PROCESSING MECHANISMS OF EYE-HEAD CUES AND EYE-FINGER-POINTING CUES IN THE DOT-PERSPECTIVE TASK

BY

CONG FAN

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

in Psychology

Victoria University of Wellington

2021

Abstract

Calculating others' visual perspective automatically is a pivotal ability in human's social communications. In dot-perspective task, the ability is shown as consistency effect: adults respond more slowly to judge the number of discs that they can see when a computer-generated avatar sees fewer discs. The implicit mentalising account attributes the effect to relatively automatic tracking of others' visual perspective. However, the submentalising account attributes the effect to domain-general attentional orienting. Accordingly, three studies were conducted to elucidate the ongoing implicit mentalising vs. submentalising debate.

Study 1 (comprising Experiments 1 and 2) replicated consistency effect either when real-human-face or spatial layout of discs was considered. Study 2 (comprising of Experiments 3 and 4) dissociated two accounts by manipulating real human's facial cues. In Experiment 3, using a new visual access manipulation (i.e., a black rectangle placed on an agent's eyes for rendering an invisible condition), a consistency effect was induced for eyes-opened but not eyes-covered faces with head direction, suggesting implicit mentalising. Experiment 4 firstly compared implicit mentalising (via consistency effect in dot-perspective task) with attentional orienting (via cue-validity effect in Posner task) when manipulating eye-head cues (head-front-gaze-averted versus head-turned-gaze-maintained). Neither effect was modulated by eye-head-related directional cue, but cue-validity effect's elicitation seemed to be related to directional cue's dynamic property. Overall, implicit mentalising as revealed in consistency effect cannot be purely reduced to attentional-orienting-related submentalising processes.

Study 3 (comprising of Experiments 5 to 7) further clarified the debate by considering the agent's different body cues. Experiment 5 extended findings of Experiment 4 by generating a new eye-head-cue comparison (head-front-gaze-averted vs. head-turned-gaze-averted). Directional cue modulated cue-validity effect but not consistency effect, favouring Study 2's conclusion. Experiment 6 adopted a new body-cue-manipulation (gaze-averted vs. finger-pointing). Both cue-validity and consistency effects were elicited for finger-pointing but not gaze-averted agents, supporting submentalising. Experiment 7 combined finger-pointing with visual access's manipulation (eyes-opened vs. eyes-covered) on the dot-perspective task. Visual access did not modulate the consistency effect when finger-pointing was simultaneously displayed, supporting submentalising. Altogether, gaze aversion cues appear to play a dominant role in moderating implicit mentalising on the dot-perspective task, but the process may be interfered by the easily-discriminable finger-pointing cues via an attentional orienting mechanism.

Acknowledgements

The thesis could not be completed without an army of remarkable people who have supported me throughout my PhD-degree study. First and foremost, I would like to thank my parents' spiritual and financial support. Without their understanding, encouragement and unconditional love, I could not accomplish my study.

I am extremely grateful to my supervisors Dr. Jason Low and Dr. Tirta Susilo. Dr. Jason Low guided me to find a research topic, do the presentations and write several manuscripts with his wisdom, patience, and insightful comments. Dr. Tirta Susilo guided me with manuscript writing, and his expertise on human/face perception as well as quick thinking about the logic of the research has saved me when I was stuck in writing.

Many thanks to my friends who have provided helpful advice for conducting the experiments. Thanks Lin Hu and Xinglong Yao for their help with uploading the experimental procedures online; Rui Fan for her assistance with image editing; Lisa Woods, Geraldine Smieszhala for their help with power analysis; Yanzhu Chen and Mingjie Xu for their assistance with photographs taken.

I appreciate my best friends' and teammates' spiritual support, which makes me always feel warm. They listen to me talking about the struggles of my study and life. Many thanks for sharing their useful and funny experiences about how to get through the difficulties. Thanks to my Master Supervisors Dr. Wenbo Luo and Dr. Shunsen Chen for their spiritual support and encouragement. Finally, thanks to the hard-working actor and singer Jackson Yi and the soul singer Tia Yuan as their performances have enriched my life.

Table of Contents

Statement of Authorship and Copyright

I am the primary investigator and author on the co-authored articles presented in this thesis. I developed the research questions, designed the studies, collected and managed the data, conducted and interpreted the analyses, and wrote the first drafts. Jason Low and Tirta Susilo helped with conceptualising the experiments and provided critical revisions of written drafts.

I confirm that the permission to include the copyrighted published material (Chapter 4, pages 49-72) in this thesis has been granted by Taylor & Francis Group in Visual Cognition: '**We will be pleased to grant permission to reproduce your authors accepted manuscript from our Journal in your thesis and to be posted in the university's repository.' (Email received on 22/02/2021)**

Prologue: Introduction Overview

Humans regularly infer others' mental states (intentions, desires, emotions, etc.) in social interaction. The capacity is known as theory of mind (ToM) (Low, Apperly, Butterfill, & Rakoczy, 2016). As one basis of ToM, Level-1 visual perspective taking (L1VPT) refers to the ability to assess what someone else can and cannot see (Samson, Apperly, Braithwaite, Andrews, & Bodley-Scott, 2010; Bukowski, Hietanen, & Samson, 2015). Since Masangkay et al. (1974) and Flavell, Everett, Croft and Flavell (1981) found that 3-year-old children passed tasks requiring "level-1 perspective-taking knowledge", the processing mechanism of L1VPT in human beings has been investigated intensely over the past few decades. However, there is significant debate over whether and to what extent the underlying mechanism of L1VPT is domain-specific (implicit mentalising account, e.g., Apperly, 2010) or domain-general (submentalising account, e.g., Heyes, 2014). The implicit mentalising account holds that the process can be specific to the calculation of others' visual perspectives in a fast, relatively automatic, and implicit way, whereas the submentalising account posits that the process may be attributed to domain-general attentional orienting. Although the debate has been investigated for several years, there are still no consistent findings on the debate. Accordingly, the current thesis attempted to pin down the mechanism of L1VPT and clarify the debate between the implicit mentalising account and the submentalising account.

Much of the pivotal work has been conducted with computer-generated avatars with unclear eye-gaze. However, how real faces with clear eye-gaze would be processed in L1VPT-related paradigm remains incomplete. An accumulating number of researchers have manipulated the agent’s line-of-sight to clarify the debate but with potential limitations, which warrants further investigations. Additionally, much less is known about manipulations of different parts of the agent’s body (i.e., eye-head cues, eye-finger-pointing cues) in modulating visual perspective-taking and attentional orienting. Considering the agent’s different body parts can be beneficial for making the theoretical basis of L1VPT more comprehensive. To address the issues, seven experiments were carried out to further contrast the two competing accounts.

The framework of the thesis is as follows. Chapters 1 and 2 outline theories, paradigms, and review previous studies related to L1VPT and attentional orienting, respectively. Then, chapter 3 describes two pilot experiments conducted for replicating Samson et al.’s (2010) Experiment 3, which is a classic experiment to measure the potential implicit mentalising in L1VPT. The following two chapters (i.e., Chapters 4-5) address five experiments together with the corresponding findings. Finally, Chapter 6 shows a general discussion of the studies by presenting explanations of the findings to provide new insights into the implicit mentalising vs. submentalising debate.

Chapter 1: Level-1 Visual Perspective Taking (L1VPT)

Compared with the more complex Level-2 visual perspective taking capability to judge how someone else sees a particular stimulus, L1VPT capability refers to the simple judgement about whether another person can see a stimulus. Samson and colleagues (2010) measured L1VPT processing by designing the dot-perspective paradigm. Since then, the paradigm has become a widely used method for exploring the processing mechanisms underlying L1VPT. Furthermore, the paradigm was created to spotlight and dissect the following effects: the egocentric bias, the altercentric bias and the altercentric interference/intrusion effect. Accordingly, Chapter 1 begins with the detailed explanations of those effects. Following that, as the debate regarding L1VPT processing is the focus of the research, the studies that sought to resolve the debate by manipulating gazer's line-of-sight, creating new versions of the dot-perspective paradigm or comparing the corresponding effects triggered in a visual-perspective-taking-related task with effects generated in an attentional-orienting-related task are described.

1.1 Perspective-related Intrusions

1.1.1 Egocentric Intrusion and its Resistance

A line of ToM investigations has found that both children and adults showed a strong bias towards their own thoughts or beliefs when they were reasoning about others' mental states, termed as egocentric bias (Apperly, Samson, & Humphreys, 2009; Bernstein, Atance, Loftus, & Meltzoff, 2004; Birch & Bloom, 2007; Keysar, Lin, & Barr, 2003; Moore et al., 1995; Royzman, Cassidy, & Baron, 2003). For instance, in a communicative game adopted by Keysar et al. (2003), a confederate asked a participant to move a target object (e.g., a tape) in a grid. Before the confederate's instruction of moving the target, participants hid a tape in a brown bag that was only known by themselves. The authors found that instead of moving the tape that was visible to both the confederate and participants (i.e., correct response), the participants often moved the bag containing the tape that the confederate did not know. The findings demonstrated that it was comparatively demanding to resist the intrusion of one's own perspective even when adults were aware of someone else's perspectives. They interpreted that adults performed egocentrically even though their own knowledge of object location was different from the confederate's knowledge, showing that considering people's own knowledge could compromise or even cancel their perspective-taking-related processes. Additionally, overriding the egocentric bias in ToM processes has been suggested to be effortful as the ability to inhibit the bias is strongly correlated with executive function abilities in children (e.g., Carlson & Moses, 2001). Altogether, succeeding in overriding the egocentric bias seems to be demanding and effortful when reasoning about others' perspectives.

1.1.2 Altercentric Intrusion in L1VPT

Even though human beings' tendency to be egocentric highlights that some ToM processes can be cognitively effortful, there is evidence showing that people can easily and effortlessly compute others' visual perspectives. For instance, Sodian, Thoermer, and Metz (2007) tracked infants' eye movements to find it was easy to understand another person's discrepant visual experience. Specifically, when the old and new targets were simultaneously presented on a table, 14-month-old infants looked longer at an actress's goal-directed action for a novel target when the old target was visible than when the old target was invisible to her (Both targets were visible to the infants). The looking behaviours were evoked under the circumstance of passively picture-viewing without any other task instruction. Thus, their looking time patterns suggest that 14-month-old infants can easily compute adults' visual perspectives independently of their own perspectives (i.e., L1VPT ability). The findings fit with the speculation that perspective computation reflects infants' apparently sophisticated ToM as indirectly measured by their looking time responses (e.g., Baillargeon, Scott, & He, 2010) (nonetheless, it is important to acknowledge that infants' success on non-verbal tasks are subject to replication problems, and their success can also be explained by a range of sub-mentalistic processes (e.g., Ruffman, Taumoepeau, & Perkins, 2012).

Similar to infants' visual computation, adults can effortlessly track others' visual perspectives (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010), which was reflected by the finding that adults were slower and less accurate to judge the number of dots they saw when an avatar saw a different number of dots (i.e., consistency effect) on self-perspective trials. The consistency effect elicited without explicit judgement about the avatar's visual perspective was interpreted as an effect of altercentric intrusion. That is, adults' computation of the avatar's visual perspective was task-irrelevant and yet appeared to be undertaken in a way that interfered with judgements on their own perspectives. Additionally, effortless calculation of others' visual perspectives in L1VPT processing was found to be independent of executive-function resources (e.g., Qureshi, Apperly, & Samson, 2010).

1.2 The Dot-perspective Paradigm and Generalization of Consistency Effect

1.2.1 The Dot-perspective Paradigm

The dot-perspective paradigm was created by Samson and colleagues (2010) as an experimental paradigm for measuring L1VPT processing. The essential idea behind the task is that if adults can implicitly track another person's visual perspective, then participants should do so even when they do not need to. In the paradigm (see Figure 1), disc(s) were presented on the left- or right-side wall of a room with a computer-generated human avatar standing in the centre of the room and facing to one side of the walls. Two kinds of visual perspectives — "You see N" on 'Self-perspective' trials ('N' ranges from 0 to 3 dots), "S/he sees N" on 'Other-perspective' trials — were presented before the scene. The participants were required to judge whether the picture matched the given perspective, leading to matching and mismatching trials. On half of the trials (Congruent condition), the avatar and the participant could see the same disc(s). On the remaining trials (Incongruent condition), the participant could see the disc(s) that were invisible to the avatar.

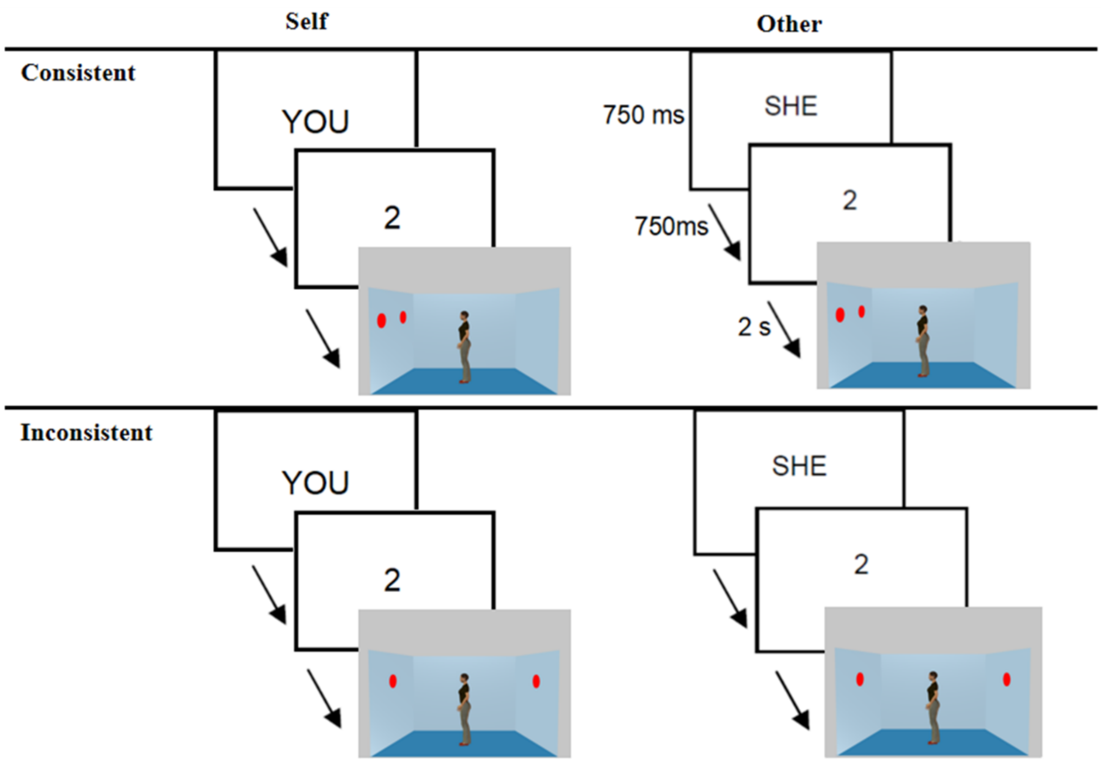

FIGURE 1 | Overview of a representative test trial of the dot-perspective paradigm created by Samson et al. (2010)

Consistent with previous studies of egocentric biases (i.e., strong biases towards the participants' own perspectives on other-perspective judgements) (e.g., Birch & Bloom, 2004), findings indicated egocentric intrusions when the participants were instructed to take the avatar's visual perspective. For instance, Samson et al. (2010) found slower response times and lower accuracy in an incongruent condition compared to a congruent condition (i.e., Consistency effect) on other-perspective trials. Additionally, the novel and key finding of the study was that the participants made more errors and responded more slowly in the inconsistent condition compared to consistent ones when making self-perspective judgements. The findings suggested that adults could rapidly and effortlessly take the avatar's visual perspective even when not required to do so. The researchers explained the result as an effect of altercentric interference/intrusion, namely, the participant's own visual perspective was interfered by the implicit computation of the avatar's visual perspective that was task-irrelevant (Samson et al., 2010).

1.2.2 Generalization of the Self-perspective-related Consistency Effect

The consistency effect on self-perspective trials, considered as an effect of altercentric intrusion in L1VPT processing, has been replicated and extended. Further, the effect has even persisted under secondary task conditions where cognitive-resource tasks are added. In the original study, Samson et al. (2010) firstly observed that adults performed more slowly and made more errors on inconsistent trials compared to consistent trials when they were asked to take their own perspectives (i.e., the consistency effect on self-perspective trials). Later, a line of L1VPT-related experiments using the dot-perspective task has generalized the consistency effect under self-judgement circumstances. In addition to Surtees, Samson and Apperly's (2016) replication of Samson et al.'s (2010) findings, Surtees and Apperly (2012) extended the altercentric intrusion effect from adults to 6-10-years-old children. More importantly, the effect persisted when considering its relationship with cognitive load. The consistency effects remained when participants judged their own perspectives regardless of time pressure for responses (i.e., shorter-deadline of 600 ms compared with a long-deadline of 1200 ms) (e.g., Todd, Cameron, & Simpson, 2017), and also persisted when the dot-perspective task was performed together with a secondary task requiring executive-function resources (e.g., Qureshi et al., 2010). The two studies, then, demonstrated that tasks requiring cognitive resources did not influence the elicitation of efficient computation of others' visual perspectives on self-perspective judgements.

1.3 The Mentalising vs. Submentalising Debate

1.3.1 The Content of the Debate

Recent work has suggested that the processing of others' minds depends on two cognitive systems. One is a flexible system that enables us to explicitly reason about how others' mental states (beliefs, intentions, emotions, etc.) influence their behaviours. The other is an efficient system that allows us to automatically track others' mental states (Kovács, Téglás, & Endress, 2010; Low et al., 2016; Schneider, Nott, & Dux, 2014). Samson et al. (2010) suggested that the efficient system enables us to automatically track what someone else sees (i.e., L1VPT). However, there is obvious debate over whether and to what extent the efficient system is specialized (implicit and automatic mentalising account: e.g., Apperly, 2010) or domain-general (submentalising account, e.g., Heyes, 2014). Specifically, the debate was sparked by the adults' performance in the dot-perspective task of L1VPT. That is, the implicit mentalising account claims that the consistency effect on self-perspective trials is elicited by implicit and effortless computation of the avatar's visual perspective via connecting her/his line-of-sight with the discs. The submentalising account, alternatively, holds the view that the effect is evoked by merely attentional orienting produced by the directional but not agentive features of the avatar (e.g., head and/or body directions).

1.3.2 Implicit Mentalising Account

Researchers supporting the implicit mentalising account claim that the consistency effect on the self-perspective judgement is invoked by implicit computation of others' mental state of seeing. Specifically, if participants could easily understand others' visual information, then they would rapidly and efficiently track the avatar's visual perspective in the dot-perspective task even without the explicit judgements of others' perspectives.

With respect to the implicit mentalising account, researchers regard eye gaze as the key factor, which can be supported by the evidence showing the important role of eye gaze in mentalising-related processes. For example, Baron-Cohen, Campbell, Karmiloff-Smith, Grant, and Walker (1995) observed that 3 and 4-year-old children can infer a person's mental state of wanting a chocolate bar (i.e., desires) by tracking others' eye-gaze direction to the target. More importantly, some researchers have found a strong relationship between the eye gaze and L1VPT processing (Sodian et al., 2007). They reported that 14-month-old infants can make expectations about an agent's goal-directed action based on understanding whether or not the line of sight between the agent's eyes and an object is physically unblocked (i.e., L1VPT ability). Based on these studies, eye gaze can convey information about others' visual perspectives. Therefore, manipulation of visual access has been created to measure L1VPT processing tapped into the mentalising-related process.

Several studies have manipulated the gazer's line-of-sight as a novel way to clarify the mentalising versus submentalising debate. Among those studies, only one has found evidence for implicit mentalising (Furlanetto, Becchio, Samson, & Apperly, 2016).

Specifically, Furlanetto et al. (2016) manipulated the avatar's visibility to explore the debate by adopting transparent goggles (i.e., visible condition) and opaque goggles (i.e., invisible condition). The participants were checked to be able to associate different coloured goggles with the corresponding avatar's ability to see (i.e., in the seeing condition, the red goggles worn by the avatar were transparent, whereas in the non-seeing condition, the orange goggles worn by the avatar were opaque). The participants were instructed to judge their own perspectives or perspectives of the avatar wearing the different coloured goggles. The authors found that participants judged their own perspectives more slowly and less accurately in the inconsistent condition compared with the consistent condition, but the consistency effect was present in the seeing but not non-seeing conditions. The explanation of the discrepancy was that participants had different beliefs of the avatar's epistemic state of seeing via understanding the transparent and opaque features of the goggles, which then led to the connection of the gazer's line-of-sight with the disc(s) on the wall(s) in the seeing condition but the disconnection in the non-seeing condition.

The findings of this study cast doubt on the submentalising account claiming the role of directional information as it would predict the consistency effects in both visible and invisible conditions due to the identical directional features. Thus, participants can implicitly and efficiently compute the visual perspective of the avatar wearing transparent goggles even when they were not required to do so, which lent support to the implicit mentalising account. However, the study is limited on its own. Specifically, the study cannot rule out the carry-over effect between self- and other-perspective conditions as the two conditions presented in the intermixed block. Therefore, the consistency effect on self-perspective judgements may be contaminated by explicit judgements about others' perspectives. To explore whether L1VPT processing is implicit mentalising, it would be better to separate the seeing condition from the non-seeing condition.

1.3.3 The Submentalising Account

Researchers supporting the submentalising account have claimed that the consistency effect is elicited by domain-general mechanisms that are not specialized for processing of others' minds (e.g., attentional orienting, Heyes, 2014). Specifically, it is the directional features of the avatar that modulate participants' attentional shifts towards the number of dots on one side of the room. Therefore, on consistent trials of the dot-perspective task, the directional property of the centrally presented avatar oriented participants' attention towards the dot(s) on the target wall; whereas on inconsistent trials of the dot-perspective task, the directional property of the avatar oriented participants' attention neither to the dot(s) on the target wall nor to all the dots on both targeted walls. Then, the directional property of the centrally presented avatar may trigger more errors and slower response times in the inconsistent condition compared to the consistent condition. Related studies that tried to cast light on the debate by adding arrows as control stimuli relative to avatars and by manipulating the agent's line of sight via opaque barriers will be described in the next two subsections.

1.3.3.1 Submentalising and Arrows

Advocates of the submentalising account have cited literature showing that in addition to social stimuli, semi-social and/or non-social stimuli can also generate self-consistency effects (Nielsen, Slade, Levy, & Holmes, 2015; Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014; Todd et al., 2017; Wilson, Soranzo, & Bertamini, 2017). For example, Santiesteban et al. (2014) modified the dot-perspective task by adding new trials where the avatar was replaced with an arrow with similar low-level directional features (e.g., height and position). Self-consistency effects of comparable size were found in the avatar and arrow conditions, suggesting that the consistency effect in the dot perspective task may be triggered by domain-general processes such as attentional orienting. Furthermore, the attentional-orienting mechanism of L1VPT processing is also reflected by the findings of significant consistency effects in the dot-perspective task regardless of the sociality of the centrally presented stimuli (i.e., all the consistency effects are significant but different in magnitude: directional avatars (social stimuli) > directional arrows (semi-social stimuli) > directional, dual-coloured blocks (nonsocial stimuli)) (Nielsen et al., 2015). Even though arrows have been found to be able to trigger self-consistency effect as avatars could, there are limitations with such approaches (Nielsen et al., 2015; Santiesteban et al., 2014; Todd et al., 2017; Wilson et al., 2017). First, participants' expertise with arrows (from previous experiences of being exposed to arrows) may make them treat arrows as purposefully designed (by the experimenters) to prioritise some perspective on the scene that may be similar to the avatar-triggered L1VPT. Furthermore, whilst the arrow that Santiesteban et al. created has a directional property, it also potentially has animacy because its height, shape, colour distribution and area were matched to the avatar (also in Experiment 1 of Conway, Lee, Ojaghi, Catmur, and Bird's (2017) study). Indeed, studies show that adults may attribute mental states to simple geometric shapes (e.g., Surian & Geraci, 2012), suggesting that, rather than being submentalisers, adults may be supermentalisers. Furthermore, Pavlidou, Gallagher, Lopez, and Ferrè (2019) found compared to mismatching, matching the participants' body posture (i.e., facing left or right) with that of avatar's triggered greater consistency effect in the dot-perspective task, whereas arrow direction could not modulate the effect. These findings indicate that the avatar but not the arrow can lead to the computation of others' visual-spatial perspective as only the avatar is human-like. Thus, it is not clear that arrow-related findings can rule out the implicit mentalising account for L1VPT processing.

1.3.3.2 Submentalising and Visual Barriers

In addition to comparing avatar- and arrow-related self-consistency effects in the dot-perspective task, other researchers attempted to manipulate the gazer's visibility by using barriers to explore the mentalising vs. submentalising debate (Cole, Atkinson, Le, & Smith, 2016; Conway et al., 2017; Langton, 2018; Wilson et al., 2017). Failing to replicate Furlanetto et al.'s (2016) findings, these studies have found that the self-perspective-related consistency effect even persists when agents' 'non-seeing' conditions are imposed by using barriers. The lack of difference between the visible and invisible conditions demonstrates that directional information of the avatar instead of mentalistic processing of seeing elicited the consistency effect, supporting the submentalising account.

To render discs in the dot-perspective paradigm visible and invisible, Conway et al. (2017) manipulated an avatar's visual access by using a cloaking device or goggles that were worn by the avatar. In Experiment 1, the authors adopted visible and invisible telescopes within a cloaking device to render the seeing condition and non-seeing condition, respectively. In Experiments 2 and 3, the avatar wearing goggles with a transparent internal lens could see whereas the avatar wearing goggles with an opaque internal lens (i.e., the lens was covered by a blackout material) could not. Inconsistent with Furlanetto et al.'s (2016) findings, they found the consistency effects in both the visible and invisible conditions even though they had ruled out the carry-over effect by intermixing self- and other-perspectives. It may be because it was relatively difficult to grasp connections between certain barriers and epistemic state of seeing, which made it hard for participants to recognize the connections during the dot-perspective task. Additionally, it may be relatively hard for participants to regard the barrier scenario as the non-seeing condition, especially when they only had a limited time-period to grasp the novel scenario. To address the potential issue of the aforementioned barriers, Wilson et al. (2017) employed easily-recognizable blindfolds to render the discs invisible. However, the following points may be regarded as being potential interpretations for the finding of the consistency effect in the non-seeing condition. Self- and other-perspective trials were intermixed, and the alternate presentation of these two types of trials may lead to a carry-over effect. Consequently, participants may be explicitly tracking the avatar's visual perspective even though no related instruction was displayed.

Furthermore, these barriers may not evoke effective non-seeing scenarios as they occupied a relatively small part of the avatar.

Instead of using relatively small eyes-covered devices, Langton (2018) displayed a pair of big opaque boards between the gazer and the target discs to create the invisible scene and, additionally, replaced computer-generated avatars with photographs of real humans (Experiment 1) or with a gazer sitting face-to-face with participants (Experiment 2). The findings of both experiments spoke against implicit mentalising but supported the submentalising accounts by observing a significant consistency effect in the invisible condition of the dot-perspective task. However, the study also had the following limitations. First, the barrier manipulation in Experiment 1 may not have effectively created a non-seeing scenario. Specifically, in the non-seeing condition, the lengthy distance between the centrally presented gazer and the peripherally presented barrier could have led participants to perceive that the target discs still fell within the gazer's visual field, particularly under a limited response duration (2 seconds). Second, the revised dot-perspective task in Experiment 2 was distinct from the classic Samson et al.'s (2010) task. Specifically, an arrow cue that appeared behind the participant's head instructed the gazer to turn his head towards one of the two laterally presented monitors. Then, two seconds followed by the presentation of the arrow, the targeted discs were displayed on one or two lateral monitor(s) (see Figure 2). In the situation, the head turn may be accomplished before the appearance of the target discs, which may trigger an SOA variable. Importantly, the 'SOA' factor made the classic dot-perspective task similar to the stimulus-presentation mode of the classic Posner task (i.e., a well-known task tapping attentional orienting), which then, may trigger an attentional orienting effect instead of visual-perspective-taking-related processing.

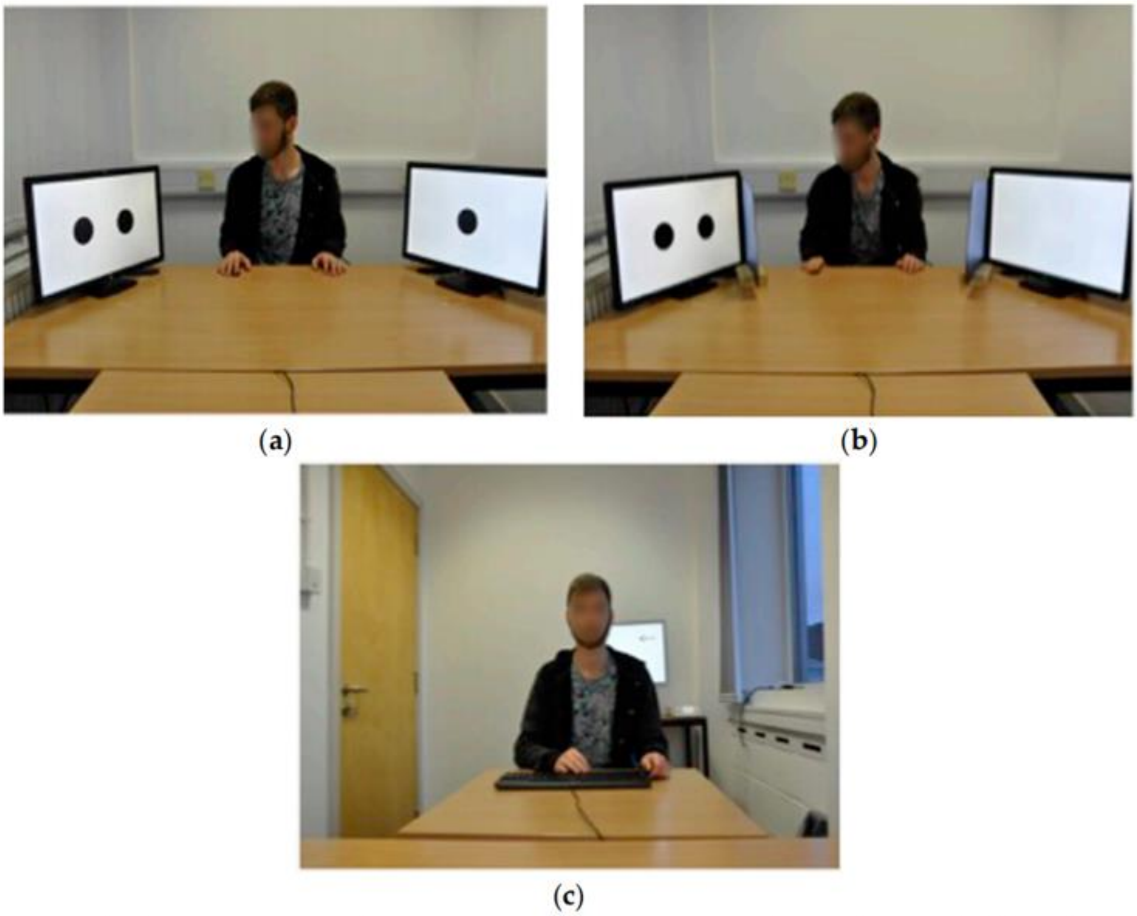

FIGURE 2 | The arrangement of the apparatus used in Langton's (2018) Experiment 2: (a) from the point of view of a participant in the "seeing" condition; (b) from the point of view of a participant in the "non-seeing" condition; (c) from the point of view of the gazer where an arrow cue instructing the gazer to direct his gaze to the left can be seen above the participant's left shoulder.

1.3.4 Dissociating the Mentalising from the Submentalising Accounts

Some researchers have attempted to dissociate the mentalising account from the submentalising account by contrasting relevant paradigms. Gardner, Hull, Taylor, and Edmonds (2018) attempted to dissociate the competing accounts by contrasting effects in a modified dot-perspective task with the Posner task. The Posner task has been widely used to measure attentional orienting, in which a directional cue is first presented, followed by a target with a stimulus onset asynchrony (SOA), and then, participants are required to detect the target location. Directional information conveyed by the cue oriented the participants' attention closer to the target location in the valid condition compared with the invalid condition, thereby, participants detected the target more accurately and quickly in the valid condition (i.e., cue-validity effect). In Experiment 1, Gardner and colleagues examined if reflexive attention orienting can sufficiently induce the self-perspective-related consistency effect that would be considered as automatic visual perspective-taking in the dot-perspective task. They adopted a revised dot-perspective task by eliminating the 'YOU-perspective' instruction from the original Samson et al. (2010) study. The novel dot-perspective task made the participants unaware that they were completing a perspective-taking task. Thus, in Experiment 1, removal of the 'YOU-perspective' instruction resulted in a non-significant consistency effect, demonstrating that the effect cannot be evoked merely by reflexive attention orienting. In Experiment 2, they used the Posner task with dot-perspective-task's stimuli to investigate whether the attentional orienting property of the avatar contributed to the consistency effect that was previously induced in the dot-perspective task. They found the cue-validity effect only for longer SOAs. Specifically, adults were faster to detect a target when the avatar was directed to the target (valid trials) compared to when the avatar was directed away from the target (invalid trials) when SOA was 600 ms but not 100 ms or 300 ms. The findings demonstrated that a voluntary rather than reflexive attention shift contributed to the consistency effect in the dot-perspective task. Taken together, the consistency effect in the classic dot-perspective task might be less automatic than first reported. Nonetheless, the discrepancy between visual-perspective-taking and attention-shift processes cannot be directly distinguished.

Previous findings revealed that compared to stance-maintained avatars (i.e., avatar's head and torso faced to the same wall, see Figure 3a), stance-averted avatars (i.e., avatar's head was oriented to one wall whereas the torso faced to the participant, see Figure 3b) induced an increased attentional orienting effect (e.g., Hietanen, 2002). Accordingly, Gardner, Bileviciute, and Edmonds (2018) hypothesized that avatar-stance may modulate attentional orienting but not visual perspective-taking. They attempted to distinguish the implicit mentalising from the submentalising accounts by manipulating avatar stance (stance-averted vs. stance-maintained, see Figure 3). Specifically, they explored whether avatar-stance could differently modulate the effect from visual-perspective-taking tasks (i.e., consistency effect in the dot-perspective task) and from attentional-orienting tasks (i.e., cue-validity effect in Posner task). Experiment 1 used the Posner cueing task to examine the cue-validity effect, finding that the target was more slowly to be detected in the invalid condition compared to the valid condition. The attentional orienting effect was modulated by avatar stance, which is reflected by the significant effect for stance-averted rather than for stance-maintained avatars. Experiment 2 adopted the dot-perspective task to replicate the classic consistency effect. More importantly, avatar-stance did not moderate the magnitude of the consistency effect in the classic visual-perspective-taking task. Accordingly, the dissociation between attentional orienting and visual-perspective-taking processes casts doubt on the submentalising hypothesis regarding the role of the directional cue but supports the implicit mentalising hypothesis.

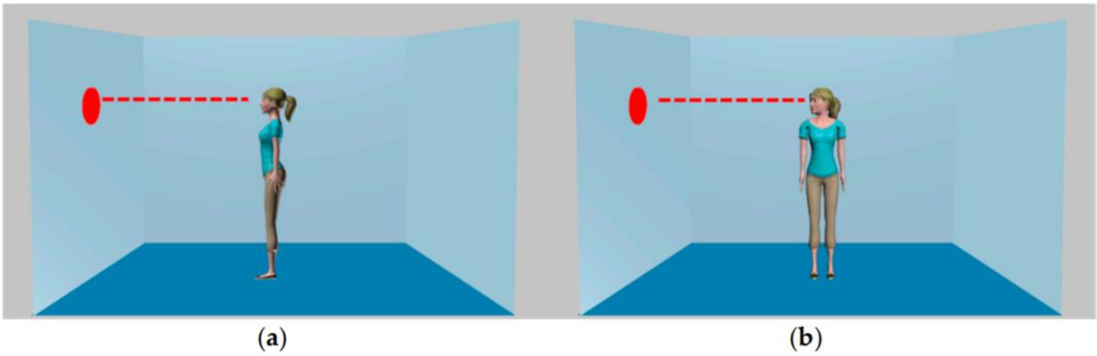

FIGURE 3 | Gaze-maintained (a) and gaze-averted (b) avatars used in Gardner et al.'s (2018a) Experiments 1 and 2, illustrating that there is unbroken light of sight between the target dot and the avatar in both cases (right facing avatars not illustrated).

1.3.5 The Implications of Investigating the Debate

The implicit mentalising vs. submentalising debate has important methodological and theoretical implications. In the methodological aspect, the debate challenges the effectiveness of the dot-perspective paradigm as a measure of L1VPT ability. It is important to find a universally recognised way to measure and clarify L1VPT processing, which can lay a foundation for connecting L1VPT with the later more complex ToM processes and understanding the related social communications. Theoretically, the debate raises the question about the efficient part of ToM processing system (i.e., whether people can effortlessly track others' mental states, e.g., Meert, Wang, & Samson, 2017). Practically, resolution of the debate is important because it is beneficial to further, understand related dysfunction in social behaviours in atypical individuals. For example, psychopathic patients have been found to have deficits in L1VPT ability, and their dysfunction in effortlessly taking others' visual perspective have been demonstrated to be correlated with their callous and criminal behaviours in real-world (Drayton, Santos, & Baskin-Sommers, 2018).

Chapter 2: Attentional Orienting

Tracking the spatial location of someone else's attention is critical for successful social interactions. People shift their attention frequently to allocate resources for more efficient processing. Numerous studies have examined the mechanisms of spatial attention using the Posner cueing task. Among these, many researchers have compared how well different directional cues (i.e., averted eye-gaze, head orientation, finger-pointing) shifts attention towards the target location.

2.1 Posner Cueing Task

Posner cueing task was created by Posner (1980) for measuring attentional orienting (see Figure 4). The crux of the task is that the directional cue can shift participants' attention towards the location directed by the cue (e.g., an arrow in Figure 4), even when the cue is irrelevant for the target. To examine the time-course of attentional orienting, SOA, the duration between the onset of one stimulus, S1, and the onset of another stimulus, S2, is also presented in the Posner cueing task. Specifically, the directional cue was presented with an SOA before the presentation of the target. Incorporation of SOA in the Posner task is important for distinguishing between voluntary attention and reflexive attention. Volitional/voluntary attention, a controlled/top-down process, happens at the later temporal stage. Reflexive attention, an automatic/bottom-up process, happens at the early temporal stage (Egeth & Yantis, 1997; Hill et al., 2010; Müller & Findlay, 1988; Müller & Rabbitt, 1989). For instance, by measuring gaze cueing effect (i.e., gaze-triggered orienting effect), volitional attention was observed for SOAs ranging from 500 to 1000 ms. In contrast, reflexive attention was observed for SOAs ranging from 100 to 500 ms (Hill et al., 2010).

The relationship between the direction of the cue and the position of target determines the ‘Validity’ of the cue. As is shown in Figure 4, on the valid trials, the cue was directed to the target location. On the invalid trials, the cue was directed to the opposite direction of the target location. On the neutral trials, the plus sign did not show any directional information.

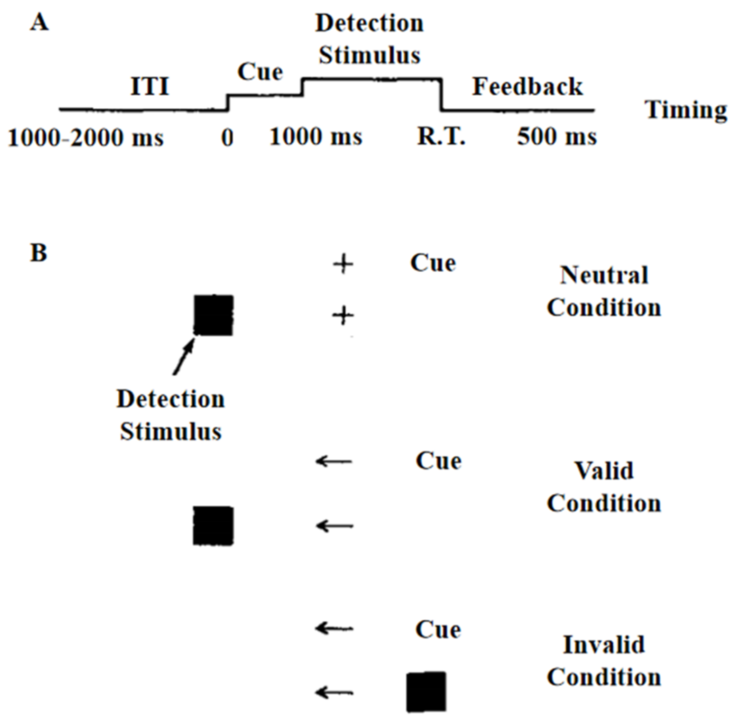

FIGURE 4 | (A) Overview of a representative test trial of the Posner cueing task created by Posner (1980). (B) Stimulus reprentations in netural, valid and invalid conditions.

The basic finding of the Posner task is that participants are slower and less accurate to detect the target on invalid trials compared to valid trials, which is known as the cue-validity effect (Posner, Nissen, & Ogden, 1978; Posner, Snyder, & Davidson, 1980). In contrast, the neutral cue did not influence response times for target detection on neutral trials.

The interpretations for the above findings are as follows. Compared with the invalid condition, in the valid condition, the participants' attention was oriented closer to the targeted object by the cue's directional information, which made participants respond more accurately and quickly in the valid condition. Accordingly, the findings demonstrate that attentional orienting can be modulated by the cue's directional information.

There are two main differences between the Posner task and the dot-perspective task. First is the stimulus presentation. The peripheral target is presented with an SOA after the central cue onset in the Posner task, whereas the central stimulus and the peripheral disc(s) are simultaneously displayed in the dot-perspective task. Second is the task instruction. Participants are required to simply detect the target in the Posner task, whereas they have to match their own or the agent's visual perspective with the subsequent scene in the dot-perspective task.

2.2 Attentional Orienting Triggered by Eye Direction and Head Orientation

Facial features of a person can capture another person's attention in social communications due to the face's social and biological functions. With regard to the directional cue, both eye direction (e.g., Driver et al., 1999) and head orientation (e.g., Langton & Bruce, 1999) represented in a human face have been shown to evoke attentional orienting.

2.2.1 Attentional Orienting Triggered by Eye Direction

Gaze direction is an effective communicative channel in social interactions, and people can discriminate the eye direction from early infancy (Vecera & Johnson, 1995). Psychological researchers have observed that people can track someone else's social attention by following another's eye-gaze, in both infants (e.g., Willen, Hood, & Driver, 1997) and adults (e.g., Baron-Cohen, 1995; Macrae, Hood, Milne, Rowe, & Mason, 2002).

Several studies have used the gaze cueing paradigm (a modified paradigm based on the Posner task) to examine the orienting effect elicited by the eye direction of a centrally presented frontal-view face. Evidence is accumulating that observing averted eye-gaze can trigger reflexive attention in adults (e.g., Driver et al., 1999; Friesen, Moore, & Kingstone, 2005; Frischen, Bayliss, & Tipper, 2007; Kuhn & Kingstone, 2009). When investigating gaze-cue orienting effect induced by averted eye-gaze of a centrally presented face, the target is often presented 100, 300 and 700 ms after gaze-cue onset (stimulus onset asynchrony, SOA) in some studies (Bayliss, Di Pellegrino, & Tipper, 2005; Driver et al., 1999). These durations can cover the operational durations of reflexive and volitional attention orienting effects, if they exist (Egeth & Yantis, 1997; Kingstone, Tipper, Ristic, & Ngan, 2004; Müller & Rabbit, 1989; Posner & Cohen, 1984).

The gaze cueing paradigm produces a gaze cueing effect. For example, in a classic experiment (Driver et al., 1999), the frontal-view face with eyes looking left or right was presented 100, 300, or 700 ms before the presentation of the target letters ('T' or 'L'). Participants were instructed to detect the letter as quickly and accurately as possible. They performed faster in detecting the letter on the valid trials compared to invalid trials, even at relatively short SOAs (i.e., 300 ms), suggesting reflexive attention. Moreover, the gaze cueing effect remained stable even when the target was presented contrary to intentions, namely, participants were told that the target on the opposite side of the eye direction appeared four times as often as the target on the same side of the eye direction. In other studies, a gaze cueing effect was also observed with more strict instructions (e.g., Friesen et al., 2005). That is, participants were instructed to focus their eyes on the nose of the central face throughout the experiment; and were told that the central cue was a non-predictive cue for the target location and should be ignored. Nonetheless, participants were still aware of the central cue, which was reflected by the finding of slower target detection in the invalid condition than in the valid condition.

Furthermore, findings from other studies suggested the generalization of the orienting effect triggered by the eye direction in a face. Specifically, the orienting effect can be triggered by the averted eye-gaze from a central face, be it a photograph of a real human's face (e.g., Driver et al., 1999), or a cartoon human-like face (e.g., Bayliss et al., 2005), or a schematic-drawing face (e.g., Friesen & Kingstone, 1998). Additionally, some studies found 6-14-year-old children's gaze-elicited orienting effect regardless of another person's age (e.g., Van Rooijen, Junge, & Kemner, 2018). Taken together, tracking a person's gaze direction can produce another person's reflexive attentional orienting.

2.2.2 Attentional Orienting Triggered by Head Orientation

In addition to gaze direction, head orientation provides another pivotal cue for observing your partner's attention during the conversation. People can use head orientation as an effective cue to track adults' attention from early infancy (Corkum & Moore, 1995, 1998; Lempers, 1979). In a social environment, head orientation is usually presented simultaneously with the eye direction as when a person's eyes are looking to a direction; a person's head is automatically turned to a lateral-view to adapt to the eye direction change. In the initial account, Perrett and colleagues have claimed that the role of the eye direction overrides that of the head orientation (Perrett & Emery, 1994; Perrett, Hietanen, Oram, & Benson, 1992). Specifically, when a person's eye-gaze is clear, eye direction rather than head orientation induces someone else's attentional orienting; whereas head orientation plays an important role in triggering attentional orienting when eyes are obscured or faces appear in a distance.

However, evidence from subsequent studies has suggested that head orientation can modulate eye-direction processing. In the beginning, Wollaston (1824) noted that head orientation of the gazer could affect the observer's discrimination of the eye direction based on the portrait painting, which was supported by later behavioural findings (e.g., Anstis, Mayhew, & Morley, 1969; Gibson & Pick, 1963). For example, a person standing in front of the observer turned his head 30 degrees to the observer's right with his eyes looking back to the observer (i.e., frontal-view eyes), and then, the person's eye direction was perceived a bit left than it really was. Moreover, a study by Langton (2000) sought to directly investigate the relationship between eye direction and head orientation discriminations. Langton used a

Stroop-type interference paradigm, in which a person's head orientation and eye direction were opposite, and participants were required to judge either head orientation or eye direction. It was observed that inconsistent head orientation triggered slower judgement for eye direction, and inconsistent eye direction triggered slower judgement for head orientation. Additionally, the interaction effect between head orientation and eye direction remained stable even when the above paradigm was cross-modally modified (i.e., participants were asked to respond to spoken directional words rather than directional feature from face stimuli). These findings suggested the mutual influence between the two cues.

An investigation regarding attention orienting has suggested a person's head cue can also work as an effective visual attention cue for an observer to produce reflexive attentional orienting). In Langton and Bruce's (1999) Experiments 1 and 2, the central cue, a person's head directing to left, right, up or down, was presented with an SOA (i.e., 100, 500, 1000 ms), which was followed by the target letter. Participants performed the Posner's cueing task based on instructions similar to a gaze cueing experiment. The authors observed that participants were slower to detect the target letter when it was followed by an invalid head cue than when it was followed by a valid head cue. The head-cue validity effect appeared immediately after the head cue onset (i.e., 100 ms), and additionally, happened even when the cue was non-predictive and should be ignored. These findings were in accordance with a reflexive attention mechanism, indicating that head cue can trigger a reflexive orienting effect like gaze direction. Moreover, these observations were obtained under a limited circumstance, namely, when head was centred but not when head was inverted (Langton &

Bruce, 1999, Experiment 4).

Other studies have focused on the relationship between the head orientation and eye direction in attentional orienting. Different from the combination of head orientation and eye direction in Langton and Bruce's (1999) study, Hietanen (1999) manipulated the two cues independently. There were four types of facial stimuli (see in Table 1) set in the Posner cueing task: (1) a frontal-view face with straight eye-gaze (neutral condition), (2) a frontal-view face with averted eye-gaze (i.e., a head-front-gaze-averted face), (3) a lateral-view face with a compatible gaze direction (i.e., a head-turned-gaze-maintained face), (4) a lateral-view face with eyes looking back to an observer (i.e., a head-turned-gaze-averted face, the face was coined based on the mutual relationship between eye direction and head orientation). Hietanen obtained three main findings. First, the result replicated the classic gaze cueing effect, namely, a cue-validity effect for head-front-gaze-averted faces. Second, a significant cue-validity effect was observed for head-turned-gaze-maintained faces. Third, an effect was triggered by head-turned-gaze-averted faces, which was contrary to the classic cue-validity effect. Namely, the lateral-view face with eyes looking back to the viewer induced longer response times to detect the asterisk on congruent trials relative to incongruent trials. The author explained that the viewer's attention seemed to be oriented according to someone else's gaze direction with reference to head orientation, but only when gaze direction was incompatible with head orientation. When gaze direction and head orientation were compatible, head-turned-gaze-maintained faces appeared to be irrelevant to the observer.

Table 1. Stimuli used in Hietanen's (1999) study and Qian et al.'s (2013) study.

Type of Faces	Head-front-gaze-averted	Head-turned-gaze-maintained	Head-turned-gaze-averted
Hietanen's (1999) study			
Qian et al.'s (2013) study			

Qian, Song and Shinomori (2013) also adopted the Posner task to measure cue-validity effects generated by eye-head cues (see in Table 1), but result patterns were different from Hietanen's (1999) findings. The magnitude of cue-validity effects triggered by faces was sorted as follows: head-turned-gaze-maintained faces > head-front-gaze-averted faces > head-turned-gaze-averted faces. The possible interpretation was that compared with head-front-gaze-averted faces, compatible gaze and head directions in head-turned-gaze-maintained faces increased the perceived angle of gaze direction towards the left or right, whereas incompatible gaze and head directions in head-turned-gaze-averted faces reduced the perceived angle of gaze direction more close to direct gaze. Thus, the bigger the perceived angle of gaze direction was, the stronger cue-validity effect was evoked. The findings demonstrated that gaze cueing effect was modulated by gaze perception with reference to head orientation. The inconsistent findings in the two studies may be due to the following two distinctions. First, SOA durations in Hietanen's (1999) study are shorter than those in Qian et al.'s (2013) study. In Hietanen's (1999) study, SOAs were short (100, 150, and 200 ms in Experiments 1; 170 and 220 ms in Experiments 2 and 3). In contrast, SOAs (i.e., 300 and 600 ms) were longer in Qian et al.'s (2013) study. Second, the perceived angle of gaze direction seems to be different between the two studies as illustrated in Table 1, especially for head-front-gaze-averted and head-turned-gaze-averted faces. For head-front-gaze-averted faces, the perceived angle of gaze direction in Hietanen's (1999) study looks bigger than that in Qian et al.'s (2013) study. For head-turned-gaze-averted faces, the gazer's eyes seem to look left to a small extent in Hietanen's (1999) study whereas the gazer appears to be looking ahead to the observer in Qian et al.'s (2013) study. Accordingly, results may change if time-related information and/or the perceived angle of gaze direction are different. Additionally, the classic gaze cueing effect in some prior studies (Bayliss et al., 2005; Driver et al., 1999) is observed when a head-front-gaze-averted face is presented immediately after a frontal-view face without eye information. Thus, an apparent motion appears to be generated for the head-front-gaze-averted face. Whether dynamic property of facial cues modulate cue-validity effect when cue type (i.e., head-turned-gaze-maintained, head-front-gaze-averted and head-turned-gaze-averted faces) is involved, and whether the dynamic interacts with the aforementioned time information and/or the perceived angle of gaze direction need further investigations.

In summary, a number of studies have indicated that averted eye-gaze is an effective spatial code to produce reflexive attention orienting. Nonetheless, investigations about the role of head orientation in shifting others' attention are relatively limited. Several research projects have connected the head orientation with other directional cues (e.g., eye gaze), suggesting that eye direction and head orientation can be mutually influenced. However, the role of a lateral-view head with compatible and conflict eye directions in attentional orienting remains unclear, which warrants further investigation.

2.2.3 Attentional Orienting and Perspective-taking Using Finger-pointing Cues

In addition to eye and head directions, the orientation of finger-pointing can also provide a crucial and effective signal in social interactions. Deictic gestures show or present a reference by arousing attention to an object (e.g., Liszkowski, 2008). As the most important component of the deictic gesture, finger-pointing appears early in development and infants at 11-months of age are already able to use it for communicative purposes (Crais, Douglas, & Campbell, 2004).

In addition to spatial attention induced by implied body gestures (e.g., throwing the ball), a line of research projects focusing on the specific gesture (i.e., hand gesture) have found the finger-pointing gesture can automatically convey directional information. For instance, Langton, O'Malley, and Bruce (1996) used the Stroop-type interference paradigm to explore both the influence of hand gesture and the influence of spoken words on utterance comprehension. In the study, there were four kinds of hand gesture (i.e., gesturing to up, down, left, right) and four spoken words corresponding to the direction of the four hand gestures, and participants were instructed to identify the direction of the hand gesture or the spoken word in half of the trials, respectively. Importantly, the to-be-ignored hand gesture modulated recognition of the spoken directional words, which was reflected by the cross-modal interference (i.e., response time for spoken directional words was slower when conflicting hand direction was simultaneously presented compared with when compatible hand direction was simultaneously presented) (Langton et al., 1996). However, the limitation of that study was that hand gesture was not controlled. Specifically, left and right gestures were hand-pointing gestures while up and down gestures were index-finger-pointing gestures. More importantly, Langton et al.'s (1996) conclusion might also be qualified by the fact that the gesturer's eyes' and head's directional information influenced the processing of the spoken directional words in addition to pointing gestures or work independently.

To address the issue, Langton and Bruce (2000) considered head/eye-gaze cues (i.e., head with compatible eye direction, looking up or down) by connecting their roles with the effect of a finger-pointing gesture. The researchers used a Stroop-type interference paradigm to find that head/eye-gaze and finger-pointing gestures can be mutually influenced. Participants responded more slowly to judge the direction of pointing gestures when the direction of the gesture was opposite to that of head/eye-gaze cues compared with when all cues' directions were the same. Reciprocally, participants were slower to judge head orientation when its direction was incompatible with pointing direction relative to when its direction was compatible with pointing direction. In summary, these findings indicated that people could automatically analyse the directional information of head, eyes as well as gestures, and the recognition of head/eye-gaze and gesture directions could evoke bi-directional interference effects. The findings, altogether, demonstrate the directional feature of hand-pointing action plays an important role in attentional shift.

However, in the aforementioned studies, eye direction was combined with a head orientation, meaning that effects related to the relationship between eye direction and finger-pointing direction cannot be clarified. In a study by Doherty and Anderson (1999), the researchers directly compared the perception of gaze direction with that of finger-pointing direction in preschool children. An experimenter sat on the opposite side of a child, and four objects (i.e., a cake, a balloon, a cup, and an airplane) were displayed at his top-left, top-right, bottom-left or bottom-right, respectively. In a looking-where task, the experimenter moved his eyes to look at an object, and the child was then required to answer the question 'Which object am I looking at'. In a pointing-direction task, the experimenter pointed to one of four objects, and the child was then instructed to answer the question 'Which one am I pointing to'. The participants responded by naming or pointing the targeted object. Interestingly, the authors found superior performance for the pointing-direction task compared with the looking-where task. Many children passed the pointing-direction task but failed the looking-where task, but no one demonstrated the opposite result pattern. The findings underscored the superiority of finger-pointing relative to eye-gaze as a directional cue.

Gregory, Hermens, Facey, and Hodgson (2016) also found an advantage of finger-pointing compared to eye-gaze for young children but in the same task. In their task, the uninformative cue was centrally presented before the presentation of the target (i.e., Buzzy

Bee) with the SOA (i.e., 100 or 500 ms), which was similar to the Posner task. During the task, the participants were required to track Buzzy Bee, and their eye movements were recorded. Although the cues were non-predictive, finger-pointing triggered a stronger cue-validity effect (i.e., slower eye movement responses to invalid cues compared with valid ones) than eye-gaze did in 3-5-year-old children. One possible explanation was that young children learned the adult's hand cues' connection with the targets earlier than other cues such as eye-gaze cue because hand gestures were more salient. However, whether the superiority of finger-pointing perception relative to gaze perception can be found in adults, and how directional information conveyed by eye-gaze and finger-pointing cues work in the Posner task tapping attentional orienting remain unknown. Thus, these issues warrant further explorations.

Beyond the many studies highlighting the attentional orienting properties of finger-pointing, there are a few studies that have connected perspective-taking with hand-related cues. Fischer and Szymkowiak (2004) used a target detection task to find a significant cue-validity effect triggered by finger-pointing, and the authors explained that finger-pointing may convey another person's intention of actions (Fischer & Szymkowiak, 2004). However, whether the inferred intention of someone else's actions represented by finger-pointing can be observed in the mentalising-related task (i.e., dot-perspective task) is largely unknown. Recently, von Salm-Hoogstraeten, Bolzius, and Müsseler (2020) investigated whether perspective-taking or referential coding mechanism played a role in triggering spontaneous response tendencies of human beings from someone else's view via the reference of hand position. They adopted the avatar-Simon task, in which the participants were required to categorize the two-coloured discs, and the avatar's hand position was required to be ignored even if it was the cue for one disc. It was observed that the participants took the avatar's spatial perspective even though it was task-irrelevant (i.e., the avatar-Simon effect). Furthermore, the effect seemed to be present only when the hand position of the participant corresponded to that of the avatar. However, the avatar in the study lacked facial features (e.g., gaze direction). It is possible that the findings may turn out differently if researchers adopted a real human's face with representations of having a mind like clear gaze direction, and compared its effects on both attentional orienting and implicit mentalising with the corresponding effects triggered by finger-pointing. Furthermore, finger-pointing with manipulations of human's visual access (finger-pointing with eyes-opened faces vs. finger-pointing with eyes-covered faces) may provide a novel and effective way to cast light on the implicit mentalising vs. submentalising debate as line-of-sight manipulation has been widely used explore the mechanism of L1VPT processing.

Chapter 3: Study 1 (Pilot): Experiments 1 and 2

Overall, there is an ongoing debate between the implicit mentalising account and the submentalising account. This pilot study attempted to explore whether the consistency effect found in the classic dot-perspective task designed by Samson et al. (2010) can be replicated and robustly produced when replacing computer-generated avatars with real human's faces. Experiment 1 was the first to replicate and reconfirm whether the consistency effect is an artefact of processing biases arising from the spatial layout of the discs on the wall. Experiment 2 explored if the self-consistency effect was robust, occurring even when a photograph of a real human face was used.

3.1 Experiment 1: Overview and Hypotheses

Experiment 1 looked into the possibility that the consistency effect documented by researchers using the dot-perspective task might have resulted from the spatial layout of the discs on the wall. In the inconsistent trials of the task, discs were often spread across two walls, whereas in the consistent condition discs were always on one wall. Differences in layout could mean that the consistency effect may be merely due to slower processing of discs spread across two walls. If the consistency effect is solely due to discs' spatial layout rather than irrelevant perspective interference, there should be a consistency effect both when the avatar is in the middle of the room and when a rectangle (a non-directional cue) is in the middle of the room. To date, Samson et al. (2010, Experiment 3) have been the only team to show that—even when the disc layout is the same when an avatar or a rectangle was in the middle of the room—the consistency effect was observed only when an avatar was present. Given that such selectivity in the consistency effect relative to spatial layout information is a critical indicator that distinguishes the predictions of the implicit mentalising account from the submentalising account, it is noteworthy that this experiment was the first to attempt a direct replication of Samson et al.'s important findings. According to the implicit mentalising account, the consistency effect on self-perspective judgments should transcend biases due to spatial configuration, and show up only when an avatar was depicted in the room.

3.1.1 Participants

Twenty-three female participants (mean age: 22.75 years; age range: 18-29 years) from the general student population of Victoria University of Wellington (VUW) volunteered to participate in the current study. All participants had a normal or corrected-to-normal vision. They were naïve to the purpose of the study. The full research (application ID #0000026509) was approved by the School of Psychology Human Ethics Committee under the delegated authority of Victoria University of Wellington's Human Ethics Committee, and was conducted in accordance with the World Medical Association's Declaration of Helsinki. Each participant was taken to a quiet room and asked to complete the information sheet and consent form firstly. Participants whose accuracy was less than 60% in any experimental condition were excluded. Based on the standard, seven participants were eliminated, leaving data from 16 participants for further analysis. The sample size of 16 was the same as that of Samson et al.'s (2010) third experiment.

3.1.2 Stimuli and Procedure

The same stimuli as used in Samson et al.'s (2010) third experiment were presented on a 19-inch monitor with Open-Sesame 3.2.5 software (Psychology Software Tools). Stimuli were downloaded from an open-access webpage (please see https://figshare.com/articles/ Level_1_ Visual_ Perspective_Taking_Task/1455943 for details). The stimulus was an image showing a room (width × height, 11.2° × 5.6°) with left, back and right walls. There was either a female avatar (0.6° × 4.6°, facing either the left or right wall) or a rectangle (0.6° × 4.6°) presented in the middle of a room. There were red discs (0.5° × 1.0° each, 4.0° from the avatar) on one or two walls (see examples in Figure 5A and 5B). In the scene, the avatar matched the rectangle in terms of height and area, and the eyes of the avatar were as high as the discs. There were 7 kinds of disc distributions (see Figure 6); each distribution was equivalently combined with an avatar (50% of trials facing to the left wall and 50% of trials facing to the right wall) or a rectangle (50% of trials had left green side and right purple side and 50% of trials had left purple side and right green side). On half of the trials, the participant and the avatar could see the same discs (consistent condition), whereas the participant could see the discs that were invisible to the avatar on the other half of trials (inconsistent condition).

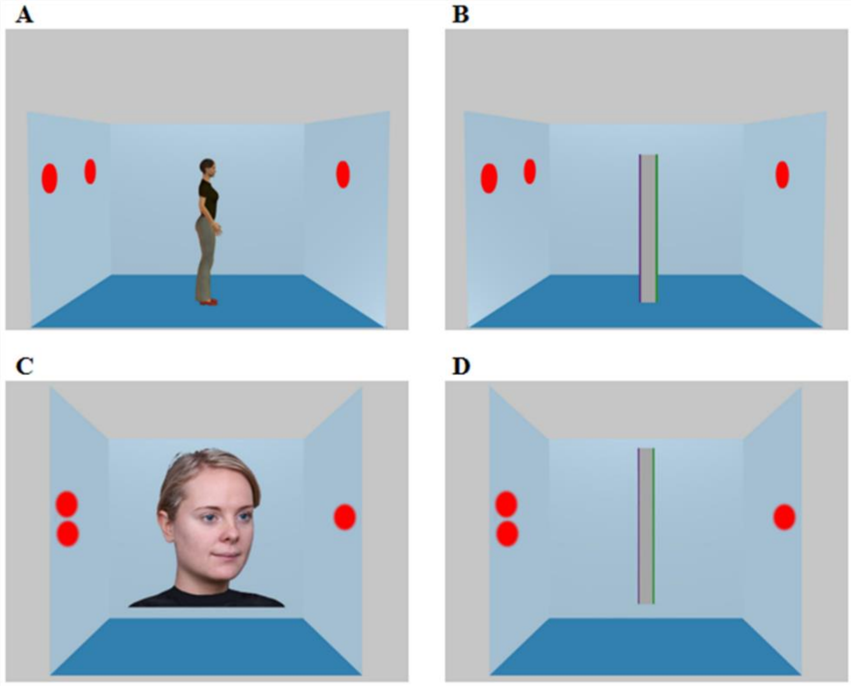

FIGURE 5 | Examples of room images used in Experiment 1 (A: Avatar, B: Rectangle), Experiment 2 (C: Face, D: Rectangle).

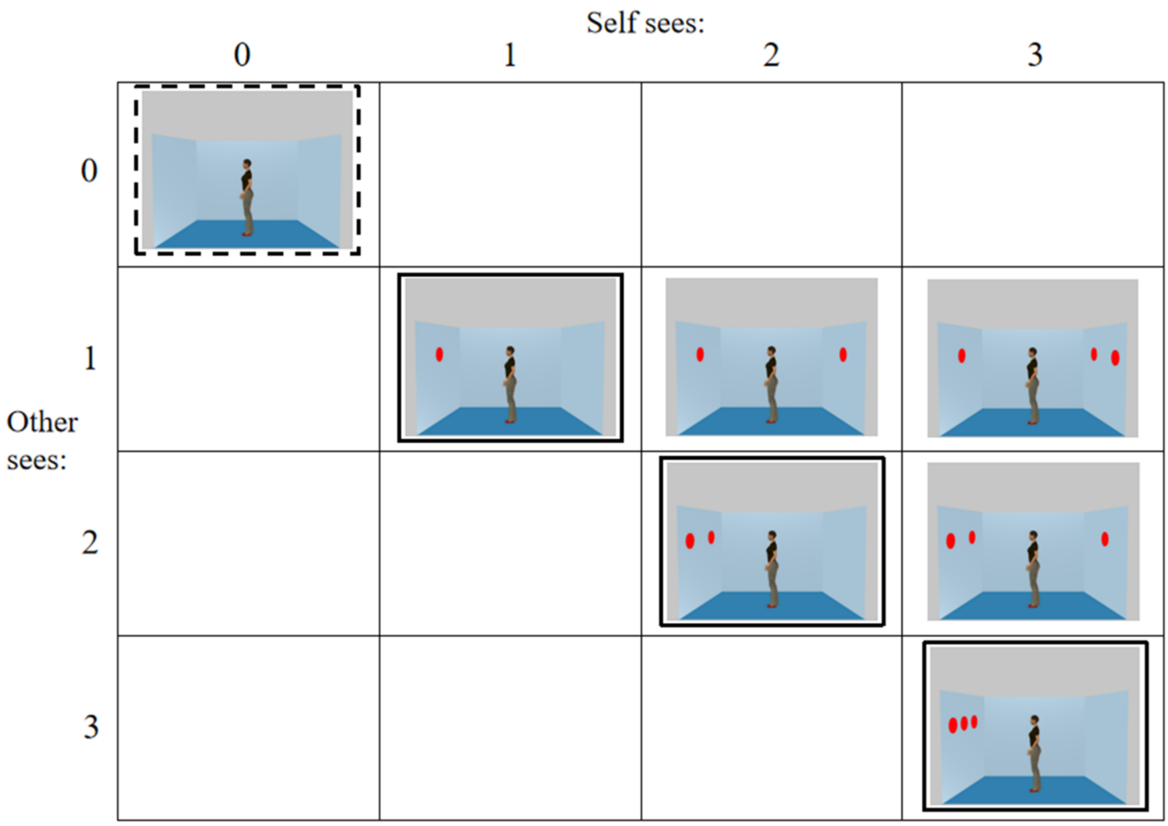

FIGURE 6 | The disc distribution when the avatar was facing towards the left wall in Experiment 1. Images circled with a dashed line were presented as stimuli of filler trials; images circled with a full line were presented as scenes on consistent trials; and the remaining images were presented as scenes on inconsistent trials.

Participants were seated approximately 72 cm from the computer screen. In the formal experiment, each participant was asked to complete the dot-perspective task shown as per Figure 7. At the beginning of each trial, a fixation cross was presented for 750 ms. Then, 500 ms later, the word “you” (i.e., meaning to judge self-perspective and ignore the central stimuli of the room) appeared for 750 ms, followed by a 500-ms-blank. Then, a digit (0-3) was presented for 750 ms. Then, 500 ms later, a scene with an avatar standing in the middle of a room and dot(s) presented on one or two lateral wall(s) was presented for 2 seconds. When the scene appeared, the participants were asked to judge whether the previously presented word-and-digit matched the scene that the participants saw. The participants needed to answer the question by pressing the relevant number key on the computer keyboard (pressing ‘1’ for “yes” response/matching using the forefinger of the left hand, pressing ‘2’ for “no” response/mismatching using the forefinger of the right hand). They were instructed to respond as quickly and accurately as possible.

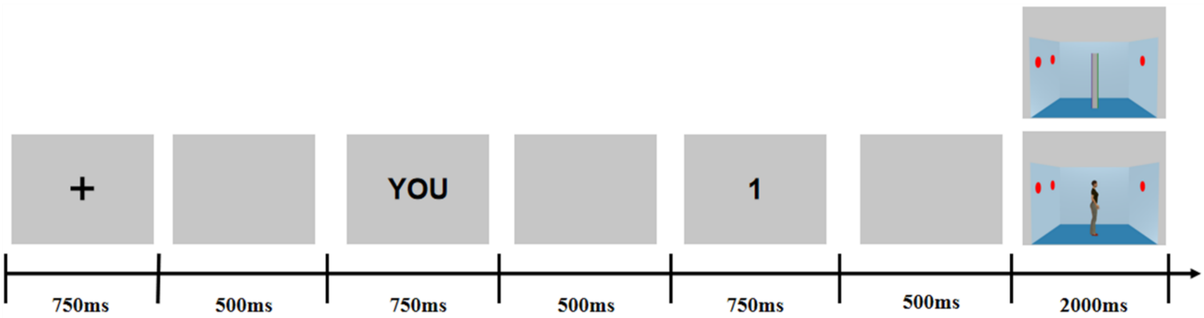

FIGURE 7 | Overview of a representative test trial of Experiment 1.

Samson et al.’s (2010) results indicated that ‘time’ (comparing the first half of the trials vs. second half of the trials) did not modulate the consistency effect. Bukowski and

Samson (2017) also found that the effect could be successfully revealed using an abbreviated version of the task. Thus, Experiment 1 adopted the short version of the task (104 trials, half of a total number of trials of previous experiments) for the replication, and would be believed to still detect automatic L1VPT. Additionally, only self-perspective trials were used to exclude potential carry-over effects generated by switching between self- and other-perspective.

Every participant completed 48 trials (12 consistent with the avatar, 12 inconsistent with the avatar, 12 consistent with a rectangle, 12 inconsistent with a rectangle) in both matching ("yes" response) and mismatching ("no" response) trials. There were 4 filler trials in addition to every 48 trials. In filler trials, there would be an avatar or rectangle in the centre of the room but without discs on the wall, which consisted of an equal number of matching and mismatching trials. Overall, there was a block of 26 practice trials and 2 blocks of 52 test trials (48 test trials and 4 filler trials). In order to avoid having more than 3 consecutive trials of the same condition in a block, the order of the trials was pseudo-randomised, which was fixed across participants. The presentation order of the blocks was counterbalanced across participants, and the avatar and the rectangle were randomly intermixed in each block.

3.1.3 Data Analysis

A 2 (Central stimulus type: Avatar vs. Rectangle) × 2 (Consistency of the perspective: Consistent vs. Inconsistent) repeated measure ANOVA analysis was performed, and accuracy and reaction time in the dot-perspective task were recorded as the key dependent variables. In keeping with Samson et al.'s (2010) study, only data from matching trials were analysed; the researchers argued that on mismatching consistent trials, the number cue did not correspond to anyone's perspective, making them easier to be processed than matching trials.

3.1.4 Results

In terms of accuracy, there were no significant main effect of Central stimulus type ($F(1,15) = 0.032$, $p = 0.86$, $\eta_p^2 = 0.002$), of Consistency ($F(1,15) = 0.055$, $p = 0.82$, $\eta_p^2 = 0.004$), or their interaction effect ($F(1,15) = 1.77$, $p = 0.20$, $\eta_p^2 = 0.11$). The mean accuracy of participants was over 90% in each experimental condition (see Table 2).

Table 2. Behavioural Results [Mean (Standard Error)] in Conditions of Experiments 1.

Experiment 1	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent
Avatar	96.40	94.30	687.46	772.30*** >
	(1.90)	(1.80)	(33.92)	(38.96)
Rectangle	94.80	96.40	727.63	748.23
	(1.80)	(1.30)	(35.01)	(35.60)

> Indicates a significant increase from the consistent condition; < Indicates a significant decrease from the consistent condition.
* p < .05; ** p < .01; *** p < .001.

In terms of response time, three participants were eliminated because they made more than 30% errors at least in one condition. ANOVA analysis showed a significant Central stimulus type x Consistency interaction effect ($F(1,15) = 11.61$, $p = 0.004$, $\eta_p^2 = 0.44$). Paired-sample t tests indicated that there was a significant Consistency effect when the avatar was in the middle of the room ($t(15) = -4.69$, $p < 0.001$), with an 85 ms advantage in the consistent condition, but no significant Consistency effect when the rectangle was in the middle of the room ($t(15) = -1.73$, $p = 0.105$). Specifically, when the avatar was in the middle of the room, participants were slower to respond in the inconsistent condition (mean ± standard error = 772.30 ± 38.96 ms) compared to the consistent condition (687.46 ± 33.92 ms). In contrast, when the rectangle was in the middle of the room, participants judged their own perspective as quickly in the consistent condition (727.63 ± 35.01 ms) as in the inconsistent condition (748.23 ± 35.60 ms) (see Table 2 and Figure 8A).

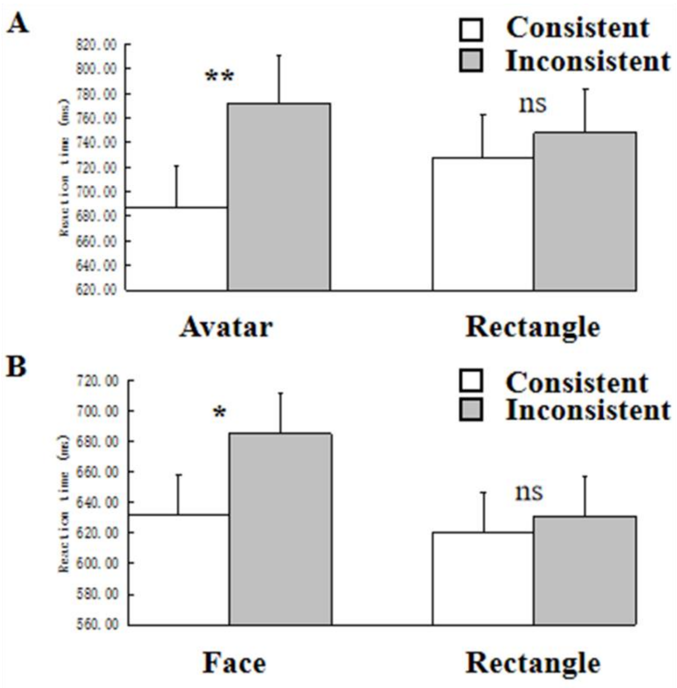

FIGURE 8 | Reaction times in each condition of Experiments 1 (A) and 2 (B). Means and standard errors are shown; error bars refer to +1 standard error of mean; * $p < .05$, ** $p < .01$.

3.1.5 Summary

The finding of Experiment 1 showed the consistency effect on self-perspective judgments only when the avatar was centrally presented in the room, which replicated

Samson et al.'s (2010) results. These findings rule out two alternative explanations of the consistency effect. First, the consistency effect did not result from the spatial layout of disc distribution, which was reflected by the rectangle-elicited non-significant difference between performance when all the discs were presented on the same wall (consistent condition) and when the discs were presented on separate walls (inconsistent condition). It could be interpreted that the rectangle, a non-directional cue, triggered neither attentional orienting nor automatic visual perspective-taking, thereby, only the spatial layout of disc distribution left in the dot-perspective task. Thus, the above non-significance ruled out the possibility that the spatial layout of disc distribution might contribute to the generation of the consistency effect. Second, the consistency effect on self-perspective trials was elicited without the carry-over effect of other-perspective trials. The resulting pattern, to a certain extent, indicated that participants automatically computed the avatar's visual perspective even though they were not asked to do so, which cast doubt on a perspective-switching explanation suggesting that a self-consistency effect is triggered by other-perspective trials. In summary, the findings' resulting pattern of Experiment 1 may support the implicit mentalising account for automatic L1VPT proposed by Samson et al. (2010). Furthermore, the above findings were documented in the short version of the task, which was consistent with previous studies (Samson et al., 2010; Bukowski & Samson, 2017), demonstrating that the dot-perspective paradigm has integrity even in an abbreviated form.

3.2 Experiment 2: Overview and Hypotheses

Advocates of implicit mentalising account propose that automatic L1VPT may be triggered by an agent's line of sight. It is important to gauge whether automatic L1VPT can be detected using real stimuli such as photographs of actual human faces where there is clear line-of-sight information to trigger automatic computation of the gazer's visual perspectives. Experiment 2 used images of a real face as the central stimuli in the dot-perspective task. If Samson et al.'s (2010) findings are robust and there is explanatory breadth in the implicit mentalising account, the consistency effect should be present when a face but not a rectangle is presented in the room.

3.2.1 Participants

A new set of 19 female individuals (mean age: 21.19 years; age range: 18-28 years) from the general student population of VUW volunteered to participate in the experiment. All of them had normal or corrected-to-normal vision. They were naïve to the purpose of the study. Three participants who did not take part in the formal experiment were eliminated from the analyses, and data from the remaining 16 participants were entered for further analysis. The sample size of Experiment 2 was selected to correspond to that of Experiment 1.

3.2.2 Stimuli and Procedure

Experiment 2 selected two face images (a female facing to the left or right side) with a half-profile view from ESRC 3D-Face Database (with usage consent from the copyright holder; see http://pics.psych.stir.ac.uk for details). In the experiment, participants (sitting at a distance of approximately 72 cm from the screen) saw the scene with a room (12.3° × 8.0°)

where a face (4.0° × 6.2°) or a rectangle (0.6° × 6.2°) was presented in the centre, and discs (0.8° × 1.0°, 3.6° - 3.8° from the face) were arranged vertically on one or two lateral walls (see examples in Figure 5C and 5D). The disc distribution pattern of the current experiment was the same as that of Experiment 1. There is an exception in procedure when compared to Experiment 1: before the formal experiment, 20 participants who did not take part in the experiment were asked to judge the room with the face presented in the centre and discs (0, 1, 2 or 3) arranged vertically on one or two walls, and all of them reported that they believed all the gazers in the images could see the dot(s) on the wall that they were facing to and could not see the dot(s) on the wall that they were not facing to. Similar to Experiment 1, the face and the rectangle trials were randomly intermixed in each block, and the presentation order of blocks were counterbalanced across participants.

3.2.3 Results

Similar to Experiment 1, a two-way repeated-measures ANOVA for accuracy and response time in the dot-perspective task was conducted with "Central stimulus type" (Face vs. Rectangle) and "Consistency of the perspective" (Consistent vs. Inconsistent) as within-subject factors, and accuracy and response times were recorded as the key dependent variables. In keeping with Experiment 1, only data from matching trials were analysed.

The Consistency effect and the Central stimulus type x Consistency interaction effect did not significantly modulate accuracy on the dot-perspective task ($F(1,15) = 2.46$, $p = 0.14$, $\eta_p^2 = 0.14$; $F(1,15) = 2.29$, $p = 0.15$, $\eta_p^2 = 0.13$); whereas the main effect of Central stimulus type did reach significance ($F(1,15) = 5.00$, $p = 0.041$, $\eta_p^2 = 0.25$) showing that participants performed better when the central stimulus was a face (97.70 ± 1.20 %) rather than a rectangle (95.60 ± 1.00 %) (see Table 3).

Table 3. Behavioural Results [Mean (Standard Error)] in Conditions of Experiments 2.

Experiment 2	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent
Face	96.90	98.40	632.36	685.06$^{**>}$
	(1.80)	(0.80)	(21.09)	(30.75)
Rectangle	97.90	93.20	620.31	631.18
	(1.20)	(2.20)	(24.04)	(26.10)

$^{>}$ Indicates a significant increase from the consistent condition; $^{<}$ Indicates a significant decrease from the consistent condition.
* $p < .05$; ** $p < .01$; *** $p < .001$.

In terms of response time, ANOVA analysis showed a significant Central stimulus type x Consistency interaction effect ($F(1,15) = 5.27$, $p = 0.037$, $\eta_p^2 = 0.26$). Paired-sample t tests indicated that there was a significant Consistency effect when the face was in the middle of the room ($t(15) = -3.81$, $p = 0.002$), rather than when the rectangle was in the middle of the room ($t(15) = -0.71$, $p = 0.49$). Specifically, when the face was in the middle of the room, participants were quicker to respond in the consistent condition (mean ± standard error = 632.36 ± 21.09 ms) compared to the inconsistent condition (685.06 ± 30.75 ms). In contrast, when the rectangle was in the middle of the room, participants judged their own perspective as quickly in the congruent condition (620.31 ± 24.04 ms) as in the incongruent condition (631.18 ± 26.10 ms) (see Table 3 and Figure 8B on page 43).

3.2.4 Summary

Experiment 2 internally replicated the self-consistency effect found in Experiment 1, and the result pattern were consistent with automatic L1VPT. More importantly, the consistency effect was replicated when face images of real humans were used on the dot-perspective task. As the number of participants was low in Experiments 1 and 2, the resulting patterns of Experiments 1 and 2 but not the findings themselves indicated that Samson et al.'s (2010) findings are robust and there is explanatory breadth in the implicit mentalising account.

Chapter 4: Study 2: Experiments 3 and 4

Chapter note:

The content of Chapter 4 (Study 2: Experiments 3 and 4) has been accepted and published by the journal 'Visual Cognition'. Doi: 10.1080/13506285.2020.1857488

4.1 Introduction

The implicit mentalising account of L1VPT originates from findings generated by Samson and colleagues' (2010) dot-perspective task. According to proponents of the implicit mentalising account, the consistency effect on self-trials aligns with claims that adults have a cognitively efficient ToM system which tracks what someone else is seeing (or not seeing) via the computation of the other person's line-of-sight (Apperly & Butterfill, 2009).

Challenging the implicit mentalising account, other researchers suggest that domain-general attentional-orienting processes may just as well contribute to the consistency effect (e.g., Cole, Atkinson, D'Souza, & Smith, 2017; Langton, 2018; Santiesteban et al., 2014; Wilson et al., 2017). Rather than discussing the computation of the avatar's eye gaze in terms of implicit mentalising about what the person sees or knows about her surroundings, this alternative – submentalising – account contends that directional cues (extracted from the avatar's eyes, head or body) oriented attention to the certain locations in the dot-perspective task. Indeed, Santiesteban et al. found that the consistency effect could even be replicated when the avatar was replaced with an arrow (which has directional property but no biological significance).

Several researchers have attempted to test the submentalising account by manipulating the gazer's mental state of seeing through the use of physical barriers (Cole et al., 2016; Conway et al., 2017; Furlanetto et al., 2016; Langton, 2018; Wilson et al., 2017). Furlanetto and colleagues found that adults were slower to judge their own perspective on inconsistent trials compared to consistent trials only when participants believed that the avatar could actually see the dots on the wall (i.e., when the avatar was wearing transparent but not opaque coloured goggles), supporting the implicit mentalising account rather than submentalising account. Unfortunately, Conway et al. (2017) and Wilson et al. (2017) failed to replicate Furnaletto et al.'s findings in conceptual replications where the agent's visual access was manipulated using a cloaking device, coloured goggles, or blindfolds. However, these barriers may not be effective in making the avatar's modified epistemic state salient because they tend to frame and occupy a comparatively small part of the computer-generated avatar.

Recently, Gardner, Bileviciute, et al. (2018) manipulated avatar stance (averted vs. stance-maintained) as a new way to discriminate the implicit mentalising account from the submentalising account. Previous studies found that compared to avatar-stance-maintained stimuli (i.e., avatar's head and torso were oriented to the same wall), avatar-stance-averted stimuli (i.e., avatar's head was oriented to one wall while the torso was frontally presented) triggered an enhanced orienting effect (e.g., Hietanen, 2002). Importantly, the avatars of both stimuli had equivalent visual access to the target. The authors aimed to explore the competing accounts by examining whether effects from visual-perspective-taking tasks (e.g., consistency effect in the dot-perspective task) and attention-shifting tasks (e.g., cue-validity effect in

Posner target-detection task) could be differently influenced by avatar-stance. They found that for the Posner task, attentional orienting was modulated by avatar stance: there was a significant cue-validity effect for avatar-stance-averted avatars but not for avatar-stance-maintained avatars. When they employed the classic dot-perspective task, they replicated the consistency effect but, importantly, the magnitude of the consistency effect was not modulated by avatar stance. The dissociation between attentional orienting and perspective-taking casts doubt on directional cue-related predictions of the submentalising account but lends support to the implicit mentalising account.

Overall, the ToM field remains fiercely divided over the extent to which evidence over the consistency effect reflects social perspective-taking (implicit mentalising) or domain-general orienting (submentalising) processes. A pressing challenge is that many current methods used to disrupt the avatar's visual access (e.g., using barriers, or goggles with unusual visual properties) can be complex and ambiguous, which may make it difficult to fully differentiate the visibility of the avatar in the non-seeing condition from the one in the seeing condition. Two experiments (Experiments 3 and 4) were carried out to further contrast the implicit mentalising account from the submentalising account.

4.2 Experiment 3: Overview and Hypotheses

Experiment 3 manipulated eye cues – blocked or not blocked. As a clear and effective non-seeing manipulation, Experiment 3 placed a black opaque rectangle onto the gazer's eyes, and had the manipulation checked before and after the experiment. It is important to gauge whether the consistency effect can be detected using real stimuli such as photographs of an actual human being, rather than a cartoon avatar. Specifically, Wiese, Wykowska, Zwickel, & Müller (2012) claimed that effects to do with the triggering of mentalising or perspective-taking may be clearer if participants perceive that the agent is imbued with the property of having a mind. Compared with computer-generated avatars, photographs of a real person's face may have more characteristic representations of having a mind (i.e., clear facial features such as eye-gaze). Accordingly (and bolstered by positive findings from Experiment 2 in Chapter 3), photographs of a real person's face seem to be relatively ideal materials to explore L1VPT processing. Therefore, in Experiment 3, images of a real face were employed as the central stimuli in the dot-perspective task. The implicit mentalising account emphasizes the contribution of tracking another person's visual perspective to the consistency effect, whereas the submentalising account claims that attentional orienting contributes to the consistency effect. The implicit mentalising account would predict consistency effects when the eyes were not blocked. The submentalising account would predict consistency effects regardless of whether the person's eyes were being blocked or not blocked, because head direction cues (Nakashima & Shioiri, 2015) were still present to trigger orienting of attention.

4.2.1 Participants

A set of 44 participants (40 females, mean age: 20.1 years; age range: 18-28 years) from the general student population and Introduction to Psychology Research Programme (IPRP) system of Victoria University of Wellington (VUW) participated in this experiment. All the participants had normal or corrected-to-normal vision. They were naïve to the purpose of the study. The full research (application ID #0000026509) was approved by the School of

Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee and was conducted in accordance with the World Medical Association's Declaration of Helsinki. Each participant was taken to a quiet room and asked to complete the information sheet and consent form.

Three participants were eliminated based on the criteria (see Section 2.3), resulting in a sample size of 41. The sample size was selected based on Samson et al.'s (2010) third experiment and a priori power analysis using G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). The sample size of 41 exceeded the threshold of 37 participants needed to detect the effect size for the consistency effect on self-judgements in Samson et al.'s (2010) third experiment, Cohen's $d = 0.61$ with 95% power and Type 1 error at $\alpha < .05$.

4.2.2 Stimuli and Procedure

The stimuli were presented on a 19-inch monitor with Open-Sesame 3.2.5 software (Psychology Software Tools). Experiment 3 selected two face images (a female facing to the left or right side) with half-profile view (i.e., horizontal 45°) from ESRC 3D-Face Database (see http://pics.psych.stir.ac.uk for details). In the experiment, participants (sitting at a distance of approximately 72 cm from the screen) saw the scene with a room (12.3° × 8.0°) where an eyes-opened face (4.0° × 6.2°) or an eyes-covered face (black rectangle on eyes of the originally eyes-opened face, 2.2° × 0.6°) was presented in the centre, and discs (0.8° × 1.0°, 3.6° - 3.8° from the face) were arranged vertically on one or two lateral walls (see examples in Figure 9A and 9B). There were 10 different kinds of disc distributions (see Figure 10); each distribution was equivalently combined with a sighted or an eyes-covered face (50% of trials facing to left wall and 50% of trials facing to right wall).

On half of trials, the participant and the face could see the same discs (consistent condition), whereas the participant could see the discs that were invisible to the face on the other half of trials (inconsistent condition). Before the procedure was conducted, all the participants were asked to judge sighted and eyes-covered faces that faced to a wall, and all of them reported that sighted faces could see (seeing condition) whereas eyes-covered faces could not (non-seeing condition); and additionally, after completing the procedure, all of the participants reported that they maintained belief in the aforementioned difference between seeing and non-seeing conditions throughout the experiment.

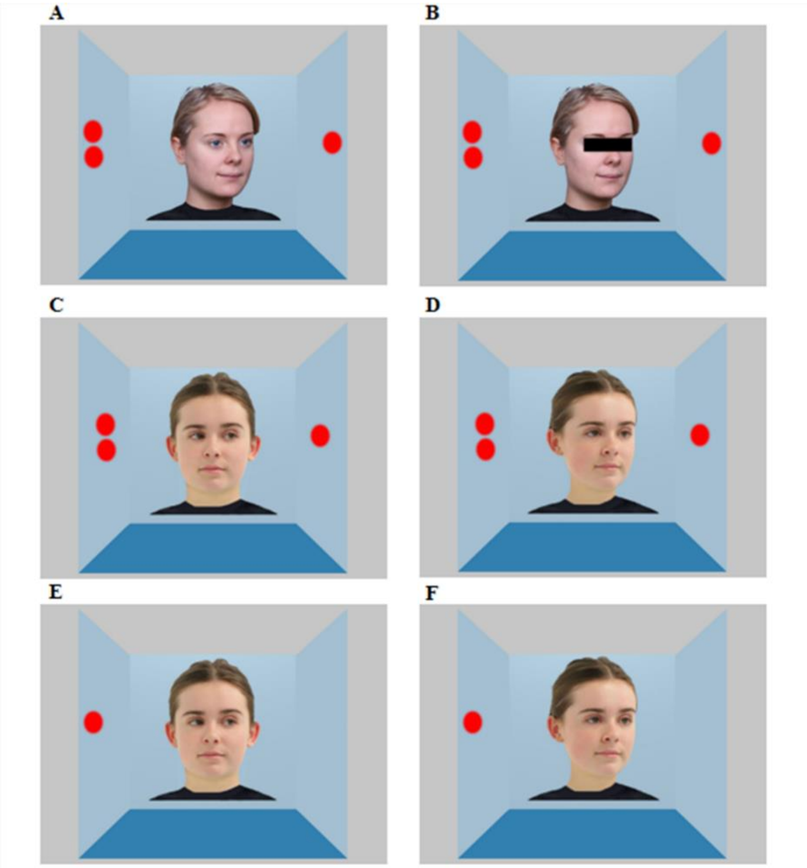

FIGURE 9 | Examples of room images used in Experiment 3 (A: Sighted face, B: Eyes-covered face), Experiment 4A (C: Head-front-gaze-averted face, D: Head-turned-gaze-maintained face), and Experiment 4B (E: Head-front-gaze-averted face, F: Head-turned-gaze-maintained face).

FIGURE 10 | The disc distribution when the real human's face was facing towards the left wall in Experiment 3. Images bordered with a dashed line were chosen as stimuli of filler trials; images bordered with a full line were chosen as stimuli on consistent trials; the remaining images were chosen as stimuli on inconsistent trials. * These stimuli were presented twice as often as the remaining stimuli to balance the overall number of consistent and inconsistent conditions.

Participants were seated approximately 72 cm from the computer screen. In the formal experiment, each participant was asked to complete the dot-perspective task shown as per Figure 11A. At the beginning of each trial, a fixation cross was presented for 750 ms. Then, 500 ms later, the word "you" (i.e., meaning to judge self-perspective and ignore the central stimuli of the room) appeared for 750 ms, followed by a 500-ms-blank. Then, a digit (0-3) was presented for 750 ms. Then, 500 ms later, the scene with central face and disc(s) on the wall(s) of a room was presented for 2 seconds, and participants had to judge whether the word-and-number matched with the scene of the room that the participants saw. On the matching trials, pressing '1' for "yes" response by using the forefinger of the left hand; on the mismatching trials, pressing '2' for "no" response by using the forefinger of the right hand) as quickly and accurately as possible.

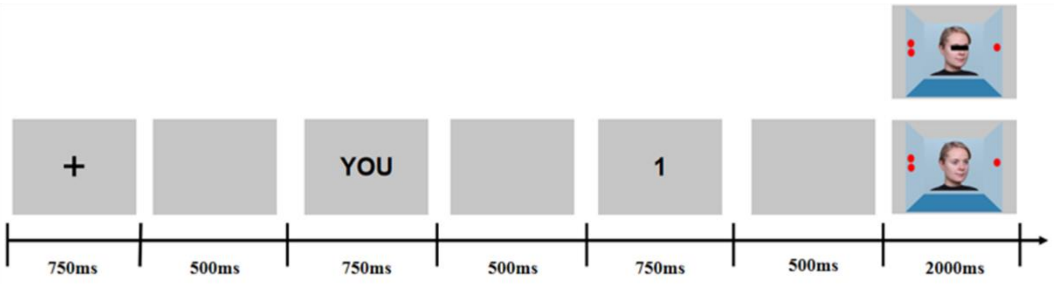

FIGURE 11A | Overview of a representative test trial of Experiment 3.

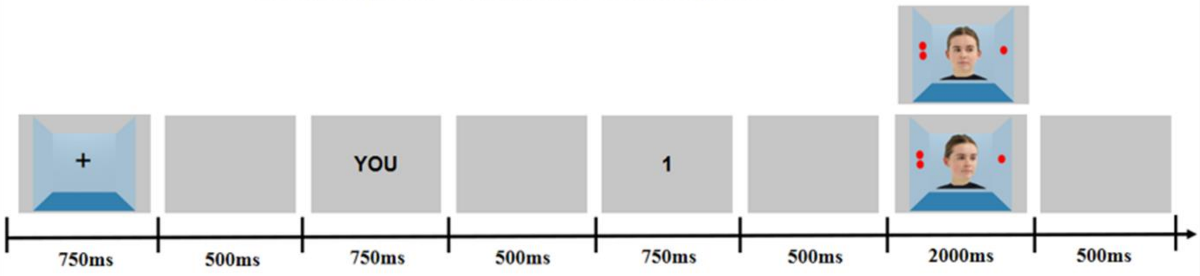

FIGURE 11B | Overview of a representative test trial of Experiment 4A.

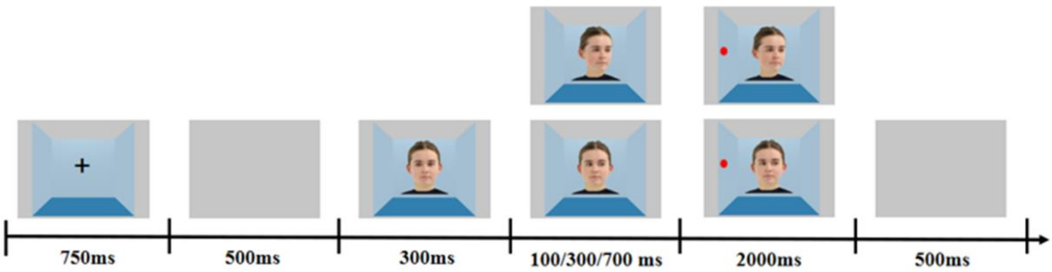

FIGURE 11C | Overview of a representative test trial of Experiment 4B.

Samson et al.'s (2010) results indicated that 'time' (comparing first half of the trials vs. second half of the trials) did not modulate the consistency effect. Bukowski and Samson (2017) also found that the consistency effect could be successfully revealed using an abbreviated version of the task. Thus, Experiment 3 used the short version of the task (104 trials, half of total number of trials of previous experiments), and expected that the modification would still detect relatively automatic L1VPT. Additionally, only self-perspective trials were used to exclude potential carry-over effects generated by switching between self- and other-perspective.

Every participant completed 48 trials (12 consistent with sighted face, 12 inconsistent with sighted-face, 12 consistent with eyes-covered face, 12 inconsistent with eyes-covered face) in both matching ("yes" response) and mismatching ("no" response) trials. There were 4 filler trials in addition to each 48 trials. In filler trials, there would be a sighted or an eyes-covered face in the centre of the room but without discs on the wall, which consisted of an equal number of matching and mismatching trials. Overall, there was a block of 26 practice trials and 2 blocks of 52 test trials (48 test trials and 4 filler trials). In order to avoid having more than 3 consecutive trials of the same condition in a block, the order of the trials was pseudo-randomised, which was fixed across participants. Sighted and eyes-covered faces were presented in separate blocks to exclude the potential carry-over effect induced by switching between different kinds of faces, and the presentation order of blocks were counterbalanced across participants.

4.2.3 Results

A two-way repeated-measures ANOVA for accuracy and response time in the dot-perspective task was performed with "Central stimulus type" (Sighted Face vs. eyes-covered Face) and "Consistency" (Consistent vs. Inconsistent) as within-subject independent variables, and accuracy and reaction time in the dot-perspective task were recorded as the key dependent variables. In keeping with Samson et al.'s (2010) study, only data from matching trials were analysed; the researchers argued that on mismatching consistent trials, the number cue did not correspond to anyone's perspective, making them easier to be processed than matching trials. Filler trials were excluded from the analysis. The dot-perspective task in Experiment 3 (and also in Experiment 4) applied the following criteria for participants' data to be subject to formal analysis. First, participants whose accuracy was less than 70% on average or 60% in any experimental condition would be eliminated. Additionally, participants whose reaction time was not within the range 'mean ± 2.5 standard deviation (SD)' would also be excluded. In Experiment 3, one participant was excluded from further analyses because the participant had low accuracy, and the remaining two participants were excluded due to excessively slow response times. Experiment 3 was based on data from 41 participants.

In terms of accuracy, there was a significant interaction effect between Central stimulus type and Consistency ($F(1,40) = 4.18$, $p = 0.047$, $\eta_p^2 = 0.095$), with participants performing better when the eyes-covered face was centrally presented (mean ± standard error = 96.70 ± 0.90 %) compared to when the sighted face was centrally presented (93.10 ± 1.30 %) in the consistent trials ($t(40) = -2.56$, $p = 0.014$) but not in the inconsistent trials ($t(40) = 0.15$, $p = 0.88$). Main effects of Central stimulus type and of Consistency did not reach significance (all *p*s > 0.05). There were no speed-accuracy trade-offs. Overall, accuracy was high; 41 participants produced accurate responses over 90% of the trials in all experimental conditions (see Table 4).

Table 4. Behavioural Results [Mean (Standard Error)] in Conditions of Experiments 3.

Experiment 3	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent
Sighted Face	93.10 (1.30)	94.90 (1.10)	657.60 (16.79)	722.97$^{**>}$ (23.34)
Eyes-covered Face	96.70$^{*>}$ (0.90)	94.70 (1.10)	673.08 (19.83)	681.25 (18.26)

$^{>}$ Indicates a significant increase from the consistent condition; $^{<}$ Indicates a significant decrease from the sighted face condition.
* p < .05; ** p < .01; *** p < .001.

In terms of response time, only correct trials were analysed. ANOVA analysis showed a significant Central stimulus type x Consistency interaction effect ($F(1,40) = 7.93$, $p = 0.008$, $\eta_p^2 = 0.17$). Paired-sample t tests and the calculation of the value of Cohen's d effect size showed that there was a significant Consistency effect when the sighted face was in the middle of the room ($t(40) = -3.73$, $p = 0.001$, Cohen's d = 0.58, medium effect size), but no significant consistency effect when the eyes-covered face was in the middle of the room ($t(40) = -0.76$, $p = 0.45$). Specifically, when the sighted face was in the middle of the room, participants were slower to judge their own perspective in the inconsistent condition (mean ± standard error = 722.97 ± 23.34 ms) compared to the consistent condition (657.60 ± 16.79 ms). In contrast, when the eyes-covered face was in the middle of the room, participants judged their own perspective as quickly in the consistent condition (673.08 ± 19.83 ms) as in the inconsistent condition (681.25 ± 18.26 ms) (see Table 4 and Figure 12A). In addition, there was a significant main effect of Consistency ($F(1,40) = 12.53$, $p = 0.001$, $\eta_p^2 = 0.24$), showing that RTs were slower in the inconsistent trials (702.11 ± 17.55 ms) than in the consistent trials (665.34 ± 16.33 ms). No other effect was significant ($p > 0.4$). The above findings were documented in the short version of the task, which was consistent with previous studies (Samson et al., 2010; Bukowski & Samson, 2017), demonstrating that the dot-perspective paradigm has integrity even in an abbreviated form.

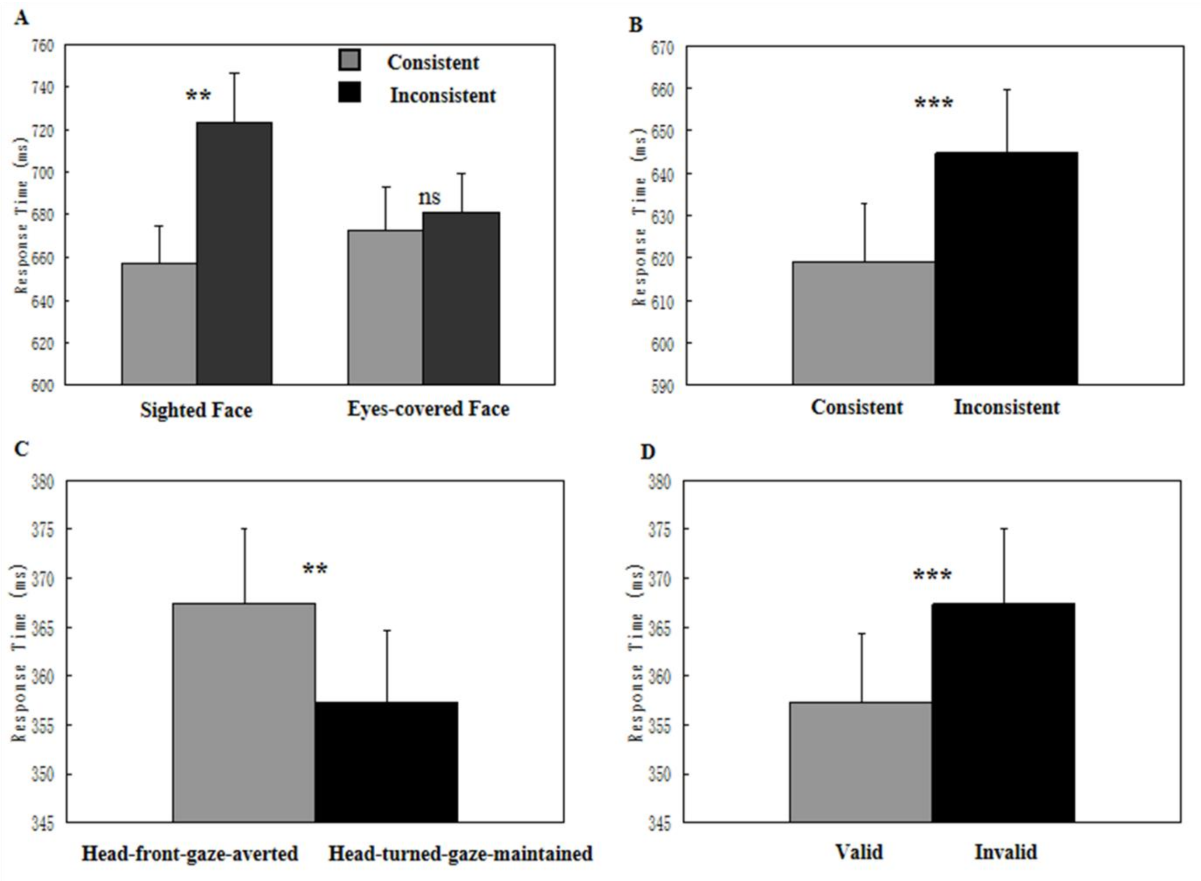

FIGURE 12 | Reaction times in each condition of Experiments 3 (A), 4A (B), and 4B (C and D). Means and standard errors are shown; error bars refer to +1 standard error of mean; * $p < .05$, ** $p < .01$, *** $p < .001$.

4.2.4 Summary

The consistency effect was replicated when face images of real humans were used on the dot-perspective task (and also in Experiment 2), indicating that Samson et al.'s (2010) findings are robust and there is explanatory breadth in the implicit mentalising account.

Similar to Furlanetto et al.'s (2016) finding of a consistency effect only in the avatar's seeing condition, the consistency effect was observed for sighted faces but not for eyes-covered faces when participants were judging their own visual perspectives.

It is important to emphasize that the stimuli in the non-seeing condition still presented head direction cues; according to the submentalising account, a consistency effect should be elicited whenever attentional orienting cues (which include head direction) are present. Therefore, to a certain extent, the fact that the consistency effect was absent in the non-seeing condition supports the implicit mentalising account and casts doubt on the submentalising account.

Nonetheless, the findings of Experiment 3 were limited on their own because only implicit mentalising (via the consistency effect) was measured. Experiment 3 did not directly measure potential attentional orienting on the basis of the submentalising account. This motivated Experiment 4, where both implicit mentalising (consistency effect) and attentional orienting (Posner's cue-validity effect) were measured. The approach in Experiment 4 (A and B) is novel because it does not manipulate eye cues but, rather, it manipulated how much direction cue was available. In one condition, Experiment 4 provided eye direction cues only (face with head-front and gaze-averted). In the other condition, Experiment 4 provided eye direction plus head direction cues (face with head-turned and gaze-maintained).

4.3 Experiment 4: Overview

Gardner, Bileviciute, et al. (2018) manipulated directional cue (i.e., avatar-stance) to dissociate attentional orienting from visual perspective-taking. Thus, an important question is whether manipulating eye and head directional cues (i.e., head-front-gaze-averted vs head-turned-gaze-maintained) would modulate the cue-validity effect in the Posner task but not the consistency effect in dot-perspective task. Specifically, Hietanen (1999) showed that the cue-validity effects in the Posner task are modulated by eye + head cues – the effects occurred for head-front-gaze-averted but not for head-turned-gaze-maintained faces. The authors interpreted that head-turned-gaze-maintained but not head-front-gaze-averted faces might make the observer feel that the other person was unrelated to himself/herself. This finding suggests that attentional orienting could be modulated by someone else's gaze direction relative to his/her head direction. In contrast, Experiment 3's result revealed that head direction cue of head-turned-gaze-maintained faces did not influence the consistency effect when eyes were covered, which raises the possibility that implicit mentalising induced by averted eye-gaze would not be modulated by head direction cue. Together, the face-related directional cue (head-front-gaze-averted vs. head-turned-gaze-maintained) may modulate attentional orienting but not implicit mentalising. Therefore, Experiment 4 adopted the dot-perspective and Posner tasks and manipulated the face-related directional cue to dissociate the implicit mentalising account from the submentalising account.

4.3.1 Experiment 4A (Dot-perspective Task): Hypotheses

Experiment 4A measured implicit mentalising via the consistency effect in the dot-perspective task. Experiment 3's findings would predict comparable effects for both head-front-gaze-averted and head-turned-gaze-maintained conditions because eye cues were always available. In contrast, based on Hietanen's (1999) findings, the submentalising account would predict effects for head-front-gaze-averted condition but not for head-turned-gaze-maintained condition, consistent with the role of face-related directional cue modulation in attentional orienting.

4.3.1.1 Participants

A new set of 54 first-year students (39 females, mean age: 20.35 years; age range: 18-38 years) from the IPRP system of VUW participated in the experiment. All of them had normal or corrected-to-normal vision. They were naïve to the purpose of the study. Each participant was taken to a quiet room and asked to complete the information sheet and consent form. Three participants were excluded (see Section 3.1.3), leaving 51 participants for further analysis. The sample size of 51 exceeded 37 participants required to detect Samson et al.'s (2010) third experiment's effect size with Cohen's d being 0.61, power being 0.95, and Type 1 error at $\alpha < .05$.

4.3.1.2 Stimuli and Procedure

Facial photographs of a volunteered female student from VUW from a frontal view and from a half-profile view (the volunteer provided signed informed consent for her photograph to be taken and used for the research) were taken for Experiment 4A. Based on these images, the head-front-gaze-averted face (i.e., front facing face with 45° left or right averted eyes, head direction was not aligned with gaze direction) and the head-turned-gaze-maintained face (i.e., half-profile face with 45° left or right averted eyes, head direction was aligned with gaze direction) were generated. Both of the faces could see the dot(s) where her eyes were directed to. In the experiment, participants (sitting at the distance of approximately 72 cm from the screen) saw the scene with a room (12.3° × 8.0°) where a head-front-gaze-averted face (4.0° × 6.2°) or a head-turned-gaze-maintained face (4.0° × 6.2°) was presented in the centre with the width of eyes (0.5°) kept same in the two different kinds of faces. Additionally, discs (approximately 0.8° × 1.0°, 3.6° - 3.8° from the face) were arranged vertically on one or two lateral walls (see examples in Figure 9C and 9D). The disc distribution pattern of the current experiment was the same as that of Experiment 3. There were several differences in procedure when compared to Experiment 3. Firstly, the number of trials was twice as many as that of Experiment 3. Secondly, only sighted faces were employed. Specifically, head-front-gaze-averted faces and head-turned-gaze-maintained faces were presented in separate block, and the presentation order of blocks were counterbalanced across participants. Thirdly, a fixation cross was presented in the centre of the room rather than having the presentation of the fixation cross itself at the beginning of each trial. Finally, at the end of each trial, an inter-trial interval of 500-ms-grey-screen was added. The latter two changes were matched with the corresponding parts of the target detection task in Experiment 4B (see Figure 11B).

4.3.1.3 Results

A two-way repeated-measures ANOVA for accuracy and response time in dot-perspective task was conducted with “Directional cue” (head-front-gaze-averted vs. head-turned-gaze-maintained), “Consistency” (Consistent vs. Inconsistent) as within-subject factors, and accuracy and reaction time in the dot-perspective task were recorded as the key dependent variables. In keeping with Experiment 3, only data from matching trials were analysed. Filler trials were excluded from the analysis. One participant whose overall accuracy was less than 70%, one participant who had made more than 40% errors in one condition, and one participant whose response times exceeded the range 'mean + 2.5 SD' were eliminated from the analyses, and data from the remaining 51 participants were entered for further analysis.

The Central stimulus type and the Directional cue x Consistency interaction effect did not significantly modulate accuracy on the dot-perspective task ($F(1,50) = 0.009$, $p = 0.93$, $\eta_p^2 = 0$; $F(1,50) = 2.72$, $p = 0.11$, $\eta_p^2 = 0.052$); whereas main effect of Consistency did reach significance ($F(1,50) = 8.32$, $p = 0.006$, $\eta_p^2 = 0.14$) showing that participants performed better in the consistent trials (97.30 ± 0.40 %) than in the inconsistent trials (95.70 ± 0.60 %) (see Table 5).

Table 5. Behavioural Results [Mean (Standard Error)] in Conditions of Experiments 4A.

Experiment 4A	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent$^{**<}$	Consistent	Inconsistent$^{***>}$
Head-front-gaze-averted Face	97.70 (0.50)	95.30 (0.70)	617.28 (13.64)	639.93 (16.47)
Head-turned-gaze-maintained Face	96.90 (0.60)	96.10 (0.70)	620.69 (15.32)	649.36 (16.18)

$^{>}$ Indicates a significant increase from the consistent condition; $^{<}$ Indicates a significant decrease from the consistent condition.
* p < .05; ** p < .01; *** p < .001.

Only response times of correct trials were analysed. ANOVA analysis showed a significant main effect of Consistency ($F(1,50) = 35.09$, $p < 0.001$, $\eta_p^2 = 0.41$). Post hoc test and the calculation of the value of Cohen's d effect size indicated that participants judged their own perspective faster in the consistent trials (618.99 ± 13.80 ms) than in the inconsistent trials (644.65 ± 14.97 ms) (Cohen's d = 0.83, large effect size) (see Table 5 and Figure 12B on page 61). Furthermore, the main effect of Directional cue ($F(1,50) = 0.42$, $p = 0.52$, $\eta_p^2 = 0.008$), and the Directional cue x Consistency interaction effect ($F(1,50) = 0.33$, $p = 0.57$, $\eta_p^2 = 0.007$) were not significant.

4.3.1.4 Summary

Consistent with Experiment 3's finding, Experiment 4A found a consistency effect with centrally presented sighted faces. More importantly, head-front-gaze-averted (i.e., gaze direction cue only) and head-turned-gaze-maintained (i.e., gaze + head direction cues) evoked similar consistency effects, in accord with the predictions of the implicit mentalising account.

4.3.2 Experiment 4B (Posner Task): Hypotheses

Experiment 4B measured attentional orienting via the cue-validity effect in the Posner task. Following Hietanen's (1999), Experiment 4B expected to find cue-validity effects for head-front-gaze-averted but not for head-turned-gaze-maintained faces.

4.3.2.1 Participants

A new set of 50 participants (36 females, mean age: 19.38 years; age range: 18-40 years) from the IPRP system of VUW took part in Experiment 4B, and each participant was 0.5 credit awarded. All participants had normal or corrected-to-normal vision. They were naïve to the purpose of the study. Each participant was taken to a quiet room and asked to complete the information sheet and consent form firstly. Forty-eight participants remained for further analysis after eliminating two participants (see Section 3.2.3). The sample size of Experiment 4B was selected to correspond to that of Experiment 4A and Gardner, Bileviciute, et al.'s (2018) work indicating validity effect for avatar images using a similar paradigm.

4.3.2.2 Stimuli and Procedure

Compared to Experiment 4A, there were two differences. First, only one disc (approximately 0.8° × 1.0°, 3.6° - 3.8° from the face) was arranged on one of the lateral walls (see examples in Figure 9E and 9F). Second, Experiment 4B added a frontal face (4.0° × 6.2°) with direct gaze (0.5°) (see the example in Figure 11C). Participants were seated approximately 72 cm from the computer screen. In the formal experiment, each participant was asked to complete the target-detection task. A fixation cross appeared for 750 ms in the centre of the room at the start of each trial, followed by a 500-ms-blank. After that, a frontal face with direct gaze appeared for 300 ms in the aforementioned room. Immediately after, a head-front-gaze-averted or a head-turned-gaze-maintained face in the same room appeared with a variable SOA (100, 300, or 700 ms), followed by a target dot on the left or right side of the wall of the room appearing for at most 2 seconds. The participants were required to press the letter "h" as accurately and as quickly as they could once they detected the dot. Finally, there was a 500 ms grey-screen delay before the beginning of the next trial (see Figure 11C).

Every participant completed 40 practice trials, followed by 4 experimental blocks. Each block included 80 trials, comprising 24 trials (12 valid trials with head-front-gaze-averted-face/head-turned-gaze-maintained-face, 12 invalid trials with head-front-gaze-averted-face/head-turned-gaze-maintained-face) in each of three SOAs (100, 300, 700 ms), and 8 catch trials. In catch trials, there would be a face in the centre of the room but without a disc on the wall, which consisted of an equal number of head-front and head-turned trials. The order of the trials was pseudo-randomised in order not to have more than 3 consecutive trials of the same condition in a block and was fixed across participants. The head-front-gaze-averted faces and head-turned-gaze-maintained faces were presented in separate blocks, and the presentation order of the blocks was counterbalanced across participants.

Before the formal experiment, a pilot experiment was conducted based on Hietanen's (1999) work. The only difference between the pilot experiment and the formal experiment was that the pilot experiment did not include the initial presentation of a frontal face with direct gaze, prior to the face-related directional cues. Similar to Hietanen's (1999) paradigm, the pilot experiment employed a Posner task where a head-front-gaze-averted or head-turned-gaze-maintained face was directly displayed as a directional cue followed by an SOA (i.e., 100, 300 or 700 ms) before the appearance of the target (i.e., a disc). In the Posner task of the pilot experiment, the performance of a group of 16 participants (who did not participate in the formal experiment) was not modulated by Validity or its interactions with SOA or/and Directional cue (all ps > 0.05). A potential reason is that there may have been no classic cue-validity effect in the pilot experiment because the head-front-gaze-averted or head-turned-gaze-maintained face may lack the dynamic property of directional cue without being first set off by the presentation of a frontal face with direct gaze. Indeed, research on the gaze cueing effect (i.e., a cue-validity effect generated by eye direction) usually starts with a frontal face with direct gaze or no eyes before the presentation of a frontal face with averted eye-gaze (e.g., Bayliss et al., 2005), making the directional cue dynamic. Therefore, in the formal experiment, a frontal face with direct gaze was presented before the appearance of the head-front-gaze-averted or head-turned-gaze-maintained face.

4.3.2.3 Results

A 3 (SOA: 100 ms vs. 300 ms vs. 700 ms) × 2 (Directional cue: head-front-gaze-averted vs. head-turned-gaze-maintained) × 2 (Validity: Valid vs. Invalid) repeated measure ANOVA analysis was performed, and accuracy and reaction time in the target detection task were recorded as the key dependent variables. Two participants whose response times were made in excess of 2.5 standard deviations of mean RT were excluded, leaving data from 48 participants for further analysis. On average, participants made response errors in 2.34% of the catch trials (i.e., pressed the response button without the presentation of the target).

Only correct trials of non-catch trials were selected for analysis of response time. ANOVA analysis showed a significant main effect of SOA ($F(2,94) = 34.27$, $p < 0.001$, $\eta_p^2 = 0.42$), with participants responding more slowly when the SOA was 100 ms (mean ± standard error, 374.54 ± 8.10 ms) compared to when the SOA was 300 ms (353.56 ± 7.14 ms) or 700 ms (358.97 ± 6.95 ms). More importantly, there was a significant main effect of Directional cue ($F(1,47) = 9.01$, $p = 0.004$, $\eta_p^2 = 0.16$). Post hoc test and the computation of the value of Cohen's d effect size revealed that participants performed faster in the head-turned-gaze-maintained condition (357.32 ± 7.22 ms) than the head-front-gaze-averted condition (367.40 ± 7.67 ms) (Cohen's d = 0.43, between small and medium effect size) (see Table 6 below and Figure 12C on page 61). In addition, the analysis revealed a main effect of Validity ($F(1,47) = 42.80$, $p < 0.001$, $\eta_p^2 = 0.48$), showing that participants performed more slowly on the invalid trials (367.38 ± 7.65 ms) than on the valid trials (357.33 ± 6.92 ms) (Cohen's d = 0.94, large effect size) (see Table 6 below and Figure 12D on page 61). Analysis of the main effect of validity at each SOA indicated that participants performed more slowly in the invalid condition than in the valid condition when SOA is 100 ms ($t(47) = -3.30$, $p = 0.002$), 300 ms ($t(47) = -3.80$, $p < 0.001$) or 700 ms ($t(47) = -4.65$, $p < 0.001$). No significant interaction effects were found (all *ps* > 0.1).

Table 6. Behavioural Results [Mean (Standard Error)] in Conditions of Experiments 4B.

Experiment 4B	Reaction time (ms)			
	Head-front-gaze-averted Face		Head-turned-gaze-maintained Face$^{**<}$	
	Valid	Invalid$^{***>}$	Valid	Invalid
SOA (100 ms)$^{***>}$	375.36 (8.54)	387.50 (9.73)	366.33 (7.94)	368.97 (7.84)
SOA (300 ms)	352.86 (7.64)	362.56 (7.55)	344.83 (7.20)	353.99 (7.77)
SOA (700 ms)	355.80 (6.36)	370.29 (8.20)	348.80 (6.93)	360.97 (8.06)

$^{>}$ Indicates a significant increase from the other condition(s); $^{<}$ Indicates a significant decrease from the other condition(s).
* p < .05; ** p < .01; *** p < .001.

4.3.2.4 Summary

Similar to previous findings demonstrating attentional orienting towards faces with averted eye-gaze (e.g., Driver et al., 1999; Van Rooijen et al., 2018), Experiment 4B found cue-validity effect for head-front-gaze-averted faces. However, contrary to Hietanen's (1999) results, Experiment 4B also found a cue-validity effect for head-turned-gaze-maintained faces. A speculative explanation might be that the initial presentation of a frontal face with direct gaze (in the formal experiment but not Hietanen's (1999) study) followed immediately by a head-turned-gaze-maintained face makes the stimulus more dynamic. In this circumstance, it is possible that the head-turned-gaze-maintained face could transmit a directional cue (eye + head) that became particularly salient and was powerful enough to influence participants' attention to the target. Additionally, when comparing the pilot experiment with the formal experiment, the lack of the cue-validity effect in the pilot experiment might be owing to a lack of dynamic property in the directional cue. Nonetheless, one should be mindful that cross-study comparisons and conclusions are difficult to draw and do not provide definitive evidence (e.g., lack of randomisation to conditions).

Overall, pilot and formal results from Experiment 4B lead to a speculation that the cue-validity effect may require dynamic processing of the face-related directional cues. This dynamic property did not appear to be required for triggering the consistency effect in Experiment 4A, which presented face-related directional cues without prior presentation of a frontal face with direct gaze. These findings, together, suggested that attentional orienting (Experiment 4B) might be dissociated from visual perspective taking (Experiment 4A).

Chapter 5: Study 3: Experiments 5, 6 and 7

5.1 Introduction

Study 2 did not find the dissociation between the implicit mentalising account and the submentalising account by manipulating an eye-head directional cue (i.e., head-front-gaze-averted vs. head-turned-gaze-maintained). As a follow-up experiment, Experiment 5 of Study 3 tried to find another eye-head directional cue (i.e., head-front-gaze-averted vs. head-turned-gaze-averted) to clarify the two competing accounts. Furthermore, Experiments 6 and 7 explored whether manipulations of different body parts could open new avenues for the debate. Theoretical bases for the following experiments are shown as follows.

The manipulation of the virtual avatar's visual access (i.e., visible and invisible conditions) has been widely used to explore the ongoing debate but triggering inconsistent findings. For instance, Furlanetto et al. (2016) explored the debate by manipulating the avatar's visual access via using transparent (i.e., red goggles) and opaque goggles (i.e., orange goggles). After grasping the associations between coloured goggles and visibility conditions, the participants were asked to complete the dot-perspective task in both seeing and non-seeing conditions. The authors observed a significant consistency effect in the seeing condition but not the non-seeing condition. They interpreted that altercentric interference effect was observed based on the beliefs that avatars wearing transparent goggles could see the scenarios. The findings supported the implicit mentalising account but cast doubt on the submentalising account. Later, Conway et al. (2017) also adopted coloured goggles to manipulate the avatar's line-of-sight but failed to replicate Furlanetto et al.'s (2016) findings by showing the comparable magnitude of consistency effects in both the visible and the invisible conditions. The authors interpreted that the consistency effect in the dot-perspective task resulted from participants' attentional shift to the avatar's directional property, supporting the submentalising account.

However, there may be a shortcoming in these two studies. Although a manipulation check was conducted to ensure that participants have grasped the associations between coloured goggles and corresponding seeing conditions, whether the connection can be represented in the dot-perspective task cannot be determined. Specifically, the manipulation of the agent's visual access is not easy to be recognized as eye region occupies a relatively small part of the avatar, and meanwhile, participants should complete the dot-perspective task in a very short duration (i.e., 2 s). To address the issue, Experiment 3 (see Chapter 4) manipulated the agent's line-of-sight by employing an easily-identifiable barrier (i.e., a black rectangle covered on the agent's eyes). All the participants reported that eyes-opened faces could see whereas eyes-covered faces could not see in an evaluation before the formal experiment. Experiment 3 discovered a significant consistency effect in the visible condition but not the invisible condition. Consistent with Furlanetto et al.'s (2016) work, the findings lend support to the implicit mentalising account.

Gardner, Bileviciute, et al. (2018) directly compared attentional orienting with visual perspective-taking to dissociate the two competing accounts. The authors adopted the Posner task (a classic task tapping attentional orienting) and dot-perspective task, and manipulated avatar-stance to find whether avatar-stance could modulate attentional orienting but not visual perspective-taking. Additionally, the directional cue ‘Avatar-stance’ was considered on the basis of the previous finding of stronger orienting effect for stance-averted avatars (i.e., avatar’s head was directed to one lateral wall whereas its torso was frontally displayed) than for stance-maintained avatars (i.e., avatar’s head and torso were directed to the same lateral wall) (e.g., Hietanen, 2002). In Experiment 1, they used the Posner task to observe a cue-validity effect, namely, participants detected the targets more slowly at the non-cued locations relative to cued-locations. More importantly, attentional orienting was affected by avatar-stance, which was reflected by an apparent cue-validity effect for stance-averted avatars but not for stance-maintained avatars. However, the consistency effect in the dot-perspective task of Experiment 2 was not modulated by avatar-stance. Thus, the dissociation between attentional orienting and visual perspective-taking supported the implicit mentalising account but cast doubt on the submentalising account.

Nevertheless, Gardner, Bileviciute, et al.’s (2018) study has an issue that the virtual avatars comprise clear head and torso directions but unclear information about other facial features, especially eye-gaze. As eye-gaze is a window into another person’s mental states (Grossmann, 2017), lack of eye information for the avatars may make Gardner, Bileviciute, et al.’s (2018) findings less credible. Instead, pictures of a real person with obvious facial features, eye direction included, appear to be more ideal materials to examine L1VPT processing compared to the virtual avatar.

Notwithstanding the replacement of a computer-generated avatar with images of a real human with clear facial features, it remains important to determine whether a face-related directional cue could distinguish attentional orienting from implicit mentalising like avatar-stance did in Gardner, Bileviciute, et al.'s (2018) study. Qian et al. (2013) provided a potential eye-head directional cue: head-front-gaze-averted faces vs. head-turned-gaze-averted faces (i.e., lateral view head with the gazer's eyes looking back to the observer), separating eye direction from head orientation. The participants firstly evaluated that the gaze direction of head-front-gaze-averted faces was significantly higher than that of head-turned-gaze-averted faces. In the gaze-cueing task (a modified Posner task), participants without the participation of the evaluation showed a stronger cue-validity effect for head-front-gaze-averted faces than for head-turned-gaze-averted faces. The researchers explained that the head-turned-gaze-averted face, a weaker cue, triggered a weaker gaze-cueing orienting effect. The findings demonstrated that gaze-cueing attentional orienting was modulated by perceived gaze direction with reference to head orientation in the modified Posner task.

By contrast, the implicit mentalising account may claim that the gaze perception cannot modulate relatively automatic tracking of others' visual perspective as visual access is always available regardless of the directional cue. Altogether, perceived averted-gaze with the consideration of head orientation may modulate attentional orienting but not implicit mentalising. Experiment 5 attempted to firstly shed light on the implicit mentalising versus submentalising debate by creating eye-head cues (i.e., head-front-gaze-averted and head-turned-gaze-averted faces) and measuring the cues' effects in both the Posner task and the dot-perspective task.

In addition to eye and head cues, manipulation of cues from other parts of the human body may also provide insights into clarifying the implicit mentalising vs. submentalising debate. Kendon (1994) and McNeill (1985) claimed that gestures like finger-pointing play a pivotal role in social communications. Additionally, finger-pointing seemed to be a more accurate spatial cue than averted-gaze for children to redirect their visual attention to the target (e.g., Butterworth, 1991; Butterworth & Itakura, 2000). One natural question is whether eye-finger-pointing cues (i.e., gaze-averted vs. finger-pointing) can distinguish the implicit mentalising from the submentalising accounts.

To date, only two studies directly compared the role of gaze direction with that of finger-pointing direction on attentional orienting, discovering an advantage of perceiving finger-pointing relative to eye-gaze for young children. Unlike Doherty and Anderson's (1999) usage of different tasks measuring the corresponding cues (i.e., the pointing-direction task and the looking-where task), Gregory et al. (2016) adopted only one task, namely, the eye-tracking task. They observed that a greater cue-validity effect was elicited by the non-predictive finger-pointing compared to eye-gaze in 3-5-year-old children. The researchers interpreted that the advantage of finger-pointing may arise from young children's earlier learning of the adult's hand cues connecting with targets and, further may arise from finger-pointing's being more salient than other cues such as eye-gaze cues.

Gregory et al.'s (2016) task is the same as the Posner task in the pattern of stimulus presentation. That is, there is an SOA between the appearances of the cue and the target. However, whether the superiority of finger-pointing relative to eye-gaze for young children could also emerge in adults, and whether the superiority could still exist when adults completed the Posner task (a classic task directly measuring attentional orienting), are scarcely understood.

Despite finger-pointing being a potentially effective cue to modulate attentional orienting, little is known about the role of finger-pointing in L1VPT and its comparison with the role of eye-gaze in the dot-perspective task. Accordingly, Experiment 6 investigated if eye-finger-pointing cues in the upper body could distinguish the implicit mentalising account from the submentalising account by measuring those cues' effects in both the Posner task (in Experiment 6A) and the dot-perspective task (in Experiment 6B). Due to the effectiveness of the manipulation of visual access in Experiment 3, another question was raised: whether visual access could still predominate the processes in the dot-perspective task when finger-pointing was considered. Accordingly, Experiment 7 explored whether the combination of line-of-sight's manipulation with finger-pointing can provide new insights into the debate.

Overall, the ongoing debate on whether the consistency effect in the dot-perspective task arises from specific implicit mentalising or domain-general attentional orienting processes appeals to further investigations. The current study sought to open new avenues for clarifying the debate by measuring effects in both the Posner task and the dot-perspective task triggered by the agent's different body parts.

5.2 Experiment 5: Overview and Hypotheses

Gardner, Bileviciute, et al. (2018) observed a dissociation between attentional orienting and visual perspective-taking by manipulating a virtual avatar's directional cue (i.e., avatar-stance). Likewise, Experiment 5 tried to clarify the debate by examining whether manipulating gaze and head directions (i.e., head-front-gaze-averted vs. head-turned-gaze-averted) could modulate attentional orienting but not visual perspective-taking (Head-front-gaze-averted means a frontal-view face with averted eye-gaze while head-turned-gaze-averted means a lateral-view face with eyes looking back to an observer). The real human's face-related manipulation was based on Qian et al.'s (2013) work revealing that head-front-gaze-averted faces induced a stronger cue-validity effect than head-turned-gaze-averted faces in a modified Posner task. The researcher explained that head-turned-gaze-averted faces evoked weaker gaze-cueing orienting effect because of their weaker directional information, demonstrating that gaze-cueing orienting effect was modulated by gaze perception with reference to head orientation in the modified Posner task. Therefore, eye-head cues' manipulation might modulate the cue-validity effect in the Posner task. In contrast, the implicit mentalising account emphasizing the role of visual access would posit another possibility: the consistency effect might not be influenced by manipulations of the agent's directional property due to visible eyes for both head-front-gaze-averted and head-turned-gaze-averted faces. Overall, attentional orienting rather than implicit mentalising might be modulated by eye-head directional cue, which fitted with Gardner, Bileviciute, et al.'s (2018) finding of a modulated effect of the directional cue's manipulation (i.e., avatar-stance) on attentional orienting but not visual perspective-taking. Accordingly, Experiment 5 used the Posner and dot-perspective tasks and manipulated the real human's eye-head cues, attempting to disentangle the submentalising account from the implicit mentalising account.

5.2.1 Experiment 5A (Posner task): Overview and Hypotheses

Experiment 5A evaluated whether eye-head cue manipulation (i.e., head-front-gaze-averted vs. head-turned-gaze-averted) would modulate attentional orienting via the cue-validity effect in the Posner task. In order to separate gaze direction from head orientation, gaze angle of head-turned-gaze-averted faces was made to be direct to a maximum extent. Following Qian et al.'s (2013) work, it was predicted that compared with head-turned-gaze-averted faces, head-front-gaze-averted faces would induce a stronger cue-validity effect.

5.2.1.1 Participants

Thirty-nine undergraduates were recruited through the IPRP system of Victoria University of Wellington (VUW) and obtained 0.5-course credit for participation. Thirty-three participants (25 females, mean age: 19.3 years; age range: 18-27 years) were remained for further analysis after the exclusions of six participants (see 'Results' section). So far, R-package can do power analyses for repeated measures designs with more than one within-subject variable whereas G*power cannot. Instead of adopting G*power, a priori power analysis conducted using an R package SIMR indicated that the sample size of 27 allowed for the examination of the Validity × Avatar-stance interaction effect found by Gardner, Bileviciute, et al. (2018), at the power of being about 80%. The sample size of 33 in the current experiment exceeded the required sample size. The research (application ID #0000026509) was approved by the School of Psychology Human Ethics Committee under the delegated authority of Victoria University of Wellington's Human Ethics Committee, which was carried out in line with Declaration of Helsinki. All participants in the full study reported normal or corrected-to-normal vision and were right-handed. Every participant was given informed consent before participation and debriefed after participation.

5.2.1.2 Stimuli

Photographs of a volunteered female undergraduate from VUW were taken (the volunteer signed the informed consent to allow her pictures to be taken and employed for the study). Head-front-gaze-averted face (i.e., a frontal-view head with 45° left or right averted eyes) and the head-turned-gaze-averted face (i.e., a half-profile head with eyes looking at the observer) are included. Sitting at the distance of about 72 cm from a 14-inch monitor, the participant could observe the scene in a room (13.0° × 12.0°) where a head-front-gaze-averted face (3.5° × 4.6°) or a head-turned-gaze-averted face (3.5° × 4.6°) was centrally presented with the same width of eyes (0.2°) in the two distinct faces. Furthermore, the room consisted of left, back, and right walls with one disc (about 0.7° × 0.9°, 3.1° from the face) displayed vertically on one of the lateral-view walls (see examples in Figure 13A and 13B). On half of the trials, the faces were oriented to the left wall, whereas they were oriented to the right wall on the other half of trials. On 50% of trials, the lateral side where the eyes were looking or the head was oriented was the same as the side where the target would be presented (valid condition), but the side where the cues were directed was different from the side where the target would be displayed on the remaining trials (invalid condition).

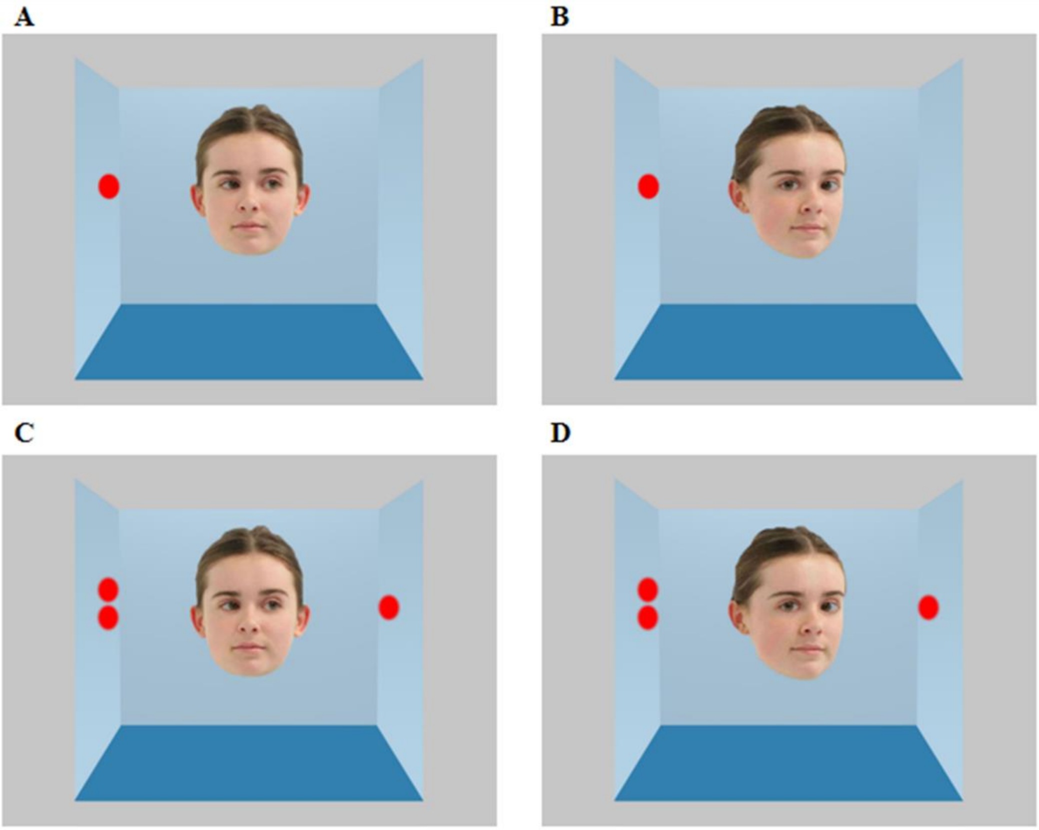

FIGURE 13 | Examples of room images used in Experiment 5A (A: Head-front-gaze-averted Face, B: Head-turned-gaze-averted Face), and Experiment 5B (C: Head-front-gaze-averted Face, D: Head-turned-gaze-averted Face).

Before the formal experiment, 20 adults (who would not participate in the formal experiment) were instructed to judge whether the centrally presented face could see the dot in the valid and invalid conditions. Eight images involving one disc on the left/right wall with a face with left/right-viewing eye-gaze (i.e., head-front-gaze-averted) or left/right-oriented head (i.e., head-turned-gaze-averted) were displayed. All the participants reported that the eyes of the head-front-gaze-averted faces could see the dot whereas the eyes of the head-turned-gaze-averted faces could not in the congruent condition (i.e., a valid condition in Experiment 5A and consistent condition in Experiment 5B), whereas all the faces could not see the disc in the incongruent condition (i.e., an invalid condition in Experiment 5A and inconsistent condition in Experiment 5B).

In addition, a new sample of 22 adults without participation in the formal experiment evaluated the gaze angle of the four faces shown in Table 7. The gaze angle was rated from 0 (i.e., direct gaze) to 5 (i.e., extremely left or right gaze). The average ratings for four faces were 2.9° (head-front-gaze-averted) and 2.1° (head-turned-gaze-averted) when evaluating Qian et al.'s (2013) stimuli; 4.2° (head-front-gaze-averted) and 1.6° (head-turned-gaze-averted) when evaluating the stimuli. A paired samples *t*-tests showed that rating of head-front-gaze-averted faces was significantly higher than that of head-turned-gaze-averted faces in these two studies (all *p*s < 0.05). Among the results, the ratings of Qian et al.'s (2013) stimuli are similar to those in Qian et al.'s (2013) study (i.e., 2.6 for head-front-gaze-averted faces and 1.9 for head-turned-gaze-averted faces).

Table 7. Stimuli used in Qian et al.'s (2013) study and the present Experiment 5A.

Type of Faces	Head-front-gaze-averted	Head-turned-gaze-averted
Qian et al.'s (2013) study		
The present study		

5.2.1.3 Procedure

Every participant was instructed to complete the online target-detection task shown as per Figure 14A. The full study was conducted as an online study (due to the impact of Covid-19 lockdowns in New Zealand), and studies 1 and 2 were conducted in person. The online study was created in OpenSesame software and uploaded through JATOS server. Trials commenced with a fixation cross appeared for 750 ms in the centre of a virtual room with left, back and right walls, followed by a 500-ms-interval. Then, a face was centrally presented in the same room with a variable delay (SOA = 100, 300 or 700 ms). After that, a target (dot) on one of the lateral walls of the room was displayed for up to a maximum of 3000 ms, and the participants executed their responses by pressing the letter “h” as quickly and accurately once they detected the dot. Finally, an inter-trial interval of 500-ms-blank was presented with the virtual room kept being presented.

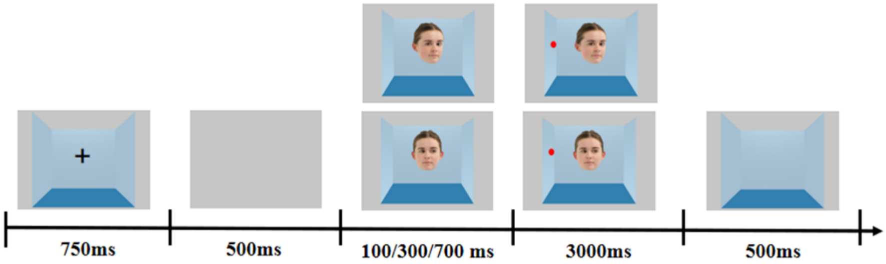

FIGURE 14A | Overview of a representative test trial of Experiment 5A.

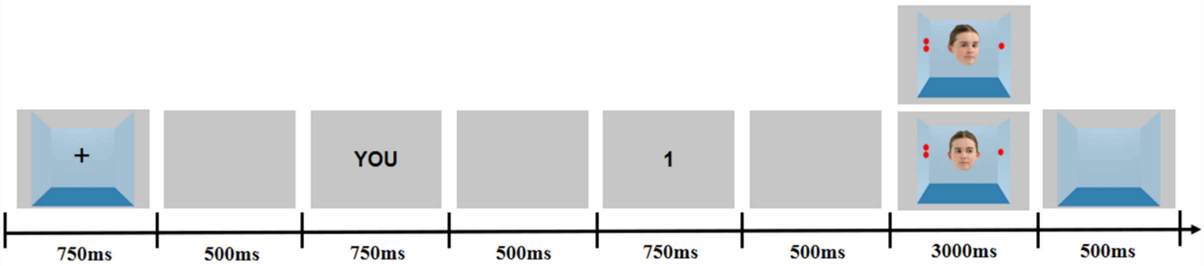

FIGURE 14B | Overview of a representative test trial of Experiment 5B.

Following a practise session of 40 trials, the participants completed 4 experimental blocks with the presentation order of the trials being pseudo-randomised and fixed across participants. Each experimental block of 80 trials contained 24 trials (12 valid trials with head-front-gaze-averted-face/head-turned-gaze-averted-face, 12 invalid trials with head-front-gaze-averted-face/head-turned-gaze-averted-face) in every SOA (100, 300, 700 ms), and 8 catch trials. In catch trials, a face centrally appeared in the virtual room but with no dot on the wall, which comprised an equal number of head-front-gaze-averted-face/ head-turned-gaze-averted trials. The head-front-gaze-averted and head-turned-gaze-averted faces were displayed in different blocks with the presentation order counterbalanced across participants.

5.2.1.4 Results

The factors “SOA” (100 ms vs. 300 ms vs. 700 ms), “Directional cue” (Head-front-gaze-averted vs. Head-front-gaze-averted) and “Validity” (Valid vs. Invalid) were included in the analysis and formed a 3×2×2 within-subject design. SOAs were considered to measure that the attentional-orienting effect was either reflexive (e.g., 100, 300 ms) or volitional (e.g., 700 ms). Percentage of error and response time as the key dependent variables. Participants whose accuracy was less than 90% or response time was not within the range ‘mean ± 2.5 standard deviations (SDs)’ were be eliminated from the data set. Three participants with low accuracy and another three participants whose response times were made in excess of 2.5 standard deviations of mean RT were excluded, leaving data from 33 participants for further analysis. On average, the error rate of catch trials was 3.98% (i.e., pressing the response button when no target was presented).

In terms of response time, only accurate trials of non-catch trials were analysed. ANOVA analysis showed a significant interaction effect between SOA and Validity (F (2,64) = 5.44, p = 0.007, η_p^2 = 0.15). Paired-sample t tests and the calculation of Cohen’s d effect size revealed that there was a significant Validity effect when SOA was 700 ms (t(32) = -3.06, p = 0.004, Cohen’s d = 0.53, medium to large effect size), but not 100 ms (t(32) = 0.88, p = 0.38) or 300 ms (t(32) = 0.023, p = 0.98). Specifically, when SOA was 700 ms, participants were slower to detect the target in the invalid condition (mean ± standard error = 376.37 ± 7.87 ms) when compared to the valid condition (365.68 ± 8.39 ms). In contrast, participants made judgements on target position as quickly in the valid condition as in the invalid condition both when SOA was 100 ms and when SOA was 300 ms (see Table 8 and Figure 15A). Additionally, there was a significant interaction effect between Central stimulus type and Validity (F (1,32) = 9.61, p = 0.004, η_p^2 = 0.23). Paired-sample t-tests and the computation of the value of Cohen's d effect size indicated a significant Validity effect when a head-front-gaze-averted face (t(32) = -3.16, p = 0.003, Cohen's d = 0.55, medium to large effect size) but not a head-turned-gaze-averted face (t(32) = 1.59, p = 0.12) was centrally presented. When the directional cue was a head-front-gaze-averted face, the participants detected the target (dot) more slowly in the invalid condition (385.95 ± 8.32 ms) when compared to the valid condition (377.00 ± 7.82 ms). By contrast, when the directional cue was a head-turned-gaze-averted face, the participants performed as quickly on valid trials (388.68 ± 10.22 ms) as on invalid trials (385.44 ± 9.20 ms) (see Table 8 and Figure 15B).

Table 8. Behavioural Results [Mean (Standard Error)] in Conditions of Experiment 5A.

Experiment 5A	Reaction time (ms)			
	Head-front-gaze-averted Face		Head-turned-gaze-averted Face	
	Valid	Invalid$^{** >}$	Valid	Invalid
SOA (100 ms)	401.84	403.66	416.93	410.98
	(9.77)	(9.62)	(11.93)	(11.13)
SOA (300 ms)	368.22	375.70	378.69	371.10
	(7.98)	(8.54)	(9.90)	(8.96)
SOA (700 ms)	360.94	378.49$^{** >}$	370.42	374.25
	(7.79)	(8.20)	(10.08)	(8.67)

$^{>}$ Indicates a significant increase from the other condition(s); $^{<}$ Indicates a significant decrease from the other condition(s).

* $p < .05$; ** $p < .01$; *** $p < .001$.

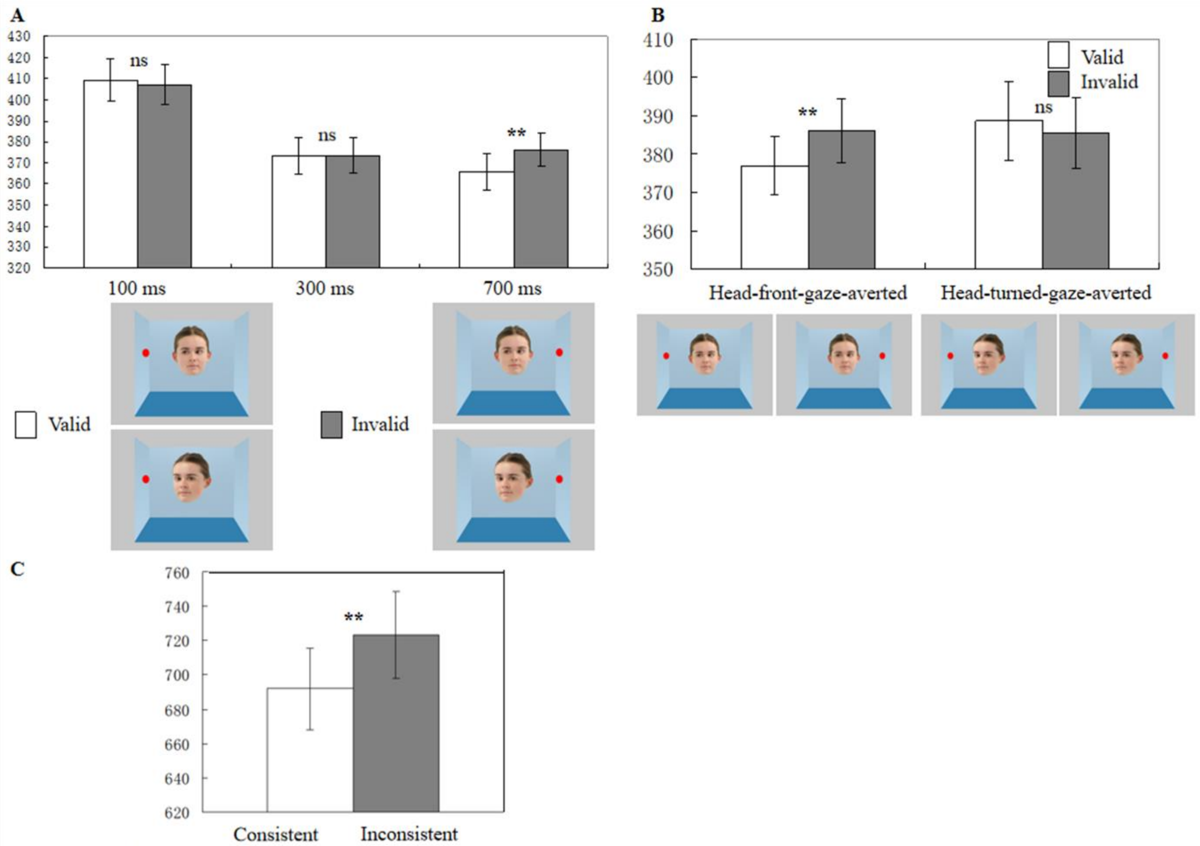

FIGURE 15 | Reaction times in each condition of Experiments 5A (A and B) and 5B (C). Means and standard errors are shown; error bars refer to ±1 standard error of mean; * $p < .05$, ** $p < .01$, *** $p < .001$.

5.2.1.5 Summary

In support of the predictions, there was a stronger cue-validity effect for head-front-gaze-averted faces than for head-turned-gaze-averted faces. Consistent with Qian et al.'s (2013) conclusion, further, the finding demonstrates attentional orienting is influenced by gaze perception with reference to head orientation. Furthermore, the significant cue-validity effect that appeared only in a longer SOA of 700 ms suggests that orienting effect was volitional, which is similar to a previous finding of the avatar-stance being as a modulator of volitional attentional orienting (Gardner, Bileviciute, et al., 2018; Gardner, Hull, et al., 2018).

5.2.2 Experiment 5B (Dot-perspective Task): Overview and Hypotheses

Experiment 5B sought to examine whether the manipulation of eye-head cues (i.e., head-front-gaze-averted vs. head-turned-gaze-averted) could moderate implicit mentalising in the dot-perspective task. There were two hypotheses corresponding to two competing accounts. The implicit mentalising account would expect comparable effects for both head-front-gaze-averted and head-turned-gaze-averted faces as visible eye cues persisted. By contrast, the submentalising account claiming the contribution of attentional orienting to the consistency effect of the dot-perspective task would result in a prediction that was in accordance with Experiment 5A's findings; specifically, weaker effects would be elicited for head-turned-gaze-averted faces than for head-front-gaze-averted faces in the dot-perspective task.

5.2.2.1 Participants

A new sample of 42 undergraduates from the IPRP system participated in the experiment. Eight participants were excluded (see Results), leaving 34 volunteers (27 females, mean age: 19.6 years; age range: 18-31 years) for further analysis. G*power can run the priori power analysis for the one within-subject factor 'Consistency'. The power analysis conducted using G*power 3.1 in Experiment 3 recommended that 24 participants were needed to detect the consistency effect found by Samson et al. (2010) with the effect size of 0.61, an alpha level of 0.05, and a power of 80%.

5.2.2.2 Stimuli

All the stimuli were the same as those of Experiment 5A with an exception of the distribution of disc(s) (about 0.7° × 0.9°, 3.1° - 3.2° from the face) that was/were displayed vertically on one or two lateral walls (see examples in Figure 13C and 13D). Each of the ten dot layouts (that were the same as Samson et al.'s (2010) disc layouts) was equivalently combined with a head-front-gaze-averted or head-turned-gaze-averted face. On 50% of trials, the number of disc(s) that the participant could see was the same as the number to which the central face was directed (congruent condition), but the above comparison was different on the remaining trials (incongruent condition).

5.2.2.3 Procedure

Participants completed the dot-perspective task as an online experiment. As illustrated in Figure 14B, trials began with a fixation cross presented for 750 ms in a virtual room. A 500-ms-blank later, the word "you" (i.e., the participants were instructed to judge self-perspective and ignore the centrally presented faces in the room) was presented for 750 ms. After a 500-ms-interval, a digit (0-3) appeared for 750 ms. Then, 500 ms later, the scene of a room with a central stimulus and dot(s) on the lateral wall(s) was displayed for 3000 ms. The participants were asked to judge whether the word-and-number corresponded to the scenario of the room as quickly and accurately as possible. Specifically, pressing '1' for "Yes" response (i.e., matching response) by the forefinger of the left hand, pressing '2' for "No" response (i.e., mismatching response) by the forefinger of the right hand. There would be a 500-ms-interval between trials with the room remaining present.

Every block consisted of 48 test trials (6 congruent with head-front-gaze-averted face, 6 incongruent with head-front-gaze-averted face, 6 congruent with head-turned-gaze-averted face, 6 incongruent with head-turned-gaze-averted face) and 4 filler trials equally organised into matching and mismatching trials. In filler trials, a face was centrally presented in a room but without dots on the wall. Together, the experiment commenced with a block of 26 practise trials followed by 4 blocks of 52 test trials. Trial presentation order was pseudo-randomised in order not to have more than 3 consecutive trials of the same condition throughout the experiment, and the order was fixed across participants. Trials were arranged into two consecutive blocks for head-front-gaze-averted faces and two consecutive blocks for head-turned-gaze-averted faces, with the presentation order of blocks counterbalanced across participants.

5.2.2.4 Results

Response time and accuracy were submitted to a 2×2 repeated-measures analysis of variance (ANOVA) with "Central stimulus type" (Head-front-gaze-averted vs. Head-turned-gaze-averted) and "Consistency" (Consistent vs. Inconsistent) as within-subject independent variables. In accordance with Samson et al.'s (2010) study, only matching trials were analysed. Filler trials were eliminated from further analysis. The dot-perspective task in Experiment 5B (also in Experiment 6B and Experiment 7) applied the following criteria for data analysis. First, participants with an accuracy less than 70% on average or 60% in any experimental condition would be excluded from the data set. Furthermore, participants with response times exceeding the range 'mean ± 2.5 SDs' were also be eliminated for further analysis. In Experiment 5B, six participants were eliminated from further analysis due to low accuracy, and the remaining two participants were eliminated because of excessively slow response times. Accordingly, Experiment 5B was based on data from 34 participants.

In terms of accuracy, neither main effect nor interaction effect reached a significant difference (all *p*s > 0.05). Overall, there was a high accuracy in the dot-perspective task; 34 participants performed accurately more than 90% of the trials in all experimental conditions (see Table 9).

Only accurate trials were chosen for the analysis of response time. Data analysis showed a significant main effect of Consistency (F (1,33) = 9.40, p = 0.004, η_p^2 = 0.22). Post hoc test and the computation of Cohen's d indicated slower response times for the inconsistent trials (mean ± standard error, 723.28 ± 25.27 ms) when compared with the consistent trials (691.71 ± 23.47 ms) (t(33) = -3.07, p = 0.004, Cohen's d = 0.53, medium to large effect size) (see Table 9 below and Figure 15C on page 88). No other effects were significant (p > 0.15).

Table 9. Behavioural Results [Mean (Standard Error)] in Conditions of Experiment 5B.

Experiment 5B	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent$^{** >}$
Head-front-gaze-averted Face	94.40 (1.20)	94.50 (1.20)	701.85 (31.32)	720.33 (29.69)
Head-turned-gaze-averted Face	96.60 (0.80)	95.20 (1.30)	681.58 (20.44)	726.23 (24.64)

$^{>}$ Indicates a significant increase from the other condition; $^{<}$ Indicates a significant decrease from the other condition.
* $p < .05$; ** $p < .01$; *** $p < .001$.

5.2.2.5 Summary

Contrary to the submentalising account's prediction, there were comparable consistency effects for both head-front-gaze-averted and head-turned-gaze-averted faces. Instead, the findings extend previous findings of implicit mentalising triggered by a virtual avatar (e.g., Samson et al., 2010) to a real human's face, suggesting that Samson et al.'s (2010) findings are robust. Similarly, in Experiment 5B, the implicit mentalising was reflected by the altercentric interference effect－participants' automatic calculation of the agent's visual perspective interfered with their own perspectives, and thereby they performed more slowly in the incongruent condition compared with the congruent condition when they were required to judge their own perspectives. More intriguingly, unlike modulation of attentional orienting by eye-head-cues in Experiment 5A, there was no modulation of visual perspective-taking by eye-head cues in Experiment 5B. Rather than eye-head directional cuing, the agent's visual access seems to be crucial for affecting implicit mentalising in the dot-perspective task. Because both head-front-gaze-averted and head-turned-gaze-averted faces have visible averted-gaze, the two kinds of faces trigger implicit mentalising. In sum, the findings in Experiments 5A and 5B favour the implicit mentalising account by casting doubt on the submentalising account.

Apart from eye and head directions, finger-pointing direction is another effective and important cue in social interactions. For eye-head cues of the current experiment, visual access predominates the implicit mentalising in the dot-perspective task. Experiments 6 and 7 explored whether the influence of visual access on the implicit mentalising persisted when eyes were presented simultaneously with finger-pointing, which might provide new insight into the implicit mentalising vs. submentalising debate.

5.3 Experiment 6: Overview and Hypotheses

The dissociation of the two competing accounts via using eye-head cues provided impetus to find a new effective way—manipulating the agent's different body parts and comparing their influences on cue-validity effect in the Posner task with those on consistency effect in the dot-perspective task—to clarify the debate. Experiment 6 aimed to explore whether manipulation of eye-finger-pointing cues (i.e., gaze-averted vs. finger-pointing) could dissociate the two competing accounts as eye-head cues did.

As described above, the agent's finger-pointing showed superiority of capturing others' attentions relative to the agent's eye-gaze for young children in the modified Posner task (Gregory et al., 2016). However, whether the superiority could also happen for adults as well as in a classic attentional-orienting task (i.e., Posner task) remained unknown. Accordingly, Experiment 6A addressed the issue by manipulating eye-finger-pointing cues (gaze-averted vs. finger-pointing) and comparing the cues' effects on attentional orienting in the Posner task. Furthermore, the role of finger-pointing in L1VPT processing and its comparison with the role of eye-gaze remained unknown. Experiment 6B measured the influence of finger-pointing on the effect of the dot-perspective task and compared it with the influence of eye-gaze.

5.3.1 Experiment 6A (Posner task): Overview and Hypotheses

Experiment 6A explored whether eye-finger-pointing cue manipulation (i.e., gaze-averted vs. finger-pointing) would moderate adults' attentional orienting via the cue-validity effect in the Posner task. Similar to Gregory et al.'s (2016) findings, it was posited that adults would show superiority of finger-pointing in modulating attentional orienting when compared with gaze-averted. Specifically, compared with gaze-averted agents, finger-pointing agents should evoke stronger cue-validity effects in the Posner task. The prediction was based on the statement that finger-pointing is a stronger directional cue than gaze-averted for adults as the former one (i.e., three-dimension) is more perceptually salient than the latter one (i.e., two-dimension).

5.3.1.1 Participants

A new group of 38 first-year undergraduates signed up to the IPRP system of VUW for experimental participation, with three participants' exclusions (see Results). The sample size of 35 participants (28 females, mean age: 19.3 years; age range: 18-30 years) was chosen based on that of Experiment 5A and Gardner, Bileviciute, et al.'s (2018) work.

5.3.1.2 Stimuli and Procedure

Apart from the pictures taken in Experiment 5, photographs of the volunteer also involved the simultaneous presentations of her face and upper part of the body. Separating eye direction from the body gesture's direction, the gaze-averted person (i.e., a person with 45° left or right averted eyes and with no body gesture) and the finger-pointing person (i.e., a person with direct eye-gaze and approximately 45° left or right averted finger-pointing) were generated. The gaze-averted person could see the dot(s) where her eyes were directed to, whereas the finger-pointing person could not see the dot(s). In fitting with the central stimuli, the sizes of all the stimuli in the scene were changed when compared with those of the stimuli in Experiment 5. Specifically, the participant (viewing at a distance of about 72 cm from a 14-inch computer screen) would observe a scene where a gazer was centrally presented and a disc (about 1.0° × 1.2°, 4.4° from the face) were vertically arranged on one lateral wall of a room (14.3° × 13.3°). An image of the gazer consisted of a frontal-view face (2.9° × 3.8°) with 0.2°-width-eyes and a frontal-view upper body (7.4° × 3.8°, see examples in Figure 16A and 16B). The scenarios of valid and invalid conditions were similar to those of Experiment 5A.

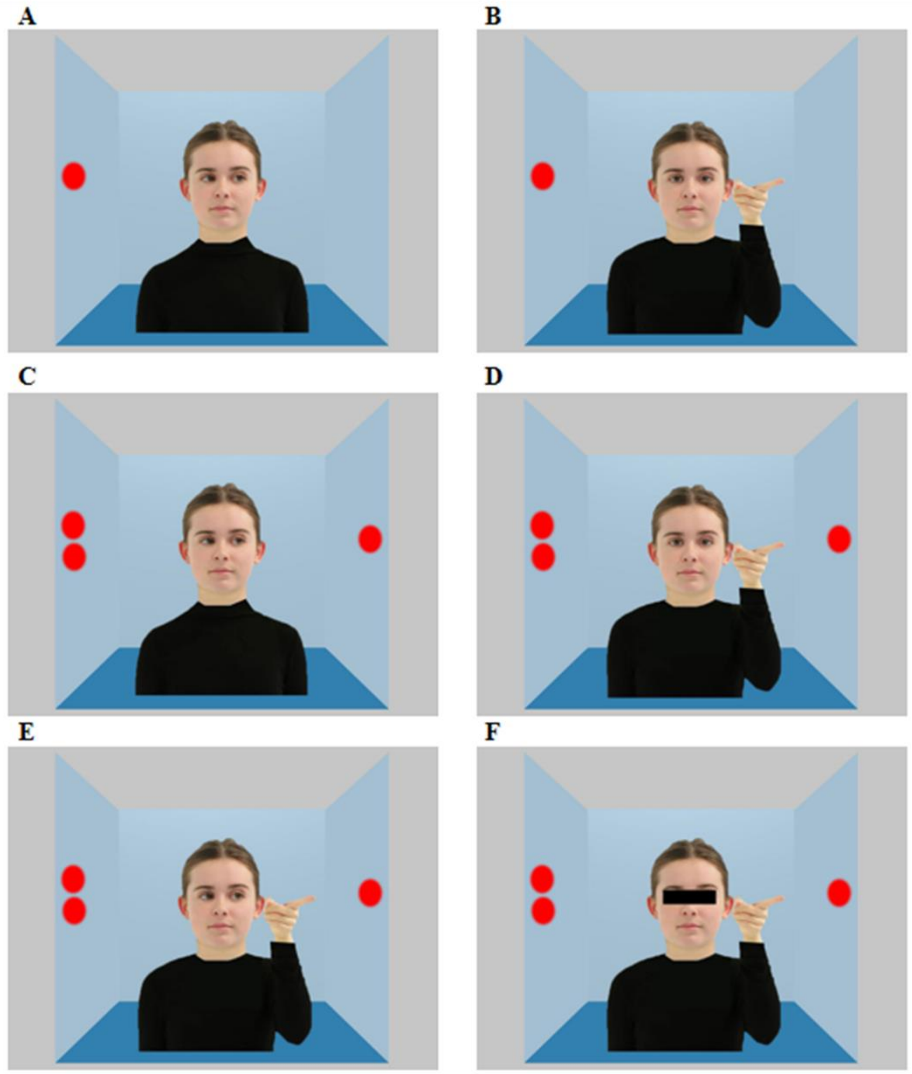

FIGURE 16 | Examples of room images used in Experiment 6A (A: Gaze-averted Agent, B: Finger-pointing Agent), Experiment 6B (C: Gaze-averted Agent, D: Finger-pointing Agent), and Experiment 7 (E: Eyes-opened-finger-pointing Agent, F: Eyes-covered-finger-pointing Agent).

20 adults who did not take part in the formal experiment completed a stimulus evaluation before the formal experiment. They were asked to judge if the centrally presented agent could see the disc in the valid and invalid conditions. There were eight images where one disc on the left/right wall with a left/right-viewing eye-gaze or a left/right-directed

finger-pointing were presented. All the participants reported that the gaze-averted agent could see the disc whereas the finger-pointing agent could not in the congruent condition (i.e., a valid condition in Experiment 6A and consistent condition in Experiment 6B), whereas all the agents could not see the disc in the incongruent condition (i.e., an invalid condition in Experiment 6A and inconsistent condition in Experiment 6B). The procedure of Experiment 6A was the same as that of Experiment 5A.

5.3.1.3 Results

The methods for data analysis as well as the standards for data exclusion were the same as those of Experiment 5A. Three participants were excluded from the data analysis because their response times were not within the range 'mean ± 2.5 SDs', leaving 35 participants' data for analysis. The 3 (SOA: 100 ms vs. 300 ms vs. 700 ms) × 2 (Central stimulus type: gaze-averted vs. finger-pointing) × 2 (Validity: valid vs. invalid) repeated-measures ANOVA was carried out for data analysis, and the percentage of error and response time were adopted as dependent variables. On average, the participants made erroneous responses in 5.67% of catch trials.

With regard to reaction time, only correct trials of no-catch trials were analysed. There was a significant SOA x Validity interaction effect (F (2,68) = 6.03, p = 0.004, η_p^2 = 0.15). Paired-sample t-tests and the computation of the value of the Cohen's d effect size showed a significant Validity effect when SOA duration was 700 ms (t(34) = -4.44, p < 0.001, Cohen's d = 0.75, medium and large effect size) but not 100 ms (t(34) = -0.52, p =

0.61) or 300 ms ($t(34) = -1.51$, $p = 0.14$). Specifically, when the SOA duration was 700 ms, the participant detected the targeted dot more slowly in the invalid condition (355.09 ± 8.03 ms) relative to the valid condition (340.94 ± 7.25 ms). In contrast, the participant completed the Posner task as quickly in the valid condition as in the invalid condition both when the SOA duration was 100 ms and when the SOA duration was 300 ms (see Table 10 and Figure 17A). Furthermore, there was a marginal interaction effect between Central stimulus type and Validity ($F(1,34) = 3.45$, $p = 0.072$, $\eta_p^2 = 0.092$). Paired-sample t-tests and the computation of Cohen's d indicated a significant Validity effect when the directional cue was a finger-pointing agent ($t(34) = -4.13$, $p < 0.001$, Cohen's d = 0.70, medium and large effect size) but not a gaze-averted agent ($t(34) = -1.29$, $p = 0.21$). Specifically, participants performed more slowly on the invalid trials (368.83 ± 8.03 ms) than on the valid trials (359.03 ± 7.46 ms) when the directional cue was a finger-pointing agent. In contrast, there was not a significant difference of response time between the valid condition (362.34 ± 7.56 ms) and the invalid condition (365.39 ± 7.44 ms) when the directional cue was a gaze-averted agent (see Table 10 and Figure 17B). Additionally, a significant main effect of SOA ($F(2,68) = 120.19$, $p < 0.001$, $\eta_p^2 = 0.78$) showed slower reaction times in an SOA of 100 ms (389.60 ± 7.00 ms) compared to an SOA of 300 ms (354.08 ± 6.82 ms, $p = 0.002$) or 700 ms (348.01 ± 7.48 ms, $p < 0.001$), but no significant difference between 300 ms and 700 ms ($p = 1.00$). No other effects were significant ($p > 0.45$).

Table 10. Behavioural Results [Mean (Standard Error)] in Conditions of Experiment 6A.

Experiment 6A	Reaction time (ms)			
	Gaze-averted		Finger-pointing	
	Valid	Invalid	Valid	Invalid*** >
SOA (100 ms)	390.98	389.06	387.02	391.35
	(8.24)	(7.94)	(7.62)	(8.22)
SOA (300 ms)	355.74	353.16	348.49	358.94
	(7.72)	(7.64)	(8.25)	(8.25)
SOA (700 ms)	340.31	353.96*** >	341.57	356.21
	(7.73)	(8.51)	(8.03)	(9.27)

> Indicates a significant increase from the other condition; < Indicates a significant decrease from the other condition.
* p < .05; ** p < .01; *** p < .001.

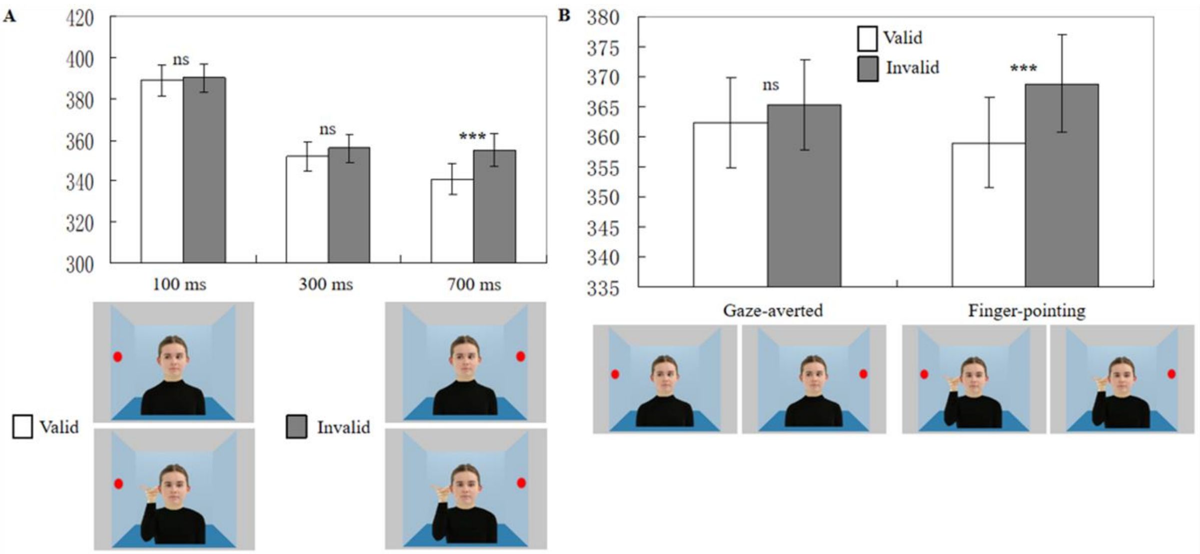

FIGURE 17 | Reaction times in each condition of Experiment 6A (A and B). Means and standard errors are shown; error bars refer to ±1 standard error of mean; * $p < .05$, ** $p < .01$, *** $p < .001$.

5.3.1.4 Summary

Consistent with Experiment 5A's, Gardner, Bileviciute, et al.'s (2018) and Gardner, Hull, et al.'s (2018) findings, there was a volitional attentional orienting effect (i.e., a significant cue-validity effect that occurred only under the condition of a longer SOA being 700 ms) that was triggered by a directional cue (i.e., finger-pointing direction). More intriguingly, there was a greater cue-validity effect induced by finger-pointing relative to gaze-averted, supporting the expectations. The resulting pattern was consistent with previous findings of preschool children's showing superior performance for lateral-oriented finger-pointing than for gaze-averted (Gregory et al., 2016). However, the non-significant cue-validity effect for gaze-averted agents is inconsistent with the classic gaze cueing effect. One possible interpretation may be the salience of the directional information. Specifically, averted-gaze may not be sufficient to trigger attentional orienting when compared to finger-pointing direction as finger-pointing is relatively salient so that averted-gaze may be ignored.

5.3.2 Experiment 6B (Dot-perspective task): Overview and Hypotheses

Experiment 6B examined whether implicit mentalising could be generated in the dot-perspective task by manipulating eye-finger-pointing cues (i.e., gaze-averted vs. finger-pointing). Two predictions were proposed based on the two competing accounts. The implicit mentalising account emphasizing the importance of line-of-sight would predict a consistency effect of the dot-perspective task for gaze-averted agents as visual access was available. Heyes's (2014) submentalising claim that attentional orienting contributes to the consistency effect of the dot-perspective task would lead to a prediction consistent with the findings of the Posner task (a classic task tapping attentional orienting) in Experiment 6B. Specifically, the effect in the dot-perspective task would be present for finger-pointing agents but absent for gaze-averted agents (given that a finger-pointing cue is more perceptually salient than a gaze-aversion cue).

5.3.2.1 Participants

Another fifty-two undergraduates from IPRP system took part in the experiment and ten participants were excluded from the data set (see Results). Power analysis was done in the same way as in Experiment 5B. The sample size of 42 (34 females, mean age: 19.9 years; age range: 18-42 years) exceeded 24 participants required to detect Samson et al.'s (2010) consistency effect.

5.3.2.2 Stimuli and Procedure

All the stimuli were the same as those of Experiment 6A with an exception that the disc(s) (about 1.0° × 1.2°, 4.4° - 4.6° from the face) were presented on one or two lateral walls (see examples in Figure 16C and 16D). The disc distributions in the present experiment were in accordance with Experiment 5B. The number of disc(s) that the participant could see and the central agent was directed to was the same in the consistent condition but were different in the inconsistent condition. The procedure used in Experiment 6B was identical to Experiment 5B.

5.3.2.3 Results

The methods for data analysis as well as the standards for data elimination were the same as the those of Experiment 5B. Nine participants' data were excluded due to low accuracy and another one participant's data was eliminated as its response time exceeded the range 'mean + 2.5 SDs'. Thus, 42 participants' data were selected for the 2 (Central stimulus type: gaze-averted vs. Finger-pointing) × 2 (Consistency: Consistent vs. Inconsistent) repeated-measures ANOVA, and accuracy as well as reaction time were used as dependent variables.

For accuracy, there was a marginally significant interaction effect between Central stimulus type and Consistency (F (1,41) = 2.97, p = 0.092, η_p^2 = 0.068). Paired-sample t-tests indicated that the participants performed better in the congruent trials (mean ± standard error = 96.10 ± 0.90 %) when compared to the incongruent trials (92.80 ± 1.30 %) when the central stimulus was a finger-pointing agent (t(41) = 2.56, p = 0.014) but not a gaze-averted agent (t(41) = 0.13, p = 0.89). Additionally, a significant main effect of Consistency (F(1,41) = 5.77, p = 0.021, η_p^2 = 0.12) was found, with higher accuracy in the consistent condition (95.60 ± 0.70 %) relative to inconsistent condition (93.80 ± 0.70 %) (see Table 11). There were no speed-accuracy trade-offs.

With respect to reaction time, only correct trials were selected for data analysis. There was a significant interaction effect between Central stimulus type and Consistency (F (1,41) = 6.41, p = 0.015, η_p^2 = 0.14). Paired-sample t-tests and the calculation of Cohen's d revealed that there was a significant Consistency effect when the central stimulus type was a finger-pointing agent ($t(41) = -5.22$, $p < 0.001$, Cohen's d = 0.81, large effect size) but not a gaze-averted agent ($t(41) = -1.06$, $p = 0.29$). Specifically, a finger-pointing agent made participants judge their own perspective more slowly in the inconsistent condition (763.74 ± 28.05 ms) relative to the consistent condition (717.89 ± 25.21 ms). In contrast, a gaze-averted agent made participants judge their own perspective as fast in the congruent condition (719.06 ± 25.23 ms) as in the incongruent condition (731.04 ± 28.17 ms) (see Table 11 and Figure 18A). Additionally, there was a significant main effect of Consistency (F (1,41) = 14.61, $p < 0.001$, η_p^2 = 0.26), showing that the participants performed more slowly on inconsistent trials (747.39 ± 25.18 ms) than on consistent trials (718.48 ± 23.01 ms).

Table 11. Behavioural Results [Mean (Standard Error)] in Conditions of Experiment 6B.

Experiment 6B	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent
Gaze-averted	95.00	94.90	719.06	731.04
	(0.80)	(0.80)	(25.23)	(28.17)
Finger-pointing	96.10	$92.80^{*<}$	717.89	$763.74^{***>}$
	(0.90)	(1.30)	(25.21)	(28.05)

$^{>}$ Indicates a significant increase from the other condition; $^{<}$ Indicates a significant decrease from the other condition.
* $p < .05$; ** $p < .01$; *** $p < .001$.

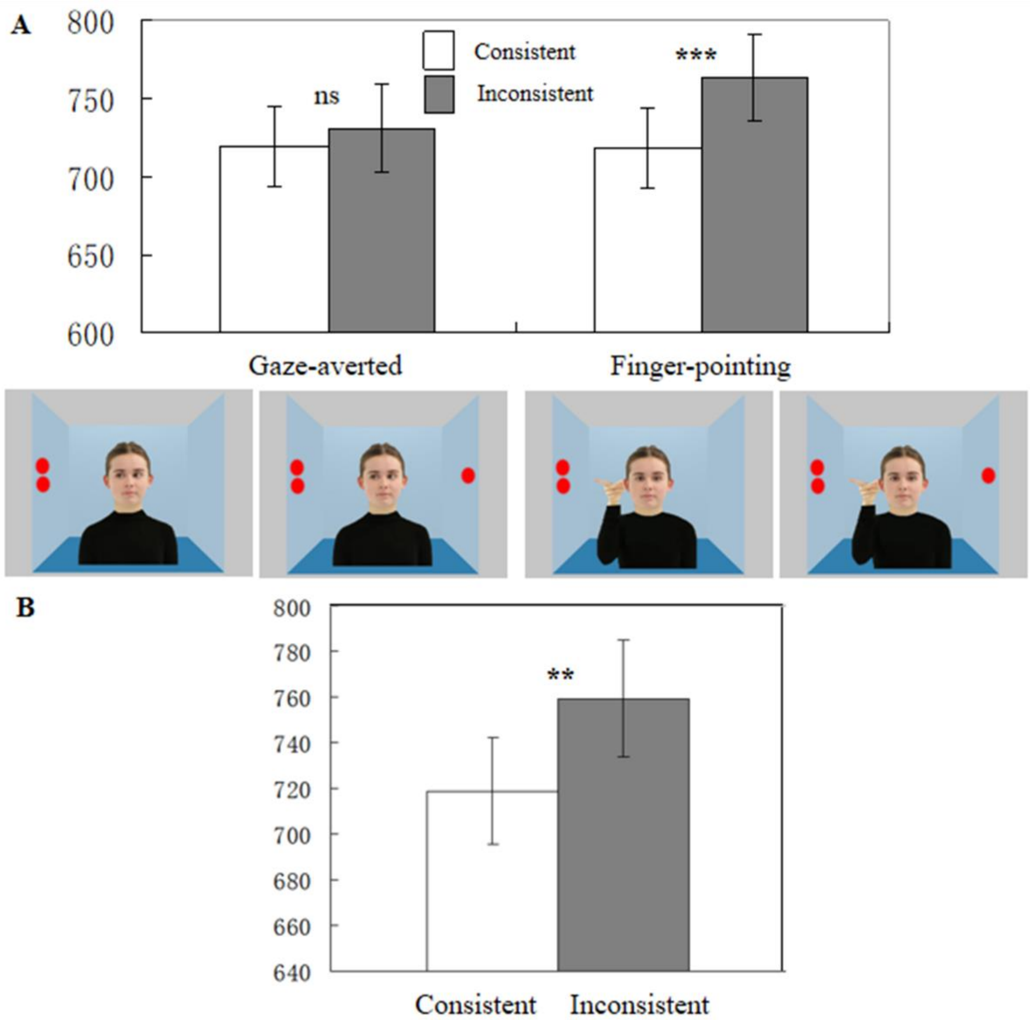

FIGURE 18 | Reaction times in each condition of Experiments 6B (A) and 7 (B). Means and standard errors are shown; error bars refer to ±1 standard error of mean; * $p < .05$, ** $p < .01$, *** $p < .001$.

5.3.2.4 Summary

Fitting with the submentalising hypothesis, there was an effect triggered by finger-pointing agents but not gaze-averted agents. The resulting pattern of Experiment 6B was the same as that of Experiment 6A. One possible interpretation is that the superiority effect of finger-pointing in the Posner task of Experiment 6A may be generalised to the dot-perspective task of Experiment 6B. Instead of visual access, finger-pointing may predominate in generating the effect in the dot-perspective task via an attentional orienting mechanism. If it was the case, then, Experiment 6's findings would support the submentalising account emphasizing the role of attentional orienting by casting doubt on the implicit mentalising account emphasizing the role of visual access. However, this is the first experiment exploring the debate by manipulating eye-finger-pointing cues and comparing the effects in both Posner and dot-perspective tasks, which needs to be further unpacked. Furthermore, finger-pointing may be so salient that the role of averted-gaze is decreased to be non-significant in both Posner and dot-perspective tasks. Thus, Experiment 7 sought to employ manipulation of the agent's visual access (a widely used manipulation to clarify the debate) and combine it with finger-pointing manipulation to further explore the processing mechanism of the dot-perspective task.

5.4 Experiment 7: Overview and Hypotheses

Experiment 7 investigated the implicit mentalising vs. submentalising debate by firstly combining a line-of-sight manipulation (i.e., eyes-opened vs. eyes-covered) with finger-pointing in the dot-perspective task. Both Furlanetto et al. (2016) and Experiment 3 manipulated the agent's visual access to find a significant consistency effect in the visible condition but not the invisible condition. The findings emphasized the importance of visual access and demonstrated people's capability to automatically calculate the agent's visual perspective, supporting the implicit mentalising account. However, in Experiment 6, finger-pointing showed superiority of capturing others' attentions relative to averted-gaze not only in the Posner task (tapping attentional orienting) but also in the dot-perspective task (tapping L1VPT). The findings raised a question about whether visual access could still predominate the processes in the dot-perspective task when finger-pointing would need to be also considered. Thus, to address the issue, manipulation of the agent's line-of-sight and finger-pointing was combined, and their influences were compared on the dot-perspective task. Based on Experiment 3's findings and the implicit mentalising account emphasizing the importance of visual access, one prediction was the elicitation of consistency effect for eyes-sighted agents but not for eyes-covered agents regardless of finger-pointing's directional information. According to Experiment 6's findings and the submentalising account emphasizing the importance of attentional orienting, another prediction was comparable effects for both eye-opened-finger-pointing and eye-covered-finger-pointing agents.

5.4.1 Participants

Fifty-seven first-year undergraduates from the IPRP system were recruited for participation with excluding 16 participants (see Results), leaving 41 participants (33 females, mean age: 19.9 years; age range: 18-33 years) for further analysis. The way to do power analysis was the same as Experiment 5B. The sample size exceeded 24 participants necessitated to detect Samson et al.'s (2010) third experiment's effect size.

5.4.2 Stimuli and Procedure

Photographs of the volunteer with averted-gaze as well as finger-pointing (i.e., a person with approximately 45° left or right averted finger-pointing that was compatible with eye direction) were taken for Experiment 7. The pictures consisted of two seeing conditions:

eyes-opened-finger-pointing (i.e., finger-pointing agent with 45° left/right-averted eye-gaze) vs. eyes-covered-finger-pointing (i.e., an original finger-pointing agent with a black rectangle on her eyes, 2.2° × 0.6°) (see examples in Figure 16E and 16F). The sizes of all the stimuli, as well as the disc distributions, were the same as those of Experiment 6B. The number of dot(s) that the participant could see and the central agent was directed to were the same in the consistent condition but were different in the inconsistent condition. Before the formal experiment, all the participants were asked to determine whether the eyes-opened-finger-pointing and eyes-covered-finger-pointing agents could see the three discs on a lateral wall where the agent was directed (i.e., where the agent looked for an eyes-opened-finger-pointing agent and where finger-pointing was directed for eyes-covered-finger-pointing agent), respectively, and all of them reported that an eyes-opened person could see (visible condition) whereas eyes-covered person could not (invisible condition). The procedure of Experiment 7 was the same as that of Experiment 6B except for the evaluations of the visibility of the central stimuli conducted before the formal experiment.

5.4.3 Results

The methods for data analysis as well as the standards for data exclusion were the same as those of Experiment 5B. A two-way repeated-measures ANOVA with Central stimulus type (Eyes-opened-finger-pointing vs. Eyes-covered-finger-pointing) and Consistency (Consistent vs. Inconsistent) as independent variables, and accuracy and response time as dependent variables was performed. Sixteen participants were excluded from further analysis, including fourteen participants' elimination because of low accuracy and another two participants' exclusion due to excessively slow response times (2.5 *SD* over the overall conditional means).

In terms of accuracy, the interaction effect between Central stimulus type and Consistency reached significance (F (1,40) = 5.16, p = 0.029, η_p^2 = 0.11). Paired-sample t-tests showed that the participants made more correct responses on consistent trials (mean ± standard error = 96.30 ± 1.10 %) relative to inconsistent trials (91.50 ± 1.40 %) when the central stimulus was an eyes-opened-finger-pointing person (t(40) = 4.23, p < 0.001) but not when the central stimulus was an eyes-covered-finger-pointing person (t(40) = 1.50, p = 0.14). Furthermore, there was a significant main effect of Consistency (F (1,40) = 17.87, p < 0.001, η_p^2 = 0.31), indicating better performance in the congruent condition (95.40 ± 0.80 %) than in the incongruent condition (92.30 ± 1.10 %) (see Table 12). There were no speed-accuracy trade-offs.

In terms of response time, only accurate trials were selected for data analysis. ANOVA analysis only showed a significant main effect of Consistency (F (1,40) = 13.33, p = 0.001, η_p^2 = 0.25). Post-hoc test and the calculation of Cohen's d indicated that the participants judged their own perspective more slowly in the incongruent condition (759.55 ± 25.55 ms) than in the congruent condition (718.75 ± 23.24 ms) (t(40) = -3.65, p = 0.001, Cohen's d = 0.57, medium to large effect size) (see Table 12 and Figure 18B on page 105). The main effect of Central stimulus type and the interaction effect between Central stimulus type and Consistency failed to reach significance (all ps > 0.05).

Table 12. Behavioural Results [Mean (Standard Error)] in Conditions of Experiment 7.

Experiment 7	Accuracy (%)		Reaction time (ms)	
	Consistent	Inconsistent	Consistent	Inconsistent** >
Eyes-opened-finger-pointing	96.30	91.50*** <	725.82	764.21
	(1.10)	(1.40)	(26.62)	(30.17)
Eyes-covered-finger-pointing	94.60	93.20	711.68	754.89
	(1.00)	(1.10)	(21.80)	(23.87)

[>] Indicates a significant increase from the other condition; [<] Indicates a significant decrease from the other condition.
* p < .05; ** p < .01; *** p < .001.

5.4.4 Summary

There was a comparable magnitude of consistency effects for both eyes-sighted-finger-pointing and eyes-covered-finger-pointing agents, favouring the submentalising hypothesis. The finding failed to replicate Experiment 3's findings of the elicitation of consistency effect in the visible condition but not invisible condition even though these two experiments used the same manipulation of visual access. It is possible to interpret that the inconsistent findings may be due to different roles of different body cues. In Experiment 3, for eye-head cues, visual access instead of head direction may predominate the processing in the dot-perspective task. In contrast, the mechanism of the dot-perspective task may be changed when finger-pointing was involved. Experiment 6's findings support the submentalising account by demonstrating finger-pointing may have superiority to trigger attentional orienting relative to averted-gaze, and the effect in the Posner task can be generalised to the dot-perspective task. Additionally, the lack of modulation of visual access in the dot-perspective task of Experiment 7, when finger-pointing was simultaneously presented, suggests that finger-pointing may override visual access to predominate the processing in the dot-perspective task via an attentional orienting mechanism. Accordingly, these findings lend support to the submentalising account by casting doubt on the implicit mentalising account.

Chapter 6: General Discussion

Previous research investigating the mechanism of L1VPT processing has suggested that there is an intense debate between the implicit mentalising account and the submentalising account. Three studies were carried out to find new ways to shed light on the debate. Study 1 replicated (Experiment 1) and extended (Experiment 2) Samson et al.'s (2010) results by finding that the self-consistency effect was induced by the original virtual avatars and persisted when the avatars were replaced by a real human's face. Study 2 indicated that implicit mentalising could be modulated by using a new manipulation of the agent's visual access (i.e., a black rectangle was placed on the agent's eyes to render a non-seeing condition) (Experiment 3). Additionally, findings of Experiment 4 result in a speculation that face-related directional cue's dynamic property might be related to the generation of attentional orienting but not implicit mentalising, providing new insight into the dissociation between the two competing accounts. Study 3 discovered that manipulation of eye-head cues (head-front-gaze-averted vs. head-turned-gaze-averted) could dissociate the two competing accounts (Experiment 5). However, the processing mechanism changed when finger-pointing was combined with eye-related manipulations. Specifically, both the cue-validity effect and the consistency effect could be triggered by finger-pointing but not averted-gaze (Experiment 6), and the modulation effect of visual access on implicit mentalising was interfered by finger-pointing (Experiment 7), supporting submentalising.

Below I discuss the overall implications of my findings. I first discuss whether manipulation of facial-related cues can dissociate the two competing accounts. I next discuss whether manipulation of different body parts can distinguish the two accounts. Finally, I address several avenues regarding the implicit mentalising vs. submentalising debate for further explorations and summarize the conclusions of the three studies.

6.1 Can the Competing Accounts be Dissociated through Facial-related Cue Manipulation?

6.1.1 Implicit Mentalising is Supported by Manipulating the Agent's Line-of-sight.

Study 2 conducted two experiments to contrast the implicit mentalising account from the submentalising account by using facial-related cues. Experiment 3 found that when the computer-generated avatar – which is typically used in almost all previous studies – was substituted with a photograph of a real human face to be the central stimulus, the consistency effect was still robustly produced. The consistency effect was absent when the stimuli involved an eyes-covered face. An intuitively appealing account of the findings is that participants efficiently tracked the agent's visual perspective by computing the agent's unoccluded line-of-sight to the disc(s) on the wall(s), which slows down participants' own judgements in the inconsistent condition of dot-perspective task. But the relatively automatic computation disappeared when line-of-sight was effectively blocked in the eyes-covered faces.

Of course, the eyes-covered faces still possessed directional features like head orientation (i.e., head oriented to the left- or right-side wall). The submentalising account claims that attentional orienting happens with any directional cue. The absence of the consistency effect in the non-seeing condition even when head orientation still existed would cast doubt on the submentalising account that general attentional orienting occurs in the dot-perspective task. Directional cues were presented in both classes of stimuli and yet the consistency effect was only present when the agent's mental state of seeing was not disrupted. The consistency effect in the seeing but not non-seeing conditions of dot-perspective task supports the implicit mentalising account.

Advocates of the submentalising account may argue that more directional cues in the seeing condition (i.e., eye direction + head orientation) than the non-seeing condition (i.e., head orientation only) triggered a greater attentional-orienting effect in the dot-perspective task. However, Bukowski et al. (2015) have found that the mere presence of avatar (with directional cue) was not sufficient to generate the consistency effect in the dot-perspective task, and that the effect was triggered due to the perspective-related task instruction. Thus, the generation of the consistency effect was related to mental states even though attentional orienting may exist in L1VPT processing.

It is important to be mindful of the possibility that the absence of the consistency effect in the non-seeing condition, even when there was head orientation, could be due merely to the head-turn cue being insufficiently strong in that context. Future research will need to delineate the boundary conditions in which head turn is a powerful attentional cue.

6.1.2 Dissociation between Competing Accounts May be Related to Dynamic Property of Directional Cue

Experiment 4A found consistency effects in the dot-perspective task, which is in line with a group of studies investigating processing mechanism of L1VPT (e.g., Surtees & Apperly, 2012; Surtees et al., 2016). Experiment 4B found a cue-validity effect in the

Posner's target detection task, which accords with prior studies exploring the processing mechanism of attentional orienting (e.g., Driver et al., 1999; Schuller, & Rossion, 2001; Friesen, & Kingstone, 1998; Van Rooijen et al., 2018). Paradigms used for measuring visual perspective-taking and attentional orienting are different in the following aspects. Firstly, "YOU" perspective instruction in the dot-perspective task is related to visual perspective selection whereas the target detection instruction in Posner task is relevant to attention shift. Secondly, the consistency effect was induced when the central stimulus and disc(s) were simultaneously presented, whereas the cue-validity effect was evoked when the central stimulus and a disc were presented with SOA. Therefore, the effects in Experiment 4A and 4B may be considered as tapping into different processes since distinct paradigms were adopted.

Most importantly though, the dot-perspective task yielded consistency effects of comparable size for head-front-gaze-averted and head-turned-gaze-maintained faces. This indicates that compared with directional information, visual access may play an important role in understanding the mental state of seeing. One explanation is that eyes are regarded as primary carrier of information into others' mental states (Grossmann, 2017), which may make eyes more important when people are computing others' visual perspectives. Because visual access is always available in both conditions and efficient mentalising (i.e., L1VPT) has been suggested to be tightly linked to eye cues (Sodian et al., 2007), adults can track others' visual perspectives via computation of line of sight in both conditions. Therefore, the findings of Experiment 4A fit with the implicit mentalising account regarding the role of visual access but not the prediction of submentalising based on Hietanen's work.

In Experiment 4B, head-front-gaze-adverted and head-turned-gaze-maintained faces induced cue-validity effects of comparable size. Head-front-gaze-adverted faces triggered attentional orienting similar to previous studies (e.g., Driver et al., 1999; Van Rooijen et al., 2018). However, the cue-validity effect for head-turned-gaze-maintained faces is inconsistent with Experiment 4B's hypothesis that is based on Hietanen's (1999) findings. In Hietanen's (1999) study, faces were presented for a relatively short duration (i.e., 50 ms), and furthermore, SOA durations (i.e., 170, 220 ms) were short. Situated within a limited time-window, certain directional features of faces may be too complex to be recognised or processed. In the current experiment, the presentation duration of face (i.e., 300 ms) and SOAs (i.e., 100, 300, 700 ms) may provide some time or be even long enough for participants to orient their attention towards the directional property of head-front-gaze-maintained faces.

Of course, the suggestion is not that the time course of stimuli presentation is the only factor at play to trigger the cue-validity effect for the head-turned-gaze-maintained condition. A tentative interpretation is the dynamic properties of the head-turned-gaze-maintained faces. That is, the head-turned-gaze-maintained face was presented immediately after a frontal face with direct gaze, which gives the faces a dynamic quality (an apparent motion, the illusion of motion, appears to be present when two still images are shown in quick succession of the formal experiment). Consequently, the dynamic head-turned-gaze-maintained face may have turned out to transmit powerful direction cuing that was salient enough to trigger the cue-validity effect. The explanation was, to a certain extent, supported by a finding of a cue-validity effect in the formal Experiment 4B with the aforementioned dynamic but not in the pilot experiment without that dynamic even though it is a cross-study comparison.

Altogether, in Experiment 4, cue-validity effect may be sufficiently triggered when the directional cue of the central face is conveyed through a dynamic display, which is not needed for triggering the consistency effect in the dot-perspective task. This suggests attentional orienting in Experiment 4B may, to a certain extent, be dissociated from visual perspective-taking in Experiment 4A. Further, the speculative interpretation regarding the dynamic of the directional cue provides a new insight into the potential dissociation, even though there was a lack of evidence to support the submentalising prediction consistent with Hietanen's (1999) pattern of cue-validity effects being elicited for head-front-gaze-averted but not for head-turned-gaze-maintained faces. One might potentially argue that, compared to head-front-gaze-averted faces (averted-eyes cue only), the head-turned-gaze-maintained faces could have plausibly served up more powerful directional cuing (i.e., eye + head), which may make head-turned-gaze-maintained faces more salient and easy to be processed. Then, head-turned-gaze-maintained faces should have triggered greater cue-validity effect than head-front-gaze-averted faces. However, in the current study, the apparent motion may make both head-front-gaze-averted and head-turned-gaze-maintained faces sufficiently salient, reducing the difference in direct comparisons between the two types of faces. Thus, both head-front-gaze-averted and head-turned-gaze-maintained faces trigger comparable magnitude of cue-validity effect.

Intriguingly, the findings of validity effects on all SOAs indicate reflexive attentional orienting plays the role in Posner task of Experiment 4B, and orienting may be specific to faces. Specifically, the validity effect in the Posner task of Experiment 4B emerges at a short SOA (i.e., 100 ms), which is similar to previous findings of the reflexive shift of spatial attention when SOA is 105 ms (Friesen, & Kingstone, 1998; Friesen et al., 2005). However, the validity effects were apparent only at longer SOAs in other studies (i.e., 300 ms, Bukowski et al., Experiment 1, 2015; 300 and 600 ms, Gardner, Bileviciute, et al., Experiment 1, 2018; 600 ms, Gardner, Hull, et al., Experiment 2, 2018), indicating a voluntary but not reflexive attention shift (Egeth & Yantis, 1997; Muller & Rabbit, 1989). The discrepancy between Experiment 4B and prior studies may be the stimuli employed. The directional information of faces (with clear averted gaze and/or head orientation of real humans) employed in Experiment 4B might have been more salient than compared to the informational features found in avatars (with computer-animated body direction and more ambiguous faces) used by previous studies; the former but not the latter stimuli may trigger participants' reflexive attention shift. Another potential explanation is that the apparent motion of the directional cue may make the attentional orienting emerge earlier when compared with the cue presented without the apparent motion. Thus, instead of volitional attentional orienting, reflexive attentional orienting was generated under the condition of the apparent motion.

Besides, the absence of the validity effect for avatar-stance-maintained avatars in Posner task of Gardner, Bileviciute, et al.'s (2018) study is contrary to the positive effects in prior studies (e.g., Bukowski et al., 2015; Gardner, Hull, et al., 2018). The authors have considered it as a consequence of the different low-level features of the avatar characters used. Additionally, other researchers may think that intermixed presentation of avatar-stance-maintained and avatar-stance-averted avatars can make the directional information of avatar-stance-maintained avatars more ambiguous, which may affect detection of the cue-validity effect. But whether blocking avatar-stance-maintained and avatar-stance-averted trials matters to the cue-validity effect in Posner task of Gardner, Bileviciute, et al.'s (2018) study warrants further investigation.

6.2 Can the Competing Accounts be Distinguished by Manipulating Different Body Parts?

6.2.1 Competing Accounts May be Dissociated by Manipulating the Eye-head Directional Cue.

Three experiments were carried out for Study 3, attempting to cast light on the implicit mentalising vs. submentalising debate. With respect to Experiment 5A, there was a significant cue-validity effect for head-front-gaze-averted faces, which was consistent with both the hypothesis and previous findings showing attentional orienting towards a real human's face with averted-gaze (e.g., Driver et al., 1999; Qian et al., 2013; Van Rooijen et al., 2018).

More importantly, consistent with Qian et al.'s (2013) work, there was a greater cue-validity effect for head-front-gaze-averted than for head-turned-gaze-averted faces, demonstrating that attentional orienting can be modulated by an eye-head directional cue in the Posner task. Compared with Hietanen's (1999) and the current Experiment 4B's work, Experiment 5A introduced a methodological advance. Specifically, Experiment 5A added the evaluation of gaze angle with reference to head orientation before the cueing experiment, ensuring that the levels of the 'Cue' factor were being clearly differentiated by its property. Unlike Qian et al.'s (2013) finding of a weaker effect for head-turned-gaze-averted faces, this thesis did find a non-significant effect for head-turned-gaze-averted faces. The discrepancy may be because of a bigger difference in gaze angle evaluation between head-front-gaze-averted and head-turned-gaze-averted faces of the present experiment (i.e., 2.6°) compared with Qian et al.'s (2013) study (i.e., 0.8°). In the present study, the greater difference may increase the salience of head-front-gaze-averted faces and decrease the salience of head-turned-gaze-averted faces, which raises the possibility of the absence of the cue-validity effect for head-turned-gaze-averted faces. Nevertheless, it should be mindful that cross-study comparisons are difficult to provide definitive evidence (e.g., lack of randomisation to conditions).

Moreover, the cue-validity effect was observed only when SOA is 700 ms, namely, when voluntary attentional orienting effect is regarded to occur (Müller & Findlay, 1988; Müller & Rabbitt, 1989). The volitional orienting effect fits with orienting effects in previous investigations examining the avatar-stance as a modulator of attentional orienting (Gardner, Bileviciute, et al., 2018), and avatars as spatial cues (Bukowski et al., 2015; Cole et al., 2017; Gardner, Hull, et al., 2018). Overall, Experiment 5A's findings suggest that eye-head directional cue moderate the volitional attentional orienting effect. However, Experiments 5A and 4B found different kinds of attentional orienting elicited by faces. One possible interpretation is that apparent motion of the directional cue in Experiment 4B may make the cue more salient, leading to the earlier appearance of the cue-validity effect (i.e., the reflexive attentional orienting).

Experiment 5B found that when the widely used virtual avatar was substituted with a photograph of a real human's face to be the central stimulus, the consistency effect was robustly yielded. Moreover, a comparable magnitude of consistency effects were observed for both head-front-gaze-averted and head-turned-gaze-averted faces in the dot-perspective task, lending support to the implicit mentalising hypothesis but speaking against the submentalising hypothesis. An intuitively appealing account of elicitation of the consistency effect in both types of faces is that the always available visual access instead of eye-head directional information may allow the participants' relatively automatic computation of the agent's visual perspective, thereby, yielding consistency effect for both the two kinds of faces under Self-perspective-instruction.

In summary, the findings demonstrate that attentional orienting but not implicit mentalising is modulated by gaze direction with reference to head orientation, and visual access may play a crucial role in triggering implicit mentalising in the dot-perspective task. These findings, to a certain extent, lend support to the implicit mentalising account by dissociating it from the submentalising account.

6.2.2 A New Viewpoint to Understand the Two Competing Accounts by Using Eye-finger-pointing Cues.

Compared with eye-head cues, eye-finger-pointing cues show a different mechanism in processes of both the Posner task and dot-perspective task, providing a new viewpoint to understand the implicit mentalising vs. submentalising debate.

Experiment 6A supports the hypothesis by finding a greater cue-validity effect triggered by finger-pointing agents than by gaze-averted agents in the Posner task. The superiority of finger-pointing was, to a certain extent, compatible with Gregory et al.'s (2016) findings showing young children's better performance for finger-pointing relative to eye-gaze in the modified Posner task. The results extend previous findings by revealing finger-pointing's superiority in adults as well as in the classic Posner task. These findings can be accounted for by a possibility that the greater perceptual variation of finger-pointing (i.e., three-dimension) than that of averted-gaze (i.e., two-dimension) may make finger-pointing gestures more recognizable, leading to the superiority of perceiving finger-pointing cues relative to gaze-averted cues. Thus, finger-pointing agents, in providing a stronger directional cue, triggered greater attentional orienting compared with gaze-averted agents.

Additionally, there was a non-significant cue-validity effect for gaze-averted agents. One possibility may be that compared with the apparent finger-pointing agents, the perceptual variation of gaze-averted agents may not be salient enough to generate the potential weaker attentional orienting. The explanation also fits with Butterworth's (1991) statement that finger-pointing may be a more accurate spatial cue relative to eye direction. Compatible with Experiment 5A's, Gardner, Bileviciute, et al.'s (2018) and Gardner, Hull, et al.'s (2018) findings, Experiment 6A's findings indicate that finger-pointing agents can trigger a volitional attentional orienting by observing a significant cue-validity effect only in a longer SOA (i.e., 700 ms).

Experiment 6B tried to associate eye-finger-pointing cues with visual perspective-taking via a consistency effect of the dot-perspective task. A greater consistency effect for finger-pointing agents than for gaze-averted agents was observed, supporting the submentalising hypothesis. The same resulting pattern of Experiments 6A and 6B may reveal the same mechanism for processing both the Posner and the dot-perspective task. Specifically, the superiority of finger-pointing relative to averted-gaze in generating volitional attentional orienting can be generalised from the Posner task to the dot-perspective task. Findings of Experiment 6, together, suggest that finger-pointing direction rather than gaze direction triggers volitional attentional orienting but not visual perspective-taking. Different from stressing the importance of visual access in eye-head cues of the dot-perspective task, finger-pointing may predominate in generating the effect in the dot-perspective task via an attentional orienting mechanism for eye-finger-pointing cues. Overall, Experiment 6's findings challenge the implicit mentalising account emphasizing the importance of the agent's visual access.

However, it is a novel experiment exploring the effect of finger-pointing in both the Posner task and the dot-perspective task and comparing the effects with those triggered by averted-gaze. Further, it is possible to question whether the manipulation of eye-gaze in Experiment 6 was salient enough to trigger the potential effects when compared with the finger-pointing. Experiment 7 adopted an apparent, easily recognizable gaze-related manipulation (i.e., manipulating the agent's visual access to create visible and invisible conditions) as the manipulation has been found to modulate the consistency effect in the dot-perspective task. Furthermore, Experiment 7 compared the manipulation of visual access with that of finger-pointing in modulating the effect in the dot-perspective task.

Experiment 7 observed comparable magnitude of consistency effects for both eyes-sighted-finger-pointing and eyes-covered-finger-pointing agents, supporting the submentalising hypothesis. The findings fit with neither Furlanetto et al.'s (2016) nor Experiment 3's findings of a significant consistency effect in the visible condition but not in the invisible condition. The contradictory findings may be because processing mechanism of the dot-perspective task is different when considering finger-pointing. In the two previous experiments, manipulation of the agent's visual access alongside beliefs of clearly distinguishing visible condition from the invisible condition can modulate the consistency effect. The effect was absent even though directional cue(s) (i.e., head + torso directions or only head direction) is identical for both seeing and non-seeing conditions. In these two studies, visual access is superior to directional cuing (head + torso/head directions) in modulating implicit mentalising in the dot-perspective task, which supports the implicit mentalising account but casts doubt on the submentalising account. However, in the current experiment, the dot-perspective task's effect induced by finger-pointing direction cannot be modulated by the manipulation of visual access. The line-of-sight manipulation was the same as Experiment 3's manipulation. Nevertheless, the comparable magnitude of effects for the two visibility conditions was observed. Thus, finger-pointing direction may be superior to visual access in moderating the effect in the dot-perspective task when finger-pointing is considered, casting doubt on the implicit mentalising account emphasizing the role of visual access. The interpretation also fits with finger-pointing vs. gaze-averted superiority's contribution to the effect in the dot-perspective task of Experiment 6B via an attentional orienting mechanism. Additionally, averted-gaze was more salient than head orientation when considering gaze and head directions, whereas averted-gaze was less salient than finger-pointing-direction when considering gaze and pointing directions. Therefore, the hierarchy in the salience of directional cue could explain the aforementioned superiority of visual access and of finger-pointing.

In addition to the two competing accounts regarding human being's L1VPT capability, animal-related studies have provided new insight into understanding the mechanism of L1VPT processing. For instance, jackdaws took human eye gaze into account and were loathe to retrieve their hidden food in front of a watching human, suggesting that birds have learned that retrieving food in front of a watching individual results in their food being stolen at some point in time (von Bayern & Emery, 2009). In other words, there is a behavioural consequence (i.e., food stealing) if jackdaws retrieve their food in front of another and so they refrain from doing so. It is likely just that we human beings are primed to take into account others' attentional cues (head turns, eye gaze, pointing) because these cues signal important things in the real world (danger, attraction, food sources, etc.). Therefore, we can't help but compute these cues, yet we don't need to attribute mental states when we do so. During the dot-perspective task, participants automatically computed another person's gaze or pointing direction, which made them look towards the gazed-at/pointed-to stimulus. Having attended to that stimulus, it then interfered with the participant's own judgement as to how many dots were present if the participant and someone else saw a different number of dots. Obviously, we adults can compute others' mental states but this doesn't mean we automatically do so when we see such events. Indeed, that would be laborious and would slow down responding when evolution has likely called for an immediate response to enhance survival like jackdaws do. Altogether, the consistency effect in the dot-perspective task may be generated through an animal survival mechanism in evolution.

In addition to the existence of attentional-orienting-related process, mentalising still plays a role in L1VPT processing. Firstly, task instruction of perspective judgement but not the presence of someone else with directional information generated the consistency effect in the dot-perspective task. Additionally, moths with "eyes" on their wings are thought to exist because the eyes look like the eyes of the moth's predators. Thus, someone may think that moth predators are sensitive to eyes but don't need to mentalise, and neither do participants in the present studies. However, mental state attributions can modulate the gaze-elicited attentional orienting effect (Morgan, Freeth, & Smith, 2018). Accordingly, in Experiment 3, manipulation check of the agent's seeing and non-seeing conditions could result in attributions of the epistemic state of seeing, which mediated the attentional orienting triggered by the agent's directional information. Finally, the findings of Experiments 4 and 5 indicated that the manipulations of directional cues (i.e., apparent motion, gaze + head directions) could modulate attentional orienting in the Posner task, whereas the manipulations did not influence the stability of the consistency effect in the dot-perspective task. Thus, we could not rule out the existence of mentalising in the dot-perspective task. Taken together, it may be time to integrate the implicit mentalising account with the submentalising account.

6.3 Future Directions

Future directions could address the following issues. Firstly, it is possible that Experiment 4B's effects may be specific to the ways that the stimuli were constructed, and future research may need to examine whether Hietanen's (1999) findings may be replicated with the stimuli used here. Further research will also be needed to determine exactly how presentation time of faces and SOA durations (e.g., temporal information of Hietanen's (1999) study, as well as from the current study) might interact with the dynamics of head-turned-gaze-maintained faces to trigger the cue-validity effect. This, then, may be beneficial for the investigations of the potential modulation of the face-related directional cue on attentional orienting. Additionally, instead of using static stimuli, dynamic stimuli (e.g., video clips involving a face with head turning or the body with pointing movement) can be used for further elucidating the processing mechanism of L1VPT as dynamic stimuli are better indicators of attention than static stimuli.

Secondly, researchers have considered that the orienting effect may be larger when making calculations over the gaze of a real person than for schematic gaze cues (e.g., Cole, Smith, and Atkinson, 2015). It may be useful to compare which configuration, the real face, the computer-generated face, eyes of the real face, eyes of the computer-generated face, contributes more to L1VPT. This can be helpful for unpacking further the processing mechanisms of L1VPT, which may then provide more evidence for resolving the longstanding debate between the implicit mentalising and submentalising accounts of human social cognition.

Thirdly, researchers could strive to find other effective ways to manipulate a gazer's epistemic visual access other than by blocking a gazer's line-of-sight per se. For instance, there is accumulating evidence suggesting that anthropomorphic characters whose face appears human but is not quite lifelike enough (thus dipping into the 'uncanny valley') evoke in us cold and eerie feelings (Waytz, Gray, Epley, & Wegner, 2010). Researchers could manipulate the eeriness (lifelikeness) of real human beings' eyes, making them inanimate so that participants believe inanimate eyes may not see objects around them (non-seeing condition), and compare the consistency effect of this condition with that of seeing condition (normal faces), providing new evidence for the implicit mentalising versus submentalising debate.

Fourthly, researchers may compare whether normal adults and psychopathic patients show different performance in the processes related to the aforementioned two points, as psychopathic individuals have deficits in efficiently and implicitly taking others' visual perspective (Drayton et al., 2018). This may be even beneficial to further understand the deficit in social behaviour related to dysfunction in L1VPT ability.

Fifthly, although finger-pointing appears not to trigger implicit mentalising in the dot-perspective task, this cannot rule out the possibility that finger-pointing may generate mentalising-related processes. Tomasello and Camaioni (1997) claimed that using finger-pointing for declarative purposes depends on a person's mentalising ability, and further, production of declarative finger-pointing has been found to be positively correlated with the comprehension of intentions (Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004). Consequently, it invites further work examining whether finger-pointing can trigger mentalising-related processes and comparing the contribution of finger-pointing with that of the agent's visual access by using other mentalising tasks.

Finally, more factors should be considered for further analyses. When analysing RT, mean RT latencies were used in all the analyses, and the analyses were conducted on the raw latencies. Future analysis should judge whether the distributions of latencies were normal or skewed, and attempt to use median RT latencies in the analyses. This is because skewness of the distribution of the difference can essentially destroy the power of inferential tests using the mean whereas tests on median are hardly influenced. Additionally, future research can also study L1VPT effects more thoroughly by assessing the Bayes Factors associated with the consistency effect for varying manipulations and conditions. When an interaction (or any other effects for that matter) is not significant ($p > .05$), this does not mean that the opposite conclusion can be drawn (i.e., one cannot conclude that the consistency/cueing effect is comparable). This is because a frequentist approach of reading the p-value allows us to reject the null hypothesis when a p-value is smaller than .05, but when a p-value is larger than .05, it cannot distinguish between the possibility of there being no effect (i.e., the null hypothesis should be accepted) versus there being insufficient evidence to favour either the null hypothesis or the experimental hypothesis.

6.4 Conclusions

The present series of studies attempted to clarify the implicit mentalising vs. submentalising debate. Study 1 replicated and confirmed that the self-consistency effect was not an artefact of the spatial layout of disc distribution in the classic dot-perspective task (Experiment 1). The self-consistency effect was also robustly detectable when a real human face (presenting clearer line-of-sight information) was used as the central stimulus (Experiment 2). The findings of Study 1 provided the basis for evaluating the implicit mentalising versus submentalising debate by using real human's stimuli in the following studies.

In Study 2, the consistency effect was robustly detected when a real human face (presenting clearer line-of-sight information) was adopted as the central stimulus and disappeared when a black rectangle was placed on the eyes of the face. The elicitation of the consistency effect for sighted faces but not for eyes-covered faces (Experiment 3) alongside the consistency effect being stable and not modulated by directional cue manipulation (Experiment 4A) lend support to the implicit mentalising account. Additionally, neither consistency effect (Experiment 4A) nor cue-validity effect (Experiment 4B) was modulated by face-related directional cue (eye vs. eye + head). However, the impact of the dynamic property of the directional cue being relevant for triggering the cue-validity effect of Posner task and being irrelevant for demonstrating the consistency effect provides some evidence for the dissociation between attentional orienting and visual-perspective taking. Even though these findings, to some extent, qualify the submentalising account's view that automatic reorientation of attention are solely driven by directional cues, much more work needs to be done before advocates of implicit mentalising can confidently rule out submentalising explanations.

In Study 3, the findings of Experiments 5, 6, and 7 indicated that the implicit mentalising and submentalising accounts can explain the effect of the dot-perspective task in different circumstances, respectively. This thesis does not try to speak against the original versions of the two competing accounts. Instead, this thesis offers up a middle-ground view that the findings of three experiments may be better accommodated to the two accounts with the consideration of different cues. That is, when eye-head cues are automatically processed, visual access is superior to head or head + torso directions in playing a predominate role of modulating the implicit mentalising. However, when eye-finger-pointing cues are automatically processed, the role of visual access may be interfered by finger-pointing that can predominate the effect of the dot-perspective task via an attentional orienting mechanism.

Both implicit mentalising and submentalising interpretations may not be mutually exclusive, further work is encouraged before advocates of both accounts can be confidently accepted. The findings make new headway for examining the implicit mentalising vs. submentalising debate, which may give scientists further insights into understanding L1VPT and ToM processes.

References

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *The American journal of psychology*, *82*(4), 474-489. doi: 10.2307/1420441

Apperly, I. A. (2010). Mindreaders: The cognitive basis of 'theory of mind'. Psychology Press.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review*, *116*(4), 953. doi: 10.1037/a0016923

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, *14*(3), 110-118. doi: 10.1016/j.tics.2009.12.006

Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore & P.J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 41–59). Lawrence Erlbaum Associates, Inc.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*(4), 379-398. doi: 10.1111/j.2044-835X.1995.tb00687.x

Bayliss, A. P., Di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology Section A*, *58*(4), 631-650. doi: 10.1080/02724980443000124

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

Birch, S. A. J. & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. Psychological Science, 18, 382–386. doi: 10.1111/j.1467-9280.2007.01909.x

Bukowski, H., Hietanen, J. K., & Samson, D. (2015). From gaze cueing to perspective taking: Revisiting the claim that we automatically compute where or what other people are looking at. *Visual cognition*, *23*(8), 1020-1042. doi: 10.1080/13506285.2015.1132804

Bukowski, H., & Samson, D. (2017). New insights into the inter-individual variability in perspective taking. *Vision*, *1*, 8. doi: 10.3390/vision1010008

Butterworth, G. E. (1991). Evidence for the geometric comprehension of manual pointing. Paper presented at the meeting of the Society for Research in Child Development, Seattle. Cited in Butterworth, G. (1995). Origins of mind in perception and action. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, *18*(1), 25-50. doi: 10.1348/026151000165553

Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnesi, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy*, *5*(3), 291-308. doi: 10.1207/s15327078in0503_3

Cole, G. G., Atkinson, M., D'Souza, A. D., & Smith, D. T. (2017). Spontaneous perspective taking in humans? *Vision*, *1*, 17. doi: 10.3390/vision1020017

Cole, G. G., Atkinson, M., Le, A. T., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others? *Acta psychologica*, *164*, 165-168. doi: 10.1016/j.actpsy.2016.01.007

Cole, G. G., Smith, D. T., & Atkinson, M. A. (2015). Mental state attribution and the gaze cueing effect. *Attention, Perception, & Psychophysics*, *77*, 1105-1115. doi: 10.3758/s13414-014-0780-6

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. (2017). Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 454. doi: 10.1037/xhp0000319

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P.J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61–83). Hillsdale, NJ: Erlbaum.

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. Developmental Psychology, 34, 28–38. doi: 10.1037/0012-1649.34.1.28

Crais, E., Douglas, D. D., & Campbell, C. C. (2004). The intersection of the development of gestures and intentionality. *Journal of Speech, Language, and Hearing Research*. doi: 10.1044/1092-4388(2004/052)

Doherty, M. J., & Anderson, J. R. (1999). A new look at gaze: Preschool children's understanding of eye-direction. *Cognitive Development, 14*(4), 549-571. doi: 10.1016/S0885-2014(99)00019-2

Drayton, L. A., Santos, L. R., & Baskin-Sommers, A. (2018). Psychopaths fail to automatically take the perspective of others. *Proceedings of the National Academy of Sciences*, *115*(13), 3302-3307. doi: 10.1073/pnas.1721903115

Driver IV, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual cognition*, *6*(5), 509-540. doi:10.1080/135062899394920

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual review of psychology*, *48*(1), 269-297. doi: 10.1146/annurev.psych.48.1.269

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191. doi: 10.3758/BF03193146

Fischer, M. H., & Szymkowiak, A. (2004). Joint attention for pointing but not grasping postures. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*. doi: 10.1016/S0010-9452(08)70937-7

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

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, *5*(3), 490-495. doi: 10.3758/BF03208827

Friesen, C. K., Moore, C., & Kingstone, A. (2005). Does gaze direction really trigger a reflexive shift of spatial attention?. *Brain and cognition*, *57*(1), 66-69. doi: 10.1016/j.bandc.2004.08.025

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

Gardner, M. R., Bileviciute, A. P., & Edmonds, C. J. (2018). Implicit Mentalising during Level-1 Visual Perspective-Taking Indicated by Dissociation with Attention Orienting. *Vision*, *2*, 3. doi: 10.3390/vision2010003

Gardner, M. R., Hull, Z., Taylor, D., & Edmonds, C. J. (2018). 'Spontaneous' visual perspective-taking mediated by attention orienting that is voluntary and not reflexive. *Quarterly Journal of Experimental Psychology*, *71*, 1020-1029. doi: 10.1080/17470218.2017.1307868

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *The American journal of psychology*, *76*(3), 386-394. doi: 10.2307/1419779

Gregory, N. J., Hermens, F., Facey, R., & Hodgson, T. L. (2016). The developmental trajectory of attentional orienting to socio-biological cues. *Experimental Brain*

Research, 234(6), 1351-1362. doi: 10.1007/s00221-016-4627-3

Grossmann, T. (2017). The eyes as windows into other minds: An integrative perspective. *Perspectives on Psychological Science*, *12*(1), 107-121. doi: 10.1177/1745691616654457

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

Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *Neuroreport*, *10*(16), 3443-3447. doi: 10.1097/00001756-199911080-00033

Hietanen, J. K. (2002). Social attention orienting integrates visual information from head and body orientation. *Psychological Research*, *66*(3), 174-179. doi: 10.1007/s00426-002-0091-8

Hill, J. L., Patel, S., Gu, X., Seyedali, N. S., Bachevalier, J., & Sereno, A. B. (2010). Social orienting: reflexive versus voluntary control. *Vision Research*, *50*(20), 2080-2092. doi: 10.1016/j.visres.2010.07.020

Kendon, A. (1994). Do gestures communicate? A review. *Research on language and social interaction, 27*(3), 175-200. doi: 10.1207/s15327973rlsi2703_2

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it!: An fMRI investigation. *Brain and cognition, 55*(2), 269-271. doi: 10.1016/j.bandc.2004.02.037

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

Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception, & Psychophysics*, *71*(2), 314-327. doi: 10.3758/APP.71.2.314

Langton, S. R. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(3), 825-845. doi: 10.1080/713755908

Langton, S. R. (2018). I don't see it your way: The dot perspective task does not gauge spontaneous perspective taking. *Vision*, *2*, 6. doi: 10.3390/vision2010006

Langton, S. R., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual cognition*, *6*(5), 541-567. doi: 10.1080/135062899394939

Langton, S. R., & Bruce, V. (2000). You must see the point: automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 747. doi: 10.1037/0096-1523.26.2.747

Langton, S. R., O'Malley, C., & Bruce, V. (1996). Actions speak no louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural information. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1357. doi: 10.1037/0096-1523.22.6.1357

Lempers, J. D. (1979). Young children's production and comprehension of nonverbal deictic behaviors. *The Journal of Genetic Psychology*, *135*(1), 93-102. doi: 10.1080/00221325.1979.10533420

Liszkowski, U. (2008). Before L1: A differentiated perspective on infant gestures. *Gesture, 8*(2), 180-196. doi: 10.1075/gest.8.2.04lis

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: A primer on the two-systems account. *Child Development Perspectives*, *10*(3), 184-189. doi: 10.1111/cdep.12183

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological science*, *13*(5), 460-464. doi: 10.1111/1467-9280.00481

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child development*, 357-366. doi: 10.2307/1127956

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological review*, *92*(3), 350. doi: 10.1037/0033-295X.92.3.350

Meert, G., Wang, J., & Samson, D. (2017). Efficient belief tracking in adults: The role of task instruction, low-level associative processes and dispositional social functioning. *Cognition*, *168*, 91-98. doi: 10.1016/j.cognition.2017.06.012

Morgan, E. J., Freeth, M., & Smith, D. T. (2018). Mental state attributions mediate the gaze cueing effect. *Vision*, *2*(1), 11. doi: 10.3390/vision2010011

Müller, H. J., & Findlay, J. M. (1988). The effect of visual attention of peripheral discrimination thresholds in single and multiple element displays. *Acta psychologica,* *69*(2), 129-155. doi: 10.1016/0001-6918(88)90003-0

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental psychology: Human perception and performance*, *15*(2), 315. doi: 10.1037/0096-1523.15.2.315

Nakashima, R., & Shioiri, S. (2015). Facilitation of visual perception in head direction: visual attention modulation based on head direction. *PloS one*, *10*(4), e0124367. doi: 10.1371/journal.pone.0124367

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

Pavlidou, A., Gallagher, M., Lopez, C., & Ferrè, E. R. (2019). Let's share our perspectives, but only if our body postures match. *Cortex*, *119*, 575-579. doi: 10.1016/j.cortex.2019.02.019

Perrett, D. I., & Emery, N. J. (1994). Understanding the intentions of others from visual signals: neurophysiological evidence. *Current Psychology of Cognition, 13,* 683-694.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical transactions of the royal society of London. Series B: Biological sciences*, *335*(1273), 23-30. doi:

10.1098/rstb.1992.0003

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology 32*(1), 3-25. doi: 10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bowhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Erlbaum, Hillsdale, NJ.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137-157). Erlbaum, Hillsdale, NJ.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160. doi: 10.1037/0096-3445.109.2.160

Qian, Q., Song, M., & Shinomori, K. (2013). Gaze cueing as a function of perceived gaze direction. *Japanese Psychological Research*, *55*(3), 264-272. doi: 10.1111/jpr.12001

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

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. British Journal of Developmental Psychology, 30, 87–104. doi:

10.1111/j.2044-835X.2011.02045.x

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. doi: 10.1037/a0018729

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

Schneider, D. Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition, 133,* 43-47. doi: 0.1016/j.cognition.2014.05.016

Schuller, A. M., & Rossion, B. (2001). Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Neuroreport*, *12*(11), 2381-2386. doi: 10.1097/00001756-200108080-00019

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

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*, 452-460. doi: 10.1111/j.1467-8624.2011.01730.x

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

10.1016/j.cognition.2015.12.010
Todd, A. R., Cameron, C. D., & Simpson, A. J. (2017). Dissociating processes underlying level-1 visual perspective taking in adults. *Cognition*, *159*, 97-101. doi: 10.1016/j.cognition.2016.11.010
Tomasello, M., & Camaioni, L. (1997). A comparison of the gestural communication of apes and human infants. *Human Development, 40*(1), 7-24. doi: 10.1159/000278540
Van Rooijen, R., Junge, C., & Kemner, C. (2018). No own-age bias in children's gaze cueing effects. *Frontiers in psychology*, *9*, 2484. doi: 10.3389/fpsyg.2018.02484
Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual cognition*, *2*(1), 59-87. doi: 10.1080/13506289508401722
von Bayern, A. M., & Emery, N. J. (2009). Jackdaws respond to human attentional states and communicative cues in different contexts. *Current Biology*, *19*(7), 602-606. doi: 10.1016/j.cub.2009.02.062
von Salm-Hoogstraeten, S., Bolzius, K., & Müsseler, J. (2020). Seeing the world through the eyes of an avatar? Comparing perspective-taking and referential coding. *Journal of experimental psychology: human perception and performance*, *46*(3), 264. doi: 10.1037/xhp0000711
Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in cognitive sciences*, *14*, 383-388. doi: 10.1016/j.tics.2010.05.006
Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: how

attentional selection is shaped by ascribing intentions to others. *PloS one*, *7*(9), e45391. doi: 10.1371/journal.pone.0045391

Willen, J. D., Hood, B. M., & Driver, J. R. (1997, March). An eye direction detector triggers shifts of visual attention in human infants. In *INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE* (Vol. 38, No. 4, pp. 313-313). 227 EAST WASHINGTON SQ, PHILADELPHIA, PA 19106: LIPPINCOTT-RAVEN PUBL.

Wilson, C. J., Soranzo, A., & Bertamini, M. (2017). Attentional interference is modulated by salience not sentience. *Acta psychologica*, *178*, 56-65. doi: 10.1016/j.actpsy.2017.05.010

Wollaston, W. H. (1824). Xiii. on the apparent direction of eyes in a portrait. *Philosophical Transactions of the Royal Society of London*, (114), 247-256. doi: 10.1098/rstl.1824.0016