Psychon Bull Rev (2015) 22:844–849
DOI 10.3758/s13423-014-0716-0

BRIEF REPORT

# Mentalizing or submentalizing in a communication task? Evidence from autism and a camera control

**Idalmis Santiesteban · Punit Shah · Sarah White · Geoffrey Bird · Cecilia Heyes**

Published online: 23 August 2014


**Abstract** In the director task (DT), participants are instructed to move objects within a grid of shelves while ignoring those objects that cannot be seen by a human figure, the "director," located beyond the shelves. It is widely assumed that, since they are explicitly instructed to do, participants use mentalizing in this communicative task; they represent what the director can see, and therefore the DT provides important information about how and when mentalizing is used in adult life. We tested this view against a "submentalizing" hypothesis suggesting that DT performance depends on object-centered spatial coding, without mentalizing. As predicted by the submentalizing account, we found that DT performance was unchanged when the director was replaced by an inanimate object, a camera, and that participants with autism spectrum disorders were unimpaired, relative to matched control participants, in both the director and camera conditions. In combination with recent critical analyses of "implicit mentalizing," these findings support the view that adults use mentalizing sparingly in psychological experiments and in everyday life.

**Keywords** Perspective-taking · Mentalizing · Submentalizing · Object-centered spatial coding · Director task · Autism

**Electronic supplementary material** The online version of this article (doi:10.3758/s13423-014-0716-0) contains supplementary material, which is available to authorized users.

I. Santiesteban (✉) · P. Shah
Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK
e-mail: i.santiesteban@bbk.ac.uk

P. Shah · G. Bird
MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, De Crespigny Park, London SE5 8AF, UK

S. White · G. Bird
Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

C. Heyes
All Souls College, University of Oxford, Oxford OX1 4AL, UK

The study of mentalizing has reached maturity. After decades in which studies of children and animals dominated the field, research on mentalizing (also known as "theory of mind" and "mindreading") is now investigating how and when mature adult humans use representations of what others see, think, know, and intend in order to navigate the social world. One school of thought suggests that mentalizing is pervasive; we use it constantly to predict what others are going to do in nonverbal social interactions (e.g., receiving a pass on the football field, avoiding a collision on the dance floor) and to decode utterances in communicative contexts (e.g., When he says "Give me the spanner," is he referring to the large spanner or the small one?). Another school suggests that mentalizing is used sparingly: Routine prediction and decoding are done by domain-general psychological processes operating on representations of observable features of the social environment, whereas the specialized and cognitively demanding processes that represent mental states are saved for the explanation and justification of behavior—for instance, to negotiate a deal or excuse an error (Heyes & Frith, 2014).

Supporting the view that mentalizing is pervasive, a number of recent studies have suggested that adults often represent what others can see, know, and intend "automatically"—they do it when mentalizing is not required by the task, and even when mentalizing is detrimental to performance. For example, in the dot perspective task, judgments about the number of dots on a computer screen are impaired when a human figure on the screen is facing a subset of the dots (e.g., Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). This suggests that I represent what you see even when doing so interferes with my judgments about what I can see.

However, consistent with the view that mentalizing is used sparingly, other studies have provided evidence that purported examples of automatic mentalizing are due to domain-general processing rather than to mentalizing. In the dot perspective task, number judgments are impaired just as much by an arrow as by a human figure facing a subset of the dots, indicating that the effect may be due to automatic attentional orienting rather than automatic mentalizing (Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2014). Experiments of this kind suggest that many behaviors that appear to be guided by mentalizing could instead be based on "submentalizing"—on domain-general cognitive processes that simulate the effects of mentalizing in social contexts (Heyes, 2014).

We attempted to distinguish mentalizing from submentalizing using the director task (DT; Keysar, Barr, Balin, & Brauner, 2000). This task is widely assumed to assess perspective-taking, the capacity to represent what another person can see (e.g., Apperly et al., 2010; Dumontheil, Apperly, & Blakemore, 2010; Keysar, Lin, & Barr, 2003). The DT is playing an increasingly important role in the debate about adult mentalizing because, unlike most other tasks that are suitable for adults, it aims to assess mentalizing in a communicative context. For example, data from the DT have been presented as evidence that mentalizing skills are still developing in late adolescence (Dumontheil, Apperly, & Blakemore, 2010) and that people from interdependent cultures are better at adopting the perspective of their interacting partners than are those from independent cultures (Wu & Keysar, 2007).

In each trial of a typical DT, the participant is presented with an array of objects located in a grid of shelves, and a human figure—"the director"—is standing beyond the shelves (see Fig. 1a). Although the participant has a full view of the objects on the shelves, occluders prevent the director from seeing some of the objects. In all trials, the participant is instructed to move an object (e.g., the "small cube"). In the experimental trials, the array contains three objects of the type specified in the instruction (e.g., a large, a medium-sized, and a small cube), only two of which are visible to the director (e.g., the large and medium-sized cubes), and the participant is required to discount the "conflicting object"—that is, the object that the director cannot see (e.g., to ignore the smallest cube and to move the medium-sized cube, the smaller of the two cubes that are visible to the director). In control trials, there is no conflicting object. Experiments using the DT typically find "egocentric bias": Responding is accurate (participants select the correct object) less often in experimental than in control trials.

It is commonly assumed that accurate performance in the experimental trials of the DT requires participants to represent what the director can see and, via controlled processing, to isolate this representation from the content of the participant's own visual representation of the scene (e.g., Dumontheil, Küster, Apperly, & Blakemore, 2010). We tested this mentalizing interpretation of DT performance against an

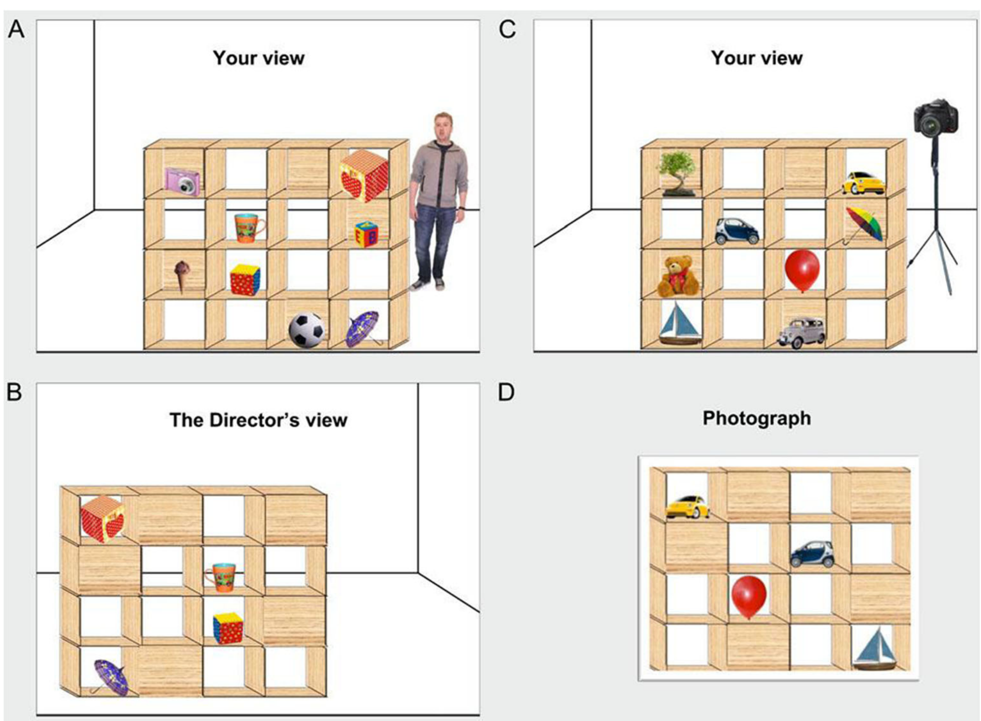


**Fig. 1** Example of the stimuli used in the director (**a & b**) and camera (**c & d**) conditions. The bottom panels represent images shown to the participants during the instructions

alternative, *submentalizing* interpretation. The submentalizing hypothesis also assumes that errors in the experimental trials are caused by failure to isolate a critical representation from the content of the participant's own, egocentric visual representation of the scene. However, the submentalizing interpretation suggests that the critical representation does not typically have mental content. It is a spatial but not a specifically visual representation (Surtees, Apperly, & Samson, 2013)—a representation of what is in front of the director, rather than a representation of what the director can see. In other words, the submentalizing hypothesis suggests that performance in the DT depends on the use of an object-centered spatial coding, in which the *object* is the director.

We tested the submentalizing against the mentalizing hypothesis in two ways. First, we introduced an inanimate control condition in which a camera took the place of the director. Cameras have directional but not agentive properties. In everyday life, we often use object-centered spatial coding to make judgments about a camera—about what is in front of the camera and whether there is an unobstructed straight line between a camera lens and an object. However, cameras are not appropriate objects for the attribution of mental states. Therefore, the submentalizing hypothesis, which suggests that the DT tests object-centered spatial coding but not the representation of seeing, predicts that performance in the camera condition would be the same as in the director condition. In contrast, the mentalizing hypothesis suggests that, as compared with the camera condition, the director condition involves an additional processing step; the participant must not only represent the locations of objects relative to the director, but also represent the fact that the objects in front of the director are the objects that the director can see. Therefore, if the typically observed egocentric bias is due to a specific failure of mentalizing, one would expect more accurate performance in the camera than in the director condition.

Second, we gave the director and camera versions of the DT to individuals diagnosed with an autism spectrum disorder (ASD), as well as to typically developing adult participants. It is widely believed that people with ASD are impaired in their capacity to represent the mental states of others (for a review, see Frith, 2012). Therefore, if the standard version but not the camera version of the DT involves the representation of mental states (mentalizing hypothesis), one might expect people with ASD to be more impaired, relative to neurotypical controls, in the director condition than in the camera condition. In contrast, if neither the director nor the camera version of the DT involves the representation of what others see (submentalizing hypothesis), one would expect no impairment in the ASD group, or any impairment in the ASD group to be comparable across the director and camera conditions.

## Method

### Participants

Due to the difficulties inherent in recruiting individuals with a low-prevalence condition such as ASD, we used an opportunity sampling method. Effect size measures are therefore included in the Results section, to avoid reliance on null hypothesis significance testing. The ASD group therefore consisted of 20 adults (18 males, two females; mean age: 36 years, $SD = 10.8$; mean FSIQ: 110, $SD = 17$, Wechsler Adult Intelligence Scale [Wechsler, 1997]). Eighteen neurotypical adults (16 males, two females; mean age: 41, $SD = 14.7$; mean FSIQ: 108, $SD = 9$) comprised the control group. The groups were matched for age [$t(36) = 1.15$, $p = .25$)] and IQ [$t(36) = 0.46$, $p = .65$]. All of the ASD participants met the criteria for ASD ($N = 10$) or for autism ($N = 10$) on the Autism Diagnostic Observation Schedule (ADOS-G; Lord et al., 2000). All participants completed the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), on which the ASD group obtained a significantly higher score than the control group (ASD, $M = 34$, $SD = 8.9$; controls, $M = 15.8$, $SD = 6.4$) [$t(34) = 6.96$, $p < .001$, $d = 2.39$]. The mentalizing abilities of 15 participants from the ASD group had previously been tested using the strange stories task (Happé, 1994) and compared to a matched typical control group. Intact mentalizing is necessary for accurate performance on the strange stories task, and the ASD group performed significantly worse on this measure than did the control group (maximum score 16; ASD group, $M = 8.8$, $SEM = 1.06$; control group, $M = 13.25$, $SEM = 0.34$) [$t(53) = 5.20$, $p < .001$], confirming the mentalizing impairment in this group.

### Materials and procedures

All participants (control and ASD) performed a computerized version of the DT under both conditions (director and camera).

In both conditions, the visual stimuli consisted of a 4 × 4 grid ("shelves") containing eight different objects. Five slots were occluded from the view of the director/camera lens, located on the other side of the shelves. Participants listened to auditory instructions from the director, or saw the instructions in text format in the camera condition, asking them to move specified objects in a particular direction. We created three different trial types: one experimental—described above—and two control (C1, C2). In C1, the instructions specified a type-unique object that the director could see (e.g., the mug in Fig. 1a). In C2, the instructions were matched with those of the experimental trials, but the competitor object was replaced by an irrelevant object (e.g., the smallest cube by a key). Prior to performing the task, participants were shown an example of what the shelves looked like from the director's perspective (Fig. 1b) or the camera's position (Fig. 1d). They were also asked to select objects that the director could and could not see (or that would or would not appear in the photograph), to ensure that they understood the task requirements before the experimental session started.

The director and camera conditions were completed in separate blocks, with the block order counterbalanced across participants. In the director condition, participants were asked to take into account the perspective of the director, not only when choosing the object, but also when moving it to the correct location (e.g., if the director asked the participant to move an object left, the participant was required to take the director's perspective and therefore move the object to the right from their own perspective). In order to equate the spatial task demands, participants were required to perform the same left/right switch for the camera condition (e.g., if the instruction was to move an object left, the participant was required to move the object to the left as it would appear in the photograph, and thus to move the object to the right from their own perspective). The arrangement of objects was counterbalanced across both conditions and participants.

A short practice session was followed by three blocks of 36 mixed (experimental, C1, and C2) trials presented in pseudorandom order. Accuracy of the selection and movement of the target object was recorded. Eye movement data were also recorded using an Eyegaze Edge System eyetracker (sampling rate: 60 Hz). The eyetracking measure consisted of the number of 100-ms fixations on both the competitor object, in the experimental condition, and the irrelevant object placed in the same slot, in the C2 condition.

## Results

### Accuracy

The accuracy scores were coded: 1 = *correct selection and movement of the target object*, .5 = *correct selection but incorrect movement*, 0 = *incorrect selection*. The proportions of correct responses were analyzed using a 2 × 2 × 3 analysis of variance with Group (ASD vs. control) as a between-subjects factor and Condition (director vs. camera) and Trial Type (experimental vs. C1 vs. C2) as within-subjects factors. Where sphericity assumptions were not met, Greenhouse–Geisser-corrected values are reported. Bonferroni corrections were used for post hoc multiple comparisons.

The accuracy data are presented in Fig. 2. A main effect of trial type [$F(1.01, 36.51) = 45.10$, $p < .001$, $\eta^2_p = .56$] indicated that participants showed the egocentric bias typically observed in the DT, with less accurate performance in experimental ($M = .53$, $SEM = .06$) than in control trials: C1 ($M = .94$, $SEM = .02$), C2 ($M = .92$, $SEM = .02$). As is

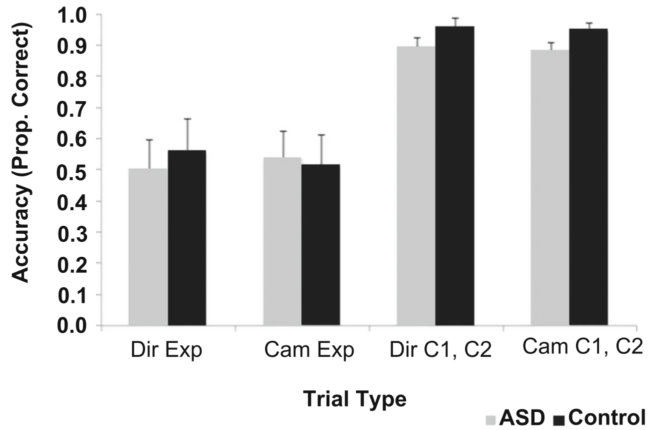


**Fig. 2** Accuracy data from the director (Dir) and camera (Cam) versions of the director task (error bars represent *SEM*s). Performance on control trials (C1 and C2) did not differ; therefore, these results are collapsed for easier graphical representation. ASD, autism spectrum disorder

predicted by the submentalizing hypothesis, the main effect of condition [$F(1.01, 36.51) = 0.034$ , $p = .86$, $\eta^2_p = .001$] and the Group × Condition interaction [$F(1.01, 36.51) = 1.75$, $p = .20$, $\eta^2_p = .046$] were not significant. Performance in the director condition ($M = .79$, $SEM = .03$) was almost identical to performance in the camera condition ($M = .80$, $SEM = .03$). The main effect of group [$F(1, 36) < 1$, $\eta^2_p = .010$] and the Group × Trial Type interaction [$F(2, 72) < 1$, $\eta^2_p = .001$] also failed to reach significance. The performance of the ASD group during experimental trials for both conditions ($M = .52$, $SEM = .09$) was not impaired relative to that of controls ($M = .54$, $SEM = .09$), [$t(36) = 0.12$, $p = .91$, $d = 0.04$]. The analysis of accuracy in object selection alone (ignoring movement selection) revealed the same pattern of results. The eyetracking data (presented in the supplementary materials) revealed no differences between the groups or conditions.

## Discussion

Consistent with the predictions of the submentalizing hypothesis, we found (1) that the egocentric bias typically observed in the DT is equally strong when an inanimate object, a camera, takes the place of the director, and (2) no evidence that the performance of people with ASD is impaired in either the camera or the director version of the DT. Neither of these findings provides unequivocal evidence that performance in the DT involves object-centered spatial coding without mentalizing. However, in combination they indicate that, without new evidence that DT performance depends on representation of what the director can see, this task cannot be assumed to assess mentalizing or to support the view that mentalizing plays a pervasive role in communicative contexts.

In defense of the mentalizing hypothesis, it could be argued that participants did not respond more accurately in the camera condition than in the director condition because—although they knew that cameras do not have mental states, and although it was detrimental to their performance—the participants not only used object-centered spatial coding, but also represented what both the director and the camera could see. This kind of "overmentalizing" is possible in principle, but an overmentalizing interpretation would need independent empirical support to make it more than an ad hoc hypothesis. Support of this kind could be sought by replacing the director with objects to which, on the basis of independent evidence, participants are judged likely or unlikely to attribute mental states. The former category might include inanimate objects, but if it were unconstrained, there would be a risk that mentalizing hypotheses would become untestable.

It could also be argued that our failure to find a difference between the ASD and control groups does not undermine the mentalizing hypothesis, because DT performance depends on explicit mentalizing—on instructed or deliberate representation of mental states—whereas people with ASD have an impairment in implicit mentalizing—in the spontaneous or automatic representation of mental states (Senju, Southgate, White, & Frith, 2009; see Schwarzkopf, Schilbach, Vogeley, & Timmermans, 2014, for conflicting findings). However, there are three obstacles to this view. First, some researchers have argued that the DT assesses implicit mentalizing (e.g., Dumontheil, Küster, et al., 2010). Second, the evidence that implicit mentalizing is impaired in people with ASD has come from tasks that are, like the DT, subject to a submentalizing interpretation (Heyes, 2014). Third, Begeer, Malle, Nieuwland, and Keysar (2010) found that participants with ASD showed no impairment in the standard version of the DT, but used fewer mental-state terms than did controls when asked to retell a story in a traditional, narrative test of explicit mentalizing. In combination with our finding that performance is no more accurate in the camera than in the director version of the DT, this dissociation suggests that DT performance does not depend on explicit mentalizing.

Our suggestion that DT performance depends on object-centered spatial coding but not on mentalizing is consistent with findings that were previously thought to fortify the mentalizing hypothesis. For example, Dumontheil, Apperly, and Blakemore (2010) found that responding was less accurate in experimental trials with the director than in control trials in which the director was absent and participants were instructed to ignore objects that appeared against a particular background color. They suggested that performance was superior in their control trials because participants could use a simple rule rather than represent what the director could see, but it is at least equally likely that performance was more accurate in the control trials because they did not require object-centered spatial coding.

The submentalizing interpretation of the DT is also consistent with the results of a previous study using a camera control in a different but related task. Aichhorn, Perner, Kronbichler, Staffen, and Ladurner (2006) found the same effects on reaction times, as well as patterns of activity in the posterior superior temporal sulcus and the temporoparietal junction, when participants were asked to judge the arrangement of two objects from the perspective of a doll and a camera.

In this study, we have focused on object-centered spatial coding, but this is just one example of a class of submentalizing processes that may contribute to variation in performance on the DT. Another possibility, mentioned above, is that participants ignore the objects with a colored background and are less successful at doing so when distracted by the presence of the director. Instead, or in addition, participants may ignore objects that can be characterized as extremes when three exemplars are presented (e.g., the largest, topmost, etc.). All of these strategies are plausible, none involve mentalizing, and all are consistent with previous reports of medial prefrontal cortex (mPFC) activation during performance of the DT (Dumontheil, Küster, et al., 2010). Activity in the mPFC has been associated not only with mentalizing, but with conflict monitoring, error detection, error prediction, outcome evaluation, and uncertainty during decision making (for a review, see Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

The view that mentalizing is a pervasive feature of everyday life has been challenged both by researchers who embrace a standard cognitive-science model of social understanding (e.g., Heyes, 2014) and by researchers who seek to replace this model with a "second-person" approach (Schilbach et al., 2013). Previous evidence in support of these challenges—showing that there are contexts in which mentalizing could be used but is not—has come from tasks such as the dot perspective task, which do not involve communication and in which participants are not asked to judge or consider the mental states of others (e.g., Dolk et al., 2011; Santiesteban et al., 2014). In contrast, in the DT participants are asked explicitly to interpret the director's instructions as communicative acts, and to take into account what the director can see and what he intends. Our results suggest that, even in this context, participants eschew mentalizing in favor of submentalizing processes, and our study therefore provides strong support for the view that, even in maturity, mentalizing is used sparingly in psychological experiments and in everyday life.

**Author note** The authors thank Iroise Dumontheil for providing helpful feedback on an earlier version of the manuscript. This work was supported by an Economic and Social Research Council studentship [ES/H013504/1] awarded to I.S.

## References

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage, 30,* 1059–1068. doi:10.1016/j.neuroimage.2005.10.026

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology, 63,* 1201–1217. doi:10.1080/17470210903281582

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5–17. doi:10.1023/A:1005653411471

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology, 7,* 104–122. doi:10.1080/17405620903024263

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbacck, S., Prinz, W., & Liepelt, R. (2011). How "social" is the social Simon effect? *Frontiers in Psychology, 2,* 84. doi:10.3389/fpsyg.2011.00084

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13,* 331–338. doi:10.1111/j.1467-7687.2009.00888.x

Dumontheil, I., Küster, O., Apperly, I. A., & Blakemore, S.-J. (2010). Taking perspective into account in a communicative task. *NeuroImage, 52,* 1574–1583. doi:10.1016/j.neuroimage.2010.05.056

Frith, U. (2012). Why we need cognitive explanations of autism. *Quarterly Journal of Experimental Psychology, 65,* 2073–2092. doi:10.1080/17470218.2012.697178

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24,* 129–154. doi:10.1007/BF02172093

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

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science, 344,* 1243091. doi:10.1126/science.1243091

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11,* 32–38. doi:10.1111/1467-9280.00211

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

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205–223. doi:10.1023/A:1005592401947

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443–447. doi:10.1126/science.1100301

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. doi:10.1037/a0018729

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

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36,* 393–414. doi:10.1017/S0140525X12000660

Schwarzkopf, S., Schilbach, L., Vogeley, K., & Timmermans, B. (2014). Making it explicit makes a difference: Evidence for a dissociation of spontaneous and intentional level 1 perspective taking in high-functioning autism. *Cognition, 131,* 345–354. doi:10.1016/j.cognition.2014.02.003

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science, 325,* 883–885. doi:10.1126/science.1176170

Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition, 129,* 426–438. doi:10.1016/j.cognition.2013.06.00

Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: Psychological Corp.

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science, 18,* 600–606. doi:10.1111/j.1467-9280.2007.01946.x