Automatic Perspective-Taking: The Implicit Mentalizing vs. Submentalizing Controversy

Authors: Li Yi, Xiao Feng, Xiao Feng

Date: 2021-05-02T00:00:00+00:00

Abstract

The phenomenon of automatic perspective-taking has been confirmed by numerous studies, yet its underlying mechanism remains controversial. Currently, there exist two viewpoints: implicit mentalizing and submentalizing: the former posits that automatic perspective-taking is a domain-specific process that spontaneously adopts others' viewpoints; whereas the latter proposes that automatic perspective-taking is essentially a domain-general process involving reflexive attentional orienting and spatial encoding of locations, which simulates the function of mentalizing in social contexts. This study proposes a collaborative model of implicit mentalizing and submentalizing, wherein the two can operate independently or jointly. Future research should employ advanced technological methods to investigate diverse participant populations and explore the mechanism of automatic perspective-taking.

Full Text

Preamble

Automatic Perspective Taking: The Debate between Implicit Mentalizing and Submentalizing

Yi LI & Feng XIAO*
(Department of Education Science, Shanxi Normal University, Linfen 041000)

Abstract: The phenomenon of automatic perspective taking has been confirmed by numerous studies, yet its underlying mechanism remains controversial. Two competing views currently exist: the implicit mentalizing view holds that automatic perspective taking reflects a domain-specific process involving the spontaneous adoption of others' perspectives, whereas the submentalizing view proposes that it stems from domain-general processes such as reflexive attention orientation and spatial coding of locations that simulate the function of mentalizing in social contexts. This study proposes a collaborative model in which implicit mentalizing and submentalizing can operate independently or in concert. Future research should employ advanced techniques to investigate diverse populations and explore the mechanisms underlying automatic perspective taking.

Keywords: automatic perspective taking; implicit mentalizing; submentalizing

In daily life, people interact closely with one another, and understanding others' perspectives is fundamental to high-quality social interaction. During these interactions, one's own viewpoint does not always align with others' perspectives and may sometimes be diametrically opposed. For instance, when you want to draw your friend' s attention to an eyelash on her cheek, you must realize that her right side corresponds to your left, and vice versa. Despite often lacking sufficient time to consciously consider others' viewpoints, we can interact quickly and successfully. Evidence from indirect measures suggests that both adults and children may implicitly acquire others' perspectives [1,2], a phenomenon known as automatic perspective taking. Automatic perspective taking refers to situations where, despite another person' s perspective being irrelevant to the task at hand, individuals are nonetheless influenced by that perspective simply due to the person' s presence.

Currently, debate surrounding automatic perspective taking centers on whether it results from implicit mentalizing or submentalizing processes [3-6]. Implicit mentalizing represents a specialized cognitive process—a highly specific psychological mechanism shaped by natural selection during human evolution, characterized by speed and automaticity. Apperly and Butterfill (2009) proposed that adult theory of mind operates through two systems: an early-developing, cognitively efficient "low-level" automatic process, and a later-developing, flexible "high-level" selective process that operates more slowly and is associated with executive function and language abilities [7]. The first system supports the implicit mentalizing view, while the second provides the basis for explicit theory of mind. Implicit mentalizing involves understanding others' mental states, upon which humans gradually develop a more mature and flexible theory of mind [8].

The submentalizing view suggests that automatic perspective taking may be guided by less complex cognitive processes, such as orientation mechanisms that help one know what appears in another' s line of sight [9]. This process does not involve thinking about mental states but simulates the role of mentalizing in social environments. This alternative to perspective taking helps individuals grasp critical information and plays an important role in understanding others' mental states. These two views—implicit mentalizing and submentalizing—stand in opposition yet both can reasonably explain experimental results on automatic perspective taking.

Below, we review empirical research from behavioral and cognitive neuroscience examining the mechanisms underlying automatic perspective taking, and conclude with future research directions.

2.1 Dot Perspective Task and Attention Orientation

The dot perspective task is commonly used to test automatic perspective taking [10]. This task presents an image of a room (as shown in Figure 1 [Figure 1: see original paper]) with an avatar standing in the center, facing one wall. Zero to three dots appear on the two side walls. Participants must judge whether the number of dots they see matches a cued number, either from their own perspective or from the avatar' s perspective. Because the avatar' s orientation allows it to see only the dots on the wall it faces, while participants can see dots on both walls, the self-perspective and avatar-perspective yield identical counts when all dots appear on the avatar-facing side (consistent condition). When dots appear on the side behind the avatar, the two perspectives yield different counts (inconsistent condition). To avoid gender effects, male participants see male avatars and female participants see female avatars. Results show slower response times and higher error rates in inconsistent versus consistent conditions, demonstrating the consistency effect of automatic perspective taking.

The consistency effect can be further divided into: (1) egocentric intrusion effect, where judging from the avatar' s perspective is interfered with by one' s own perspective, resulting in slower responses and more errors in inconsistent versus consistent conditions (i.e., a consistency effect under others' perspective); and (2) altercentric intrusion effect, where judging from one' s own perspective is interfered with by the avatar' s perspective, resulting in slower responses and more errors in inconsistent versus consistent conditions (i.e., a consistency effect under self-perspective) [5,10-12]. The present review focuses primarily on the altercentric intrusion effect.

The implicit mentalizing view explains the altercentric intrusion effect as follows: Apperly proposed that theory of mind operates through two systems—the first being automatic processing and the second being controlled processing requiring executive function. Since the consistency effect occurs without explicit consideration of the avatar' s perspective and emerges within extremely short response times, it represents an automatic (implicit) rather than controlled (explicit) process. This implicit or automatic process indicates that participants automatically process the avatar's perspective, which then interferes with self-perspective judgments in inconsistent conditions, producing the consistency effect. Furlanetto et al. (2016) manipulated whether avatars wore transparent versus opaque glasses to create visible and invisible conditions, finding a consistency effect only in the visible condition, thereby supporting the implicit mentalizing view [13].

The submentalizing view argues that the consistency effect results from the avatar' s directional cue automatically orienting attention to the indicated side rather than from processing the avatar' s perspective [14]. Specifically, the avatar's frontal features (forehead, eyes, nose) automatically trigger participants' attention to shift toward the avatar-facing side [15], enhancing processing of dots in that direction. In inconsistent conditions, conflict arises between the number of dots in the avatar' s view and the total number of dots, requiring participants to resolve this conflict before responding, thus producing slower responses and higher error rates.

Santiesteban et al. compared dot perspective tasks using avatars versus arrows as central stimuli, finding similar results for both conditions [16]. Since arrows possess only directionality without mental state content, this provides support for the submentalizing view.

Figure 1. Dot perspective task experimental materials. Left panel shows consistent condition for female participants; right panel shows inconsistent condition for male participants (adapted from Samson et al., 2010).

2.2 Social Simon Task and Spatial Coding

Researchers also frequently use the social Simon task to investigate automatic perspective taking mechanisms. The Simon effect can be explained by spatial coding. In the standard Simon task [17], participants respond to non-spatial stimulus attributes (e.g., press left key for red stimulus, right key for green stimulus). Despite stimulus location being task-irrelevant, results show a spatial consistency effect: responses are faster when stimulus and response are on the same side (e.g., green stimulus on right, right key press). In the go/nogo Simon task [18], this spatial consistency effect disappears. For example, when participants press the right key for green stimuli and make no response for red stimuli, response speed is identical regardless of whether green stimuli appear on the left or right. However, if the social Simon task requires two people to jointly perform a go/nogo Simon task—with a confederate sitting beside the participant, the participant responding to green stimuli with the right key and the confederate responding to red stimuli with the left key—the spatial consistency effect reappears, such that participants respond faster when green stimuli appear on the right [19].

The implicit mentalizing view holds that the social Simon task represents an automatic process. Although participants respond under time pressure without needing to consider the confederate' s mental state, they nonetheless spontaneously consider the confederate' s perspective. Thus, two people performing a go/nogo Simon task effectively constitutes a standard Simon task, producing the spatial consistency effect. Freundlieb et al. [20-22] required participants and partners to complete adapted Simon tasks (as shown in Figure 2 [Figure 2: see original paper]), confirming that humans spontaneously adopt another person' s perspective. Comparing visible versus invisible confederate conditions, they found consistency effects only when the confederate was visible, demonstrating that opaque glasses prevent spontaneous perspective adoption. Boffel and colleagues [23-25] used the same paradigm, first confirming that humans spontaneously adopt avatar perspectives, then manipulating agency between participant and avatar. They found that higher agency conditions (where avatar responses corresponded to participants) produced stronger perspective-taking effects than low agency conditions (where avatars responded randomly). Similarly, Salm-Hoogstraaten et al. (2020) used a Simon task to manipulate interaction with robots, finding slower response times and higher error rates in conditions with remote-controlled robots versus no robot interaction [26]. Finally, Kuhbandner et al. (2010) found that social Simon task performance is modulated by theory-of-mind factors, with participants in positive moods showing stronger spatial consistency effects than those in negative moods, providing further evidence for implicit mentalizing [27]. In summary, because social cues modulate social Simon task performance, this task can be considered a paradigm for automatic perspective taking.

Heyes (2014) argued that the social Simon effect can also be explained by submentalizing: participants actually consider the confederate' s spatial features rather than mental states, with these spatial features modulating the spatial consistency effect [9]. Specifically, a confederate sitting to the participant' s left is just one type of stimulus; any stimulus—living or non-living—placed to the left can be automatically coded, leading participants to perceive themselves as not merely making a key press but making a "right key press" response [28]. This produces the spatial consistency effect in social Simon tasks. Research shows that spatial consistency effects disappear when confederates sit farther from participants [29], supporting the submentalizing view.

Figure 2. Social Simon task (adapted from Freundlieb et al., 2015).

2.3 Ambiguous Numbers and Spatial Coding

Surtees et al. used an ambiguous numbers task to demonstrate that adults rapidly adopt others' visual perspectives [30]. The experiment required participants to judge whether numbers (5, 6, 8, 9) on a table were greater or less than 7, either alone (solo condition) or sitting opposite another person (joint condition) (as shown in Figure 3 [Figure 3: see original paper]). In the joint condition, numbers 5 and 8 appeared identical to both individuals (consistent condition), while numbers 6 and 9 appeared differently (inconsistent condition). Results showed slower responses in inconsistent versus consistent conditions during joint performance, but no difference between conditions in the solo condition.

The implicit mentalizing view explains that in inconsistent conditions, participants automatically process the avatar' s perspective, creating conflict between self- and other-perspectives that must be resolved before responding, thus slowing reaction times. Elekes et al. used the ambiguous numbers task with added confederate task cues in the joint condition, informing participants that the confederate' s task matched their own (perspective-dependent condition). This manipulation produced consistency effects, supporting the implicit mentalizing view [31].

The submentalizing view attributes consistency effects to spatial coding, with inconsistent conditions creating conflict between egocentric and object-centered reference frames. In Millett et al.' s study, participants viewed two types of images: one showing an avatar facing an ambiguous number (appearing as 6 from the avatar' s view and 9 from the participant' s view), and another showing a chair with a computer screen facing it instead of an avatar. Participants completed the same task in both conditions. Results showed no significant difference in the probability of responding "6" between conditions, suggesting that participants processed the stimuli through spatial coding, supporting the submentalizing view [32].

Figure 3. Ambiguous numbers task (adapted from Surtees et al., 2016).

2.4 Anticipatory Looking Paradigm and Distraction

The classic false belief task measures explicit theory of mind through verbal reports about object locations from different perspectives, with results showing that children only pass false belief tasks around age four [33]. However, researchers using non-verbal anticipatory looking paradigms have demonstrated that infants and primates can also pass false belief tasks, suggesting these paradigms measure implicit theory of mind [34-36]. Schneider et al. used a Sally-Anne variant task with eye-tracking technology to measure looking time toward anticipated locations, demonstrating implicit mentalizing processes [2]. In the video (as shown in Figure 4 [Figure 4: see original paper]), an avatar could see two opaque boxes on a table, one containing a red ball. Participants then viewed two scenarios: (1) false belief scenario: after the avatar left, the ball was moved to another location, creating a false belief (the avatar thinks the ball is at location A when it is actually at location B); (2) true belief scenario: the ball' s final location matched the avatar' s belief. The avatar then returned, and participants' eye movements were recorded. Results showed that typical adult participants first looked at and spent more time fixating the location where the avatar last saw the ball. Schneider et al. tested autistic participants with the same procedure and found no anticipatory looking effect, possibly due to impaired implicit mentalizing in autism [37].

The implicit mentalizing view holds that typical participants first and longest fixate the location where the avatar last saw the ball because they automatically adopt the avatar' s belief about the ball' s location. Autistic participants, with impaired implicit mentalizing abilities, cannot understand the avatar' s false belief about the ball' s location, so their judgments are unaffected by others' beliefs [36-38], and they show no anticipatory looking effect. Kano and colleagues showed apes videos from anticipatory looking paradigms, finding that three ape species (bonobos, chimpanzees, orangutans) all showed anticipatory looking effects. When the avatar in the video was replaced with geometric shapes, apes showed anticipatory looking effects only in the avatar condition, not the geometric shape condition. Manipulating the opacity of obstacles in front of the avatar revealed no anticipatory looking effect with opaque obstacles but an effect with transparent obstacles, supporting the mentalizing view [34,35,39-41].

The submentalizing view argues that typical adults and infants are easily distracted by head orientation, fixating on the avatar' s head and gaze region, which reduces attention to and memory for the ball' s movement. Consequently, participants do not expect the ball to appear where the avatar believes it to be. Moreover, autistic participants show more focused attention on objects and are less distracted by the avatar' s departure [42], resulting in weaker anticipation of the ball' s location from the avatar' s perspective [9]. Burnside et al. used eye-tracking to record 16-month-old infants' gaze patterns in an anticipatory looking paradigm, finding that even inanimate agents (a toy crane) elicited anticipatory looking effects—infants looked longer at the location where the crane had falsely believed the ball to be—supporting the submentalizing view [43].

Figure 4. False belief task, eye-tracking recorded only during final 60-70 seconds (based on Schneider et al., 2012).

In summary, all four paradigms have been explained using both implicit mentalizing and submentalizing views. However, since Samson et al.' s research, studies have shown that the dot perspective task can demonstrate automatic perspective taking effects through low-level visual forms alone, involving fewer cognitive factors and making it easier to exclude other interfering factors [10]. Consequently, the dot perspective task has become the classic paradigm for studying automatic perspective taking, with many researchers adapting it to distinguish between implicit mentalizing and submentalizing influences.

3 Factors in the Dot Perspective Task

The debate between implicit mentalizing and submentalizing remains unresolved, partly due to inconsistent methodologies across studies. Despite using essentially the same experimental task (the dot perspective task), various factors in experimental design—such as the social relevance of cues, visual attribution status of cues, and acquisition of social perspective—make it difficult to directly differentiate between implicit mentalizing and submentalizing. We elaborate on these three aspects below.

3.1 Social Relevance of Cues

Social relevance of cues refers to how the consistency effect is influenced by the degree of social characteristics in cues, such that others' perspectives interfere with self-perspective judgments. For example, Nielsen et al. required participants to complete dot perspective tasks under three cue conditions varying in social relevance—avatar (social), arrow (semi-social), and two-color block (non-social)—while measuring explicit perspective taking with the Interpersonal Reactivity Index (IRI) [44]. Results showed altercentric intrusion effects in all three conditions, but the interference effect was stronger in the avatar condition than in the other conditions, and explicit perspective-taking scores correlated only with the altercentric intrusion effect in the avatar condition. Thus, the degree of interference depends on social relevance, with stronger social relevance producing stronger interference effects.

However, Wilson et al. argued that although Nielsen et al. compared altercentric intrusion effects across different social relevance cue conditions, their analysis combined arrows and two-color blocks before comparing them to the avatar condition, meaning they did not directly compare arrows and avatars. Wilson et al. addressed these limitations by: (1) using horizontal arrows instead of the human-shaped arrows in Santiesteban et al. [16] to reduce interference from "human-like" social properties; and (2) using a directional camera instead of two-color blocks that required learning directionality, creating avatar, camera, and arrow cue conditions. Results showed altercentric intrusion effects in all three conditions with no significant differences among them, suggesting that social relevance is not a necessary condition for producing altercentric interference effects [45].

Furthermore, Cole et al. proposed that although Santiesteban et al. found similar results with arrows, perspective effects and cueing effects might exist independently—the consistency effect in avatar conditions results from avatar perspective interference, while the effect in arrow conditions results from cue-based attentional orienting [3]. Additionally, debate continues over whether arrows are semi-social or non-social. Nielsen et al. considered arrows semi-social because they are commonly used in daily life to indicate direction, much like an avatar looking somewhere. In contrast, Gunalp et al. considered arrows non-social because they do not interact with people in social life, whereas a chair could be considered semi-social because one might imagine a person sitting in it [46]. To circumvent this issue, many researchers manipulate visual attribution status only in avatar conditions to explore automatic perspective taking mechanisms.

3.2 Visual Attribution Status of Cues

Meltzoff and Brooks claimed that we follow others' gaze because we attribute to them a "seeing" mental state [47]. Visual attribution (or mental state attribution) refers to participants' ability to directly process an avatar' s line of sight during the dot perspective task. Participants show altercentric interference effects when they see that the avatar can see the dots, but not when they believe the avatar cannot see them [10]. Researchers often use physical barriers to manipulate what agents can see [48], including glasses, barriers, or blindfolds in dot perspective tasks to create visible and invisible conditions that manipulate cue visual attribution status.

Furlanetto et al. [13] and Marshall et al. [49] used glasses tasks and found altercentric interference effects in visible but not invisible conditions. O' Grady et al. [50] and Baker et al. [51] used barrier tasks and found similar results—interference effects disappeared when barriers blocked the avatar' s view—suggesting that automatic perspective taking reflects a relatively complex computation of the avatar' s line of sight, supporting implicit mentalizing theory.

However, Conway et al. [52] found altercentric interference effects in both visible and invisible telescope conditions in Experiment 1, challenging the implicit mentalizing explanation and suggesting that submentalizing provides a more plausible account. Experiments 2 and 3 replicated Furlanetto et al.' s glasses experiment but found no evidence of automatic mentalizing. Wilson and colleagues [44] argued that both glasses and barriers introduce additional variables—glasses require a learning process to distinguish, while barriers increase perceptual complexity. Using black cloth to cover the avatar' s eyes, they obtained results consistent with Conway et al., supporting the submentalizing view.

Although these studies used essentially the same paradigm, they produced contradictory results, with some supporting implicit mentalizing [13,49,51] and others supporting submentalizing [3,52]. The reasons for these conflicting results may include···

3.3 Acquisition of Social Perspective

Acquisition of social perspective refers to participants' attempts to adopt others' perspectives when asked what others can see. Research has identified three main situations (as shown in Figure 5 [Figure 5: see original paper]):

1. **Explicit task:** Participants make judgments from both self- and avatar-perspectives within the experiment, producing both egocentric and altercentric interference effects. Judging from the avatar' s perspective shows egocentric interference, while judging from self-perspective shows altercentric interference [4,11,53-57].

2. **Implicit task:** Samson et al.' s Experiment 3 required participants to respond only from self-perspective throughout the entire task [10]. This manipulation aimed to avoid participants actually adopting the avatar' s perspective due to mixed "self" and "avatar" trials. Before each trial, participants were cued with "you" and instructed to ignore the central stimulus, explicitly requiring self-perspective judgments [4,45,54]. Results showed altercentric interference effects in this implicit task.

3. **No-cue task:** These tasks provide no social perspective information whatsoever—neither requiring participants to adopt the avatar' s perspective nor cuing "you" on each trial. These tasks show neither egocentric nor altercentric interference effects unless further modifications are made, such as increasing the stimulus onset asynchrony (SOA) between avatar and dot appearance or directing attention to the avatar, in which case results align with orienting or implicit mentalizing accounts [58-60].

Figure 5. Three experimental paradigms for social perspective acquisition.

Across occlusion tasks, studies typically employ either explicit designs throughout the entire task block or implicit designs throughout. Researchers using explicit tasks tend to find evidence consistent with implicit mentalizing theory [13,49,51], while those using implicit tasks tend to find evidence consistent with submentalizing theory [3,45,61].

Clements-Stephens et al. [62] compared the effects of simple triangular models versus triangular models given social names on visuospatial perspective taking, finding better performance in the social-name condition, suggesting that explicit tasks enhance implicit mentalizing processing. One study compared explicit and implicit tasks but used a within-subjects design that substantially changed the dot perspective task, making results difficult to interpret [52]. Since no-cue tasks have not been used in occlusion tasks, we do not discuss them further here.

By reviewing research paradigms of automatic perspective taking and factors influencing the dot perspective task, we summarize the findings in Table 1. The processing mechanisms of automatic perspective taking remain controversial between implicit mentalizing and submentalizing. We provide dual explanations for different experimental paradigms and summarize factors influencing implicit mentalizing in the dot perspective task, suggesting that similar effects may exist in other perspective-taking paradigms [21,63]. Additionally, no studies have examined factors influencing submentalizing, but based on attention orienting research, factors such as cue directional strength [44], number of cues [64], and degree of spatial coding [24,28] may also modulate consistency effects in automatic perspective taking. Future research could manipulate these submentalizing factors to provide richer evidence for the mechanisms of automatic perspective taking.

Table 1 Experimental paradigms, explanations, and influencing factors of automatic perspective taking

Paradigm	Implicit Mentalizing	Submentalizing	Influencing Factors
Dot perspective: Alter-centric intrusion effect	Automatic processing of others' perspectives interferes with self-perspective judgments	Domain-general processing without mental state consideration (e.g., attentional orienting); conflict between avatar orientation and overall scene	Social relevance; visual attribution status; social perspective acquisition

Paradigm	Implicit Mentalizing	Submentalizing	Influencing Factors
Social Simon: Spatial consistency effect	Automatic processing of others' intentions	Considering confederate' s spatial features rather than mental states	Social relevance; visual attribution status; social perspective acquisition
Ambiguous numbers: Cue consistency effect	Automatic processing of others' perspective numbers	Conflict between egocentric and object-centered spatial reference frames	Social relevance; visual attribution status; social perspective acquisition
Anticipatory looking: Distraction effect	Participants first and longest fixate where avatar last saw the ball	Participants distracted by head orientation, reducing attention to and memory for ball movement	Social relevance; visual attribution status; social perspective acquisition

---

4 Neuroscientific Exploration of Automatic Perspective Taking

Current neuroscientific research on visual perspective taking is in its early stages, focusing primarily on comparing neural mechanisms between self- and other-perspective conditions rather than directly comparing consistent versus inconsistent conditions within self-perspective—that is, automatic perspective taking itself.

Event-related potential (ERP) technology, with its high temporal resolution, can detect underlying mechanisms of automatic perspective taking, yet it has been underutilized in this research area [65-69]. For example, McCleery et al. [67] studied adults using a dot perspective task and found longer latencies for other- versus self-perspective, larger TP450 (325-525ms) amplitudes in consistent versus inconsistent conditions, and larger late frontal slow wave (LFSW) (600-800ms) amplitudes only in the right hemisphere for consistent versus inconsistent conditions. Building on this, Ferguson et al. [65] manipulated social information by comparing adult versus child avatars in automatic perspective taking. The adult avatar condition replicated McCleery et al.' s findings, showing consistency effects, but the child avatar condition eliminated the consistency effect, demonstrating that avatar age modulates automatic perspective taking and supporting implicit mentalizing.

Although these studies compared consistency effects, only Ferguson et al. distinguished consistency differences within self-perspective. Future research could use ERP technology to explore factors influencing automatic perspective taking to address the implicit mentalizing versus submentalizing debate.

Functional magnetic resonance imaging (fMRI), with its high spatial resolution, can reveal brain activation differences between third-person and first-person perspectives. fMRI studies have identified key brain regions involved in perspective taking, including the temporoparietal junction (TPJ), medial prefrontal cortex (mPFC), ventral precuneus, and posterior dorsal precuneus (dpPC) [70-74]. To date, only Ramsey et al. [72], Schurz et al. [73], and Vogeley et al. [74] have used dot perspective paradigms to study automatic perspective taking. Ramsey et al. and Vogeley et al. found that avatar-perspective conditions showed less activation in dorsolateral prefrontal cortex (dlPFC) and regions extending from right inferior parietal lobe (IPL) to TPJ compared to self-perspective conditions, reflecting that avatar perspective engaged fewer cognitive control resources than self-perspective, indicating that avatar perspectives can be automatically processed during self-perspective judgments. Schurz et al. [75] noted that previous studies comparing avatar versus self-perspective brain activation could not resolve the debate between implicit mentalizing and submentalizing theories. Therefore, they compared brain activation in avatar versus non-social (arrow, brick wall, desk lamp) conditions, finding greater activation in TPJ, ventromedial prefrontal cortex, and ventral precuneus for avatar versus non-social conditions. These brain regions were specifically modulated by avatars, supporting implicit mentalizing theory.

However, Catmur et al. [76] argued that fMRI results cannot distinguish between stimulus properties and mentalizing processes, as merely presenting a human-like stimulus activates regions including mPFC and TPJ [75]. Therefore, future research should attempt to separate perspective-taking stimuli from processing to accurately isolate brain regions and cognitive mechanisms involved in automatic perspective taking. Subsequent researchers have directly stimulated relevant brain regions to investigate effects on automatic perspective taking. For example, Santiesteban et al. [5] used transcranial magnetic stimulation (TMS) to inhibit right TPJ (rTPJ) involvement in mentalizing, finding that rTPJ inhibition impaired self-perspective performance in both avatar and arrow conditions. This suggests that social and non-social stimuli share directional or action-related information, activating similar neural mechanisms to produce automatic perspective taking effects. Martin et al. [54] used transcranial direct current stimulation (tDCS) to excite or inhibit dorsomedial prefrontal cortex (dMPFC), finding that excitation enhanced altercentric interference effects while inhibition reduced them. However, Martin et al. did not compare altercentric interference effects across social and non-social tasks under different stimulation conditions, leaving the mechanism of automatic perspective taking unresolved. Future research could compare altercentric interference effects across different social tasks to dissociate implicit mentalizing from submentalizing mechanisms.

5 A Collaborative Model of Implicit Mentalizing and Submentalizing

The debate over automatic perspective taking mechanisms features two mainstream views, with some supporting implicit mentalizing and others supporting submentalizing. Yet evidence for one view cannot completely negate the other. As summarized in Table 1, altercentric interference effects in dot perspective tasks, spatial consistency effects in social Simon tasks, cue consistency effects in ambiguous numbers tasks, and distraction effects in anticipatory looking paradigms cannot rule out either view. Moreover, manipulations of social factors have not directly refuted submentalizing' s role. Neuroscientific research also shows that both social and non-social information activate similar brain regions (rTPJ) during automatic perspective taking, indicating that submentalizing cannot be excluded [5]. Conversely, some studies show that social conditions specifically activate brain regions associated with theory of mind (e.g., mPFC), suggesting that implicit mentalizing cannot be excluded either [75]. Therefore, in different variants of the dot perspective task, domain-general submentalizing processes and domain-specific implicit mentalizing processes are not mutually exclusive. The two explanations can be integrated within a single framework to understand how submentalizing and implicit mentalizing influence automatic perspective taking [77].

Teufel, Fletcher, and Davis [78] proposed a social perception model in which explicit theory of mind modulates implicit theory of mind. They argued that when visual stimuli are presented, individuals engage in perceptual processing. Initially, neural impulses undergo early perceptual processing through the lateral geniculate nucleus (LGN) and primary visual cortex (V1), followed by social perceptual processing through the superior temporal sulcus (STS) and association areas. After perceptual processing, if social information is not sufficiently processed and fails to reach the theory-of-mind activation threshold, automatic gaze following occurs through the parietal attention system. If social information is fully processed, mentalizing occurs through the mirror system or mPFC and TPJ. Additionally, factors such as cue evaluation, personal knowledge, and prior experience modulate explicit theory of mind, which in turn influences implicit theory of mind and automatic gaze following—that is, top-down mentalizing processes can regulate bottom-up perceptual processing. However, this model emphasizes explicit rather than automatic (implicit) theory of mind and cannot fully explain the possible mechanisms of automatic perspective taking.

Drawing on Teufel, Fletcher, and Davis' s model, we propose a processing model of implicit mentalizing and submentalizing, shown in Figure 6 [Figure 6: see original paper]. Observed consistency effects can be achieved through three pathways:

1. After visual stimulus presentation, individuals engage in early perceptual processing. If only clear directional cues are present without social cues, consistency effects arise through directional cue processing. For example,

researchers using arrows as central stimuli have found consistency effects in arrow conditions [16,44,45].

2. If social cues are present but fail to reach the social perception activation threshold, they are processed as directional cues, activating automatic perspective taking through submentalizing. Researchers using semi-social stimuli such as lamps, chairs, and cameras have found consistency effects, suggesting that these effects may arise from independent submentalizing processes [16,45,46].

3. If social cues are present and exceed the social perception activation threshold, automatic perspective taking is activated through implicit mentalizing. As researchers using glasses, barriers, or black cloth to manipulate visible versus invisible conditions have found inconsistent results despite using essentially the same paradigm: some show consistency effects only in visible conditions, suggesting that processing of the avatar' s line of sight reached the social perception threshold and was processed through implicit mentalizing [13,49,51,79]; while others find consistency effects in both visible and invisible conditions, suggesting that both submentalizing and implicit mentalizing may operate jointly when avatar stimuli are present [52].

The activation threshold for implicit mentalizing remains undefined, but our analysis suggests several distinguishing factors: (1) whether stimulus materials involve social components (e.g., lamps, cameras, chairs), with more social components making threshold activation easier [5,45,46]; (2) the degree of attention to social information, which may regulate this threshold—for example, adding numbers to the avatar to focus attention on it may more easily reach the social perception activation threshold, producing larger consistency effects [58]; (3) demonstrative pronouns may also influence social cue perception, with pronouns like "you" or "he/she" reaching activation thresholds more easily than general property words like "red" or "all" [58,60]. Therefore, future research can build on previous work to further explore the processing threshold of implicit mentalizing, thereby distinguishing the contexts in which implicit mentalizing and submentalizing occur to resolve the debate between these two automatic perspective-taking mechanisms.

Figure 6. Collaborative model of implicit mentalizing and submentalizing. : Directional cue processing produces consistency effects; : Social cues fail to pass social perception threshold and are processed only as directional cues (submentalizing); : Social cues pass social perception threshold and undergo social cue perception processing (implicit mentalizing). Dashed line indicates social perception threshold.

6 Summary and Outlook

The debate between implicit mentalizing and submentalizing views of automatic perspective taking continues, with researchers exploring new angles to investigate its mechanisms, such as applying dot tasks to Flanker tasks, Simon tasks, and gaze cueing effects [25,58]. Submentalizing' s most important contribution is to reduce explanations of implicit mentalizing theory and analyze whether altercentric interference results solely from simple domain-general cognition. Existing research still has several limitations requiring further investigation.

First, the relationship between domain-general and domain-specific processes needs exploration. Submentalizing theory implies that mentalizing emerges from domain-general cognitive functions integrated through learning and culture [80]. Teufel and Alexis [81] also emphasize the important role of domain-general cognitive functions in implicit mentalizing, while noting that cue evaluation, individual knowledge, and prior experience influence explicit mentalizing, thereby affecting social information processing. However, the boundary between domain-general and domain-specific processes remains unclear. Therefore, future research attempting to separate submentalizing and implicit mentalizing should also examine their collaborative interactions.

Second, researchers have focused primarily on how factors influencing implicit mentalizing affect automatic perspective taking, neglecting factors that influence submentalizing. Through dot perspective paradigms, researchers have manipulated factors affecting implicit mentalizing and found that social relevance, visual attribution status, and social perspective acquisition all influence automatic perspective taking [3,44,56]. However, almost no studies have manipulated submentalizing factors. If factors influencing submentalizing can modulate automatic perspective taking results, this would provide additional evidence for the implicit mentalizing versus submentalizing debate.

Third, methodological details differ across studies, making direct comparisons difficult. As discussed in the factors influencing dot perspective tasks, some researchers use explicit perspective tasks [58,82], while others use only self-perspective judgments [44]; some use completely random designs for social relevance tasks [3], while others use block designs [10,52]. Additionally, many behavioral experiments have small sample sizes, poor replicability, and cannot guarantee the reliability of null and positive results. Therefore, future research should conduct adequately powered studies with careful differentiation and comparison of different cues in automatic perspective taking to verify result validity and reliability.

Fourth, the mechanisms of automatic perspective taking require more neuroscientific evidence. Social neuroscience evidence indicates that implicit and explicit theory of mind share a common neural network, with both processes activating TPJ and mPFC [83]. However, these regions are not exclusively social—TPJ is involved in attention, memory, spatial cognition, and language, while mPFC is associated with working memory [84]. As modern techniques continue to improve, eye-tracking, tDCS, TMS, and functional near-infrared spectroscopy (fNIRS) can further explore the causes of automatic perspective taking.

Finally, research on automatic perspective taking in special populations remains limited. Infants are born with only non-social cognition and gradually learn about others' mental states through social environments. Using infant participants may provide stronger evidence for the automatic perspective taking debate. For example, manipulating mentalizing factors such as social relevance or visual attribution status, or using Posner paradigms to compare non-social versus social cue effectiveness as in Ji et al. [85], could test whether infants show implicit mentalizing or only submentalizing. Autism spectrum disorder involves impaired general social functioning, yet most research has focused on explicit theory of mind, with little investigation of implicit mentalizing [86,87]. Research with autistic populations should select or adapt paradigms suitable for this group. Deaf individuals show delayed theory-of-mind development compared to typical individuals and are often used as comparison groups in theory-of-mind research [88]. Automatic perspective taking research could also be extended to deaf populations [89]. Theory-of-mind research in special populations has important practical implications for treatment and represents a key direction for future investigation.

References

[1]Kovacs, A. M., Teglas, E., & Endress, A. D. The social sense: Susceptibility to others' beliefs in human infants and adults[J]. Science, 2010, 330(6012): 1830–1834.

[2]Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. Cognitive load disrupts implicit theory-of-mind processing[J]. Psychological Science,2012, 23(8): 842–847.

[3]Cole, G. G., Atkinson, M., Le, A. T. D., & Smith, D. T. Do humans spontaneously take the perspective of others? [J]. Acta Psychologica, 2016, 164(2016): 165–168.

[4]Michaelab, J., Wolfb, T., Letessonb, C., Butterfilla, S., Skewesc, J., & Hohwy, J. Seeing it both ways using a dobule cueing task to investigate the role of spatial cueing in level-1 visual perspectiving taking[J]. Journal of Experimental Psychology: Human Perception and Performance, 2017, 44(5): 693.

[5]Santiesteban, I., Kaur, S., Bird, G., & Catmur, C. Attentional processes, not implicit mentalizing, mediate performance in a perspective-taking task: Evidence from stimulation of the temporoparietal junction[J]. Neuroimage,2017, 155(2017): 305–311.

[6] 潘威, 汪寅, 陈巍. (2017). 心智化社会认知观的演变及发展–来自潜心智化的思考 [J]. 心理科学,2017, 40(5):1274–1279.(Pan, W., Wang, Y., & Chen, W. Evolvement and Development Trend of Mentalizing Accounts of Social Cognition: Reflections from Submentalizing[J]. Journal of Psychological Science, 2017, 40(5): 1274–1279.)

[7]Apperly, I. A., & Butterfill, S. A. Do humans have two systems to track beliefs and belief-like states? [J]. Psychological Review, 2009, 116(4): 953-970.
[8]Apperly, I. A. What is "theory of mind" ? Concepts, cognitive processes and individual differences[J]. Quarterly Journal of Experimental Psychology, 2013, 65(5): 825-839.
[9]Heyes, C. Submentalizing: I am not really reading your mind[J]. Perspectives on Psychological Science : A Journal of the Association for Psychological Science, 2014, 9(2): 131-143.
[10]Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. Seeing it their way: Evidence for rapid and involuntary computation of what other people see[J]. Journal of Experimental Psychology: Human Perception and Performance, 2010, 36(5): 1255-1266.
[11]Wang, J. J., Tseng, P., Juan, C. H., Frisson, S., & Apperly, I. A. Perspective-taking across cultures: shared biases in Taiwanese and British adults[J]. Royal Society Open Science, 2019, 6(11): 190540.
[12] 王雨晴，游旭群，焦健，谌鹏飞. 观点采择：基于自我的推理及其个体差异 [J]. 心理学报, 2015, 47(8), 1039-1049.(Wang, Y. Q., You, X. Q., Jiao, J., & Zhan, P. F., Perspective Taking: Making Inferences Based on Oneself and Related Individual Differences[J].Acta Psychologica Sinica, 2015, 47(8): 1039-1049.)
[13]Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing[J]. Journal of Experimental Psychology: Human Perception and Performance, 2016, 42(2): 158-163.
[14]Marotta, A., Lupianez, J., Martella, D., & Casagrande, M. Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection[J]. Journal of Experimental Psychology: Human Perception and Performance, 2012, 38(2): 326-335.
[15] 鲁上, 刘烨, 傅小兰. 头部朝向在社会性注意转移中的作用 [J]. 心理科学进展, 2013, 21(2): 211-219.(Lu, S., Liu, Y., & Fu, X. L. The Role of Head Orientation in Social Attention Shift[J]. Advances in Psychological Science, 2013, 21(2): 211-219.)
[16]Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. Avatars and arrows: Implicit mentalizing or domain-general processing?[J]. Journal of Experimental Psychology: Human Perception and Performance, 2014, 40(3): 929-937.
[17]Simon, J. R. Reactions toward the source of stimulation[J]. Journal of Experimental Psychology, 1969, 81(1):
[18]Hommel, B. S-R compatibility effects without response uncertainty[J]. The Quarterly Journal of Experimental Psychology, 1969, 49(3): 546-571.
[19]Sebanz, N., Knoblich, G. N., & Prinz, W. Representing others' actions: Just like one' s own?[J]. Cognition, 2003, 88(3): B11-21.
[20]Freundlieb, M., Kovacs, A. M., & Sebanz, N. When do humans spontaneously adopt another' s visuospatial perspective?[J]. Journal of Experimental Psychology: Human Perception and Performance, 2016, 42(3):
[21]Freundlieb, M., Kovács, Á. M., & Sebanz, N. Reading your mind while you are reading—evidence for spontaneous visuospatial perspective taking during a semantic categorization task[J]. Psychological Science, 2018, 29(4): 614-622.

[22]Freundlieb, M., Sebanz, N., & Kovács, Á. M. Out of your sight, out of my mind: Knowledge about another person' s visual access modulates spontaneous visuospatial perspective-taking[J]. Journal of Experimental Psychology: Human Perception and Performance, 2017, 43(6): 1065–1072.
[23]Boffel, C., & Musseler, J. Perceived ownership of avatars influences visual perspective taking[J]. Frontiers in Psychology, 2018, 9, 743.
[24]Boffel, C., & Musseler, J. Action effect consistency and body ownership in the avatar-Simon task[J]. Plos One, 2019, 14(8): e0220817.
[25]Boffel, C., & Musseler, J. Visual perspective taking for avatars in a Simon task[J]. Attention, Perception, & Psychophysics, 2019, 81(1): 158–172.
[26]Salm-Hoogstraeten, S., & Musseler, J. Human cognition in interaction with robots: Taking the robot' s perspective into account[J]. Human Factors The Journal of the Human Factors and Ergonomics Society,
[27]Kuhbandner, C. The role of positive and negative affect in the â mirroringâ of other persons' actions[J]. Cognition & Emotion, 2010, 24(7): 1182–1190.
[28]Dolk, T., Hommel, B., Colzato, L. S., Schutz-Bosbach, S., Prinz, W., & Liepelt, R. How "social" is the social Simon effect?[J]. Frontiers in Psychology, 2021, 2(84): 1–9.
[29]Guagnano, D., Rusconi, E., & Umiltà, C. A. J. C. Sharing a task or sharing space? On the effect of the confederate in action coding in a detection task[J]. Cognition , 2010, 114(3): 348–355.
[30]Surtees, A., Samson, D., & Apperly, I. A. Unintentional perspective-taking calculates whether something is seen, but not how it is seen[J]. Cognition, 2016, 148(2016): 97–105.
[31]Elekes, F., Varga, M., & Király, I. Evidence for spontaneous level-2 perspective taking in adults[J]. Consciousness and Cognition, 2016, 41(2016): 93–103.
[32]Millett, A. C., D' Souza, A. D. C., & Cole, G. G. Attribution of vision and knowledge in spontaneous perspective taking[J]. Psychological research. Advance online publication doi:10.1007/s00426-019-01179-1
[33]Wimmer, H., & Perner, J. Beliefs about beliefs: Representation and constraircing function of wrong bekfs in young children' s understanding of deception[J]. Cognition, 1983, 13(1983): 103–128.
[34]Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., & Call, J. Great apes use self-experience to anticipate an agent' s action in a false-belief test[J]. Proceedings of the National Academy of Sciences, 2019, 116(42):
[35]Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. Great apes anticipate that other individuals will act according to false beliefs[J]. Science, 2016, 354(6308): 110–114.
[36]Senju, A., Southgate, V., White, S., & Frith, U. Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome[J]. Science, 2009, 325(5942): 883–885.
[37]Schneider, D., Slaughter, V. P., Bayliss, A. P., & Dux, P. E. A temporally sustained implicit theory of mind deficit in autism spectrum disorders[J]. Cognition, 2013, 129(2): 410–417.
[38]Katherine, J., Kimberley, S., Andrew, B., & Patric, B. Looking ahead:

Anticipatory cueing of attention to objects others will look at[J]. Cognitive Neuroscience, 2015, 7(1-4).pp. 74-81.
[39]Kano, F., Krupenye, C., Hirata, S., Call, J., & Tomasello, M. Submentalizing cannot explain belief-based action anticipation in apes[J]. Trends in Cognitive Sciences, 2017, 21(9): 633-634.
[40]Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. A test of the submentalizing hypothesis: Apes' performance in a false belief task inanimate control[J]. Communicative & Integrative Biology, 2017, 10(4): e1343771.
[41]Kano, F., Call, J., & Krupenye, C. Primates pass dynamically socialAnticipatory-looking false-belief tests[J]. Trends in Cognitive Sciences, 2020, 24(10): 777-778.
[42] 王伟平, 苏彦捷. 孤独症儿童基于眼睛注视的社会性注意 [J]. 中国特殊教育, 2006, 6: 12-17. (Wang, W. P., & Su, Y. J. A Review on Autism Children' s Social Attention Based on Eye Gaze[J]. Chinese Journal of Special Education, 2006, 6: 12-17.)
[43]Burnside, K., Se[verdija, V., & Poulin-Dubois, D. Infants attribute false beliefs to a toy crane[J]. Developmental Science, 2019, 23(2).
[44]Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. Inclined to see it your way: Do altercentric intrusion effects in visual perspective taking reflect an intrinsically social process?[J]. Quarterly Journal of Experimental Psychology, 2015, 68(10): 1-21.
[45]Wilson, C. J., Soranzo, A., & Bertamini, M. Attentional interference is modulated by salience not sentience[J]. Acta Psychologica, 2017, 178(2017): 56-65.
[46]Gunalp, P., Moossaian, T., & Hegarty, M. Spatial perspective taking: Effects of social, directional, and interactive cues[J]. Memory Cognition, 2019, 47(5): 1031-1043.
[47]Meltzoff, A. N., & Brooks, R. Eyes wide shut: The importance of eyes in infant gaze following and understanding of other minds[M]. In R. Flom, K. Lee & D. Muir (Eds.), Gaze following: Its development and significance(2007, pp.217-241). Mahwah, NJ: Erlbaum.
[48]Hare, B., Call, J., & Tomasello, M. Do chimpanzees know what conspecifics know?[J]. Animal Behaviour, 2001, 61(1): 139-151.
[49]Marshall, J., Gollwitzer, A., & Santos, L. R. Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks[J]. Plos One, 2018, 13(3): e0194101.
[50]O' Grady, C., Scott-Phillips, T., Lavelle, S., & Smith, K. Perspective-taking is spontaneous but not automatic[J]. 2019. Advance online publication.doi:10.31219/osf.io/wzcqs.
[51]Baker, L. J., Levin, D. T., & Saylor, M. M. The extent of default visual perspective taking in complex layouts[J]. Journal of Experimental Psychology: Human Perception and Performance, 2015, 42(4): 508.
[52]Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test[J]. Journal of Experimental Psychology: Human Perception and Performance, 2016, 43(3): 454-465.

[53]Bukowski, H., & Samson, D. Can emotions influence level-1 visual perspective taking?[J]. Cognitive Neuroscience, 2015, 7(1-4): 1-10.
[54]Martin, A. K., Dzafic, I., Ramdave, S., & Meinzer, M. Causal evidence for taskspecific involvement of the dorsomedial prefrontal cortex in human social cognition[J]. Social Cognitive and Affective Neuroscience, 2017, 8(8): 1209-1218.
[55]Simpson, A. J., & Todd, A. R. Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation[J]. Cognition, 2017, 166(2017): 371-
[56]Todd, A. R., Cameron, C. D., & Simpson, A. J. Dissociating processes underlying level-1 visual perspective taking in adults[J]. Cognition, 2017, 159(2017): 97-101.
[57]Todd, A. R., & Simpson, A. J. Anxiety impairs spontaneous perspective calculation: Evidence from a level-1 visual perspective-taking task[J]. Cognition, 2016, 156(2016): 88-94.
[58]Bukowski, H., Hietanen, J. K., & Samson, D. From gaze cueing to perspective taking: Revisiting the claim that we automatically compute where or what other people are looking at[J]. Visual Cognition, 2015, 23(8):
[59]Capozzi, F., & Ristic, J. How attention gates social interactions[J]. Annals of the New York Academy of Sciences, 2018, 1426(1): 179-198.
[60]Gardner, M. R., Bileviciute, A. P., & Edmonds, C. J. Implicit mentalising during level-1 visual perspective-taking indicated by dissociation with attention orienting[J]. Vision, 2018, 2(3): 1-10.
[61]Langton, S. R. H. I don' t see it your way: The dot perspective task does not gauge spontaneous perspective taking[J]. Vision, 2018, 2(6): 1-17.
[62]Clements-Stephens, A. M., Vasiljevic, K., Murray, A. J., & Shelton, A. L. The role of potential agents in making spatial perspective taking social[J]. Frontiers in Human Neuroscience, 2013, 7: 497.
[63]Gobel, M. S., Tufft, M. R. A., & Richardson, D. C. Social beliefs and visual attention: How the social relevance of a cue influences spatial orienting[J]. Cognitive Science, 2017, 42: 161-185.
[64]Milgram, S., Bickman, L., & Berkowitz, L. Note on the drawing power of crowds of different size[J]. Journal of Personality and Social Psychology, 1969, 13(2):79-82.
[65]Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. Age of avatar modulates the altercentric bias in a visual perspective-taking task: ERP and behavioral evidence[J]. Cognitive, Affective, & Behavioral Neuroscience. 2018. Advance online publication. doi:10.3758/s13415-018-0641-1
[66]Kuhn-Popp, N., Sodian, B., Sommer, M., Dohnel, K., & Meinhardt, J. Same or different? ERP correlates of pretense and false belief reasoning in children[J]. Neuroscience, 2013, 248(2013): 488-498.
[67]McCleery, J. P., Surtees, A. D. R., Graham, K. A., Richards, J. E., & Apperly, I. A. The neural and cognitive time course of theory of mind[J]. Journal of Neuroscience, 2011, 31(36): 12849-12854.
[68]Meinhardt, J., Sodian, B., Thoermer, C., Dohnel, K., & Sommer, M. True- and false-belief reasoning in children and adults: An event-related potential

study of theory of mind[J]. Developmental Cognitive Neuroscience, 2011, 1(1): 67-76.
[69]Sabbagh, M. A., & Taylor, M. Neural correlates of theory-of-mind reasoning: An event-related potentialstudy[J]. Psychological Science, 2000, 11(1): 46-50.
[70]Apperly, I. A., Samson, D., & Humphreys, G. W. Domain-specificity and theory of mind: Evaluating neuropsychological evidence[J]. Trends in Cognitive Sciences, 2005, 9(12): 572-577.
[71]Kaiser, S., Walther, S., Nennig, E., Kronmüller, K., Mundt, C., Weisbrod, M., . . . Vogeley, K. Gender-specific strategy use and neural correlates in a spatial perspective taking task[J]. Neuropsychologia, 2008, 46(10):
[72]Ramsey, R., Hansen, P., Apperly, I. A., & Samson, D. Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes[J]. Journal of Cognitive Neuroscience, 2013, 25(5): 670-684.
[73]Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies[J]. Neuroscience and Biobehavioral Reviews, 2014, 42(2014): 9-34.
[74]Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. Neural correlates of first-person perspective as one constituent of human self-consciousness[J]. Journal of Cognitive Neuroscience, 2004, 16(5): 817-827.
[75]Schurz, M., Kronbichler, M., Weissengruber, S., Surtees, A., Samson, D., & Perner, J. Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity[J]. Neuroimage, 2015, 117(2015): 386-396.
[76]Catmur, C., Santiesteban, I., Conway, J. R., Heyes, C., & Bird, G. Avatars and arrows in the brain[J]. Neuroimage, 2016, 132(2016): 8-10.
[77]Capozzi, F., & Ristic, J. Attention and mentalizing? Reframing a debate on social orienting of attention[J]. Visual Cognition, 2020, 28: 97-105.
[78]Teufel, C., Fletcher, P. C., & Davis, G. Seeing other minds: Attributed mental states influence perception[J]. Trends in Cognitive Sciences, 2010, 14(8): 376-382.
[79]Seow, T., & Fleming, S. M.. Perceptual sensitivity is modulated by what others can see[J]. Attention, Perception, & Psychophysics, 2019, 81(6): 1979-1990.
[80]Heyes, C., & Frith, C. D. The cultural evolution of mind reading[J]. Science, 2014, 344(6190): 1357-1364.
[81]Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. Mental-state attribution drives rapid, reflexive gaze following[J]. Attention, Perception & Psychophysics, 2010, 72(3): 695-705.
[82]Todd, A. R., Forstmann, M., Burgmer, P., Brooks, A. W., & Galinsky, A. D. Anxious and egocentric: How specific emotions influence perspective taking[J]. Journal of Experimental Psychology: General, 2015, 144(2): 374-391.
[83]Frank, V. O., & Marie, V. Implicit and explicit social mentalizing: dual processes driven by a shared neural network[J]. Frontiers in Human Neuroscience, 2013, 7(560),1-6.
[84]Carter, R. M., & Huettel, S. A. A nexus model of the temporal-parietal

junction[J]. Trends in Cognitive Sciences, 2013, 17(7): 328–336.
[85]Ji, H., Wang, L., & Jiang, Y. Cross-category adaptation of reflexive social attention[J]. Journal of Experimental Psychology: General, 2020, 149(11): 2145–2153.
[86]Eack, S. M., Wojtalik, J. A., Keshavan, M. S., & Minshew, N. J. Social-cognitive brain function and connectivity during visual perspective-taking in autism and schizophrenia[J]. Schizophrenia Research,2017:
[87]Rosenblau, G., Kliemann, D., Heekeren, H. R., & Dziobek, I. Approximating implicit and explicit mentalizing with two naturalistic video-based tasks in typical development and autism spectrum disorder[J]. Journal of Autism and Developmental Disorders, 2015, 45(4): 953–965.
[88] 陈友庆，郭本禹．聋儿的心理理论发展特点及影响因素．心理科学进展, 2006, 14(2): 382–388. (Chen, Y. Q.,&Guo, B. Y. Features of the development of theory of mind for deaf children and its influencing factors.Advances in Psychological Science, 2006, 14(2): 382-388.)
[89]Peterson, C. C., & Siegal, M. Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children[J]. Psychological Science, 1999, 10(2): 126–129.

(Corresponding author: Feng XIAO E-mail: xiaofeng19850328@gmail.com)

Author Contribution Statement:
Yi LI: Proposed the research question, drafted the manuscript;
Feng XIAO: Proposed the cognitive model, revised the final version.

Note: Figure translations are in progress. See original paper for figures.