be addressed to Lewis Baker, Department of Psychology and Human Development, Vanderbilt University, Peabody #552, 230 Appleton Pl, Nashville, TN 37203. E-mail: lewis.j.baker@vanderbilt.edu

A Two-Systems Account of Perspective Taking

The two-system theory of perspective taking emerged from evidence differentiating relatively early developing infant perspective taking and apparently later-developing, cognitively taxing adult theory of mind tasks. A number of studies demonstrated that infants under 14 months can discriminate between objects an adult can and cannot see (Sodian, Thoermer, & Metz, 2007), follow gaze behind objects and others (Moll & Tomasello, 2006), respond to goal-oriented visual displays (Luo & Baillargeon, 2005), and may

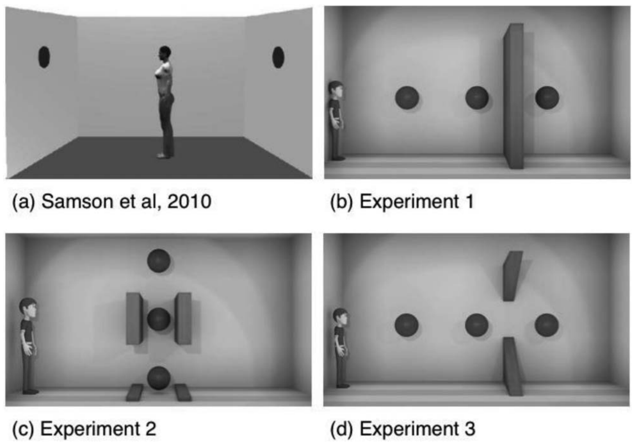

Figure 1. Comparison of stimuli across all experiments. Stimuli used by Samson et al. (a) that potentially conflates gaze direction and spatial grouping with perspective. Gaze cues were controlled using a wall in Experiment 1 (b). Perceptual grouping effects were controlled using a partial occlusion of sets Experiment 2 (c). The effect of visual access was directly tested using a "window" in Experiment 3 (d), in which avatar and self perspectives were always consistent relative to matched trials that mirrored Experiment 1.

even register the presence of false beliefs (Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007). Even so, older children and adults frequently display *egocentrism*, a failure to reason from perspectives different from one's own (Flavell, Everett, Croft, & Flavell, 1981; Piaget & Inhelder, 1956). Participants display egocentric errors when experiencing cognitive load (Lin, Keysar, & Epley, 2010), time constraints (Epley, Morewedge, & Keysar, 2004), increasing misalignment of perspectives (Shelton & McNamara, 1997), or a lack of motivation (Klein & Hodges, 2001). Such evidence has led to the view that adults and children have to overcome a default egocentricity to take others' perspectives (e.g., Birch & Bloom, 2004; Epley, Keysar, Van Boven, & Gilovich, 2004).

Given the contrast between infants' success on some perspective-taking judgments and the cognitive cost of more nuanced tasks, an emerging body of research suggests that multiple systems contribute to perspective taking. Such evidence supports an early developing system that quickly and efficiently tracks perspectives, while more complex theory of mind relies upon an effortful process working in parallel to the early system (Apperly & Butterfill, 2009).

Although previous authors have proposed multiple-systems accounts of theory of mind skills (e.g., Baron-Cohen, 1995; Leslie, Friedman, & German, 2004), the most relevant for present purposes stems from the work of Samson, Apperly, and colleagues (McCleery et al., 2011; Qureshi, Apperly & Samson, 2010; Samson et al., 2010; Surtees & Apperly, 2012). They argued that an early developing system directs the perception and representation of goals and belief-like states, while a complementary elaborative system selects between mental states (Ramsey, Hansen, Apperly, & Samson, 2013). The researchers hypothesized that the early system should respond to the presence of another's perspective even in circumstances where the late system generally ignores irrelevant perspectives. To demonstrate this, the researchers asked participants to report whether they or an avatar saw a cued number of items (see Figure 1a). Each trial began with a perspective prime (either YOU or HIM), and then a number prime within the range of subitization (0–3; Trick & Pylyshyn, 1994). After this, participants were shown a display containing an avatar and several targets. Participants were faster to respond when taking their own perspective, indicative of a default egocentrism while selecting between perspectives. Crucially, participants also demonstrated *self-judgment interference*. Even when they were prompted to take their own perspective and ignore the avatar, participants were slower and less accurate when the avatar's perspective differed from their own, supporting the existence of an early system.

Samson and colleagues suggested that participants implicitly encoded another's mental state, even when it was disadvantageous to do so. Samson and Apperly (2010) have since theorized that infant studies reflect an early stage of perspective taking that guides attention to the focal point of another's perspective. In contrast, a late system effortfully selects one perspective while inhibiting the other (Apperly & Butterfill, 2009; for a similar idea see Leslie, Friedman, & German, 2004). The Samson-Apperly task reflects both aspects of the system: the early system leading to self-judgment interference while the late system leads to egocentric interference when explicitly taking the avatar's perspective.

The Scope of Default Perspective Taking

As discussed in Samson et al. (2010), default perspective taking may result from perceptual, attentional, or cognitive cues embedded in the stimuli. Focusing on the simplest possible variant of this hypothesis, Heyes and colleagues recently argued that processing costs associated with another's differing perspective reflect the operation of a well-known form of domain-general directional spatial cueing (Heyes, 2014; Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014a; Santiesteban, Shaw, White, Bird, & Heyes, 2014b). For example, Santiesteban et al. (2014a) demonstrated that self-judgment interference could be replicated using arrow cues instead of avatars, suggesting a general cognitive effect instead of an effect specific to visual perspective taking. However, the authors also note that the default system theorized by Samson and Apperly might simply be promiscuous. Given that participants viewed scenes with both avatars and arrows, and given the proclivity of human beings to anthropomorphize inanimate objects, ranging from clouds (Guthrie, 1993) to robots (Baker, Hymel, & Levin, in review; Epley, Waytz, & Cacioppo, 2007), it is possible that the self-judgment interference effect readily generalizes to perceptual and symbolic forms of agency.

The possibility that self-judgment interference reflects a system that is promiscuous with regard to different specific agents makes it difficult to test the default system's limits using variations in attention-directing cues. However, other aspects of the self-judgment interference paradigm may be more effective in revealing whether the default system activates specifically to visual perspectives. The work by Santiesteban and colleagues suggests a *directional cue*, whereby attention is drawn to the target of any generalized spatial cue, including arrows or avatars. Meanwhile, Samson, Apperly, and colleagues propose a *gaze specific cue*, whereby attention is drawn to the accessible targets of another agent's gaze. Such a line-of-sight hypothesis is reinforced by decades of research demonstrating sensitivity to others' line of sight in humans (for review, see Kleinke, 1986), nonhuman primates (Hare, Call, & Tomasello, 2001), and even scrub jays (Dally, Emery, & Clayton, 2006). Subjects in these paradigms quickly assess occlusion of another's perspective by calculating the sight lines linking agents to objects. Therefore, it may be possible to assess the degree to which gaze cueing is specific to a line of sight that links an agent with a target, whereas a more general form of directional cueing would activate parts of scenes without reference to the agent's specific perceptual access to parts of a scene.

The agent-target relationships in the original Samson-Apperly task and all of its subsequent replications (McCleery et al., 2011; Qureshi, Apperly, & Samson, 2010; Samson et al., 2010; Samson & Apperly, 2010; Surtees & Apperly, 2012) leave open the possibility that the observed directional cueing effect reflects preferential activation of an entire half of the scene without regard to sightlines because the avatar was centered in the scene and looked to one half or the other. Looking at Figure 1a, which reproduces the original Samson et al. avatar-target configuration, it is apparent that an avatar's perspective is perfectly correlated with its orientation. In contrast, the avatar-target relationships illustrated in Figure 1b and 1c preclude a simple directional cue from generating interference because all targets lay to one side of the avatar. Instead, the avatar's limited perspective is derived entirely from an occluded sightline to only two of the three targets present in the scene. Therefore, the experiments reported here test for a self-judgment interference effect using these stimuli. If participants demonstrate self-judgment interference, then it is likely that interference reflects a form of attentional cueing specific to gaze rather than a more general directional cue.

Experiment 1

Experiment 1 tested whether the self-judgment interference effect associated with conflicting perspectives originates from directional cueing, or whether this interference reflects a more sophisticated tracking of gaze to targets that are in an avatar's line of sight. The design and procedure were identical to Experiment 1 of Samson et al. (2010), with one key exception. In previous experiments, the perspective of the avatar was confounded with the direction of its gaze. For instance, when an avatar faced left, it could see everything to the left and only the items to the left (Figure 1a). In the current experiment, a barrier sometimes obstructed the avatars' view. In this manner, it was possible for an avatar to see only a subset of objects in a scene even though all of the objects were located in the general direction of the avatar's gaze (Figure 1b). If self-judgment interference results from directional cues, then we should see no interference in any trial, as all targets are in the cued direction. On the other hand, participants should demonstrate self-judgment interference if they are calculating perspective from avatar's line-of-sight access to targets.

Method

Participants. Twenty-four Vanderbilt University undergraduates were recruited for class participation credit (14 women, mean age = 19.24). All participants in this study were treated in compliance with the Vanderbilt Institutional Review Board and American Psychological Association (APA) ethical guidelines.

Stimuli. Stimuli were reproduced following the method described in Samson et al. (2010) with the critical replacement of avatar gaze direction with the placement of an occlusion. Test images consisted of a three-dimensional (3D) room containing a cartoon avatar, a red wall, and zero to three blue dots (Figure 1b). In all cases, the participant could see all of the displayed dots, while the avatar could see varying numbers of the dots. Sixty unique image combinations were created using the 3D editing software, Blender (v2.66; Blender Foundation, 2013, www.blender.org), using cartoon avatars downloaded from an open source file sharing site (VMComix, 2011). Images consisted of an avatar (standing on the far left facing right or standing on the far right facing left), 0–3 dots, and a wall in one of four possible locations, occluding 0–3 dots from the avatar's view. Half of the images contained *Inconsistent* perspectives, where the avatar and a participant saw different numbers of targets, while the other half contained *Consistent* perspectives. Avatars were matched to participant gender, with males seeing male avatars and vice versa. Positions of avatars, dots and the wall were evenly balanced across all trials.

Four unique cue sets were created for each image, again following the method of Samson et al. (2010) The breakdown of stimulus combinations is shown in Figure 2. Each image was paired with each combination of perspective cues and number cues, for a total of four trials per image. Perspective cues either primed the participant's perspective (*Self* trials: "YOU") or the avatar's perspective (*Other* trials: "HIM" or "HER"). Number cues either aligned with the perspective cued (*Matched*) or were not aligned with the perspective cued (*Mismatched*). Critically, Mismatched trials for Inconsistent images displayed the number of targets seen by the irrelevant perspective. Take the example of an image where the avatar sees one target but the participant sees two targets (as in the bottom of Figure 2). In a Self-Matching trial, the cues would be "YOU" and "2," but in Self-Mismatching trials, the cues would be "YOU" and "1." In

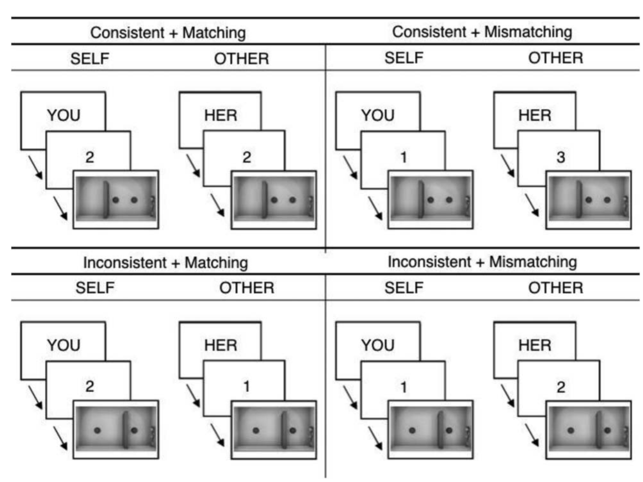

Figure 2. Trials for Experiment 1. Images consisted of 0–3 targets, a wall in any of four positions, and an avatar that could appear on either side of the display. Images showed either Consistent or Inconsistent perspectives (same or different views of targets). Participants responded from Self ("YOU") or Other ("HIM"/"HER") perspectives that either Matched or Mismatched the number of targets seen by that perspective.

Mismatching trials for the Consistent images, the number cue did not correspond to either perspective.

Procedure. Upon giving consent, participants completed the entire experiment using the psychophysics toolbox for MATLAB (Kleiner, Brainard, & Pelli, 2007) on an Apple Mac Mini with a 19.5" LCD monitor (1,600 × 900 p, 60 Hz). Participants read instructions and then completed a tutorial before beginning experimental trials. The tutorial consisted of four practice trials, which followed the experimental procedure outlined below. Each practice trial demonstrated a different condition (self-consistent, other-consistent, self-inconsistent, and other-inconsistent). Practice trials gave feedback and repeated until participants made the correct response. The experimenter verbally confirmed that participants understood the tutorial before the experiment began.

The flow of experimental trials is diagramed in Figure 3. Participants viewed each of the 60 displays four times, once for each combination of perspective cue (YOU vs. HIM/HER) and number cue (Matching vs. Mismatching), for a total 240 trials in all. Trials were completely randomized over four blocks. The entire experiment took less than 25 min.

Results

Previous studies have excluded large numbers of trials to investigate this effect, most often only analyzing Matching trials (McCleery et al., 2011; Qureshi et al., 2010; Samson et al., 2010; Surtees & Apperly, 2012). Such broad exclusion here would eliminate 2,880 of 5,760 total trials, and may potentially bias our findings. In an effort to analyze this effect using the most inclusive criterion possible, only those trials with no targets present at all were excluded (198 trials; 3.4%). Separate analyses were conducted for response times and error rates. All multiple comparison tests were adjusted using the Benjamini and Hochberg method (Benjamini & Hochberg, 1995). We have provided supplementary material for all experiments that adhere to the broader exclusion principle for comparison to previous research.

Response time. Response times are visualized in Figure 4. A 2 (Perspective: Self vs. Other) × 2 (Consistency: Consistent vs. Inconsistent) within-subjects analysis of variance (ANOVA) revealed a significant main effect of perspective type ($F(1, 23) =$ 56.224, $\eta^2 = .678$, $p < .001$), with response times increasing from Self trials to Other trials by 65.15 ms. There was also a main effect of consistency ($F(1, 23) = 81.340$; $\eta^2 = .780$, $p < .001$), with a net increase of 48.8 ms from Consistent to Inconsistent trials. There was a Perspective*Consistency interaction effect ($F(1, 23) = 5.591$, $\eta^2 = .196$, $p = .027$), with greater response time delays between Other-Consistent and Other-Inconsistent trials relative to Self-Consistent and Self-Inconsistent trials. Of critical importance is the presence of self-judgment interference, when participants reported their own perspective in the face of an avatar's contrasting perspective. Self-Inconsistent trials were significantly slower than Self-Consistent trials (mean difference = 42.9 ms, $t_{23} = 5.242$, $d = 1.07$, $p < .001$). This interference effect was also seen between Other-Consistent and Other-Inconsistent trials (mean difference = 69.6 ms, $t_{23} = 6.770$, $d = 1.382$, $p < .001$).

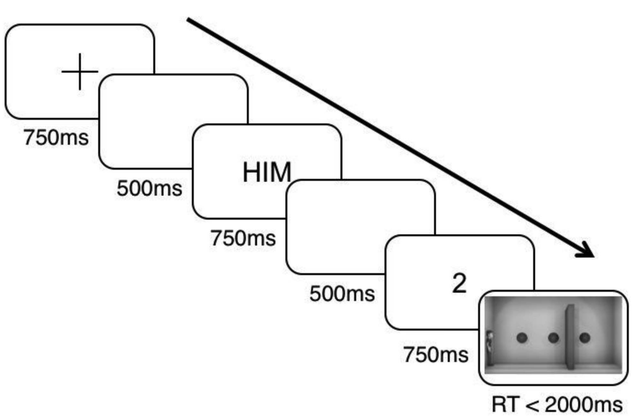

Figure 3. Sample timing of an Other-Inconsistent-Matching trial. Participants responded whether the perspective cued could see the number of objects cued in the scene. All experiments followed this time-course.

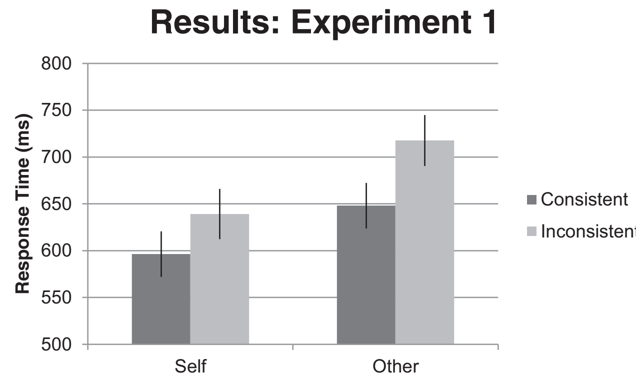

Figure 4. Response times for Experiment 1. Participants were significantly slower for inconsistent perspectives, even when taking their own perspective. Error bars denote *SEM*.

Error rate. Error rates followed the same trends as response times. A 2 × 2 ANOVA revealed a significant main effect of Consistency (mean difference = 7.6% correct; $F(1, 23) = 27.64$, $\eta^2 = .546$, $p < .001$), but no main effect of Perspective (mean difference = 2.5%; $F(1, 23) = 2.887$ $p = .103$) and no interaction effect ($F(1, 23) = 1.88$, $p = .183$). Self-Inconsistent trials were significantly more error prone than Self-Consistent trials (mean difference = 9.16%, $t_{23} = 9.04$, $d = 1.847$, $p < .001$). Participants were also more error prone in Other-Inconsistent trials compared to Other-Consistent trials (mean difference = 6.04%, $t_{23} = 1.797$, $d = .367$, $p = .043$).

Discussion

These results demonstrate a self-judgment interference effect, even when avatar direction was not predictive of avatar perspective. If the interference effect stems from simple directional cues, we would expect no interference for the inconsistent trials in Experiment 1, as all targets were consistent with the avatar's direction. However, participants were slower and more error prone when an avatar's perspective differed from their own because of occlusion, suggesting interference caused by some accounting of the avatar's sightlines.

It is still possible that the interference effects of Experiment 1 were at least partially influenced by another general cognitive cue: perceptual grouping. Every inconsistent trial in Experiment 1 featured a perceptual division (i.e., a wall) that segmented targets in space. It may be that interference effects are limited to situations where this form of perceptual support is strong. Issues such as these suggest the need to test for interference using multiple wall-target configurations to ensure that the effect is robust enough to serve as a model for a substantial range of real-world situations. Experiment 2, therefore, tested whether processing delays would occur using a new configuration of occluders and targets that was less easily grouped.

Experiment 2

Experiment 1 demonstrated that directional cueing alone could not account for interference from an avatar's perspective. This supports the hypothesis that participants calculated the contents of an agent's visual perspective and were not simply guided by directional cues. However, it is possible that the specific arrangement of targets and occluders increased the complexity of the task on critical inconsistent trials. In Experiment 1, as well as other studies of this design, objects in every critical incongruent trial were segmented in space by a wall (in our experiment) or by the avatar itself (in previous experiments). This division of targets into subgroups may have elicited processing delays by interrupting subitization of targets (Starkey & McCandliss, 2014). Effectively, this perceptual division might require participants to subitize two groups of targets rather than one holistic chunk, thus, increasing response times and reducing accuracy for inconsistent trials. Therefore, it is important to disrupt the avatar's line-of-sight access to targets without perceptually separating targets from each other. If the self-judgment interference effect is limited to situations involving subdivision of targets, then ensuring that occluders are not interposed among targets should eliminate the effect.

To investigate this, we created scenes in which some targets were blocked from the avatar's perspective without the set of targets being subdivided by a wall (Figure 1c). Targets were arranged vertically, with wall segments blocking line-of-sight access to some of the targets on inconsistent trials. Because the occluders were placed between the avatar and the entire set of objects there was no opportunity to perceptually group the objects with the wall acting as a divider.

Method

Participants. Eighteen Vanderbilt University undergraduates were recruited for class participation credit (14 women, mean age = 19.24).

Stimuli and procedure. The design was identical to Experiment 1, with the exception of the alignment of targets and occluders (Figure 1c). Experiment 2 controlled for perceptual grouping effects by using two walls segmented into three parts each to occlude a subset of targets on a vertical axis. Whereas the targets in Experiment 1 were arranged horizontally with a wall dividing the space, Experiment 2 featured vertically arranged targets with six wall segments that were toggled to permit the avatar's line of sight to one or more of the targets. In consistent trials, there were no walls between any targets and the avatar. In inconsistent trials, walls blocked a number of targets from the avatar's line of sight. Wall segments were always shown in pairs, to ensure target occlusion regardless of avatar position. Each stimulus image was transposed from the horizontal layouts of Experiment 1, so all design parameters were also identical. The procedure was the same as depicted in Figure 3.

Results

Response times. Analysis of response times (Figure 5) again revealed a significant self-judgment interference effect. A 2 (Perspective: Self vs. Other) × 2 (Consistency: Consistent vs. Inconsistent) within-subjects ANOVA revealed a significant main effect of perspective type ($F(1, 17) = 61.560$, $\eta^2 = .784$, $p < .001$), with response times increasing from Self trials to Other trials by 94.66 ms. There was also a main effect of consistency ($F(1, 17) = 35.65$; $\eta^2 = .677$, $p < .001$), with an net increase of 43.28 ms from Consistent to Inconsistent trials. The interaction was not significant ($F(1, 17) = 2.722$, $p = .117$). Self-Inconsistent trials were again slower than Self-Consistent trials (mean difference = 31.9 ms, $t_{17} = 2.593$, $d = .611$, $p = .009$). Once again, this interference effect was seen between Other-Consistent and Other-Inconsistent trials as well (mean difference = 54.7 ms, $t_{17} = 5.479$, $d = 1.291$, $p < .001$).

Error rate. Error rates trended in the same direction as Experiment 1. A 2 × 2 ANOVA revealed a main effect of Consistency (mean difference = 3.75% correct; $F(1, 17) = 9.026$, $\eta^2 = .347$, $p = .008$), but no significant effect of Perspective (mean difference = 2.01%; $F(1, 17) = 2.700$, $p = .119$) or interaction ($F(1, 17) = 1.242$, $p = .281$). Participants were again more error prone in Self-Inconsistent trials than in Self-Consistent trials (mean difference = 4.92%; $t_{17} = 4.46$, $d = 1.052$, $p < .001$), but not between Other-Inconsistent and Other-Consistent trials (mean difference = 2.56%; $t_{17} = 0.38$, $p = .354$).

Discussion

Participants again demonstrated a self-judgment interference effect when responding to stimuli containing a conflicting perspective. The stimulus design of Experiment 2 ensured that this effect was not because of perceptual grouping principles that might delay response time when enumerating targets separated in space. The

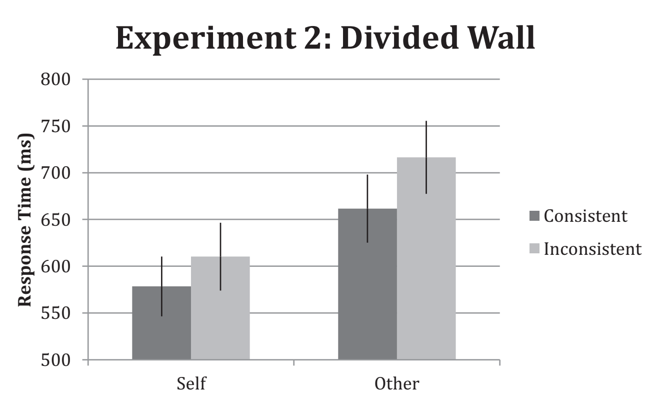

Figure 5. Response times for Experiment 2. There was a significant cost for Inconsistent trials from both self and other perspectives. Error bars denote *SEM*.

results of Experiments 1 and 2 successfully rejected two general spatial cue alternatives to the default perspective-taking claim.

The interaction effect for response times in Experiment 1 might suggest that egocentrism and perspective interference are multiplicative in terms of cognitive cost, with the difference between Other-Consistent and Other-Inconsistent trials being greater than the difference between Self-Consistent and Self-Inconsistent trials. It is interesting that this interaction is not significant in Experiment 2, although behavior follows similar trends. It is therefore possible that the need to subdivide targets into different subspaces induces some form of complexity that in turn leads to interference between processes needed to focus on the "other" perspective and processes needed to suppress one's own knowledge. Because these experiments were not designed to test this hypotheses, factors moderating variability in the amount of interference between the self and other conditions warrant further investigation.

Experiments 1 and 2 found that the self-interference effect appeared even while controlling for directional cues and perceptual grouping. The most likely alternative hypothesis is that individuals are sensitive to the line-of-sight access of agent to targets of attention. Experiment 3 further tested whether the avatar's line of sight to the targets was responsible for the interference effect, while also directly replicating Experiment 1.

Experiment 3

Experiment 3 further tested the hypothesis that default perspective taking reflects line of sight calculations by directly comparing responses in conditions with and without constant line-of-sight access to targets. The Wall condition directly replicated the design of Experiment 1, where an avatar's visual access to targets was potentially occluded by a wall. Participants in the Window condition were presented the exact same displays as the Wall condition, except a notch in the wall provided avatars with a line of sight to all potential targets (Figure 1d). This allowed us to test the robustness of default line of sight calculations. Experiment 2 demonstrated that an interference effect could be produced based on participants interpretation of the specific access afforded by occluders of varying form and location. Experiment 3 was designed as a complement to this, by testing whether similar variation in the form of occluders (in this case, the addition to notched windows) could eliminate interference effects by allowing avatars a view of all targets through the window.

Method

Participants. Thirty-two Vanderbilt University undergraduates were recruited for class participation credit. One participant was excluded for failing to follow instructions. This data was substituted by another participant, leaving 16 participants in each condition (mean age = 19.0, 30 women).

Stimuli and procedure. Participants were tested in one of two conditions. Stimuli for the Wall condition completely replicated the design and procedure of Experiment 1. Stimuli for the Window condition used the exact same combination of targets, avatar locations and ratio of consistent to inconsistent images as Experiment 1, save for one key difference shown in Figure 1d: the walls in Experiment 3 were notched, leaving a large "window." Avatars maintained line of sight of all targets through the window. Because of this, Self and Other perspectives in all trials in the Window condition are technically consistent. "Inconsistent" trials in the Window condition refer to those displays that would be Inconsistent in the Wall condition (Figure 6). Before beginning the experiment, all participants verbally confirmed that they perceived that the avatar could see through the window. All other parameters were exactly as in Experiment 1.

Results

Analysis of variance. Mixed-effects ANOVAs were implemented to test whether participants demonstrated significant processing costs within each condition. Response time analysis revealed processing costs in the Wall condition but not the Window condition. A 2 (Perspective: Self vs. Other) × 2 (Consistency: Consistent vs. Inconsistent) × 2 (Condition: Wall vs. Window) mixed-effect ANOVA revealed no main effect of Condition ($F(1, 30) = 1.187$, $p = .28$), but significant effects of Perspective ($F(1, 30) = 11.411$, $\eta^2 = .259$, $p = .002$) and Consistency ($F(1, 30) = 27.961$, $\eta^2 = .439$, $p < .001$). There was a significant Condition*Consistency interaction ($F(1, 30) = 5.702$, $\eta^2 = .089$, $p = .0234$) indicating greater response time delays for Inconsistent trials in the Wall condition, and a nonsignificant trend for a Condition*Perspective interaction ($F(1, 30) = 2.739$, $p = .108$), but no Perspective*Consistency ($F(1, 30) = 1.497$, $p = .231$) or three-way interaction ($F(1, 30) = 1.464$, $p = .236$). Paired t tests of critical trials revealed a significant increase in response times from Self-Consistent to Self-Inconsistent trials in the Wall condition ($t_{15} = 2.237$, $d = .559$, $p = .020$) and no significant difference in the Window condition ($t_{15} = 0.612$, $p = .275$). Between-subjects t tests revealed no significant differences between conditions in Self-Consistent ($t_{15} < .1$, $p = .99$) and Self-Inconsistent trials ($t_{15} < .1$, $p = .99$).

Identical models were then tested for increases in error rates across conditions. Error rates again also revealed effects in the Wall condition but not in the Window condition. A mixed-effects ANOVA revealed a main effect of Consistency ($F(1, 30) = 9.336$, $\eta^2 = .234$ $p = .005$), but no effect of Perspective ($F(1, 30) = 1.905$, $p = .178$), Condition ($F(1, 30) = 1.158$, $p = .29$), or any of their interactions (all $F(1, 30) < 1.606$). Paired comparisons of critical trials revealed a significant increase in error rate in Self

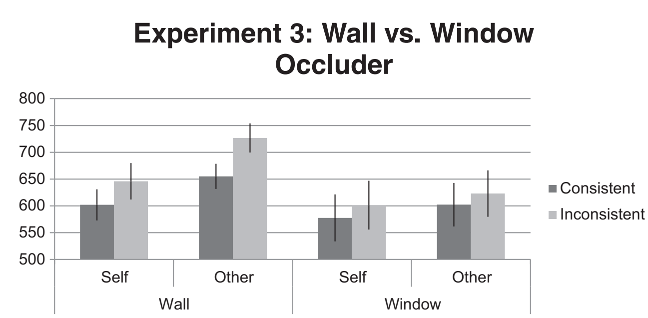

Figure 6. Processing costs across conditions for Self trials. "Inconsistent" trials in the Window condition are trials that would have been Inconsistent if avatars could not have seen through the window. Participants displayed processing costs for inconsistent trials in the wall condition but not the window condition. Error bars denote *SEM*.

Inconsistent versus Self Consistent trials in the Wall condition (mean difference = 5.7%; t_{15} = 2.576, d = .644, p = .011) but not in the Window condition (mean difference = 1.2%; t_{15} = 0.308, p = .38).

Difference scores. To directly compare the magnitude of the interference effect between conditions, we calculated the difference in response time and error rate between Consistent and Inconsistent trials from the Self perspective in each condition. Mean response times and error rates were calculated for each combination of Perspective*Consistency for each participant. Difference scores are the mean Self-Consistent values subtracted by the mean Self-Inconsistent values. If the self-judgment interference effect stems from calculation of another's line of sight, then we should see smaller difference scores in the Window condition than in the Wall condition. Consistent with this hypothesis, one-tailed tests revealed significant decreases in difference scores between Wall and Window conditions in response times (mean$_{\text{Wall}}$ = 44.17 ms difference, mean$_{\text{Window}}$ = 21.92 ms difference; $F(1, 30) = 2.961$, $\eta^2 = .089$, $p = .043$) and error rates (mean$_{\text{Wall}}$ = 5.71% error difference, mean$_{\text{Window}}$ = 1.22% error difference, $F(1, 30) = 5.314$, $\eta^2 = .150$, $p = .0142$).

Discussion

Experiment 3 replicated the perspective interference effect of Experiment 1 in the Wall condition while finding little to no effect in the Window condition. Analysis of difference scores between Self-Consistent and Self-Inconsistent perspectives for participants in each condition reveal significantly less interference in the Window condition. These results suggest that individuals encounter cognitive interference in the Wall condition but not the Window condition, implying that interference results from specific visual access of an agent to objects. This further supports the existence of a default system that calculates visual access, not just gaze direction or spatial grouping, by linking agents to targets by perceived line of sight.

It is worthwhile to note that small, nonsignificant increases in error rates and response times did appear in the Window condition's inconsistent trials, which may be because of the unusual nature of the task. For instance, responding what an avatar can see, even though one is aware that the avatar can see all targets, is a somewhat odd task, and participants anecdotally mentioned initial hesitance for fear they were being duped. This raises the possibility that on some proportion of trials, participants considered that the windowed occluder might block the avatar's view of the far targets; thus, slowing their responses slightly on what would have been inconsistent trials had the walls been complete.

These nonsignificant effects notwithstanding, we find no self-judgment interference in displays that perceptually segment targets from an avatar while also maintaining the avatar's line of sight to targets. We therefore conclude that participants monitored the presence of a different visual perspective by default by calculating line of sight access from agent to targets of the agent's attention.

General Discussion

Recent debate has questioned whether the effects documented in default perspective-taking designs are because of simple directional attentional cues. Here we provide critical evidence that self-judgment interference is elicited by another's conflicting visuospatial access calculated from their line of sight. This gaze specific cue operates selectively over targets that the avatar faces (Experiment 1) and is not limited to situations involving by the perceptual division of target objects (Experiment 2). Moreover, providing visual access from an avatar to targets eliminates interference (Experiment 3), suggesting that participants are interpreting relationships between agents and objects in relatively complex, three-dimensional, partially occluded displays. These results strongly suggest that self-judgment interference is because of visual perspective taking and not more basic perceptual processes, such as simple directional cueing.

Beyond the more concrete contrast between line of sight calculation and directional cueing, an emerging debate has questioned whether the self-judgment interference effect constitutes a modular system specific to theory of mind or a domain-general system of attention. This study finds that interference effects do not likely emerge from confounds of simple directional cuing or perceptual grouping. However, just because interference appears in the absence of these specific lower level cues, we do not know the mechanism governing attention to line of sight. For this reason we remain agnostic to the question of modularity. Whether this effect stems from a modular process specific to perspective-taking or from a domain-general attentional network that has been shaped by learning about agents to conform with specific patterns of eyeline-based spatial priming, our results provide important evidence in support of the hypothesis that default patterns of attention reflect a relatively sophisticated accounting of an agent's perceptual access to specific parts of a scene.

What these results do make clear is that the early system of perspective taking appears to reflect a relatively specific structuring of visual attention. At this point, we can conclude that individuals monitor the perspectives of other agents, likely through necessary attention to the direction of gaze and a calculation of the agents' line of sight relative to obstacles. This aligns with demonstrations chimpanzees can track the gaze of other agents and make causal inferences based on visuospatial access, but do not form intentional representations to do so (Povinelli, Eddy, Hobson, & Tomasello, 1996). Likewise, the self-judgment interference effect does not necessarily imply implicit theory-of-mind representation. Rather, this early system likely characterizes what Samson and Apperly (2010) call "belief-like states" that permit rapid attention to social events without necessarily forming costly intentional representations.

These results also demonstrate the ambiguity of the "agent" category to the perceptual system. Previous research has critiqued the plausibility of an early system, demonstrating self-judgment interference using arrows and cameras instead of human-like agents (Heyes, 2014; Santiesteban et al., 2014a, 2014b). We argue that arrows and cameras, like other objects with an identifiable front and back, may be easily interpreted as reflecting the goals of agents, or actually perceived as agents when those objects exhibit intentional motion (e.g., Gao, Newman, & Scholl, 2009; Luo & Baillargeon, 2005; Michotte, 1963) or when they can be conceptualized as having human qualities (Epley, Waytz, & Cacioppo, 2007). However, a fascinating starting point for future research may be the role of anthropomorphism in moderating the strength of interference effects. For instance, would participants demonstrate greater interference for basic shapes if those shapes had demonstrated intentional movement before trials began?

Moving forward, it would be particularly interesting to test how the default system interacts with other processes of scene and event perception. For instance, it is unknown whether self-judgment interference applies only to the focus of joint attention between self and others. In all tests of the default system thus far, participants are asked to enumerate objects that an avatar can or cannot see. This design requires that avatars only see objects that may be targets. In the real world, however, an agent may be looking at any number of objects that may or may not be the target of joint attention. Does the default system track any and all gaze cues within a scene, or is it specific to some form of task relevant gaze? When trying to grab the last slice of pizza, for example, we can imagine calculation of another's line of sight to the task-relevant food. However, would the default system necessarily calculate line of sight to nearby, task-irrelevant objects, like empty cans of soda? We can see that a simple task such as foraging for party food presents numerous problems for the default system for which we have no answers yet. Furthermore, although it is possible that the default system distinguishes perspective purely categorically (e.g., by determining whether an agent can or cannot see an object), it is also possible that perspective detection operates on a continuum, possibly moderated by the physical access of an agent to an object. If default perspective taking stems from a line-of-sight calculation, a different paradigm might capture processing costs as targets appear incrementally further from an avatar's reach.

In summary, we addressed the emerging debate over whether individuals efficiently calculate another's visual perspective or rather respond to simple directional cues embedded in the stimuli. We report evidence that perspective taking interference occurs when an agent's line of sight differs from one's own. This interference effect appears even while controlling for the directional cues for attention and perceptual grouping. We conclude that scene perception is directed by a gaze-specific cue, but that future research should test the scope and limitations of this system.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116,* 953–970. http://dx.doi.org/10.1037/a0016923

Baker, L. J., Hymel, A. M., & Levin, D. T. (in review). *Anthropomorphism and intentionality improve memory for events*.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: Bradford Books/MIT Press.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57,* 289–300.

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences, 8,* 255–260. http://dx.doi.org/10.1016/j.tics.2004.04.011

Blender Foundation. (2013). *Blender*. GNU General Public License v3.0. Retrieved from www.blender.org

Dally, J. M., Emery, N. J., & Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science, 312,* 1662–1665. http://dx.doi.org/10.1126/science.1126539

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87,* 327–339. http://dx.doi.org/10.1037/0022-3514.87.3.327

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40,* 760–768. http://dx.doi.org/10.1016/j.jesp.2004.02.002

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114,* 864–886. http://dx.doi.org/10.1037/0033-295X.114.4.864

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

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 319–329. http://dx.doi.org/10.1037/0096-1523.30.2.319

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology, 59,* 154–179. http://dx.doi.org/10.1016/j.cogpsych.2009.03.001

Guthrie, S. (1993). *Faces in the clouds*. New York, NY: Oxford University Press.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour, 61,* 139–151. http://dx.doi.org/10.1006/anbe.2000.1518

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57,* 243–259. http://dx.doi.org/10.2307/1416950

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

Klein, K. J. K., & Hodges, S. D. (2001). Gender differences, motivation and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin, 27,* 720–730. http://dx.doi.org/10.1177/0146167201276007

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? 30th European Conference on Visual Perception, Arezzo, Italy. [ECVP Abstract Supplement]. *Perception, 14*.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100,* 78–100. http://dx.doi.org/10.1037/0033-2909.100.1.78

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330,* 1830–1834. http://dx.doi.org/10.1126/science.1190792

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Sciences, 8,* 528–533. http://dx.doi.org/10.1016/j.tics.2004.10.001

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46,* 551–556. http://dx.doi.org/10.1016/j.jesp.2009.12.019

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16,* 601–608. http://dx.doi.org/10.1111/j.1467-9280.2005.01582.x

McCleery, J. P., Surtees, A. D. R., Graham, K. A., Richards, J. E., & Apperly, I. A. (2011). The neural and cognitive time course of theory of mind. *The Journal of Neuroscience, 31,* 12849–12854. http://dx.doi.org/10.1523/JNEUROSCI.1392-11.2011

Michotte, A. (1963). *The perception of causality*. Oxford, England: Basic Books.

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology, 24,* 603–613. http://dx.doi.org/10.1348/026151005X55370

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308,* 255–258. http://dx.doi.org/10.1126/science.1107621

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.

Povinelli, D. J., Eddy, T. J., Hobson, R. P., & Tomasello, M. (1996). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development, 61,* i–vi. http://dx.doi.org/10.2307/1166159

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,* 230–236. http://dx.doi.org/10.1016/j.cognition.2010.08.003

Ramsey, R., Hansen, P., Apperly, I., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes. *Journal of Cognitive Neuroscience, 25,* 670–684.

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,* 1255–1266. http://dx.doi.org/10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014a). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance, 40,* 929–937. http://dx.doi.org/10.1037/a0035175

Santiesteban, I., Shah, P., White, S., Bird, G., & Heyes, C. (2014b). Mentalizing or submentalizing in a communication task? Evidence from autism and a camera control. [Advance online publication]. *Psychonomic Bulletin & Review*.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4,* 102–106. http://dx.doi.org/10.3758/BF03210780

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,* 199–204. http://dx.doi.org/10.1111/j.1467-7687.2007.00580.x

Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *Journal of Experimental Child Psychology, 126,* 120–137. http://dx.doi.org/10.1016/j.jecp.2014.03.006

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18,* 580–586. http://dx.doi.org/10.1111/j.1467-9280.2007.01943.x

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development, 83,* 452–460.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102. http://dx.doi.org/10.1037/0033-295X.101.1.80

VMComix. (2011, July 04). *Cartoon character pack 1*. Public domain dedication under Creative Commons license 1.0 Universal. Retrieved from http://www.blendswap.com/blends/view/14445

Received March 9, 2015
Revision received September 14, 2015
Accepted September 19, 2015 ■