Three Key Questions to Move Towards a Theoretical Framework of Visuospatial Perspective Taking

Steven Samuel [1, †]

Thorsten M. Erle [2, †]

Louise P. Kirsch [3 †]

Andrew Surtees [4, 5 †]

Ian Apperly [4 †]

Henryk Bukowski [6 †]

Malika Auvray [7 †]

Caroline Catmur [8 †]

Klaus Kessler [9 †]

Francois Quesque [10, 11 †]

[†] All authors contributed equally to this work

ACCEPTED MANUSCRIPT FOR PUBLICATION IN COGNITION

1 Department of Psychology, School of Health and Psychological Sciences, City, University of London, U.K.

2 Department of Social Psychology, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

3 Université Paris Cité, INCC UMR 8002, CNRS, F-75006, Paris, France

4 Centre for Developmental Science, School of Psychology, University of Birmingham, Edgbaston, Birmingham, U.K

5 Birmingham Women's and Children's NHS Foundation Trust, Steelhouse Lane, Birmingham, UK

6 Institute of Psychological Sciences, Université catholique de Louvain, Louvain-La-Neuve, Belgium

7 Sorbonne Université, CNRS, Institut des Systèmes Intelligents et de Robotique, Paris, France.

8 Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK

9 School of Psychology, University College Dublin, Dublin, Ireland

10 Centre de Recherche en Neurosciences de Lyon CRNL, U1028, UMR5292, Trajectoires, F-69500, Bron, France.

11 Centre Ressource de Réhabilitation Psychosociale, CH Le Vinatier, Lyon, France.

Word count (excluding abstract and references only): 2998

Address for correspondence: Steven Samuel. City, University of London, Department of Psychology, Northampton Square, London, EC1V 0HB. Email: steven.samuel@city.ac.uk

Abstract

What would a theory of visuospatial perspective taking (VSPT) look like? Here, ten researchers in the field, many with different theoretical viewpoints and empirical approaches, present their consensus on the three big questions we need to answer in order to bring this theory (or these theories) closer.

Key words: visuospatial perspective taking; visual perspective; spatial perspective; perspective taking; social cognition

As cognitive scientists, we seek to generate theoretical frameworks or models that can predict and explain behaviour. Visuospatial perspective taking (henceforth VSPT) is considered by many researchers to be a key component of social cognition, self-other distinction, communication, and navigation (e.g., Clark & Brennan, 1991; Hamilton, Brindley, & Frith, 2009; Kampis & Southgate, 2020; Schurz et al., 2013; Wolbers & Hegarty, 2010). For example, taking others' perspectives facilitates trust in others (Erle, Ruessmann, & Topolinski, 2018), leads to shared spatial frames of reference (Samuel et al., 2020) and shared psychological perspectives (e.g. Erle & Topolinski, 2017). Understanding what someone else sees underpins the successful use of language, such as the grounding effect of knowing what an interlocutor sees and is therefore aware of ('copresence'), which impacts the choice of determiners the/a (Clark & Brennan, 1991). The self-other distinction that VSPT provokes has been argued to reinforce self-consciousness (see Arnold, Spence, & Auvray, 2017). In sum, a better understanding of VSPT is likely to benefit our understanding not only of wider social cognition but also spatial cognition, motor representations, communication and even philosophy of mind. However, consensus is lacking regarding key theoretical points in the field (e.g., Cole et al., 2020). In March 2023 the present ten authors, all researchers with a shared interest in VSPT, gathered at a specially organised workshop to share research and discuss ideas. Attendees were selected to represent a diversity of viewpoints and thereby promote the most vigorous debate possible. Despite our differences, we worked to achieve a consensus on the most important questions which, if answered, would help *all* of us and (we hope) the wider research community in navigating this field going forward. Here, we report these questions and explain why we believe they are crucial to furthering the goal of formulating a theory of VSPT.

What do we Mean Here by VSPT?

The task of defining VSPT is made difficult given that our questions get to the heart of what VSPT is and should therefore *precede* a definition rather than follow it. For example, we considered as a guide a definition arising out of a recent study led by Francois Quesque (under review) and generated from the collated opinions of multiple specialists in the field: "*Visuo-spatial perspective taking is the process by which one represents a scene from another person's viewpoint, by adopting their perspective.*" However, for the purposes of this paper we have opted instead for a more theory-neutral definition. The entailments around VSPT as necessarily about a *person*, *representing scenes*, and so on, are best determined by the answers to our questions rather than a priori. We define VSPT here as any process by which one infers something about the visual or spatial properties of a scene in relation to another person or position (e.g., Hamilton et al., 2014; Quesque et al., 2018).

Question 1. Which Processes, if any, are Shared Across Different Forms of VSPT?

Our first question is not theory-neutral; it assumes (that is, we are convinced) that there is no one way to solve all VSPT tasks. Broken down, this question is about two things: what processes (if any) can solve more than one type of problem, *and* what types of problems (if any) can be solved by more than one process.

The distinction between Level 1 vs. Level 2 perspective taking (Flavell et al., 1981; Masangkay et al., 1974) is well known. Within the context of VSPT, Level 1 concerns the binary question of whether an agent sees something or not, and one process to solve such a problem is a line-of-sight one, whereby one concludes that a target is visible if a line from the agent to the target is unbroken by obstacles. This has been elegantly demonstrated through increasing response times the further away from the agent the target object is, indicative of the time required to "mentally draw" the line (Michelon & Zacks, 2006). Level 2 VSPT

problems, on the other hand, concern inferences such as the relative spatial arrangement or visual appearance of objects. A line-of-sight process will not suffice to infer whether an object is on an agent's left or right, whether what appears to one person to be a number 6 appears to another to be a number 9, or whether a yellow object viewed through a blue colour filter appears green. Thus, we can say with confidence that line-of-sight is not a process that could extend to solving Level 2 problems. Instead, Level 2 problems have been linked to a variety of other processes, including (but not limited to) embodiment (Erle & Topolinski, 2017; Kessler & Rutherford, 2010; Surtees et al., 2013a, 2013b), object rotation (Langdon & Coltheart, 2001; Wraga et al., 2000), imaginatively simulating others' perception (Ward et al., 2019; Ward et al., 2020), and rule-based reasoning (Samuel, Eacott, et al., 2022; Samuel et al., 2021).

Understanding the VSPT "process toolkit" (see Box 1) and what each tool can be applied to is key to understanding what VSPT *means*. How can this be achieved? It should be possible to reverse engineer different processes by detecting the types of representation that each should elicit, such as whether they are minimal (e.g., "registrations" of agent-object-property links: Apperly & Butterfill, 2009), "classically" propositional (e.g., *she sees a 9*), quasi-perceptual (a mental image simulating others' visual experience), or rule-based (e.g. on my left + she's facing me = on her right). Detection may be aided by the fact that representations vary in the ease with which they capture different kinds of information – for example, minimal representations of agent-object properties cannot capture the difference between seeing a symbol as a "6" or a "9". Variation in demands on cognitive resources is also potentially informative, because more minimal representations may be faster and more efficient, while more flexible representations are likely also to be more complex processes and therefore slower and more costly (e.g., Apperly & Butterfill, 2009; Butterfill & Apperly, 2013).

Box 1. Examples of processes (or "tools") that may be used to solve VSPT problems

Embodiment - Rotating the perceiver's own body schema into the position of the agent.

Utilising directional cues - Utilising directional cues (e.g. canonical orientation of object, gaze, body orientation) present in a scene.

Geometrical computation - Mentally operating geometrical reasoning (e.g. something oriented at 180° from a perspective [i.e. facing me], is oriented at -90° from a perspective at 270° from the first [i.e. from my right]).

Line-of-sight drawing - Determining whether the line between the agent's eyes and the target is broken (or not) by obstacles.

Minimal registrations - The registration of agent-object-property links.

Object rotation - Creating a mental image of an object from all directions and rotating this mental image to select the one that would fit another given perspective.

Propositional representations - Representing perspectives propositionally, such as *she sees a 9*.

Rule-based reasoning - For example, using folk optics laws to determine how something would appear to another agent (e.g. looking through sunglasses provides a darker view), or applying simple rules such as 6 = 9 upside-down.

Simulating others' perception - Quasi-perceptually representing the content of another's perspective in a way that is analogous to first-person visual perception.

Different problem types often prompt different processes (e.g., Surtees et al., 2013b; Wang et al., 2006), but the relationship between problem and process is likely to be more complex. While a single problem type (such as whether someone opposite you sees a 9 when you see a 6) might itself be solvable in different ways, hence by multiple processes, so might a single process (e.g., embodiment) cover a variety of problem types. This potentially quite

fluid interplay of process and problem type, itself very possibly mediated by rich and complex individual differences (cf. Question 3), is central to understanding and predicting what will happen when an individual is presented with a given VSPT problem. Clarifying and accounting for this interplay must be central goals of any theory of VSPT and is therefore the rationale for (and benefit of) this question.

Question 2. When Might VSPT not be Specifically Social Cognition?

Understanding the different processes that function within the domain of VSPT is crucial, but to understand more about what VSPT means it is also important to identify to what extent any of these processes can be applied to other types of task or other types of representation. VSPT has often been linked with social cognitive concepts related to the inference of others' thoughts, beliefs, intentions, and affective states. This leads us to our second key question.

In a series of studies, Quesque et al. (2020) asked participants to describe the spatial relations between objects in visual scenes. A significant proportion of respondents spontaneously took the "perspective" of a seat in the scene, even when the seat was not referred to, and even when there was also a social agent in the scene (see also Quesque et al., 2018; Surtees et al., 2012). Other studies have demonstrated that effects previously considered to reflect the operation of a social-specific perspective-taking mechanism (e.g. the interference of the other's perspective when judging one's own perspective in the Dot Perspective-taking task; Samson et al., 2010) can be generated by other, non-social processes (e.g. Conway et al., 2017; Heyes, 2014; Santiesteban et al., 2014; Santiesteban et al., 2017).

The current evidence is increasingly consistent in demonstrating that VSPT tasks recruit processes that do not need to be specifically social. This observation could be interpreted in three ways: (1) some forms of VSPT recruit non-social processes, (2) there are

alternative routes or strategies afforded by some VSPT situations, and (3) there is individual variability regarding participants' propensity to use non-social heuristics rather than (potentially) more effortful social-specific processes. These options need not be mutually exclusive.

Just as for Question 1, this question also gets to the heart of what VSPT actually is. Consider chairs, which neither see anything nor have a perspective but do have a "front" and can be the centre of a "frame of reference". In these respects a person is also like a chair. If you're asked what that person sees, and you only consider her frame of reference and answer with what's in front of her, have you done VSPT? Does VSPT require the consideration of a mental state to *be* VSPT?

Such questions about the scope of VSPT will not be settled by empirical evidence. One may choose to define VSPT as *requiring* the consideration of a mental state, without which it is not VSPT but something else, such as domain-general problem solving. Alternatively, one may opt for a definition which defines any process which solves a VSPT problem as a VSPT process, social or otherwise. According to the latter view, although the debate around social vs. non-social VSPT has enabled us to gain insights about the nature of the processes underlying perspective taking and the methodological complexity required to capture it, whether a process an individual uses is also applied to arrows or chairs does not make the process less important to understand how we solve VSPT problems. A variation on this option is the view that social processes are co-opted for non-social VSPT. This is effectively the inverse of the argument that domain general or non-social, rule-based processes serve VSPT with agents. For instance, face detection is unquestionably an important social process but faces are often spontaneously perceived on inanimate objects, by adult and infant humans and by non-human apes (i.e. pareidolia; Taubert et al., 2017). Maybe considering some processes as “social but non-specific” could better describe a range of

highly efficient social processes that primarily serve social adaptation but can be applied to inanimate objects. Again, it could be that both sides of this argument are simultaneously correct (some social processes serve non-social VSPT while some domain-general processes serve VSPT with agents).

The empirical programme for social-specificity concerns which processes, if any, do *not* work (or do not work the same way) for VSPT with both agents *and* non-agents. Those which cut across both contexts could instead be classified as domain-general, not social-specific, or alternatively, social but non-specific. The answer to this question goes beyond a philosophical understanding around the nature of VSPT; it also has ramifications for clinical settings (for instance, interventions and/or training to improve a particular process would differ if that process were social-specific compared to if it were domain-general) and for our understanding of VSPT performance from an evolutionary and comparative cognition perspective.

There is a strong argument for moving beyond the standard experimental paradigms to study this question. Ideally, we should study how VSPT works in naturalistic social situations, which are typically situations with a lot of information to process, under time pressure, and with few cognitive resources to spare (Del Sette et al., 2022). In such conditions, individuals are perhaps more prone to use heuristics which could qualify as non-social. In other words, the answer to this question may well be that VSPT is partly social and partly domain-general, and any line drawn between them could shift with the varying demands of real-world performance. It could also shift inter- and intra-individually. This leads us to our third key question.

Question 3. Are There Individual Differences in VSPT, and What Accounts for Them?

Individual differences in VSPT have received significant attention only in the last decade or so (e.g., Bukowski & Samson, 2017; Erle, 2019; Erle et al., 2019; Job et al., 2021; Kessler et al., 2014; Kessler & Wang, 2012; for a review see Samuel, Cole, et al., 2022). For example, a preference for embodied perspective taking over mental rotation of scenes to fit the self perspective may be more common in women than men (Kessler & Wang, 2012). Bilingualism has been associated with precocious perspective taking, such as when judging whether a turtle is 'on its back' or 'on its feet' in a picture between a child and an experimenter (Goetz, 2003). Important individual differences on the Dot Perspective-taking task have been evidenced, where performance could be disentangled into four profiles: egocentric versus altercentric perspective takers (based on how much better one performs at adopting the self- versus other person's perspective irrespective of the presence of a viewpoint conflict) and conflicted versus flexible perspective takers (based on how much they struggle at selecting one perspective when the self- and other perspective are conflicting irrespective of the perspective to adopt; see Bukowski & Samson, 2017 for details). Another example of privileged-by-default perspective comes from the graphesthesia task, which consists of the recognition of ambiguous letters (b,d,p,q) "projected" on the stomach using vibrators (e.g., Arnold et al., 2016). Investigators have learned that perspective taking abilities are linked to interoceptive and sensory factors as well as with personality-traits. First, better interoceptive awareness is associated with a preference for a decentred (non-egocentric) perspective, likely reflecting stronger perceived boundaries between internal states and the external world (Baiano et al., 2023). In addition, perspective taking abilities are influenced by sensory impairments, as deficits in the visual modality decrease allocentric and increase egocentric perspective taking, those in the vestibular system decrease egocentric perspective taking, and those in the auditory system increase the decentred one (see Job, Kirsch, & Auvray, 2022).

Finally, lower interpersonal anxiety and higher social intelligence is associated with a stronger grounding in one's natural spatial perspective, i.e. the perspective participants spontaneously adopt when given the choice (about 80% of people spontaneously choose an ego-centred perspective and 20% a decentred one; Arnold et al., 2016; Job et al., 2021).

We have identified four distinct levels at which individual differences may operate: i) propensity to take perspectives; ii) default perspective; iii) process selection (cf. Question 1); and iv) performance. When addressing the question of individual differences, it will be important to determine the extent to which variance is due to '"true" individual differences in VSPT processes; which aspects of performance are instead due to individual differences in other, perhaps domain-general processes such as attention, inhibition, or conflict resolution, or in stable traits; and which aspects are due to transitory and/or environmental factors such as the individual's emotional state (e.g., guilt feelings make individuals more other-centred whereas feeling ashamed or being sleep-deprived impairs the capacity to select one perspective; Bukowski & Samson, 2016; Deliens et al., 2018). For clinical application, if the source is a domain-general process like conflict resolution, one would design different interventions than if the source is a VSPT-specific process (cf. Question 2).

Conclusion

These three questions were selected from among multiple candidates because we all agree that the answers to them will move us towards the goal of a theory (or theories) of VSPT. However, depending on these answers, it may yet be that the more principled case is for the fragmentation of VSPT into separate fields. For example, future work may demonstrate that the grouping together of both visual and spatial perspective taking is theoretically unwarranted. The classic Level 1 vs. Level 2 contrast may prove to be so sharp as to require independent accounts, with perhaps one or two processes alone accounting for

the former but a plethora for the latter. We are sceptical that any one of the many processes involved in Level 2 VSPT is essential. As an alternative it could be that particular instances of Level 2 VSPT instead bear a "family resemblance" to one another, with each instance involving a set of processes that only partially overlaps with others. Or instead there might be distinct clusters of tasks that involve largely dissimilar processes, in which case these clusters would replace the Level 2 category altogether. We note that similar challenges have been identified in the literature on executive functions, where they have been tackled with some success through converging evidence from task analysis, and analysis of the covariance structures of behavioural data and neural correlates (e.g., Friedman & Miyake, 2017). Finally, future data may reveal a blurring of the social vs. non-social distinction. Whether the answers to these questions culminate in greater synthesis or greater segmentation, they will advance our understanding of VSPT.

References

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

Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective. Natural stance or effortful decentring? *Cognition, 148*, 27-33.

Arnold, G., Spence, C., & Auvray, M. (2017). A unity of the self or a multiplicity of locations? How the graphesthesia task sheds light on the role of spatial perspectives in bodily self-consciousness. *Consciousness & Cognition, 56*, 100-114.

Baiano, C., Job, X., Kirsch, L. P., & Auvray, M. (2023). Interoceptive abilities facilitate taking another's spatial perspective. *Scientific Reports, 13*(1), 10064.

Bukowski, H., & Samson, D. (2016). Can emotions influence level-1 visual perspective taking? *Cognitive neuroscience, 7*(1-4), 182-191.

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

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language, 28*(5), 606-637.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition, 13*(1991), 222–233.

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

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*(3), 454-465.

Del Sette, P., Bindemann, M., & Ferguson, H. J. (2022). Visual perspective-taking in complex natural scenes. *Quarterly Journal of Experimental Psychology, 75*(8), 1541-1551.

Deliens, G., Bukowski, H., Slama, H., Surtees, A., Cleeremans, A., Samson, D., & Peigneux, P. (2018). The impact of sleep deprivation on visual perspective taking. *Journal of sleep research, 27(2), 175-183.*

Erle, T. M. (2019). Level-2 visuo-spatial perspective-taking and interoception–More evidence for the embodiment of perspective-taking. *PloS one, 14*(6), e0219005.

Erle, T. M., Barth, N., & Topolinski, S. (2019). Egocentrism in sub-clinical depression. *Cognition and Emotion*, *33*(6), 1239-1248.

Erle, T. M., Ruessmann, J.K., & Topolinski, S. (2018). The effects of visuospatial perspective-taking on trust. *Journal of Experimental Social Psychology*, *79*, 1239-34-41.

Erle, T. M., & Topolinski, S. (2017). The grounded nature of psychological perspective-taking. *Journal of Personality and Social Psychology*, *112*(5), 683-695.

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

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186-204.

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and cognition*, *6*(1), 1-15.

Hamilton, A. F. d. C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autism spectrum disorder. *Cognition, 113*(1), 37-44.

Hamilton, A. F. D. C., Kessler, K., & Creem-Regehr, S. H. (2014). Perspective taking: building a neurocognitive framework for integrating the "social" and the "spatial". *Frontiers in human neuroscience*, *8*, 403.

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

Job, X., Kirsch, L., & Auvray, M. (2022). Spatial perspective-taking: Insights from sensory impairments. *Experimental Brain Research*, *240*, 27-37.

Job, X., Kirsch, L., Inard, S., Arnold, G., & Auvray, M. (2021). Spatial perspective taking is related to social intelligence and attachment style. *Personality and Individual Differences*, *168*, 109726.

Kampis, D., & Southgate, V. (2020). Altercentric Cognition: How Others Influence Our Cognitive Processing. *Trends in cognitive sciences*, *24*(11), 945-959.

Kessler, K., Cao, L., O'Shea, K. J., & Wang, H. (2014). A cross-culture, cross-gender comparison of perspective taking mechanisms. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1785), 20140388.

Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology*, *1*, 213.

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: differences predicted by sex and social skills score. *Spatial Cognition & Computation*, *12*(2-3), 133-158.

Langdon, R., & Coltheart, M. (2001). Visual perspective-taking and schizotypy: evidence for a simulation-based account of mentalizing in normal adults. *Cognition*, *82*(1), 1-26.

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.

Quesque, F., Chabanat, E., & Rossetti, Y. (2018). Taking the point of view of the blind: Spontaneous level-2 perspective-taking in irrelevant conditions. *Journal of Experimental Social Psychology, 79*, 356-364.

Quesque, F., Foncelle, A., Chabanat, É., Jacquin-Courtois, S., & Rossetti, Y. (2020). Take a seat and get into its shoes! When humans spontaneously represent visual scenes from the point of view of inanimate objects. *Perception, 49*(12), 1333-1347.

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

Samuel, S., Legg, E. W., Manchester, C., Lurz, R., & Clayton, N. S. (2020). Where was I? Taking alternative visual perspectives can make us (briefly) misplace our own. *Quarterly journal of experimental psychology, 73*(3), 468-477.

Samuel, S., Cole, G., & Eacott, M. (2022). It's not you, it's me: A review of individual differences in visuo-spatial perspective taking. *Perspectives on Psychological Science, 18*(2), 293-308.

Samuel, S., Eacott, M. J., & Cole, G. G. (2022). Visual perspective taking without visual perspective taking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(7), 959-965.

Samuel, S., Hagspiel, K., Eacott, M. J., & Cole, G. G. (2021). Visual perspective-taking and image-like representations: We don't see it. *Cognition, 210*, Article 104607.

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 929-937.

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

Schurz, M., Aichhorn, M., Martin, A., and Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, 7, 712

Surtees, A., Noordzij, M.L. & Apperly, I.A. (2012). Sometimes losing your self in space: spontaneous use of intrinsic and relative spatial reference frames in children and adults. *Developmental Psychology*, *48*(1), 185-191.

Surtees, A., Apperly, I., & Samson, D. (2013a). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, *129*(2), 426-438.

Surtees, A., Apperly, I., & Samson, D. (2013b). The use of embodied self-rotation for visual and spatial perspective-taking. *Frontiers in Human Neuroscience*, *7*, 698.

Taubert, J., Wardle, S. G., Flessert, M., Leopold, D. A., & Ungerleider, L. G. (2017). Face pareidolia in the rhesus monkey. *Current Biology*, *27*(16), 2505-2509.

Wang, R. F., Crowell, J. A., Simons, D. J., Irwin, D. E., Kramer, A. F., Ambinder, M. S., Thomas, L. E., Gosney, J. L., Levinthal, B. R., & Hsieh, B. B. (2006). Spatial updating relies on an egocentric representation of space: Effects of the number of objects. *Psychonomic Bulletin & Review*, *13*, 281-286.

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of another's visual perspective. *Current Biology*, *29*(5), 874-880. e874.

Ward, E., Ganis, G., McDonough, K. L., & Bach, P. (2020). Perspective taking as virtual navigation? Perceptual simulation of what others see reflects their location in space but not their gaze. *Cognition*, *199*, 104241.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences*, *14*(3), 138–146.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 151-168.