Orientation and metacognition in virtual space

Tenbrink, T.; Salwiczek, L.

Journal of Experimental Psychology: Human Perception and Performance

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

Orientation and metacognition in virtual space

Thora Tenbrink

School of Linguistics & English Language, Bangor University, UK

+44 1248 382263

t.tenbrink@bangor.ac.uk

Lucie H. Salwiczek

Department of Behavioural Neurobiology, Max Planck Institute for Ornithology,

Seewiesen, 82319 Starnberg, Germany

+49 176 32832462

lsalwiczek@orn.mpg.de

Abstract. Cognitive scientists increasingly use virtual reality scenarios to address spatial perception, orientation, and navigation. If based on desktops rather than mobile immersive environments, this involves a discrepancy between the physically experienced static position and the visually perceived dynamic scene, leading to cognitive challenges that users of virtual worlds may or may not be aware of. The frequently reported loss of orientation and worse performance in point-to-origin tasks relate to the difficulty of establishing a consistent reference system on an allocentric or egocentric basis. We address the verbalisability of spatial concepts relevant in this regard, along with the conscious strategies reported by participants. Behavioural and verbal data were collected using a perceptually sparse virtual tunnel scenario that has frequently been used to differentiate between humans' preferred reference systems. Surprisingly, the linguistic data we collected relate to reference system verbalisations known from the earlier literature only to a limited extent, but instead reveal complex cognitive mechanisms and strategies. Orientation in desktop VR appears to pose considerable challenges, which participants react to by conceptualising the task in individual ways that do not systematically relate to the generic concepts of egocentric and allocentric reference frames.

Introduction

Imagine you are comfortably seated on your chair in front of your computer. On your computer screen, a virtual, winding journey unfolds through a short tunnel. At the end of the journey, your task is to point from your current position back to the journey's origin. Obviously, this is not a position currently shown on the screen. However, it is nowhere in the real world either; as a matter of fact, you never moved in your seat. Nevertheless you can provide an answer because you compensate by a mental representation of the scene as three-dimensional. Since the scene on the screen suggested a forward movement through the tunnel, you will imagine the tunnel's entrance as being somewhere behind you. Where you imagine this to be depends crucially on how your conceptualisation of navigating through the virtual space works.

Studying navigation in virtual reality (VR) is becoming increasingly popular in spatial cognition (Boccia, Nemmi, & Guariglia, 2014; Maguire et al., 1999, Darken & Peterson, 2002, Stankiewicz et al., 2006, Török et al., 2014), even for investigating primates (Dolins et al., 2014) or rodents (Harvey et al., 2009; Chen et al., 2013). There are obvious advantages; for instance, features of the environment (e.g. weather conditions, type and number of presented stimuli) can be controlled and manipulated to a much higher extent than any real world scenario could ever offer, and certain disabilities become irrelevant. VR can be combined with numerous other technologies, such as eye-trackers, fMRI, and EEG. However, these advantages come with a cost, especially if they necessitate immobility (as with most neurocognitive methods to date). Recent comparative studies consistently found that spatial orientation in even highly advanced stationary (such as desktop based) VR environments deviates from situations

involving actual movement. This affects performance in path integration tasks (Kearns, Warren, Duchon, & Tarr, 2002), map drawing (van der Ham, 2015), route learning (Ruddle et al., 2013) and wayfinding (Ruddle et al., 2011).

Specifically, Riecke & Wiener (2007) and Riecke (2008) observed that many participants lose orientation in virtual navigation environments if they are not provided with real world features like physical motion or useful landmarks (Klatzky et al., 1998; Chance et al., 1998; Wraga et al., 2004). To perform well, participants often require extensive feedback training and/or explanations (e.g., Gramann et al., 2005; Riecke & Wiener, 2007; Wiener & Mallot, 2006). Such feedback has the effect of enhancing high-level cognitive processes including metacognitive consideration of the navigation experience. This may be more demanding and computationally more expensive as compared with similar experiences in the real world that require few if any cognitive resources (Presson & Montello, 1994; Farrell & Robertson, 1998). Indeed, Riecke and Wiener (2007) suggested that participants need to *think more* before responding to a spatial updating task in virtual reality as opposed to real environments; this explains larger reaction times in virtual reality point-to-origin tasks (Goeke et al., 2015) as compared to physical motion conditions (Riecke et al., 2007). Actually, the underlying cognitive processes may even be fundamentally different, since the available evidence suggests that the sensorimotor experience of real-world navigation cannot be straightforwardly transferred to a virtual setting in which the only perceptual information is visual. Since the number of VR studies continues to increase, it is essential to gain further insight into the (meta-)cognitive processes involved in virtual navigation, and in particular with respect to cognitive strategies that are adopted to maintain orientation in virtual space.

In this paper, we elicit participants' verbalisations of concepts and strategies used to deal with the discrepancy between their actual static situation and the imagined navigation in virtual space. Narrowing down the complexity of potential issues around virtual navigation, we focus on a point-to-origin task performed in desktop VR (tunnel task, see below) that is as neutral and simple as possible and should therefore allow for effortless spatial updating during navigation. In fact, based on this very assumption, this task has been used frequently in spatial cognition research to identify basic cognitive reference systems that serve to support orientation in space (following Schönebeck et al., 2001; see e.g., Goeke et al., 2015). In the following, we briefly review pertinent findings on navigation and orientation in virtual and real environments, and the role of language in spatial cognition research, before addressing the notion of reference systems in more depth. We then introduce the tunnel task paradigm, followed by the details of our study. Our aim is to address to what extent the cognitive processes involved in virtual navigation are accessible to participants and can be externalised via language, as well as changed via language-based instruction. Furthermore, we ask how participants understand their own virtual navigation in this task, and examine the conscious strategies that they report.

Navigation and orientation in real and virtual environments

It has long been established that different types of environment evoke different spatial concepts and call for different skills. Based on a large correlational study of spatial ability types, Hegarty et al. (2006) identified fundamental limitations regarding skill transferability between different scales and task settings. For instance, small-scale spatial abilities as identified in paper-and-pencil tasks appear to be more directly related to virtual environment based learning than to real-

world experience. In contrast, self-reported measures correlated more highly with real-world experience than with VR based learning. This suggests systematic differences between VR and real-world learning that have important implications on methodological approaches in spatial cognition.

Out of the wide variety of aspects that contribute to spatial cognition performance in humans and other species, path integration is a well-researched fundamental skill needed for spatial orientation (Benhamou & Séguinot, 1995). The ability to point back to the starting point of motion, demonstrating successful orientation and path integration, can involve at least two distinct cognitive mechanisms (Loomis et al., 1999). Continuous updating is quick and automatic (Farrell & Robertson, 1998), and suffices for simple motion trajectories. In contrast, configural updating requires computation of previous turns, involving a higher amount of cognitive effort (May & Klatzky, 2000; Wiener et al., 2011). These mechanisms are at work across real-world and virtual navigation scenarios; however, performance is only at its best given full availability of sensorimotor input (Zetzsche, Galbraith, Wolter, & Schill, 2009). Longer reaction times in virtual desktop scenarios (cf. Riecke et al., 2007) appear to suggest a computation process going beyond simple continuous updating even with simple motion trajectories.

The lack of decisive sensory feedback in virtual reality navigation studies that do not involve physical movement also affects cortical alpha band activity (Ehinger et al., 2014). Altogether, it appears highly likely that the reported problems with orientation in desktop VR environments are due to the lack of idiothetic cues and proprioception. This would suggest that higher complexity and sophistication of the VR environment would not necessarily solve these issues. For example, the advantages of sensory information do not extend to inertial

information (Waller et al., 2003). Also, within desktop presentation modes not involving physical movement, Gaunet et al. (2001) found that spatial updating was not affected by the mode of presentation at all. Performance in scene recognition and estimating the homing direction did not differ between conditions using active joystick-based navigation, passive computer-imposed travel, and even sequential snapshot-views. Complementarily, Wraga et al. (2004) found that spatial updating is consistently more difficult with passive movement or display-based scene change rather than active physical movement, leading to worse performance across various task settings.

In order to gain a better understanding of the role of orientation in virtual reality settings, it is necessary to address individual components that enable humans to maintain orientation in space. One crucial aspect concerns the environment as such, which supports humans in establishing a spatial reference direction as a basis for orientation. In a study set in a sophisticated immersive VR environment, Kelly et al. (2008) demonstrated better performance in angular (even highly symmetrical) rooms than in circular ones that failed to provide any environmental cues at all.

Establishing a reference direction is not only dependent on environmental cues but also on the form in which they are experienced. Reviewing a wide range of evidence accumulated in different types of environments (VR and others), Montello et al. (2004) found that across settings, spatial knowledge is accumulated in an experience-based way; crucially, the experienced orientation fundamentally affects spatial performance in subsequent tasks. Moreover, the potential for developing extensive survey knowledge based on route experience in VR appears to be restricted. Generally, perceptual experience guides how a conceptual reference system is established in support of spatial memory (Shelton

& McNamara, 2001). Once participants have established a conceptual reference frame, they aim to adopt the same concept for subsequent tasks (Kelly & McNamara, 2010).

Correspondingly, different perspectives shown in a virtual environment lead to different kinds of cognitive representations (Shelton & McNamara, 2004; Török et al., 2014). With an egocentric perspective on the scene, it is easier to navigate; with an aerial view, path integration relative to environmental cues is easier. Shifts of perspective, such as between egocentric and aerial views, negatively affect performance (Shelton & McNamara, 2004). Problems arise in particular when the view on the scene is ground-based rather than aerial, yet not egocentric, as in a third-person perspective using an avatar with a constantly north-oriented map. The mismatch of perceived view and cognitive reference system then affects both navigation and orientation (Török et al., 2014).

However, some of the experience-based orientation cues can be counter-acted by explicit information about future demands. Mou et al. (2009) showed that some of the positive effects of self-locomotion (as discussed above) may be traced back to the ability to track the relevant reference direction while moving. Participants performed just as well in tasks not involving self-movement when they received previous information about the viewing direction to be tested subsequently.

Explicit instruction necessarily involves language, and may potentially trigger meta-cognitive considerations about how the task is performed. In the following, we briefly discuss the role of language in spatial cognition research.

The role of language in spatial cognition research

As a cognitive skill, language is necessarily involved in spatial cognition research in many different respects. Language is used across task settings and disciplines to instruct experimental participants, i.e., to explain task procedures. Language can be a vehicle for spatial memory, supporting wayfinding and navigation for instance through recalling formulations like 'second left' (Meilinger et al., 2008). Language can be used as input to accumulate spatial knowledge (Avraamides et al., 2013). Language, as a system, has been widely researched to access structures of thought, for example with regard to human concepts of space (Talmy, 2000). Language can be an external representation of thought and as such is often used as data, frequently alongside other measures (Ericsson & Simon, 1993; Gugerty & Rodes, 2007; Tenbrink, 2015). Language based spatial descriptions have often been used to address how humans conceptualise (and communicate) their environment (Allen, 2000; Denis et al., 1999; Taylor & Tversky, 1996).

Across task settings and goals, research in spatial cognition has drawn widely on the close connection between language and thought. The relationship is systematic, though not necessarily straightforward, as demonstrated by heated debates on various issues (e.g., does language affect spatial reasoning or vice versa? Levinson, 2002; Li & Gleitman, 2002). Some aspects appear to be shared across linguistic and non-linguistic tasks (Levinson, 2003), while others are strikingly dissimilar (Klippel & Montello, 2007). Corresponding to the findings on the impact of environmental cues on conceptual orientation and perspective described above, Taylor & Tversky (1996) demonstrated that features of the environment affect how people describe it verbally. However, they also shifted between perspectives in a more flexible way than would be expected based on the cognitive costs of conceptual perspective shifts reported above. Tversky, Lee, &

Mainwaring (1999) suggest that perspective shifts may be highly task-related. In some situations, shifting conceptual perspective and referring to it in language may involve less cognitive costs than maintaining the same perspective in spite of a changed situation. The next section addresses the relationship between conceptual and linguistic reference systems in more detail.

Reference systems

Reference systems appear to be essential in human everyday life, although most people may never consider the phenomenon consciously, and remain unaware of discrepancies between different ways of conceptualising a spatial scene. Language offers no grammaticalised distinction between reference systems, and this can lead to misunderstandings. Phrases like *the hat on the right* are ambiguous if the addressee adopts a different perspective than the speaker. On the whole, speakers have a wide range of options available for producing and interpreting spatial descriptions (see Tenbrink, 2011, for a systematic account). Of these options, the distinction between speaker and hearer perspective is arguably the most obvious (or consciously accessible) one, and it sometimes finds its way into language by expressions such as 'from my/your point of view'.

Generally, reference systems allow positioning locations relative to a body (own or others) or relative to other objects or locations, as well as developing consistent concepts of extended environments. The literature offers numerous accounts of reference systems with diverse, partly contradictory terminologies (see Levinson, 2003, for an insightful overview). A binary distinction widely used in psychology (O'Keefe & Nadel, 1978), brain sciences (Paillard, 1991), philosophy (Kant, 1991 [1768]), and other areas in cognitive science is that between *allocentric* and *egocentric* reference systems. A conceptualisation of the

environment relative to one's own current position is called *egocentric*, whereas a concept independent of the self is *allocentric*. Many aspects of this distinction still remain unresolved, including the precise mapping of conceptual distinctions to behavioural performance across tasks as well as to linguistic representations. Moreover, theories and findings appear to be in conflict about the underlying cognitive mechanisms for spatial updating processes based on reference systems (Wang, 2012).

To our knowledge, systematic and context-independent linguistic markers for egocentric and allocentric reference systems consistent with the binary distinction as just outlined have not been suggested in the literature. Linguistic analysis of reference systems centers on object configuration descriptions (e.g., Levinson, 2003; Carlson, 1999) and route descriptions (Couclelis, 1996; Taylor & Tversky, 1996). In large-scale environments, a distinction is often made between *route* vs. *survey* perspective. Here, *route* can be roughly associated with the *egocentric* reference system because the environment is conceived from the route traveller's perspective, and *survey* with *allocentric* because the environment is conceived from above, independent of the traveller's current perspective (Wang, 2012). Linguistically, route perspective is associated with expressions such as *left* and *right*, relating locations to the traveller's current position in the environment, while survey is based on compass directions such as *north* and *south,* relating locations to each other rather than the traveller (Taylor & Tversky, 1996).

In linguistic analysis in other contexts, a three-fold categorisation proposed by Levinson (2003) is now widely adopted. Levinson identified three kinds of conceptual reference systems expressed in language: absolute, relative, and intrinsic. Absolute reference systems are associated with environmental features (uphill, downhill) or compass directions (north, south) on the horizontal plane,

and with gravity concerning the vertical dimension. This means that they can typically be clearly identified by the linguistic choices. Relative and intrinsic reference systems, in contrast, are both expressed by *projective terms*, i.e., expressions that denote a spatial direction such as *left, right, front, back, above*, and *below*. Relative reference systems are used when an object is described in relation to another object from an observer's point of view. Intrinsic reference systems rely on an object's or an observer's intrinsic features, such as the speaker's own front in *The ball is in front of me*, or a salient object part, such as *The ball is in front of the car*. Therefore, an incomplete utterance like *The ball on the left* could be used with either type of reference system, adding to the potential complications through perspective conflicts.

Levinson (2003) points out that the binary distinction between *allocentric* and *egocentric* is neither sufficient to adequately represent the conceptual distinctions reflected in language, nor do these two terms map straightforwardly to any of the three categories identified by Levinson. For instance, intrinsic reference frames are egocentric when the speaker refers to their own intrinsic features, as in *to my left*, but allocentric when the speaker refers to another person's intrinsic features, as in *to your left*. However, *survey* perspective (as used in large scale environments) is conceptually related to the absolute reference system (Taylor & Tversky, 1996), while route perspective could be either relative or intrinsic. The crucial binary distinction lies in the role of the observer, which is irrelevant in the survey case, just as in the notion of an allocentric reference system. Therefore, we take the linguistic markers identified for route and survey perspectives as a starting point to distinguish *egocentric* and *allocentric* reference systems, keeping in mind the associated complications as just outlined. This binary reference system distinction is crucial to the tunnel task paradigm, which we describe next.

The tunnel task paradigm

The tunnel task was first introduced by Schönebeck et al. (2001), who proposed that it triggers two fundamentally distinct ways of conceptualising spatial experience in VR. Participants are seated in front of a computer screen and watch a journey through a short tunnel. Afterwards the task is to identify the direction of the start point. The only information received consists of the translational and rotational changes provided by the rate of optic flow during the imaginary journey while watching the tunnel (see Figure 1).
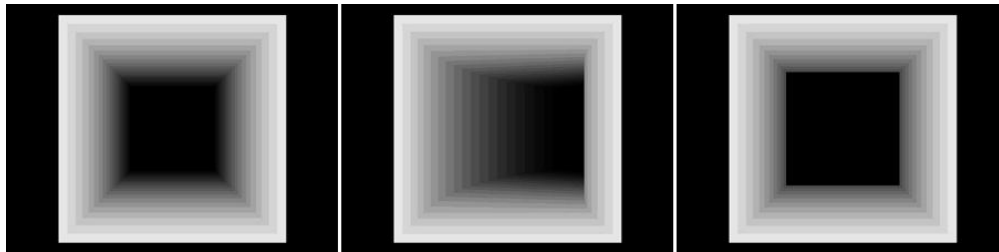
Figure 1: Tunnel sequence as observed by participants. Left: beginning of the tunnel (straight), middle: traversing tunnel (turn to the right), right: straight end segment of the tunnel.

Because there are no reference points during or at the end of the journey, the task can only be solved by relying on an internal spatial representation of the tunnel (journey). People are often quite consistent and accurate (with respect to angles) in their judgment if they receive feedback on numerous training trials. Crucially, there are generally two clearly differentiated groups of participants generating systematically different results (Goeke et al., 2015; Gramann et al., 2005; Plank et al. 2010), with different neurological correlates (Chiu et al., 2012; Gramann et al. 2009, 2010; Plank et al., 2010). These results are interpreted as consistently reflecting two distinct conceptualisations, corresponding to the conceptual spatial

reference systems termed *egocentric* vs. *allocentric* (Klatzky, 1998) as described above. Accordingly this task has frequently been used to distinguish between reference system preferences using as little input information as possible (Chiu et al., 2012; Goeke et al., 2015; Gramann et al., 2005, 2006, 2009, 2010; Gramann, 2011; Lin et al., 2009; Planck et al., 2009, 2010; Schönebeck et al., 2001; Seubert et al., 2008). According to this interpretation, considering the situation shown in Figure 1, an egocentric representation would lead a participant to point to the *right* and back to show the location of the entrance when coming out of the tunnel, whereas an allocentric representation would induce pointing to the *left* and back.

According to Gramann (e.g. Gramann et al., 2005), the two strategies are related to concepts of *turning* or *not turning* one's head or body while imagining the travel through the tunnel. The idea is that "turners" follow the virtual path through the curving tunnel by mentally adjusting their facial orientation, i.e., they imagine turning with the tunnel while navigating. For this reason, their results are associated with an egocentric reference system. In contrast, "non-turners" seem not to update their heading direction during the imagined rotation of the tunnel; it appears that they keep a stable view direction throughout the task as if the tunnel's curve did not imply a personal orientation change for them. Accordingly, their results are associated with an allocentric reference system, in which they build up a mental map of the simulated environment that is independent of their own view direction.

These proposed associations of *turning* with an egocentric reference system and *non-turning* with an allocentric reference system seem intuitive and straightforward. Nevertheless, it should be noted that the distinction of allocentric and egocentric reference systems refers to conceptualisations of the real world across a wide variety of tasks, while the idea of *turning* vs. *non-turning* is highly

task-specific for this particular VR scenario. Associating these concepts with each other is actually a hypothesis about cognitive mechanisms. In particular, it is hypothesised that the cognitive strategies activated by the perceptually sparse tunnel task are, at least to some extent, the same as those activated by tasks that involve views of decisions in the real world triggered by environmental cues (Kelly et al., 2008; Török et al., 2014), verbalisations of internalised environmental knowledge, embodied experience, and so on. However, this fairly far-reaching hypothesis has so far not been addressed explicitly. To our knowledge, independent evidence that participants' tunnel task response patterns match different strategies and performance in the real world, validating the interpretation in terms of egocentric and allocentric reference systems, has not been put forward.

As such, the interpretation of turning and non-turning in terms of egocentric vs. allocentric reference systems is up for debate. As outlined above, previous research consistently suggests a profound and lasting effect of the experienced perspective (e.g., Kelly & McNamara, 2010; Montello et al., 2004; Shelton & McNamara, 2004; Török et al., 2014). The direct navigation experience in the tunnel task, i.e., the visual perception of travelling through a tunnel (rather than viewing it from above), should therefore have a strong impact on the cognitive representation, overruling other possible reference systems that do not correspond to the experienced view.

In fact, alternative explanations have been put forward. Klatzky et al. (1998) observed a systematic failure of spatial updating of turn changes in situations involving no actual movement, which could only be overcome by explicit and systematic instruction (enhancing participants' metacognition). Riecke & Wiener (2007) and Riecke (2008) confirmed this finding in a virtual reality task

bearing much resemblance to the tunnel task; however, these authors disentangled translation and rotation in their task. While the curves of the tunnel task integrate both of these, Riecke and Wiener had participants first navigate straight ahead and then turn on the spot. The effect of this is that participants who mentally updated the view direction ("turners") would point to the left after a rotation to the left. However, participants who did not point to the left in such a situation did *not*, as would be expected for "non-turners" (or clearly allocentric) responses, simply point back behind them to the starting point following translation. Instead, they responded by pointing in the *opposite* direction (to the right). Hence, Riecke and Wiener report this effect as a left-right inversion error. This finding raises further doubt concerning the reference system based explanation. This kind of VR induced spatial updating error consistently occurs across various task variations, to varying degrees (Riecke, Cunningham, & Bülthoff, 2007; Riecke, Sigurdarson, & Milne, 2012). However, participants' underlying conception of the task, the nature of the challenge posed by spatial updating, and the cognitive strategies for dealing with them have not been directly addressed.

Since research on orientation in virtual settings has thus opened up a number of issues related to basal vs. metacognitive processing, the tunnel task paradigm provides an ideal setting to address our concerns about the role of cognitive strategies in maintaining orientation in virtual navigation. The paradigm has been used as an indicator for the *results* of cognitive processes, presupposing (rather than addressing) the actual cognitive mechanisms and strategies that are employed when solving this task. To some extent, neglecting the cognitive processes involved makes sense because the conceptual reference systems used to solve this task are *expected* to be unconscious, and, accordingly, not directly accessible to analysis. However, this does not preclude participants from

developing conscious navigation or task solution strategies – and these may or may not be associated with generic types of reference systems. Moreover, since people can be instructed and trained to successfully adopt the strategy opposing their own intuitive concepts (Gramann et al., 2005), the underlying conceptualisations can apparently be manipulated by conscious (verbalised) decision. In fact, the extent to which reference system preferences are basal, unconscious, and inaccessible to metacognitive consideration is not known to date. Recent evidence (Münzer & Hölscher, 2011) indicates that people can tell reliably what they actually do and how successful they are in wayfinding and spatial orientation. This is based on preferences that are systematically related to egocentric vs. allocentric perspectives, though these, as such, may remain inaccessible to conscious reflection. Moreover, numerous studies show that reference systems are consistently (though not explicitly) expressed in language and can therefore be derived from how descriptions are formulated, even though potential ambiguity may complicate the analysis and necessitate a direct comparison with the observed scene (Carlson, 1999; Carlson-Radvansky & Logan, 1997; Taylor & Tversky, 1996; Tenbrink, Coventry, & Andonova, 2011). Therefore, systematic features of the language used to describe reference system based behaviour should be indicative of the underlying concepts, even if the speakers are unable to produce (or do not volunteer) an explicit description of the conceptual reference system they used.

Our present study was designed to investigate the participants' verbalised methods to solve a virtual navigation task in which spatial decisions needed to be made, based on as little perceptual information as possible. We asked to what extent the underlying reference systems are verbalisable as well as accessible to conscious manipulation via explicit instruction, and to what extent participants'

verbalisations exhibit linguistic features and patterns paralleling those that have previously been found for reference systems in language (Taylor & Tversky, 1996; Tenbrink, 2007, 2011). If participants' preferences in the tunnel task reflect the same underlying allocentric and egocentric reference systems as used to conceptualise the real world, the verbalisations should resemble those that have repeatedly been found for other types of verbalisations of different reference system concepts. Moreover, if participants' preferences reflect their generic (i.e., task-independent) preferences for a particular type of reference system, their verbalisations should reflect a change of conceptualisation when asked to change consciously towards a non-preferred strategy. Following Gramann et al. (2005), we predicted that such a change is actually possible based on explicit verbal instruction that enhances metacognitive consideration of the activation of a particular reference system. In a nutshell, the following hypotheses emerge:

H1. Verbalisability. Based on the literature on reference systems, orientation, and language as a representation of spatial concepts, we expect that participants will find ways of describing cognitive strategies in spite of the scarcity of perceptual input.

   a. If the concepts triggered by the tunnel task correspond to reference systems used in more complex situations (in real world or VR settings), verbalisations will clearly relate to previously identified verbalisations of allocentric or egocentric navigation.

   b. Alternatively, if the tunnel task triggers different types of concepts or strategies other than the two basic reference systems, the verbalisations will differ in nature from these known patterns of linguistic reference system indicators.

H2. Individual differences. Previous research using the tunnel task has shown that participants provide two clearly distinct kinds of behavioural responses, which we accordingly expect to be the case in our study.

    a. To the extent that their concepts and strategies are verbalisable, verbalisations should exhibit distinct patterns as well, corresponding to consistently different ways of referring to conceptually distinct reference systems or strategies.

    b. If the tunnel task triggers basic reference systems, these patterns will correspond to previously identified linguistic patterns for allocentric and egocentric spatial reference systems. Allocentric verbalisations will be produced along with allocentric behavioural results, and egocentric verbalisations with egocentric behavioural results.

    c. If the tunnel task triggers other kinds of cognitive strategies, the verbalisation patterns will not correspond to the basal reference systems, but may reveal participants' understanding and awareness of their experience in other ways.

H3. Conceptual flexibility. Following Gramann (2005), speakers can be made aware of different ways of conceptualising the tunnel task.

    a. If the tunnel task triggers reference systems, behavioural patterns will emerge as instructed, and the ensuing verbalisations will exhibit distinct patterns, as before. This is in line with previous work showing that humans are able to switch between reference systems and use them differentially in language (Levinson et al., 2002; Taylor & Tversky, 1996).

b. If the tunnel task triggers other kinds of cognitive strategies, behavioural and linguistic patterns may not correspond to the instruction conditions systematically.

Empirical study

Our version of the tunnel task was modelled on the pre-test and Experiment 2 described in Gramann et al. (2005), with three major differences.

First, we collected verbal data to address how participants would represent reference systems and cognitive strategies in language, building on principles of Cognitive Discourse Analysis (Tenbrink, 2015). For this purpose we had participants *think aloud* (Ericsson & Simon, 1993) while travelling through the virtual tunnels and indicating the homing direction (i.e. pointing back to the origin of travel); additionally we elicited retrospective verbal reports about their thought processes after each block of tasks.

Secondly, we speculated that participants might have difficulties interpreting the simulated 3D arrows that were shown at the end of each tunnel to indicate the possible position of the entrance. We reasoned that this method of eliciting direction judgments could be confusing because the arrows were, like the tunnel, only shown on the computer screen. They simulated a direction perpendicular to the actual two-dimensional plane on which the arrows were shown (see Figure 2). We wished to rule out the possibility that people's differing answers were the result of different interpretations of quasi-3D arrows on a 2D screen, rather than reflecting different reference systems or conceptualisations of the tunnel journey. Therefore we introduced a condition in which participants did not see the arrows; instead the screen remained black, and participants were asked

to indicate the position of the tunnel entrance by *pointing back over their shoulders*.

Figure 2: Two three-dimensional answer arrows displayed after a tunnel disappears (arrow condition). In the pointing condition the screen appeared just black without arrows.

Apart from avoiding potential confusions, actual pointing in a real world direction might also serve to raise awareness of the simulated embodiment involved in this task, and thus support the participants' metacognitive consideration of the simulated navigation. This idea is related to findings showing that physical performance (e.g. walking or turning) of simulated actions improves performance in VR tasks (Klatzky et al., 1998). Additionally however, physical pointing may provide an additional cue that could enhance the salience of one or another reference frame, for instance by reinforcing the egocentric perspective. Therefore, and in order to be able to compare our results to previous work in the literature using the tunnel task paradigm with arrows, we used both conditions in a between-participants design. We predicted that results should not systematically differ between the arrow and pointing conditions, but individuals might possibly become more consistent in their judgments because they were supported by

embodiment. If there was a bias towards one type of response, we expected it to correspond to the egocentric reference system.

Thirdly, to test if participants were flexible enough to employ a non-preferred strategy, Gramann et al. (2005: Experiment 2) instructed their participants explicitly that there were two ways of solving this task, and that they should use the one they hadn't used before. They gave feedback to participants and let them practise until they were comfortable with the non-preferred strategy. This procedure opens up various avenues for solving this task in a more conscious manner. On the one hand, participants may activate a metacognitive level, i.e., become aware of the two different reference systems and make an effort to conceptualise the scene in a different way. On the other hand, once the participants were informed that there are two correct possibilities, and that they were now supposed to adopt the alternative solution to the one they had consistently taken so far, they could use a strategic logical approach. It was possible to conform to the experimenter's expectation by projecting the preferred solution to the opposite side, i.e. the second previously not chosen arrow. Participants trying to be cooperative may not perceive this strategy as cheating, because they may intellectually grasp the verbal instructions, and yet they still might not genuinely adapt to the new way of conceptualising the tunnel task. Since we were interested in the extent to which participants could be made aware of different underlying reference systems based on verbal instruction only, we neither provided feedback nor allowed for additional practice trials, nor did we specify the relation of our instruction to the participants' performances on previous sessions. Specifically, we neither told them that they should now use the opposite strategy, nor that they should continue using their previous one. Instead,

we simply informed participants that there were several ways of solving this task, and that they were now asked to solve it in a particular way.

Method

45 native English speakers (32 females, 13 males; mean age 22 years, median: 21 years, age range: 19-33 years) participated in the study for compensation of expenses. They had normal or corrected-to-normal vision and were undergraduate, graduate, or postdoctoral students of Psychology at Cambridge University (UK). The tests and their accompanying explanations were organised in a single computer program to control information about the trials across all participants. The experimenter was not present in the room for the duration of the trials. Participants were asked to *think aloud* during task completion, and they were subsequently interrogated about their cognitive processes.

The participants' task was to keep alert during the passage through 20 simulated tunnels so as to be able to determine the starting point afterwards. The participants started the task by pressing one of the mouse buttons. An asterisk would then appear for 0.5 seconds and then tunnel movement would commence. Once the path was traversed the tunnel movement ended with the display of the last frame for 0.5 seconds. The participants' response was elicited, and the next trial started after a short interval. No feedback was given about their choice.

The tunnel length was kept constant. Each tunnel started and ended with a straight segment, but the curves differed with respect to direction (left or right) as well as angle. Participants were given 3 blocks of tunnels, with three different instructions as described below. The blocks always used the same set of 20 tunnels, presented in different order across the three blocks of tunnels. Each block consisted of 17 tunnels with just one curve varying between angles of 10 and 90

degrees in either direction, and 3 distractor tunnels that had two curves. Only tunnels with one curve were used for analysis because we did not have any particular hypotheses concerning more complex navigation scenarios, and therefore wished to keep the situation as basic and simple as possible.

For the purposes of providing a response, they were given either one of the following options, depending on condition. Participants were allocated randomly to these two conditions.

In the *arrow* condition, an asterisk was presented for one second, followed by two three-dimensional response arrows. By clicking on the corresponding mouse button participants chose the arrow that indicated the direction of the tunnel entrance relative to the end position.

In the *pointing* condition, they were asked to respond by pointing "back over your shoulder to where the starting point of the tunnel (the tunnel entrance) should be, in a straight line (as the crow flies)."

Altogether this yields a 2 (arrow vs. pointing) * 3 (instructions for tunnel blocks) design, with the former pertaining to our goal of ruling out an alternative explanation of the results (misinterpretation of arrows), and the latter directly addressing our third main hypothesis (H3: Conceptual flexibility). Our first two main hypotheses did not feed into the experimental conditions because the first (H1: Verbalisability) is qualitative, and the second (H2: Individual differences) pertains to different behavioural responses to the same stimulus.

Procedure

After a short briefing, participants were trained in the typically unfamiliar task of thinking aloud, following the procedural suggestions by Ericsson & Simon (1993). Then they were seated in an unlit, windowless testing room in front of a

16" computer monitor that was surrounded by black cardboard to eliminate additional reference information, and provided with a written instruction (see Appendix 1) that explained the tunnel task in detail, modelled on the instruction given by Gramann et al. (2005). They were shown an example tunnel for demonstration and asked to read the instruction again. After two practice trials, the experimenter reminded the participant to think aloud constantly, and left the room for the duration of the first phase consisting of 20 trial tunnels.

When the participant indicated that they were done with the tunnels, the experimenter returned and asked the following retrospective question: "Please describe to me in a few sentences what you think was happening and how you made your decision at the end of each tunnel journey." Next, participants were asked to fill out a questionnaire consisting of 13 or 15 (depending on condition) closed (yes/no) questions addressing various ways of conceptualising the tunnel task (see Appendix 2). These questions were motivated from verbal reports collected in a series of pilot studies[1] as well as from theoretical considerations (see Appendix 2). Subsequently, phase 2 commenced with a verbal instruction (again modelled from Gramann et al., 2005) intended to trigger either an allocentric or an egocentric representation (in balanced order across participants). The egocentric version was as follows, henceforth referred to as *bike* instruction:

"There is more than one way of solving this task, as you may have realised from the questionnaire. No matter how you did it so far, you are now asked to do the tunnel task again, but this time you are asked to conceive of it in a particular way.

---

[1] PART II NEUROSCIENCE PROJECTS, YEARS 2005/2006 AND 2006/2007. Department of Experimental Psychology, University of Cambridge, UK.

Please imagine that you are driving a bike, leaning yourself into the turns of the tunnel.

Again, while you solve this task, please talk aloud constantly, saying anything that comes to your mind."

The allocentric version (henceforth referred to as *ball* instruction) was identical except that the bike-related sentence was replaced by the following sentence:

"Please imagine that you are looking down on the tunnel from above, watching a ball roll through it."

All participants received both of these instructions, either for phase 2 or for phase 3. The order was balanced. These phases were carried out in the same way as the neutral phase 1, with the exception of the questionnaire, which was not administered again. After having answered the retrospective question in phase 3, participants were asked a few demographic questions, debriefed, and paid.

Analysis

To address our three main research hypotheses (H1: verbalisability*,* H2: individual differences*,* H3: conceptual flexibility) along with the more general question of how participants conceptualise the tunnel task, we focused on three types of evidence that provide diverse perspectives on the cognitive processes involved: behavioural data, elicited verbalisations, and questionnaire responses. According to previous literature as outlined above participants systematically diverge in their use of either one of the two available solution strategies, as well as with respect to developing a consistent strategy at all. With this solid empirical

background there was no reason to include a control group without language elicitation in our fairly time-intensive study. The issue of possible effects of this design will be taken up critically in the Discussion section.

Behavioural analysis

1. To prepare addressing H2 (individual differences) and H3 (conceptual flexibility), based on the neutral condition we allocated participants to behavioural preference groups, based on their preference for behavioural responses at a threshold of 80% (compare the 70% threshold in Gramann et al., 2005). Participants were allocated to the "egocentric" group if at least 14 of their 17 responses corresponded to the predictions for an egocentric conceptualisation (according to the literature using the tunnel task), and to the "allocentric" group if at least 14 responses corresponded to the predictions for an allocentric conceptualisation. Participants with less consistent responses were allocated to a third group called "no preference".

2. To address our concern that the conventional arrow representation might be interpreted in different ways as it involves translation of the current spatial experience to a two-dimensional direction on the screen, we compared the behavioural results (directional responses) in our three instruction conditions (*neutral*, *bike, ball*) for "arrow" vs. "pointing" responses.

3. To address H3 (conceptual flexibility) we compared the behavioural results of the three conditions (*neutral*, *bike*, *ball*). As part of this we performed additional analyses of response times and gender (not central to our research goals), which are provided in Appendix 3.

Linguistic analysis

To address all three hypotheses H1, H2, and H3 with their specifications as detailed above, we transcribed the think-aloud data collected during virtual tunnel navigation as well as the retrospective reports, and identified occurrences of the following linguistic indicators. While the actual set of linguistic items chosen was based on the nature of our data through post-hoc examination, the majority (all except *map* and *compass*) had been previously identified as reflecting underlying reference systems (e.g., Taylor & Tversky, 1996; Tenbrink, 2007; Wang, 2012). Using these in our analysis allowed us to distinguish between H1a and H1b:

- survey (arguably corresponding to *allocentric*) representation: *north, south, east, west, map, compass, (from) above*

- route (arguably corresponding to *egocentric*) representation: *left, right, front, straight, forward, back*

The relative frequency of indicators for both categories, as well as the number of participants mentioning any indicators of a category at all, was assessed as follows. Results were assessed separately for the preference groups (egocentric, allocentric, no preference) established post-hoc in the first step of the behavioural analysis as described above, allowing us to address H2a and distinguish between H2b and H2c, as well as for the complete data set in all three instruction tasks (*neutral*, *ball*, and *bike*) so as to distinguish between H3a and H3b.

According to H2b, participants preferring allocentric solutions should employ more survey indicators than those using egocentric solutions if the concepts triggered by the tunnel task correspond to reference systems known from other contexts. Likewise, according to H3a, participants should provide more survey indicators in the *ball* (allocentric instruction) condition than in the *bike* (egocentric instruction) condition. For route indicators, the opposite may be true.

The alternatives, H1b, H2c, and H3b, allow for the possibility that verbalisations may not correspond to these patterns, and so we approached the data with an open mind with respect to the types of linguistic indicators we might find.

Moreover, it stands to reason that, in this particular scenario, people might frequently use the terms *left* and *right* independently of the reference system used, simply because the tunnel curved to the left or right (as seen from above just as well as as seen from within the tunnel). In more general terms, the distinction between route-related projective (left/right, etc.) and survey-related expressions, which are typically associated with expressions for compass directions, may not be suitable at all for this particular structure-less scenario. In this sense, the systematic linguistic distinctions between reference systems as highlighted in the literature may be of limited scope. Ultimately, the literature to date does not provide a sufficient basis for predicting how the two conceptual representations triggered by the tunnel task may be represented linguistically.

For these reasons, we extended our linguistic analysis beyond the quantitative analysis of linguistic indicators of reference systems. To gain further insights from participants' verbalisations, we searched the think-aloud data and retrospective reports for verbalised ideas and strategies that might shed light on the participants' thought processes. The results of this qualitative analysis will be presented in terms of case studies, where the verbalisations along with the associated questionnaire results serve to explain the individuals' behavioural results. This will shed light on the range of variability in participants' concepts, which to some extent accounts for the limitations of a purely quantitative approach in this context.

Questionnaire analysis

We were interested in the relationship between participants' behavioural results and retrospective rationalisations that might reflect various ways of conceptualising the task. For this purpose, we administered a questionnaire (see Appendix 2) with 13-15 questions (depending on condition) requesting yes/no answers. 10 of these were designed for specific quantifiable purposes as follows, while the remaining questions were designed to provide further background information. The first two (see table 1 below) questioned the VR experience created by a minimalistic and simplistic visual input while simultaneously excluding any other senses (e.g. propriosensoric), by asking if the participants actually experienced themselves as moving through the tunnel (as they were expected to). Questions 9 and 10 addressed the possibility of turning (re-orienting) back after emerging from the tunnel. These four questions were confirmatory in nature.

Table 1: Questionnaire answers predicted from participants with allocentric and egocentric preferences (see Appendix 2 for the complete questionnaire). The first four questions in this table (no. 1, 2, 9, 10 in the questionnaire) were confirmatory in nature and addressed the participants' general conceptualisation of the task, independent of behavioural preference or underlying reference system.

	Answers expected from an allocentric perspective	**Answers expected from an egocentric perspective**
(1) moving through tunnel	yes	yes
(2) tunnel moving past me	no	no
(9) keep orientation constant	yes	yes
(10) turn to look at direction	no	no
(3) re-orienting head	no	yes
(4) re-orienting body	no	yes
(5) changing direction	no	yes
(6) looking straight	yes	no
(7) lean body	no	yes
(8) bird's eye view	yes	no

Questions 3 to 8 tested participants' answers for allo- or egocentric preferences, and were therefore quantified for further analysis. We had clear predictions for these questions based on previous findings in the literature as outlined above. Any answer in accordance with these predictions was rewarded with 1 point, while any deviation received -1. Consequently participants could achieve a maximum value of 6 if they always responded in accordance with

expectations, and a minimum of -6 if they always gave the opposite answer than expected.

Since the data collected in this way are discontinuous, i.e. only discrete integers are possible, they were analysed using nonparametric test statistics (Kruskal Wallis test for the three-group comparison, and Mann-Whitney-U test for comparing two groups).

Results

The analysis is based on a set of 45 participants. 22 of these provided their directional answers by pointing over their shoulder (pointing condition); the others clicked on one out of two arrows (arrow condition).

Behavioural results

Behavioural preference groups. In the neutral phase 1, most (34 of the 45) participants provided consistent answers that corresponded to the expectations of an underlying preference for either an egocentric (N=22) or an allocentric (N=12) reference system. For convenience these will in the following be called 'allo' and 'ego' groups, respectively, bearing in mind that this does not necessarily imply that these response patterns were actually based on such a reference system. Eleven participants did not show a clear preference and will therefore be referred to as the 'no preference' group.

Comparison of 'arrow' and 'pointing' conditions. Under both conditions -'arrow' and 'pointing' – a similar number of participants (see Figure 3) consistently used an allocentric or egocentric reference system, or remained without preference. Therefore Fisher's exact test could not detect any significant

difference between the 'arrow' and 'pointing' conditions, and we combined data from both conditions ('arrow' and 'pointing') to one dataset for further analysis.
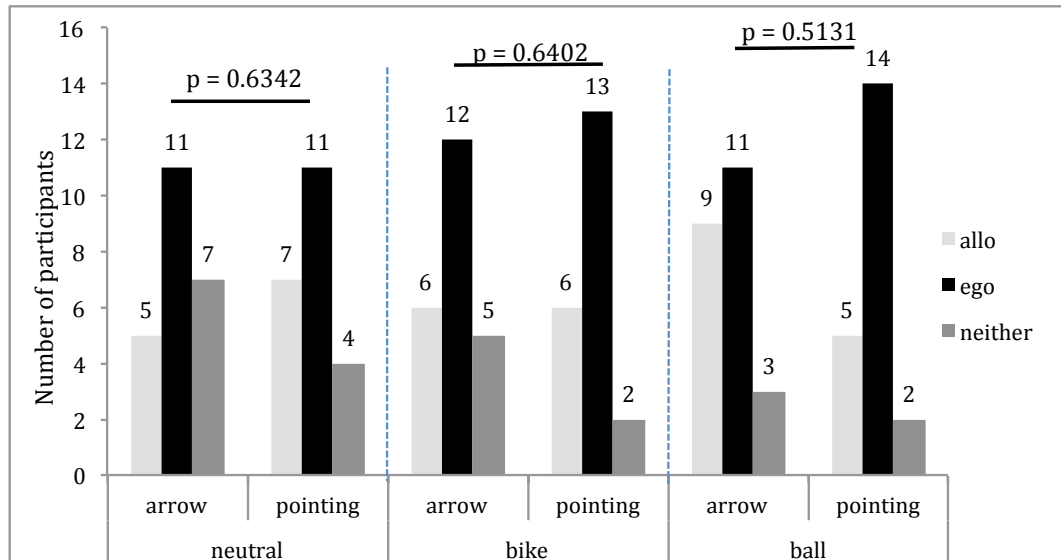

Figure 3: Number of participants with preferences for "egocentric" or "allocentric" or no clear response pattern in the arrow and pointing conditions.[2]

Influence of instruction on behavioural preferences. Addressing H3 (conceptual flexibility), the instructions in the *bike* and *ball* conditions were formulated to trigger a shift in the dominance of preferences. The *bike* condition provided instructions for an egocentric reference system, while the instructions in the *ball* condition favoured a conceptualisation based on an allocentric reference system. Surprisingly, results revealed little change in individuals' responses for those who had developed a clear preference during the neutral condition. However, in both the 'arrow' and 'pointing' conditions, the number of people without a consistent preference decreased over the course of the three sessions (see Table 2); in this respect, the "no preference" group differed significantly

---

[2] Datapoints for one participant each – different ones – are unfortunately missing for the bike and ball conditions.

(Fisher's Exact Test, p = 0.0013) from the other two groups. The majority (82%) of the participants who did not show a clear preference during the neutral phase developed a preference and favoured either egocentric (6 participants) or allocentric (3 participants) solutions. Two of the latter developed this preference only in the *ball* condition (consistent with the instruction), while the seven others kept their preference constant independent of instruction. Only two people remained without a clear preference.

Table 2: Number of participants with allocentric, egocentric, or no preference during the neutral phase who showed the same preference during all conditions (constant) or changed their preference (change). *: Fisher's Exact Test (p = 0.0013)

	constant	change	Total
Allocentric	10	2	12
Egocentric	17	5	22
No preference	2*	9*	11*

Two participants with an allocentric preference lost their preference in one of the conditions. One of them had no preference in the *bike* condition and the other had none in the *ball* condition.

Five participants with an egocentric reference system changed their preference. Four of them failed to show a preference either during the *ball* condition, the *bike* condition, or in both. Only one participant switched her preference according to the instructions from egocentric to allocentric in the *ball* condition.

In sum, only 35,6 % of all participants were influenced by the instructions and consequently changed their preference. However, there is no indication that the instructions shifted the responses of the participants in a predictable (such as the intended) direction. This unexpected result could not be due to a lack of statistical power, i.e., our data showed no tendency for the response patterns to shift according to the instructions that might emerge more clearly with a higher number of participants.

Quantitative linguistic results

To address all three of our hypotheses along with their specifications as detailed above, we investigated the use and distribution of route and survey perspective markers in the linguistic data. We found a strikingly consistent preference for route perspective indicators in the linguistic data. Across all data we only collected 43 instances of survey-type linguistic markers produced by 45 participants, i.e., less than one per speaker on average. This casts doubt on the association of tunnel task responses with generic reference system types (to be discussed below). Nevertheless, their distribution across conditions and behavioural preferences shows some interesting tendencies. Participants in the 'allo' response group used slightly more survey perspective indicators. There was also a tendency towards survey-type verbalisations following instruction in the *ball* condition, but not in the *bike* condition (corresponding to the predictions according to reference systems).

With the neutral instruction in phase 1, only 26.67% of the 45 participants used survey indicators *at all* throughout the verbalisations. 41.67% of the 12 participants in the 'allo' group used survey indicators, and 22.73% of the 22 'ego' participants used survey indicators. In contrast, all participants used route

indicators frequently throughout their verbalisations, rendering this aspect unsuitable as an indicator for differentiated cognitive processes in this setting. Altogether, 98.85% of all identified linguistic indicators in the neutral condition were route perspective indicators; i.e., only 1.15% fell into the survey category (N=43). Of these, allocentric behaviour was associated with 2.65% linguistic survey indicators, egocentric behaviour with 0.44%, and inconsistent behaviour with 0.80% (see Figure 4). A t-test revealed a marginally significant difference between the preference groups with respect to the frequency of use of survey indicators in the neutral condition (T=1.989; df=32; p=0.055), but not with respect to the use of route indicators (T=0.369; df=32; p=0.715).
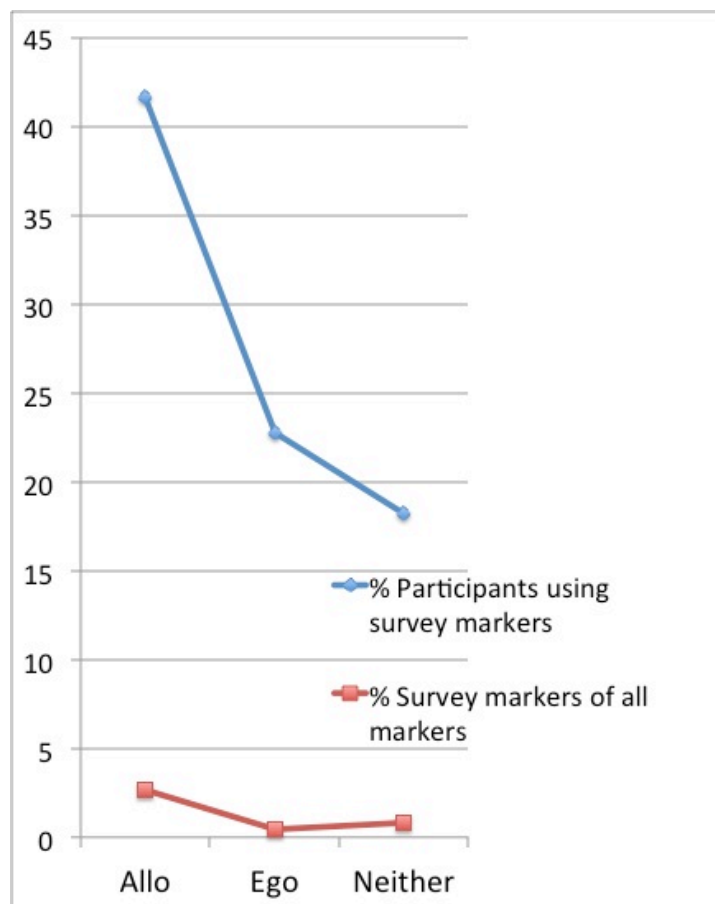

Figure 4: Distribution of linguistic survey indicators across participants according to their behavioural preference ('allo', 'ego', or 'no preference'). The top line shows

whether they used survey indicators at all, and the lower line their percentage of all markers used (i.e., relative to route indicators).

With the *bike* (egocentric) instruction there were 1.03% survey indicators, and with the *ball* (allocentric) instruction 2.73%. The use of survey indicators was highly correlated across conditions (p<0.001). It differed significantly between the *neutral* and *ball* conditions (T=2.107; df=42; p<0.05) and between the *bike* and *ball* conditions (T= -2.936; df=41; p<0.01), but not between the *neutral* and *bike* conditions (T=-0.852; df=42; p=0.399). Altogether, it seems fair to say that linguistic indicators of survey perspective may reflect behaviour to a limited extent. Nevertheless this is not a prominent feature of the think-aloud protocols, since the number of such terms used spontaneously is so low. Therefore it is necessary to take a closer look at what participants actually say, as we will do in the qualitative analysis below.

Questionnaire results

In the four questions that were asked to probe the participants' general conceptualisation of the task independent of the underlying reference system, the three groups ('allo', 'ego', no preference) did not differ in their responses (Kruskal-Wallis Test, df=2, H= 2.622, p=0.269). This indicates that participants mostly followed the instructions as intended, or at least believed they had done so. Questions 3-8 explored the participants' conceptualisations with respect to reference systems, addressing the concepts employed explicitly in the *ball* and *bike* instructions (which followed the questionnaire). Participants in the 'allo' group responded in ways consistent with our predictions significantly more often than did participants in the 'ego' group (Mann-Whitney U Test, U= 211.000, r =

10.480, p= 0.001). The 'ego' group had a tendency to answer contrary to predictions; the results of about 80% of the participants in this group equaled or were below 0 (Figure 5). In contrast, this was only the case for 25% of participants in the 'allo' group. Judging from these results, it appears that while the 12 participants in the 'allo' group might have had a (reasonably) clear conceptualisation of an allocentric reference system and kept it stable, this was not case for the 22 participants in the 'ego' group.
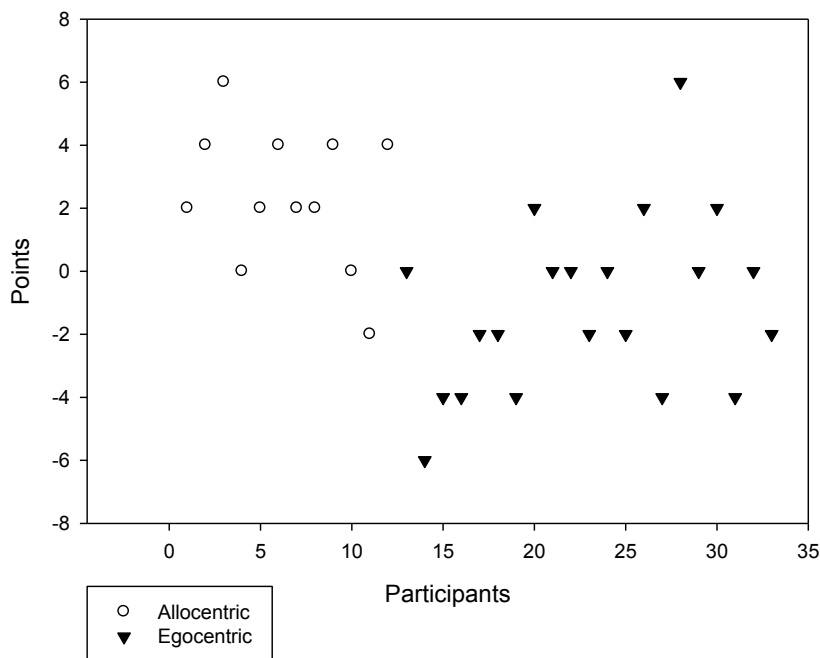

Figure 5: Questionnaire analysis for 12 participants categorised based on their behavioural preference as 'allo' (x-axis: 1 – 12) and 22 categorised as 'ego' (13 – 34). Each symbol represents a participant. Any answer in accordance with predictions was rewarded with 1 point, while any deviation received -1 (see Methods). Consequently participants could achieve a maximum value of 6 (on the y-axis) if they always responded in accordance with expectations, and a minimum of -6 if they always gave the opposite answer than expected. The distribution shows that the 'allo' participants responded mostly according to predictions (nearly all dots are above the 0 line), while the 'ego' participants did not.

Case studies

Since the questionnaire results were inconclusive and the quantitative linguistic analysis revealed only marginal informativity of linguistic indicators such as those known from the literature, we were interested in comprehending in more detail how individual participants might have understood this task. In the following, we present different cases of metacognitive considerations exhibited in verbalisation. We begin by discussing two clear cases of consistent egocentric or allocentric conceptualisations before turning to more surprising (and therefore illuminating) cases.

1) Consistent egocentric representation

The behavioural results of participant no. 12 are consistently egocentric across all three conditions, independent of the explicit instruction given. The *think-aloud* data are fairly repetitive in the neutral condition, with some variations to the format of "going straight, sharp curve off to the left, um and then straight again, so back – to my left." Astonishingly, the participant retrospectively reports "I was just kind of imagining where I was from above really", which points to an allocentric representation (contrary to the behavioural results). However, when given this same question explicitly in the questionnaire, the answer is NO, reflecting an inconsistency or lack of awareness of the employed conceptual perspective. In the *bike* task the participant frequently mentions leaning into the curves as requested, and subsequently reports "I was – imagining it more from being in the tunnel – um – this time". In the *ball* task, the participant in effect takes the ball's "perspective", as evidenced by repeated utterances such as "so the

ball's now rolling straight, starting to turn right, um he's curved to the right, and he's now going straight, so it's – uh behind him to the right." Here, both the personal pronoun "he" and the concept of localising the entrance "behind" the ball (which should not possess an intrinsic back side) suggest a personification rather than an allocentric perspective. In the associated retrospective report the participant says, "I found that one a bit harder to imagine sort of being from above just 'cause - you - you know, you're going straight down the tunnel" – reflecting on the discrepancy between the perceptual input shown on the screen, and the suggested notion of viewing the scene from above. In all three conditions, there are a few linguistic survey indicators (two, one, and three, respectively, in the *neutral*, *bike*, and *ball* condition). Altogether, this candidate appears like a prototypical participant exhibiting a strong egocentric reference system preference, with little metacognition of the underlying concepts and processes. The initial retrospective report and a higher than average use of survey indicators deviate from the egocentric pattern.

2) Consistent allocentric representation

Participant no. 51 starts out in the first few trials of the tunnel task, in the neutral phase, by verbalising egocentric representations such as "so we're going round to the right which means that the tunnel entrance should be over my right shoulder still". With the third trial, the concepts appear to shift, and the participant mutters a few almost inaudible remarks sounding like "let me think about it... how do I know". After this, she starts thinking aloud as follows: "we went left, so we want to go back to southeast". Such reflections of clear allocentric thinking were then continued throughout the trials during the neutral phase. The behavioural data show two egocentric trials followed by almost consistently allocentric responses

(with an extremely long reaction time in trial 3, corresponding to the conceptual shift), with one further exception in trial 7 (another egocentric response) where the participant's think-aloud data reflect confusion ("hmm, what is that?").

The retrospective report reflects no conscious recollection of this conceptual shift; instead, the participant simply reports, "I was thinking about in terms of a compass how you get back". Taken together, the think-aloud data and the retrospective report of the neutral phase contain 15 indicators of survey perspective, and 69 indicators of route perspective, yielding a percentage of 17.86% survey indicators.

Altogether in this neutral phase, this participants' verbalisations were precisely as expected for participants using the allocentric perspective, with the additional feature of a conceptual reorientation after the first two trials. Surprisingly however, in the questionnaire she reports reorienting both head and body during the tunnel journeys, but negates conceptualising the scene from a bird's eye view.

The explicit instructions did not affect this participant's concepts. The behavioural results remained consistently allocentric with both *bike* and *ball* instructions, the think-aloud data and the retrospective reports reflected survey perspective as before, and the percentage of survey indicators remained high (26.76% in the *bike* condition and 27.16% in the *ball* condition). The retrospective reports reflect a conscious preference for the *ball* instruction as might be expected for a person with an allocentric preference.

Altogether, this consistently allocentric representation appears to mirror the egocentric version in the previous case in that responses remain consistent and unaffected by instructions, and in that there are some discrepancies with respect to the verbalisability of underlying concepts, in accord with the assumption of

mainly unconscious processes. It could be expected that most people showing allocentric preferences would act and verbalise in similar ways. However, notably this person is *the only one* in our data set whose verbalisations reflected an allocentric perspective in a clearly identifiable and consistent manner. This fact is reflected in the percentage of linguistic survey indicators, which were about ten times higher for this person than the average (see section Quantitative Linguistic Results above). In the following we look more closely at two other participants with consistent behavioural allocentric preferences to see how these people verbalise this.

3) Logic calculation based on allocentric reference system

When asked about their use of logic rather than sense of direction in the questionnaire, only six participants responded in the negative, in spite of the fact that only six (different) participants negated using their sense of direction each time. Thus, people did not regard this as a contradiction, but combined their logical thinking (arguably a conscious effort) with their sense of direction (arguably a more unconscious process). Here we look more closely at one participant for whom logic appeared to play a major role.

The think-aloud data produced by participant no. 4 in the neutral condition contain frequent markers of causal relations such as *so,* as in "I'm going straight ahead again, then we're turning - to the left, that was a significant turn (...) so I'm gonna point more - to the right than the last time behind me." Either *so* or *therefore* appeared almost every time before indicating the start direction (in a consistently allocentric manner), sometimes reinforced by "so I'm just going to say it's...". In retrospect, this participant reports "basically if - if we went along and then turned to the right on the tunnel, I knew that the point of entrance was to

the left, so that - that was basically how I made my decision". Following the *bike* instruction the think-aloud data reflect the participant's efforts to imagine the situation as requested, as in "now this seems like quite a sharp, oh yes very sharp - I'm about to fall off my bike". Corresponding to the unaltered allocentric behavioural results, the participant reports in retrospect, "so I made the decision on exactly the same grounds, the only difference was this time I tried to move my body as if I was on the bike".

Following the *ball* instruction, the think-aloud data reflect confusion in spite of the participant's clear preference for the corresponding perspective: "what I'm trying to do now is not think about left and right and just visualise a ball, so to do it somewhat more intuitively, which in a way requires a lot more concentration and it isn't such that I can justify my answer". The causal markers become less frequent; a typical utterance before indicating the homing direction now is "I think that's over there". The final retrospective report after finishing the trials in the *ball* condition is revealing with respect to the relative roles of metacognition, intuitive conceptualisations, and logic in this situation: "whether I was a person or a bike or a ball I think I wasn't actually using the same way of doing it, so I was trying to think well what's another way of doing it. So I tried not to do the left right kind of calculating mode of thought and just to imagine it leaving a trace, a visual trace. Um, so – a more kind of intuitive thing. I did that for a while and then I found myself going back to the left right calculating thing". Nevertheless, this participant consistently produced allocentric responses throughout, except for one egocentric response in each of the neutral and bike conditions and two exceptions in the ball condition. Thus, in this case the metacognitive verbalisations revealed some awareness of the way in which responses were given (based on logic), but

no flexibility towards changing these concepts using a different kind of reference system.

4) Turning back and dropping pennies

Coming out of the tunnel, participant no. 71 apparently adopted a different kind of logic as compared to the previous example, namely a logic of turning back, as in the following think-aloud statements which are typical for this participant's utterances before making a decision: "it's reached the end of the tunnel, and because it went right, um, to go backwards you'd have to go left, so that's why I think it's the one on the left", and "because the tunnel turned right – um, to back on itself, um - you'd have to turn left, so that's why I think it's the left one." In the retrospective report, this participant says "I imagined a map in my head of - of what it - what the journey would look like", which corresponds to our expectations for a person using an allocentric perspective; and in fact all responses were allocentric. In the second condition, which for this participant was the *ball* instruction, the conceptualisation changes entirely. From the first trials, the participant starts talking about adopting a perspective from above, though not in the manner expected: "pretending like I've - dropped a penny down and it's going to the right and then it's fallen out and then the penny would have fallen out to the right. So to come back up to me it would have had to go to my left, um - so it would have to go back up to my left. So it's the one on the left". Although the direct references to dropped pennies stop after a while, the retrospective report again confirms that this unexpected perspective, which combines an egocentric concept with a view from above but nevertheless yields consistent allocentric results, was kept up during this condition. Somewhat surprisingly, another participant (no. 8) reports similar concepts following the *ball* instruction, as in "a

ball is dropping down through the hole". Following the *bike* instruction, participant no. 71 goes back to the previous logic of turning back, as evidenced by verbalisations such as "because I turned right I'd have to go back in the left direction" (yielding allocentric results in 14 out of 17 cases rather than the egocentric type expected with the *bike* instruction).

To conclude, the case studies suggest a range of metacognitive strategies that participants adopted in order to find consistent solutions to the tunnel tasks in the neutral condition, and subsequently to follow the instructions in the *ball* and *bike* conditions. However, these strategies and efforts did not lead to a change in the behavioural results, in spite of the conscious effort invested in changing perspective. During the tunnel experience itself, participants said little more than 'left' and 'right' (or more complete sentences containing these terms), but then pointed to decisive concepts that led to consistent decisions after the journey was completed. These concepts varied considerably between participants and corresponded to the notions of allocentric and egocentric reference systems only to a limited degree.

Discussion

In this study, we aimed to gain insights into human understanding of orientation in desktop virtual reality scenarios, motivated by earlier research associating worse performance and enhanced cognitive effort with desktop VR navigation as compared to settings involving physical motion. Specifically, we addressed participants' verbalisations in a simple virtual navigation scenario called the tunnel task, which has frequently been used in the literature to test orientation skills and reference system proclivity (e.g., Goeke et al., 2015; Gramann et al., 2005;

Schönebeck et al., 2001). Our results highlight a range of ways in which participants conceptualise this task in an effort to maintain orientation in the absence of idiothetic or contextual cues. Their verbalisations alongside task performance and the reported metacognitive strategies corresponded only in a limited way to the spatial reference systems humans are known to use and verbalise in settings involving contextual cues and physical motion. Instead, they reflected highly varied ways of cognitively representing the scene. In the following, we will briefly discuss our study in relation to the earlier literature using the tunnel task, before turning specifically to our three research hypotheses and their implications in light of the wider literature.

Although our empirical design differed from earlier studies in a number of respects as detailed above, our behavioural data correspond to previous results (e.g. Goeke et al., 2015; Gramann et al., 2005, 2006, 2009) in that most participants provided consistent responses in the neutral task (which did not include further instruction specific for a reference system). To be successful in VR tasks, feedback training is often critical for enabling participants to achieve the level of accuracy necessary to be accepted for participation (Riecke & Wiener, 2007). Even though they were not given such explicit feedback, the majority of our participants developed a clear preference for choosing one response significantly more often than the other. Our results in the neutral task (condition 1) fell into two groups of similar magnitude; paralleling the earlier literature, these could be interpreted as a preference for either egocentric or allocentric reference systems.[3] Eleven of our 45 participants did not develop a consistent strategy; these

_____

[3] It was not the aim of the current study to examine the extent to which the various changes in the design might lead to increased preference of either one of these consistent solutions. Distributions of preferences vary between studies in the earlier literature; furthermore, they may be influenced

were labeled the 'no preference' group. This group corresponds to the minor number of participants who would be regularly excluded from further participation because they did not reach criterion (or could not be unambiguously assigned to a category) in the pre-test or neutral phase. We, however, consider the existence of this group a natural consequence of the cognitive challenges involved in this task, as evidenced throughout the analysis in various ways. While the preference groups soon settled for a strategy to solve the tunnel task, participants with no preference remained the slowest to make a choice in all three conditions (see Appendix 3). This delay in response could be an indicator for extended thought processes and/or indecisiveness.

While in other studies physical movements (e.g. turning) improved performance (e.g. Riecke & Wiener, 2007; Wraga et al. 2004), participants requested to point over their shoulder back to the origin of the journey did not significantly differ from those who had to choose one out of two arrows, except for a tendency to produce "egocentric" results more consistently. Arguably, the experience of a physical motion is more important *during* navigation than *after* it, i.e. at the time of pointing back to the entrance. In any case, our results show that complications in conceptualising the tunnel journey and subsequently identifying the homing direction cannot be traced back to the display of arrows on the screen, as might have been speculated.

Altogether, our behavioural results mirror previous findings with respect to the general distribution of response patterns, allowing us to address the cognitive challenges concerning orientation in desktop VR scenarios. Our first hypothesis

---

by the training or feedback procedures in such studies. The result that two distinct preference groups emerged in our data is sufficient for the purposes of the current study (in terms of comparability with earlier results).

concerned verbalisability. Based on the literature on reference systems, orientation, and language as a representation of spatial concepts, we expected that participants would find ways of describing cognitive strategies in spite of the scarcity of perceptual input. In this respect, the participants' responses in questionnaires as well as verbalisations revealed a complex picture. The journey itself appeared to have been experienced in a fairly intuitive manner, leading to relative simple and uniform verbalisations mostly reflecting the tunnel's curves to the left or right. As such, the tunnel experience may be unproblematic; the tunnel's curves can simply be represented in their directionality, without any potential orientation issues or reference system divergences.

Following navigation, corresponding to previous studies, our participants were asked to make a conscious choice between arrows, or shoulders, to indicate the direction they came from. In other words, they were challenged to transfer the visually observed, but not physically experienced, navigation experience from the two-dimensional screen to an imagined three-dimensional situation. This challenge corresponds to other desktop VR navigation studies where participants are asked make a spatial orientation judgement based on a scene shown on a screen, except that the tunnel task scenario was designed to exclude the known influence of environmental cues (Kelly et al., 2008; Török et al., 2014). Our verbalisation data suggest that participants used diverse kinds of metacognitive strategies to solve this task. There are various indications that at least some of the participants employed cognitive shortcuts (such as *the tunnel went to the left, therefore the entrance must be to the right*) to identify the homing direction even *without* relying on their disembodied virtual navigation experience. Many participants furthermore reported confusion and described how they consciously

changed their procedures due to learning effects. The outcome of these changes could to some extent be traced in the behavioural results.

If participants had been able to solve the orientation task without conscious strategies, verbalisations should have been less rich and diverse. As specified in H1a, based on the literature we expected that verbalisations might clearly relate to previously identified verbalisations of allocentric or egocentric navigation. This would have yielded a more uniform verbalisation data set falling into two categories, possibly corresponding in nature to our two first case studies. However, these were exceptional cases rather than consistent patterns throughout the data. We conclude that the cognitive challenges involved in orienting in desktop VR scenarios are profound enough, in the absence of environmental cues, to trigger conscious and verbalisable strategies that support participants in making consistent judgements across trials.

Our second hypothesis concerned individual differences. Based on previous research using the tunnel task we could expect two distinct kinds of behavioural responses, which is also what we found. However, the verbalisation data did not fall into corresponding patterns associated with allocentric and egocentric reference systems, or any other reference system categorisations such as those proposed specifically for language by Levinson (2003) or Taylor and Tversky (1996). Neither did participants producing egocentric responses exclusively describe their experience from an egocentric point of view, nor did those producing allocentric responses consistently focus on the spatial trajectory from an objective, outside perspective. Instead, without consistent relations to the behavioural patterns, the verbal descriptions uniformly represented the tunnel's curves to the left and right, but were highly diverse when representing the orientation task itself. Although the verbal data showed slightly enhanced use of

linguistic survey-type indicators produced by participants with an allocentric behavioural preference, numbers of survey-type indicators were very low throughout, and many of the participants never used any such indicators at all. This is in spite of the fact that one of the linguistic indicators of the survey perspective, 'above', was used in the *ball* instruction and might therefore have primed the participants to use it directly in their verbalisations. In fact, this term was used more frequently in the *ball* condition than any of the other linguistic indicators. This marginal priming effect can account for the observation that survey indicators were used somewhat more in this condition than in others, without presupposing a survey-type representation.

As revealed by the qualitative analysis, a consistent allocentric conceptualisation was hardly ever reflected in a verbal data set (with the exception of one single participant). This result stands in contrast to the literature on survey (associated with allocentric) vs. route (associated with egocentric) representations in language (e.g., Taylor & Tversky, 1996; Wang, 2012), where different conceptualisations consistently lead to different linguistic description types, even though the linguistic indicators as such may depend on the task context (e.g., Tenbrink et al., 2010). Here, in contrast, the spatial terms used by the participants did not seem to differ according to their behavioural responses, irrespective of the analytical framework for linguistic analysis. Participants consistently preferred projective terms throughout, without any associated linguistic patterns that could reveal any differences in underlying reference systems.

One reason for these results might be as follows. The assumption that an allocentric reference system is used by a subset of participants essentially presupposes that no particular orientation is needed to assess the corresponding homing direction. In fact, allocentric reference systems are typically associated

with survey knowledge about the relative relationship of objects or entities to each other, independent of the human's current orientation (Klatzky, 1998; Levinson, 2003; Burgess et al., 2004). However, in the tunnel task participants actually depend on the stability of their head direction. Rather than being able to lose track of orientation and imagine themselves as a point rather than an intrinsically oriented entity, they need to keep their current orientation stable no matter how their path curves. Thus they derive the relative position of the entrance from their own line of sight, which remains unchanged. As an outcome, even if they develop a map-like ("survey") knowledge of the tunnel journey, in order to identify the homing direction they still need to relate to their own orientation.

Moreover, there is no way of excluding that participants who imagine themselves as aligning with the tunnel's curves (changing their orientation) also develop survey knowledge along the way. In effect, then, the main difference between the two kinds of homing directions lies in whether or not the participants imagined turning their head during travelling. This interpretation corresponds to references to "turners" vs. "non-turners" in Gramann's work (e.g., Gramann et al., 2005), but it does not necessarily imply distinct reference systems. Previous findings on reference systems suggested that shifts or inconsistencies of perspective lead to worse performance and thus involve higher cognitive effort (Shelton & McNamara, 2004; Török et al., 2014). This should induce participants to prefer the egocentric response pattern, which corresponds to the visual input. In line with this assumption, Klatzky et al. (1998) postulated that a failure to update one's head position is simply an error, rather than representing an alternative conceptualisation. Riecke & Wiener (2007) and Riecke (2008) further suggested that some of the left-right confusions cannot simply be traced back to a failure to update one's head position, since such a conceptualisation would have led to

different results in their study. Their results rather suggest that participants actually failed to understand the implications of a virtual spatial environment. All of these authors agreed, however, that participants can be trained by explicit instruction and feedback to provide correct responses.

This leads us to our third hypothesis, which concerned conceptual flexibility. When Gramann et al. (2005) made speakers aware of different ways of conceptualising the tunnel task and instructed them to use a particular one, this had consistent effects on their behavioural results. We used the same strategy in order to trace effects on verbalisations and metacognitive strategies; however, we wished to rule out that participants simply conformed to the experimenter's expectations after having been told to adopt the *other* kind of reference system than the one they had been using before. This would be possible by choosing the other arrow than the one they would have chosen without the instruction. In contrast to Gramann et al. (2005), therefore, our participants did not receive any such feedback.

While the neutral condition yielded clearcut and consistent behavioural results that generally conformed to previous studies as described above, the results of the two other conditions came as a surprise. Contrary to the results reported by Gramann et al. (2005), our participants failed to adopt the suggested conceptualisations. In fact, there was no indication (neither behaviourally nor verbally) that they changed their concepts and strategies at all. Therefore, it stands to reason that, although participants can be trained to choose a different *solution* to the one they had previously employed, they do not necessarily adopt a different conceptual *representation* on the basis of a verbal description, at least not without correcting feedback.

The idea (outlined above) that the two kinds of responses may not correspond to distinct underlying reference systems provides a suitable explanation why participants in our study consistently failed to adopt the alternative reference systems suggested by the instructions in the *bike* and *ball* conditions. These descriptions do not provide a good match with the cognitive strategies involved in (mentally) turning the head as opposed to trying to maintain the current orientation during navigation, nor are they suited to induce or remedy a left-right confusion that has its roots in a fundamental misconception of the virtual environment. Accordingly, the verbalisation data collected in our study reflected the participants' sincere attempts to adopt the recommended conceptualisation, together with their consistent failure to draw any meaningful conclusions from these instructions that would have led to a change of reference systems.

Conceivably, a different formulation of the instructions for adopting a specific reference system might have been more effective. However, these instructions were taken directly from previous research that led to far-reaching conclusions (e.g., Gramann et al., 2005). Moreover, it seems unlikely that a fairly superficial modification would have led to substantially different outcomes.

However, there is another consideration leading to an alternative explanation that cannot be ruled out by our data. Our results might be an artefact of our version of the task, triggered by our efforts to elicit metacognitive representations via concurrent verbalisation and questionnaires. Also the real-world pointing condition may have enhanced metacognitive strategies to a higher extent than arrows on the screen. Thus, in other studies that do not employ these features, participants may not necessarily use the same metacognitive strategies,

nor use cognitive shortcuts. However, we consider this interpretation unlikely for the following reasons.

First, our results are consistent with previous insights gained across VR scenarios concerning the considerable challenge of achieving reliable spatial updating results (e.g., Klatzky et al., 1998; Riecke et al., 2007, 2012; Wraga et al., 2004). Our findings provide crucial insights into the cognitive processes involved in this challenge, and participants' ways of dealing with them. In the absence of further cues, they establish a logical or imaginative concept that serves to anchor their orientation in such a way as to allow for a consistent judgement of direction. This is in line with previous findings that participants can learn to establish and track a reference direction if they know about future demands in a setting not involving active movement (Mou et al., 2009).

Second, the present study was (among other aspects) motivated by informal discussions with people who had experienced the tunnel task in versions corresponding to the published ones, and who reported a number of cognitive strategies that they (consciously) employed. In a subsequent series of pilot studies[4] participants were not asked to think aloud (i.e., they remained unbiased with respect to metacognition while performing the task itself) but were subsequently asked to report their experience. They consistently volunteered information about cognitive strategies, similar to the informal experience reports just mentioned. The linguistic data that we collected in this study were entirely within the scope of these earlier observations. Our current study design represents a more controlled extension of these experiences in that we aimed to elicit

---

[4] PART II NEUROSCIENCE PROJECTS, YEARS 2005/2006 AND 2006/2007. Department of Experimental Psychology, University of Cambridge, UK.

information about the extent to which participants were already aware about their strategies *while* (rather than only *subsequent to*) solving the tunnel task.

Third, it is common practice to explain the tunnel task (and similar ones) on a detailed (metacognitive) level to participants so as to ensure consistent performance. Thus, some amount of metacognition appears to be considered a normal part of this kind of study, and arguably in VR navigation tasks in general.

Fourth, there is no indication that the task of thinking aloud affected participants' performance. Neither do the behavioural results reveal systematic differences to previous results, nor does the extensive literature on the effects of thinking aloud (e.g., as put forward concisely by Ericsson & Simon, 1993) point to distorting effects in a task like the present one. Reactivity effects are found, for instance, when verbalisation interferes with the natural flow of thoughts while solving a problem, or when reports are elicited that exceed the range of thoughts retrieved directly from memory (Ericsson, 2002). In light of the consistency of our observations with previous findings, it seems unlikely that participants' behaviour was influenced to such a high degree that substantially different cognitive processes were activated by the task of thinking aloud.

This is particularly true for our neutral condition, which was carried out paralleling previous studies using the tunnel task paradigm. To our knowledge, no other studies are available that directly test whether participants can be persuaded to use a specific conceptualisation by verbal instruction only, rather than feedback on their performance. Our rather unexpected finding that such instruction had no effect on performance should be pursued further, shedding further light on what exactly is achieved during training. It seems advisable to test various ways of manipulating conceptualisation strategies through explicit instruction and other

means, using both think-aloud and non-think-aloud settings so as to further explore any possible effects of verbalisation.

Altogether, the possibility cannot be entirely ruled out that verbalisation triggers metacognitive considerations that lead to conceptual divergence in a task that has often been assumed (and utilised) as a neutral indicator of basal, unconscious cognitive strategies. For this reason, the outcome of our work must necessarily remain somewhat inconclusive. Two alternative explanations are available in the assessment of our results. On the one hand, the reported complicated and often contradictory metacognitive strategies may be taken as evidence that human preferences for allocentric vs. egocentric reference systems are so basal that the attempt to draw participants' attention to them directly leads to confusion and distortion of results. A consequence of this would be to keep participants unaware of the options as much as possible in order to avoid this complication caused by the metacognitive level. On the other hand, if participants use metacognition in the tunnel task even without being asked to do so via verbalisation, questionnaires and the like, the behavioural results may be a reflection of a wide variety of underlying (at least partially rational and conscious) cognitive strategies. Our results strongly suggest that at least some of these strategies are entirely unrelated to the binary conceptual distinction of egocentric vs. allocentric reference systems known from the literature. If this is the case, the tunnel task is not a suitable tool for assessing generic human preferences for reference systems on a basal, unconscious level.

More generally, our results suggest that desktop-based VR navigation involves conceptualisations and cognitive processes that differ fundamentally from real world navigation, at least with respect to basic processes of orientation and spatial updating. This resonates with the vast amount of evidence pointing to

the supportive effects of active physical movement and idiothetic cues (Klatzky et al., 1998; Wraga et al., 2004), as well as notions of embodiment (Mallot & Basten, 2009; Wilson, 2002). How we perceive our environment, and how we interact with it, is affected fundamentally by basic sensorial information derived directly from our physical immersion in the environment (Zetzsche, Wolter, & Schill, 2008; Zetzsche et al., 2009). Accordingly, there are natural limits to what VR without actual motion can achieve as a tool for research in spatial cognition (Taube, Valerio, & Yoder, 2013). Fortunately, increasingly resources exist that allow for a simulation of embodiment in virtual space, such as virtuspheres, VR goggles, caves (Waller et al., 2007), and the like, all of which however come with challenges of their own (Fröhlich & Wachsmuth, 2013; Marsh et al., 2013). In some respects, navigating in the real world will remain unique.

Conclusion and Outlook

We addressed the verbalisability, awareness, and metacognitive strategies involved in dealing with the known orientation problems in desktop VR settings. To rule out reliance on environmental cues, we used the perceptually sparse tunnel task, a well-established paradigm that has frequently been used to identify reference system proclivities and orientation skills in humans. Our results have implications for the significance of the tunnel task in this regard, as well as the nature of spatial updating in desktop VR in general.

With respect to the tunnel task, our analysis of metacognitive strategies did not provide conclusive support for the use, in this context, of the two basal reference systems that are generally referred to as allocentric and egocentric in the literature. While the behavioural results collected in the neutral phase of our study (with a neutral instruction) were comparable to previous studies using the tunnel

task, participants failed to adopt a different reference system based on verbal instruction. Moreover, the verbal representations collected during and after the task, along with a range of questionnaire items, were not consistent with the assumption of two distinct homogeneous groups of participants whose cognitive representation differed systematically along the lines of spatial representation. Instead, the participants' explicit considerations pointed to a broad range of cognitive strategies employed to reach consistent solutions across trials.

While the alternative explanation cannot be ruled out that these results are an artefact of our study design, they nevertheless correspond to previous insights in the literature pointing to a systematic human error in spatial updating tasks in virtual environments (Klatzky et al., 1998; Riecke & Wiener, 2007). We have to conclude that, as things stand, there can be no certainty concerning what the tunnel task is actually measuring. As shown by abundant previous results reported in the literature, there is little doubt that there are systematic differences between participants categorised (based on the tunnel task) as having 'allocentric' or 'egocentric' response patterns (e.g., Goeke et al., 2015; Gramann et al., 2006, 2009, 2010). However, our present results along with the earlier findings highlight the possibility that these patterns may not be directly related to conceptual representations that are as basic to spatial cognition as allocentric vs. egocentric reference systems.

With respect to spatial updating in desktop VR in general, it appears that the transfer of the visual two-dimensional information to a conceptualised three-dimensional world poses a major cognitive challenge that frequently affects performance. Previous work suggested that these difficulties can be overcome with training and explicit feedback. Our findings highlight that untrained humans

develop their own individual strategies to achieve a consistent reference direction so as to maintain orientation in virtual space. Beyond these options (training or individual solutions), reliable experiments addressing orientation and spatial updating performance in VR would require systems as sophisticated as, for instance, the HIVE (Waller et al., 2007), which involves an immersive environment allowing for natural movement almost equivalent to real world settings. Since this is impractical for most research labs and impossible for most types of neurocognitive studies to date, desktop-based navigation studies are likely to continue to play an important role in spatial cognition research. To some extent and depending on the research goals, results from desktop VR settings can provide valuable evidence about human spatial skills (Hegarty et al., 2006). However, conclusions on orientation-related performance that are collected in less sophisticated settings will need to be treated with caution.

In future research, a more direct comparison between strategies, conceptualisations, and metacognitive considerations employed in virtual reality tasks as opposed to real world navigation is needed. Experimental designs need to be developed that are suitable for a direct alignment between both kinds of tasks, testing the predictions derived from the notion of conceptual errors, diverse cognitive strategies, and basal reference systems, as well as addressing the metacognitive effects of explicit training. Given the complexity of the real world and the many ways in which humans are perceptually immersed in it, it may never be possible to create a controlled virtual representation that adequately accounts for the full range of human cognitive processes active during navigation.

Acknowledgements

References

Allen, G.L. (2000). Principles and practices for communicating route knowledge. *Applied Cognitive Psychology,* 14, 333-359.

Avraamides, M.N., Galati, A., Pazzaglia, F., Meneghetti, C., & Denis, M. (2013). Encoding and updating spatial information presented in narratives. *The Quarterly Journal of Experimental Psychology 66*(4):642-70. doi: 10.1080/17470218.2012.712147

Benhamou, S., & Séguinot, V. (1995). How to find one's way in the labyrinth of path integration models. *Journal of Theoretical Biology, 174*, 463-466.

Boccia, M., Nemmi, F., & Guariglia, C. (2014). Neuropsychology of environmental navigation in humans: Review and meta-analysis of fMRI studies in healthy participants. Neuropsychology Review 24(2): 236–251. doi: 10.1007/s11065-014-9247-8

Burgess, N., Spiers, H. J. & Paleologou, E. (2004). Orientational manoeuvres in the dark: dissociating allocentric and egocentric influences on spatial memory. *Cognition 94*, 149-166.

Carlson, L.A. (1999). Selecting a reference frame. *Spatial Cognition and Computation,* 1 (4), 365-379.

Carlson-Radvansky, L.A. & Logan, G.D. (1997). The influence of reference frame selection on spatial template construction. In: *Journal of Memory and Language* 37, 411-437.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments*, 7, 168-178.

Chen, G., King, J.A., Burgess, N., & O'Keefe, J. (2013) How vision and movement combine in the hippocampal place code. *Proceedings of the National Academy of Sciences 110* (1) 378 - 383.

Chiu, T.-C., Gramann, K., Ko, L.-W., Duann, J.-R., Jung, T.-P. & Lin, C.-T. (2012). Alpha modulation in parietal and retrosplenial cortex correlates with navigation performance. *Psychophysiology* 49(1):43-55.

Couclelis, H. (1996). Verbal directions for way-finding: space, cognition, and language. In J. Portugali (ed.), *The construction of cognitive maps*, pp. 133-53. Dordrecht: Kluwer Academic Publishers.

Darken, R.P. & Peterson, B. (2002). Spatial orientation, wayfinding, and representation. In Mahwah, K.S. *Handbook of Virtual Environments: Design, Implementation, and Applications*, pp. 493-518. NJ, US: Lawrence Erlbaum Associates Publishers.

Denis, M., Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: an analysis of route directions in the city of Venice. *Applied Cognitive Psychology* 13:2, 145 – 174.

Dolins, F.L., Klimowicz, C., Kelley, J. & Menzel, C.R. (2014). Using virtual reality to investigate comparative spatial cognitive abilities in chimpanzees and humans. *American Journal of Primatology* 76:496–513.

Ehinger, B. V., Fischer, P., Gert, A. L., Kaufhold, L., Weber, F., Pipa, G., & König, P. (2014). Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile EEG study. *Frontiers in Human Neuroscience, 8*, 71. doi:10.3389/fnhum.2014.00071

Ericsson, K. A. (2002). Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology 16:8*, 981 – 987.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: Bradford books/MIT Press.

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1):227–233.

Fröhlich, J. & Wachsmuth, I. (2013). The visual, the auditory and the haptic – a user study on combining modalities in virtual worlds. In R. Shumaker (ed.), Virtual Augmented and Mixed

Reality. Designing and Developing Augmented and Virtual Environments, pp 159-168. Heidelberg: Springer.

Gaunet, F., Vidal, M., Kemeny, A., & Berthoz, A. (2001). Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory. *Cognitive Brain Research, 11*(3), 409–420.

Goeke, C., Kornpetpanee, S., Köster, M., Fernández-Revelles, A.B., Gramann, K., & König, P. 2015. Cultural background shapes spatial reference frame proclivity. *Scientific Reports 5, 11426*; doi: 10.1038/srep11426.

Gramann, K., Müller, H.J., Eick, E.-M., & Schönebeck, B. (2005). Evidence of separable spatial representations in a virtual navigation task. *Journal of Experimental Psychology: Human Perception and Performance 31(6),* 1199-1223.

Gramann, K., Müller, H.J., Schönebeck, B. & Debus, G. (2006). The neural basis of egocentric and allocentric reference frames in spatial navigation: Evidence from spatio-temporal coupled current density reconstruction. *Brain Research, 1118*, 116-129.

Gramann, K., el Sharkawy, J. & Deubel, H. (2009). Eye-movements during navigation in a virtual tunnel. *International Journal of Neuroscience 119*, 1755-1778.

Gramann, K., Onton, J., Riccobon, D., Mueller, H. J., Bardins, S. & Makeig, S. (2010). Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience 22*, 2836-2849.

Gugerty, L. & Rodes, W. (2007). A cognitive model of strategies for cardinal direction judgments. *Spatial Cognition and Computation 7*(2), 179–212.

Harvey, C.D., Collman, F., Dombeck, D.A. & Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature 461*:7266, 941--946. dx.doi.org/10.1038/nature08499.

Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence 34*, 151– 176.

Kant, E. (1991 [1768]). Von dem ersten Grunde des Unterschiedes der Gegenden im Raume. [On the first ground of the distinction of regions in space]. In J. van Cleve and R.E. Frederick (eds.), *The philosophy of right and left: Incongruent counterparts and the nature of space*, pp. 27-34. Dordrecht: Kluwer.

Kearns, M.J., Warren, W.H., Duchon, A.P., & Tarr, M.J. (2002). Path integration from optic flow and body senses in a homing task. *Perception 31*, 349-374. DOI:10.1068/p3311

Kelly, J. W., McNamara, T. P., Bodenheimer, B., Carr, T. H., & Rieser, J. J. (2008). The shape of human navigation: how environmental geometry is used in maintenance of spatial orientation. *Cognition, 109*(2):281-286.

Kelly, J.W. & McNamara, T.P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition, 116*(3), 409-420.

Kim, B., Lee, S., & Lee, J. (2007). Gender differences in spatial navigation. *Proceedings of World Academy of Science, Engineering and Technology (PWASET) 25*, 297-300.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In Freksa, C., Habel, C. & Wender, K. F. (eds.), *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge,* pp. 1-17. Berlin: Springer.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, 9, 293–298.

Klippel, A. & Montello, D.R. (2007). Linguistic and nonlinguistic turn direction concepts. In S. Winter, M. Duckham, L. Kulik, & B. Kuipers (Eds.), *Spatial Information Theory,* pp. 354-372. Berlin: Springer.

Levinson, S.C. (2003). *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Levinson, S.C., Kita, S., Haun, D.B.M., & Rasch, B.H. (2002). Returning the tables: language affects spatial reasoning. *Cognition 84:*2, 155-188.

Li, P. & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. *Cognition 83:3*, 265-294.

Lin, C.-T., Yang, F.-S., Chiou, T.-C., Ko, L.-W., Duann, J.-R. & Gramann, K. (2009). EEG-based spatial navigation estimation in a virtual reality driving environment. *Ninth IEEE International Conference on Bioinformatics and Bioengineering*, pp. 435-438.

Loomis, J., Klatzky, R., Golledge, R., & Philbeck, J. (1999). Human Navigation by Path Integration. In R. Golledge (Ed.), *Wayfinding Behavior: Cognitive Mapping and Other*

Spatial Processes (pp. 125–152). The Johns Hopkins University Press: Baltimore and London.

Maguire, E.A., Burgess, N. & O'Keefe, J. (1999). Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Current Opinion in Neurobiology 9*(2): 171-177.

Mallot, H.A., & Basten, K. (2009). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing* 27:1658–1670.

Marsh, W.E., Hantel, T., Zetzsche, C., & Schill, K. (2013). Is the User Trained? Assessing Performance and Cognitive Resource Demands in the Virtusphere. *3DUI 2013: IEEE 8th Symposium on 3D User Interfaces*, March 16-17, 2013, Orlando.

May, M. & Klatzky, R.L. (2000). Path integration while ignoring irrelevant movement. *Journal of Experimental Psychology: Learning, Memory and Cognition 26*(1):169–186.

Meilinger, T., Knauff, M., & Bülthoff, H.H. (2008). Working memory in wayfinding: a dual task experiment in a virtual city. *Cognitive Science 32*(4), 755-770.

Moffat, S. D., Hampson, E. , Hatzipantelis, M. (1998). Navigation in a "Virtual" Maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behaviour 19*(2): 73-87.

Montello, D. R., Waller, D., Hegarty, M & Richardson, A. E. (2004). Spatial memory of real environments, virtual environments, and maps. In G. Allen (Ed.). *Human Spatial Memory: Remembering Where* (pp. 251-285). Mahwah, NJ: Lawrence Erlbaum Associates.

Mou, W., Zhang, H., and Mcnamara, T. P. (2009). Novel-view scene recognition relies on identifying spatial reference directions. *Cognition, 111*(2):175-186.

Münzer, S., Hölscher, C. (2011). Entwicklung und Validierung eines Fragebogens zu räumlichen Strategien. *Diagnostica* 57: 3, 111–125.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.

Paillard, J. (ed.), (1991). *Brain and Space*. Oxford: Oxford Science Publications.

Picucci, L., Caffò, A.O., Bosco, A. (2011). Besides navigation accuracy: Gender differences in strategy selection and level of spatial confidence. *Journal of Environmental Psychology 31*(4): 430-438.

Plank, M., Müller, H. J., Makeig, S., Onton, J. & Gramann, K. (2009). VR as promising tool for experimental research on human spatial navigation. In Welke, S., Kolrep, H. & Rotting, M. (eds.), *Biophysiologische Interfaces in der Mensch-Maschine-Interaktion*, pp. 50-621.

Plank, M., Müller, H.J., Onton, J., Makeig, S., & Gramann, K. (2010). Human EEG correlates of spatial navigation within egocentric and allocentric reference frames. In C. Hölscher, T.F. Shipley, M. Olivetti Belardinelli, J.A. Bateman, & N.S. Newcombe (Eds.), *Spatial Cognition VII*, pp. 191-206. Berlin: Springer.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception 23*(12):1447–1455.

Riecke, B.E. (2008). Consistent left-right reversals for visual path integration in virtual reality: More than a failure to update one's heading? *Presence: Teleoperators and Virtual Environments 17* (2), 143-175.

Riecke, B.E. (2012) Are left-right hemisphere errors in point-to-origin tasks in VR caused by failure to incorporate heading changes? In C. Stachniss, K. Schill, & D. Uttal (Eds.): *Spatial Cognition 2012*, pp. 143–162. Springer, Heidelberg.

Riecke, B.E., Cunningham, D.W., & Bülthoff, H.H. (2007) Spatial updating in virtual reality: the sufficiency of visual information. *Psychological Research 71*:298–313.

Riecke, B.E., Sigurdarson, S., & Milne, A. P. (2012). Moving Through Virtual Reality Without Moving? *Cognitive Processing*, *13*(1), 293–297.

Riecke, B. E., & Wiener, J. M. (2007). Can people not tell left from right in VR? Point-to-origin studies revealed qualitative errors in visual path integration. *Proceedings of IEEE Virtual Reality 2007*, 3-10.

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2013). Learning to walk in virtual reality. ACM Transactions on Applied Perception, 10, 2, Article 11, 17 pages. DOI=http://dx.doi.org/10.1145/2465780.2465785.

Ruddle, R. A., Volkova, E., Mohler, B., & Bülthoff, H. H. (2011). The effect of landmark and body-based sensory information on route knowledge. *Memory & Cognition, 39*, 686-699. DOI=10.3758/s13421-010-0054-z.

Schönebeck, B., Thanhäuser, J., & Debus, G. (2001). Die Tunnelaufgabe: Eine Methode zur Untersuchung kognitiver Teilprozesse räumlicher Orientierungsleistungen [The tunnel task: A method for the investigation of cognitive subprocesses of spatial orientation performance]. *Zeitschrift für Experimentelle Psychologie, 4*, 339-364.

Seubert, J., Humphreys, G. W., Müller, H. J. & Gramann, K. (2008). Straight after the turn: The role of the parietal lobes for egocentric space processing. *Neurocase 14*, 204-219.

Shelton, A.L. & McNamara, T.P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology 43*, 274–310.

Shelton, A.L. & McNamara, T.P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30*:1, 158–170.

Stankiewicz, B.J., Legge, G.E., Mansfield, J.S. & Schlicht, E.J. (2006). Lost in Virtual Space: Studies in Human and Ideal Spatial Navigation. *Journal of Experimental Psychology: Human Perception and Performance 32*(3):688-704.

Talmy, L. 2000. *Toward a Cognitive Semantics*, 2 vols. Cambridge, MA: MIT Press.

Taube, J.S., Valerio, S., & Yoder, R.M. (2013). Is navigation in virtual reality with fMRI really navigation? *Journal of Cognitive Neuroscience 25*:7, 1008-1019.

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371-391.

Tenbrink, T. (2007). *Space, Time, and the Use of Language: An Investigation of Relationships*. Berlin: Mouton de Gruyter.

Tenbrink, T. (2011). Reference frames of space and time in language. *Journal of Pragmatics* 43:3, 704-722.

Tenbrink, T. (2015). Cognitive Discourse Analysis: Accessing cognitive representations and processes through language data. *Language and Cognition* 7:1, 98 – 137.

Tenbrink, T., Coventry, K.R., & Andonova, E. (2011). Spatial strategies in the description of complex configurations. *Discourse Processes* 48:237–266. DOI:10.1080/0163853X.2010.549452

Tenbrink, T., Ross, R.J., Thomas, K.E., Dethlefs, N., & Andonova, E. (2010). Route instructions in map-based human-human and human-computer dialogue: a comparative analysis. *Journal of Visual Languages and Computing* 21:5, 292–309.

Török, Á., Nguyen, T. P., Kolozsvári, O., Buchanan, R. J., & Nadasdy, Z. (2014). Reference frames in virtual spatial navigation are viewpoint dependent. *Frontiers in Human Neuroscience, 8*, 646. doi:10.3389/fnhum.2014.00646

Tversky, B., Lee, P., & Mainwaring, S. (1999). Why do speakers mix perspectives? *Spatial Cognition and Computation*, 1 (4), 399-412.

Van Der Ham, I., Faber, A., Venselaar, M., van Kreveld, M., and Loffler, M. (2015). Ecological validity of virtual environments to assess human navigation ability. *Frontiers in Psychology 6*:00637, DOI=10.3389/fpsyg.2015.00637.

Vavrečka, M. & Lhotská, L. (2009). The EEG correlates of the allocentric and the egocentric spatial reference frames processing. In Dössel, O. & Schlegel, W. C. (eds), *World Congress on Medical Physics and Biomedical Engineering*, September 7 - 12, 2009, Munich, Germany, pp. 2012-2015.

Waller, D., Loomis, J., & Steck, S. (2003). Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin & Review 10*:987 – 993.

Waller, D., Bachmann, E., Hodgson, E., & Beall, A. C. (2007). The HIVE: A Huge Immersive Virtual Environment for research in spatial cognition. *Behavior Research Methods 39*, 835-843.

Wang, R.F. (2012). Theories of spatial representations and reference frames: What can configuration errors tell us? *Psychonomic Bulletin & Review 19*.4: 575-587.

Wiener, J.M., Berthoz, A., & Wolbers, T. (2011). Dissociable cognitive mechanisms underlying human path integration. *Experimental Brain Research 208*(1):61-71.

Wiener, J.M. & Mallot. H.A. (2006). Path complexity does not impair visual path integration. *Spatial Cognition and Computation 6*(4), 333–346.

Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review* **9** (4): 625–636.

Wraga, M. Creem-Regehr, S. H. & Proffitt, D. R. (2004). Spatial updating of virtual displays during self- and display rotation. *Memory & Cognition, 32*(3):399–415.

Zetzsche, C., Galbraith, C., Wolter, J. & Schill, K. (2009). Representation of Space: Image-like or Sensorimotor? *Spatial Vision 22*:5, 409-424.

Zetzsche, C., Wolter, J., & Schill, K. (2008). Sensorimotor representation and knowledge-based reasoning for spatial exploration and localisation. *Cognitive Processing 9*, 283-297.

Appendix 1: Instruction Phase 1

In this experiment, you will travel through a virtual tunnel.

A single trial will look like this: first a small cross will appear in the centre of the screen, which you should focus on. You will then see the beginning of a virtual tunnel. Each tunnel starts with a straight segment followed by a curve and ends with another straight segment. It will look as though you are "moving" forward into the depth of the tunnel through straight and curved segments. Imagine that the first segment points straight ahead into the depth of the simulated space.

When the tunnel disappears behind you, you stop moving while the screen turns black.

[ARROW CONDITION:] Then two arrows will appear, pointing towards the starting point of the tunnel, the tunnel entrance. Your task is to decide which one of the two arrows represents the correct direction towards the tunnel entrance (as the crow flies). If it is the right arrow, please press the right mouse button; if it is the left arrow, please press the left button. Take your time to make your decision – we are not interested in how fast you respond. If you lost orientation during the journey through the tunnel, choose the arrow that you feel most likely represents the correct answer.

Please think aloud about what is happening, both while travelling through the tunnel and also when you are deciding which is the correct arrow. Say anything that comes to mind about what you are thinking and doing, no matter what it is, the way you just practiced it.

[POINTING CONDITION:] During this time, please point back over your shoulder to where the starting point of the tunnel (the tunnel entrance) should be, in a straight line (as the crow flies). Please say aloud in which direction you are pointing, back to the right over your right shoulder, or back to the left over your left shoulder. Take your time to make your decision – we are not interested in how fast you respond. If you lost orientation during the journey through the tunnel, choose the direction that you feel most likely represents the correct answer. When you are done with pointing and talking, start the next tunnel by clicking on the mouse.

Please think aloud about what is happening, both while travelling through the tunnel and also when you are deciding where to point back over which of your shoulders. Say anything that comes to mind about what you are thinking and doing, no matter what it is, the way you just practiced it.

Appendix 2: Questionnaire following phase 1

Motivation: Gramann et al. (2005, and elsewhere) suggest that the systematic distinctions between participants' responses can be explained by the fact that an egocentric perspective means conceptualising turning one's head while navigating, while an allocentric perspective means keeping one's head stable. However, at a closer look, various possible conceptualisations might lead to different kinds of consistent results. For example, it is also possible that humans do not conceive themselves as moving through the tunnel, but rather conceive the tunnel as moving through or passing by them while the body remains static (corresponding to their actual proprioception, although the tunnel is not actually moving through them). The instruction explicitly states: *"During this simulation, you are "moving" forward into the depth of the simulated space through straight and curved segments".* Although this instruction emphasises *self* rather than *tunnel* movement, there is no independent way of ensuring that this corresponds to the participants' actual conceptualisations. If participants conceived of the tunnel journey in this alternative way, this may have led to a systematic distortion of the conceptualisation of the entrance location.

Another complication concerns the direction of pointing at the end of the journey. Participants may either conceptually turn around to face the tunnel's exit and retrace the tunnel's journey to return to the starting point, or keep their current orientation, representing the arrow as extending to their back or side. The distinction can be imagined by comparison with a real-world path integration experiment, in which participants travel along several path segments including curves (perhaps blind-folded) and are then asked to indicate the homing direction by a gesture. Then they may turn around to point to the start location, rather than trying to point while remaining oriented in the direction of the previous path. In

the real world, this will not have any disruptive effects because the experimenter can directly assess the pointing arrow. However, in the virtual environment the two possibilities will have two exactly opposing results. The relevant part of the instruction in Gramann et al. (2005) is: *"At the end of the tunnel, you stand still while viewing outside of the last segment (the last segment will stay on screen for a little while). Then, after a short time, two arrows will appear, pointing towards the starting point of the tunnel, the tunnel entrance. Your task is to decide which one of the two arrows represents the correct direction towards the tunnel entrance."*

This instruction suggests that the orientation of "viewing outside" should be maintained when the arrows appear. Nevertheless participants might involuntarily infer that they should turn to look back in order to assess whether the arrow points in this same direction. Thus, in theory, both interpretations are available, although the instruction advocates "standing still". Accordingly one participant in a pilot study[5] pointed out, *"on occasion I was not sure which arrow was correct and whether it should go back down the tunnel or point direct to the entrance."*

To address the possibility of such alternative conceptualisations we administered a questionnaire asking participants explicitly about what they did, or thought they did, as follows.

1. Did you feel that you were moving through the tunnel?
2. Did you feel that the tunnel was moving past you?

---

[5] Michael Flower: Strategies for Navigation in a Virtual Tunnel. PART II NEUROSCIENCE 2005/2006. Department of Experimental Psychology, University of Cambridge, UK.

3. Did you feel that you were re-orienting your head during the curves?

4. Did you feel that you were re-orienting your body during the curves?

5. Did you feel that you were changing the direction in which you were looking, as if you were turning your head along with the curves in the tunnel?

6. Or did you feel you were looking straight on in spite of the curves in the tunnel?

7. Did you lean your body into the curves (as if driving a *bike* leaning into the turns of the tunnel)?

8. Or was it more like walking through the tunnel while seeing yourself from a bird's-eye view?

9. When you came out of the tunnel, did you keep your orientation constant?

10. Or did you turn to look at the direction you were coming from?

11. When you imagined the direction to the start, was it like pointing back over your shoulder?

12. Or was it like pointing in front and to the right or left?

13. Did you imagine travelling in the tunnel each time you did the task?

14. Did you generally decide about the correct arrow based on your sense of direction?

15. Or did you (at least sometimes) decide about the correct arrow based on logic rather than sense of direction?

Questions no. 11 and 12 were only given to participants in the *arrow* condition.

Appendix 3: Additional analyses

Response times

In addition to the main analyses reported in the main text, analysis of response times was performed to complement our findings on cognitive processes activated for the various conceptualisations and instruction types. For instance, Riecke (2012) found significantly shorter reaction times for "egocentric" (turner) responses than for "allocentric" (non-turner) responses, pointing to higher cognitive effort. We investigated response times where this was feasible, namely in the 'arrow' condition in terms of the time elapsed between presentation of the two arrows and choosing one out of two arrows by mouse click. Before performing statistical comparisons, all relevant data were tested for normality using the one-sample Kolmogorov-Smirnov Lillifors test, and histograms of each variable were examined visually. Data were skewed to the left, and were therefore transformed by '1/ ln (response time)'. After transformation, all variables conformed to normality both when testing the actual data and when testing the residuals generated by parametric models.

To test whether the response time of behavioural preference groups differed across conditions, we used linear mixed models with group, condition, and group*condition interaction as fixed factors; participant, tunnel sequence, and sequence of conditions (neutral-ball-bike vs neutral-bike-ball) were used as random factors.

We used Systat version 12.0 for all statistical comparisons, and significance is reported using both the conventional $\alpha = 0.05$ and post-hoc Bonferroni test of effect correction. Transformed data are graphically represented as back-transformed means ± SEM; because of the back transformation, the

graphical representation of the standard errors is not symmetrical.

Response time differed significantly between the three preference groups (linear mixed model: $F_{2,1060}$-ratio $= 80.176, p < 0.0001$, fit statistics: AIC(corrected) : -5768,090; BIC: -5758.323), and was also significantly different between the three conditions ($F_{2,1060}$-ratio $= 3.241, p = 0.04$), but there was no significant interaction between the two factors (i.e. reference system * condition: $F_{2,1060}$-ratio $= 0.671, p = 0.612$, Figure 6).

The group categorised as "ego" (i.e. who showed response patterns corresponding to the expectations for an egocentric reference system) always responded the fastest. In all three conditions "ego" responses were significantly faster as compared to the "no preference" group (neutral: t=7.385, p $\ll 0.001$, 95% confidence interval (=CI) ranges from 0.014 (upper) to 0.006 (lower); bike: t $= 7.813$, p $\ll 0.001$, 95% CI ranges from 0.015 to 0.006; ball: t $= 6.785$; p $\ll$ 0.001, 95% CI ranges from 0.014 to 0.005), and there was a tendency for "ego" responses to be faster than "allo" responses, but only in the neutral condition (Bonferroni posthoc test of effects Group*Condition: t $= -3.230$, p $= 0.096$, 95% CI ranges from -0.010 to 0.000). Participants without a clear preference showed the longest response times, with responses significantly slower than in the "allo" group during the ball condition (t $= 4.029$; p $\ll 0.002$, 95% CI ranges from 0.0013 to 0.001) and the bike condition (t $= 4.958$; p $\ll 0.001$, 95% CI ranges from 0.0014 to 0.003).
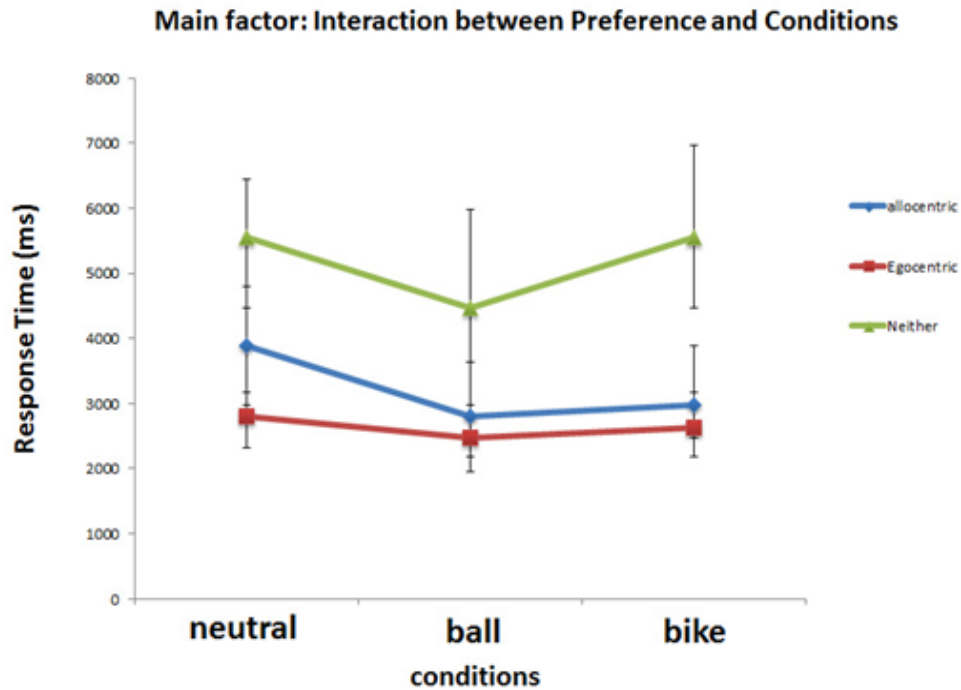
Figure 6: Preference across conditions (showing interaction for groups and condition). Back-transformed means ± SEM for all preference categories. Response time (ms) = Time between presenting arrows on the screen and time clicking on. Significant comparisons of main conditions: ego – allo: p < 0.001, ego – neither: p<0.001, allo –neither: p <0.001, neutral – bike: < 0.05.

Analysis of possible gender differences

Because of earlier results on differences between men and women in spatial orientation and strategies (e.g. Goeke et al. 2015; Hegarty et al., 2006; Moffat et al., 1998; Kim et al., 2007; Picucci et al., 2011), we addressed the distribution of response strategies according to gender. We could not detect any gender bias in our data (x^2 (df =1, N=45), p=0.219). The same proportion of male (38.5%) and female (34.4%) participants changed or lost their initial preference (Figure 7). Both sexes showed both preferences (egocentric: 13 females, 9 males; allocentric:

10 females, 2 males) or no preference (9 females, 2 males); in both sexes
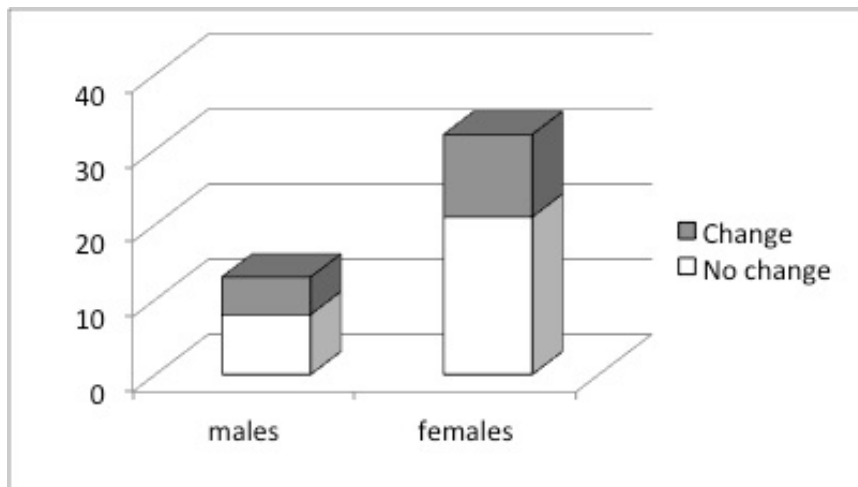egocentric preference was dominant.

Figure 7: Distribution of behavioural results according to gender. Solid white:
number of participants whose preferences were not changed; solid gray: number
of participants whose preferences changed after instructions.