Embedding presence-related terminology in a logical and functional model

Markus von der Heyde & Bernhard E. Riecke

Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Keywords: spatial presence, human spatial cognition, logic, framework, terminology

Summary

In this paper, we introduce first steps towards a logically consistent framework describing and relating items concerning the phenomena of spatial presence, spatial orientation, and spatial updating. Spatial presence can be regarded as the consistent feeling of being in a specific spatial context, and intuitively knowing where one is with respect to the immediate surround. The core idea is to try to understand presence-related issues by analyzing their logical and functional relations. This is done by determining necessary and/or sufficient conditions between related items. This eventually leads to a set of necessary prerequisites and sufficient conditions for spatial presence, spatial orientation, and spatial updating. More specifically, the logical structure of our framework allows for novel ways of quantifying spatial presence and spatial updating.

1 Introduction

This paper presents first steps towards a logically consistent framework describing and relating items associated to spatial presence, spatial orientation, and spatial updating. "Spatial presence" can be regarded as the consistent "gut" feeling of being in a specific spatial context, and intuitively and spontaneously knowing where one is with respect to the immediate surround. Hence, spatial presence might be a critical factor for achieving and understanding spatial updating and consequently also for quick and intuitive spatial orientation. The main goal is to understand underlying processes and mutual dependencies by analyzing their logical relations. Providing a coherent representation for the large number of experimental paradigms and results can furthermore allow for a unifying “big picture” that might help to structure and clarify our reasoning and discussions. Last but not least, it can suggest novel experiments and experimental paradigms, allow for testable predictions, and stimulate the scientific discussion. More specifically, the logical structure of our framework allows for novel ways of quantifying spatial presence and spatial updating.

The core idea of the framework is to try to understand issues and terms related to spatial orientation, spatial updating, and spatial presence by analyzing their logical and functional relations. This is done by trying to determine a set of necessary prerequisites and sufficient conditions for spatial orientation, spatial updating, and spatial presence. For example, it is evident that ego-motion perception cannot occur without some kind of motion perception. That is, intact ego-motion perception seems to be logically dependent on intact motion perception. Conversely, if we observe intact ego-motion perception, we can conclude that motion perception must also be intact, which can be represented as "*ego-motion perception* $\Rightarrow$ *motion perception*" using standard logical notation (see Table 1).

Name	Statement	Operator	Meaning of statement
simple statements			
assertion	A		A is true
negation	$\neg A$	not	A is false
compound statements and sentential connectives			
disjunction	$A \vee B$	or	either A is true, or B is true, or both
conjunction	$A \wedge B$	and	both A and B are true
implication (conditional)	$A \Rightarrow B$	if ..., then	if A is true, then B is true
equivalence (biconditional)	$A \iff B$	if and only if ..., then	A and B are either both true or both false

Table 1: Operators and statements as used in propositional logic.

The framework in its current state is not intended as a final model describing and connecting all related issues, but rather as a working model useful for understanding and analyzing what is happening in certain spatial situations or experiments. It covers aspects ranging from spatial perception through allocentric and egocentric spatial memory up to spatial behavior.

In the following, the framework will first be introduced by describing each item briefly, categorizing it, and stating its hypothesized functional connections. We will continue by discussing some implications for the quantification of spatial updating and spatial presence, and by hypothesizing about further logical connections. That is, this framework will be used to generate hypotheses which can guide future research and can be experimentally tested.

The framework is graphically represented in Figure 1 and will be introduced in detail below. It covers on the vertical axis items ranging from low-level processes like spatial perception at the bottom to high-level processes like spatial behavior at the top. On the horizontal axis, the range spans from reflexive to cognitive control of behavior. This model is built on experimental evidence as well as working hypotheses and tries to link spatial perception to spatial behavior in two logically connected lines of thought: The static or landmark-based approach (right branch of Figure 1) and the dynamic extension based on motion perception and path integration, which contains additional temporal aspects of perception, memory, and action (left branch of Figure 1).

Having determined the logical connections between the functionally connected items, we will hypothesize further connections that are plausible and helpful in interpreting experimental results, but not yet well-grounded on experimental data. These hypothesized connections, however, suggest novel ways of quantifying spatial updating, spatial presence, and immersion by measuring the adjacent, logically related items of the framework. An exhaustive analysis would unfortunately go beyond the scope of this paper.

Ideally, the final version of this framework should describe the **functional and logical relationships** between all related terms. As a first step, all terms introduced in this framework are grouped by their coarse classification into GOAL/DESIRED PROPERTY, DATA, and PROCESS. The logical connections (arrows) between terms are meant to be understood in the mathematical sense, and we use the syntax from propositional logic as summarized in Table 1. Note that if A implies B, this is equivalent to saying that non-B implies non-A ($A \Rightarrow B \iff \neg B \Rightarrow \neg A$). A is therefore a *sufficient* but not a *necessary* prerequisite for B. This is tantamount to saying that B is a *necessary* but *not sufficient* prerequisite for A (contraposition). Please note that the **information flow** is in most cases in the opposite direction, i.e, from B to A. That is, B is typically “more general” and does include (in the mathematical sense) the more specific A. The individual items, however, are not meant to be understood as simple yes-or-no decisions, such as "either spatial updating works, or else it does not". As human spatial orientation is like most mental processes highly complex and error-tolerant, this would oversimplify things. Rather, we would like to propose a more qualitative interpretation of the logical connections for this framework, much like a fuzzy logic approach. In this manner, $A \Rightarrow B \iff \neg B \Rightarrow \neg A$ would imply that, e.g., "if B is impaired, so is A", or "if A works well, so does B". Furthermore, "if B does not work or exist at all, A is also substantially impaired or defunct".

We will start by introducing global goal definitions and desired system properties that most researches from the field will most likely agree upon. On that basis, the logical chain from *spatial perception* up to *spatial behavior* will be sketched[1]. Finally, some debatable hypotheses are put forward to be discussed in a larger context. We are confident that this framework will on the one hand initiate fruitful scientific discussions about the functional dependencies between related terms, and on the other hand help in developing novel methods for measuring spatial updating and spatial presence. Furthermore, analyzing experimental results in its context might allow for a deeper understanding of the underlying processes and could help to adapt and refine the framework.

1.1 Overall goal guiding this framework: *Spatial Orientation*

All moving organisms have the goal of finding for example food, shelter and partners from the same species, which are all tasks critically relying on spatial orientation. Hence, our framework has to follow this global aim as a critical boundary condition for the functioning of spatial behavior in general. Homing is one prominent example from the literature. The ability to find the way back to the origin of an excursion can be found in most moving species (from ants to humans) (Klatzky, Loomis, & Golledge, 1997; Maurer & Séguinot, 1995; Mittelstaedt & Mittelstaedt, 1982). We assume that this ability to find one's way around (*spatial orientation*) is a sufficient motivation for spatial learning.

1.2 Additional goals guiding this framework: *Consistency* and *Continuity*

Perception is in many respects *continuous* in space and time. Furthermore, the different sensory modalities are typically found to contribute to one *consistent* percept of the world. That is, the relation between oneself and the surrounding world is spatio-temporally continuous and consistent. Unless we navigate computer-generated worlds, we are neither teleported in space or time (discontinuity) nor do we perceive ourselves to be at several places at the same time (inconsistency). Both consistency and continuity should therefore be additional desired properties of our framework. In general, organism might also use this continuity of perception to deduce high spatio-temporal correlations in order to statistically learn properties of the world (Bayesian approach). Hence, it seems plausible to include both *consistency* and *continuity* in the framework as prerequisites we can count on.

2 Framework

Overview In the following, we will try to guide the reader sequentially through this model in a bottom-up manner: We will start with the most fundamental processes and data structures and gradually work our way up until we have all the main ingredients enabling good spatial orientation, which is our overall guiding goal. After briefly describing and categorizing each term, we will state the

[1]To avoid potential confusion, items of the framework will be set in italics.

most relevant logical and/or functional connections to the aforementioned terms. Figure 1 shows the complete overview. As the complete model is rather complex, we advise the reader to focus on the terms and relations we have introduced up to that point. We will start be describing the dynamic left branch of the framework.

Spatial Perception [PROCESS] Physical stimuli of the surround can be perceived in multiple dimensions and modalities. We group here all kinds of perception, regardless of their sensory modality (e.g. visual, auditory, haptic, kinesthetic etc.), into *spatial perception* if the percept covers some spatial aspect of the stimulus. For the purpose of the overall framework, we do not need or intend to refine this rather coarse and low-level definition of *spatial perception*. Its main purpose is to constitute the basis and necessary prerequisite for the whole framework.

Motion Perception [PROCESS] When we see temporal changes in spatial stimuli, we can have the percept of object-motion. Examples include the perception of visual motion from optic flow using simple Reichardt-detectors. *Motion perception* depends logically on *spatial perception* in the sense that we cannot perceive any motion if we cannot perceive spatial cues: (*motion perception* $\Rightarrow$ *spatial perception)* $\Longleftrightarrow$ ($\neg$ *spatial perception* $\Rightarrow$ $\neg$ *motion perception*). Furthermore, only if *continuous* changes in space occur over time can we perceive motion. (Under certain conditions, however, very small spatial jumps can be perceived (interpreted) as “apparent motion”).

Ego-Motion Perception [PROCESS] If perceived motion is interpreted as self-motion of the observer and not just as a motion of some entity relative to the (stationary) world or observer, we call this phenomenon *ego-motion perception*. The classical example for this is visually induced vection (feeling of ego-motion) that can be achieved by presenting a rotating optic flow pattern in an optic drum for several seconds (see, e.g. Dichgans & Brandt, 1978; Mach, 1922). Obviously, without perceiving any motion in any modality, one would not feel any ego-motion. Therefore, we can state: *ego-motion perception* $\Rightarrow$ *motion perception*.

Egocentric Reference Frame [DATA] An *egocentric reference frame* can be understood as a mental model of the world in our head, as seen from the first-person perspective. This mental model is thought to contain at least the immediate surround or scene. Even if this mental model does not explicitly exist, it makes nevertheless sense to store somewhere the existing knowledge of the immediate surround from the egocentric perspective, as this is the perspective from which we interact with the environment by grasping objects, moving towards them etc.

Incoming information from several modalities can code multiple *egocentric reference frames*. The most prominent one, on which the majority of sensory inputs agree, is called the primary *egocentric reference frame*, which can be in conflict with additional (secondary) reference frames indicated by other sensory input. In most VR applications, for example, at least two competing *egocentric reference frames* are present: On the one hand, the intended or simulated one, that is, the reference frame of the virtual environment. On the other hand, participants are embedded in the physical reference frame of the simulation room. Hence, the *egocentric reference frames* depend critically on *spatial perception*: *egocentric reference frame* $\Rightarrow$ *spatial perception,* because without (typically multi-modal) perception we would not have the basis for the perceived egocentric perspective. This connection is not further specified here, but is supposed to cover the dependency on multiple modalities.

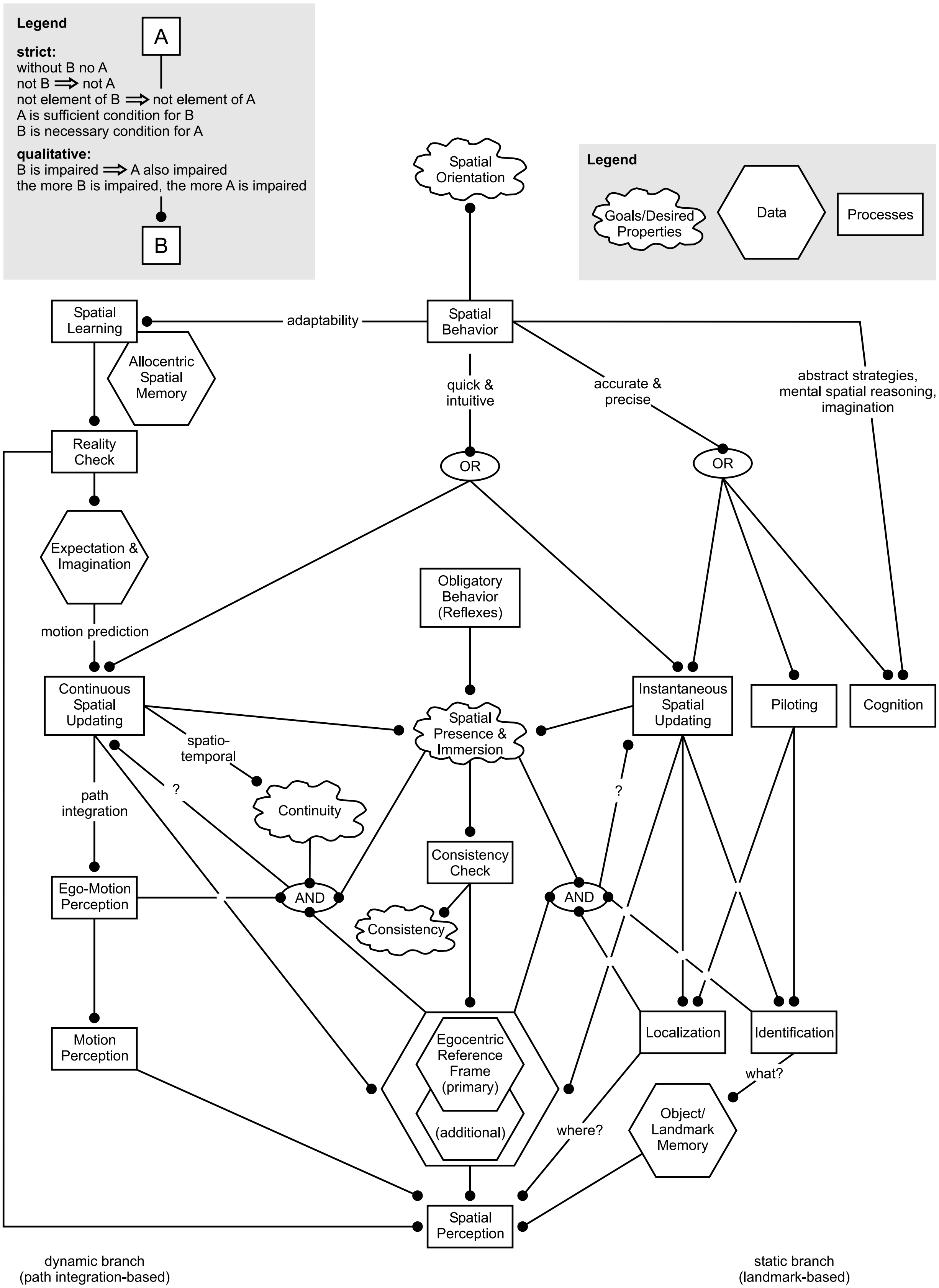

Figure 1: Conceptual framework, as described in the text.

Consistency [GOAL/DESIRED PROPERTY] As stated in the introduction, we propose the overall goal of a spatio-temporally consistent relation between oneself and the surround.

Consistency Check [PROCESS] In connection to an existing *egocentric reference frame* and the overall goal of *consistency*, we propose the notion of a *consistency check*: At any moment, we should have one and only one consistent mental reference frame that defines our perceived ego-position in the world. That is, both an *egocentric reference frame* and *consistency* are necessary prerequisites for a *consistency check*, as without the overall goal of *consistency* and the existence of the data structure (egocentric reference frame) there would be no process checking for consistency: *Consistency check* ⇒ *egocentric reference frames* and *consistency check* ⇒ *consistency*. This *consistency check* is related to spatial presence, as we will see below.

Spatial Presence & Immersion [GOAL/DESIRED PROPERTY] *Spatial presence* can be regarded as the consistent feeling of being in a specific spatial context, and intuitively knowing where one is with respect to the immediate surround. *Immersion*, on the other hand, could be seen as the subjective feeling of being fully drawn into that spatial context. For the sake of simplicity, however, we do not distinguish between *spatial presence* and *immersion* in this framework and therefore put them into the same box in Figure 1.

Spatial presence & immersion requires the functioning of the *consistency check* of the primary *egocentric reference frame*: If we do not agree on one single (consistent) reference frame at a time, we are not fully immersed in the spatial situation (*spatial presence & immersion* ⇒ *consistency check*). Furthermore, without the knowledge of some egocentric spatial reference frame, we would obviously not be able to immerse into anything (*spatial presence & immersion* ⇒ *consistency check* ⇒ *egocentric reference frame*).

In virtual reality applications, we can perceive high *spatial presence & immersion* only if the simulated world is consistently accepted as the only reference frame. That is, in order to be fully immersed in the simulated world, one has to "forget" about the physical reference frame of the simulator (which would constitute a second, conflicting reference frame) or else the consistency check would not be fulfilled.

Obligatory Behavior (Reflexes) [PROCESS] For the first time in this paper we would like to introduce something which can actually be measured directly: the process of *obligatory behavior (reflexes)*, which cannot easily be voluntarily suppressed. For example, people with a fear of heights cannot help but be afraid if they stand close to an abyss. The same is true for fear of flight or fear of narrow spaces. For example, people with arachnophobia (fear of spiders) might not like to look at pictures of spiders, but that would most certainly not elicit any spatial response like running away. Only if the spider is in a spatial context and crawling towards them would they react spatially by trying to escape. In sum, *obligatory behavior* in this context is meant to refer to compulsory behavior that is elicited by a spatial context or situation. That is, it would seem most natural for us to dodge away if an unknown object flies at high speed towards our head.

One critical point in those situations is to believe the actual danger - that is to feel immersed: *Obligatory behavior* ⇒ *spatial presence & immersion*. Without the immersion the obligatory response is not elicited. This means for example that people with fear of height do not feel that fear if the are not fully immersed into the situation of, e.g., standing at the edge of a cliff. Conversely, if we observe intact reflexive behavior, the participant was spatially present and immersed. That is, *spatial presence & immersion* can be quantified indirectly be measuring obligatory spatial behavior.

It is to be noted, however, that for phobic people, merely *imagining* a fear-inducing situation can elicit all characteristics of a panic attack. Here, we would argue that they feel fully immersed in their *imagined* environment. This suggests that in extreme cases, our framework can operate on purely imagined space, too.

Continuity [GOAL/DESIRED PROPERTY] As mentioned in the introduction, one of the overall desired properties of perception is the apparent *continuity* of the perceived stimulus in particular and the world in general (at least for self-initiated ego-motions). We propose that this property can be seen as the guiding goal of the overall system.

Continuous Spatial Updating [PROCESS] When we move, all spatial relationships between ourselves and the environment change. Nonetheless, we feel immersed in the current surround and naturally experience spatial presence. Hence, some robust process needs to continuously update these self-to-world relationships as we move: This *continuous spatial updating* process refers to the incremental transformation of our *egocentric reference frame* based on relative positional and rotational information. That is, it operates without any landmarks, by incrementally updating the *egocentric reference frame* using perceived velocity, acceleration, and relative displacements. Blindfolded walking with ears muffled is the stereotypical example for this process.

More specifically, convincing *ego-motion perception, spatial presence & immersion, egocentric reference frame* as well as *continuity* are necessary (but not sufficient) prerequisites for *continuous spatial updating*. Simply put, we cannot update any position if we cannot perceive its changes (*continuous spatial updating* $\Rightarrow$ *ego-motion perception*). This part is often understood as path integration. Furthermore, we cannot update to a new location in space if we are not already spatially present at any location beforehand and possess a corresponding *egocentric reference frame* - otherwise there would be nothing to update (*continuous spatial updating* $\Rightarrow$ *spatial presence & immersion* and *continuous spatial updating* $\Rightarrow$ *egocentric reference frame*). Finally, a continuous update is only possible if the sequential changes are continuous in time and space (*continuous spatial updating* $\Rightarrow$ *continuity*). Without *continuous spatial updating*, the *egocentric spatial reference* frame would become increasingly misaligned, which would eventually lead to a discontinuity the next time *instantaneous spatial updating* re-aligned the *egocentric reference frame*.

Expectation & Imagination [DATA] Executing all possible behaviors in order to test their potential outcome is very inefficient. A more efficient cognitive approach would be to predict and *imagine* what we would perceive if we would perform a certain action. In this manner, we generate an *expectation* of what we should perceive if we had actually performed that action. Moving in space is in this sense very predictable by the organism and therefore we hypothesize: *expectation & imagination* $\Rightarrow$ *continuous spatial updating* in the sense that without *continuous spatial updating* or imagined *continuous spatial updating* one would not be able to predict the changed percept of the world through *expectation & imagination*.

Reality Check [PROCESS] Once we have an *expectation* of what we ought to perceive for a given action or motion, we can compare the actual percept to the predicted one. Naturally we describe this by: *reality check* ⇒ *expectation & imagination*, since we have to have some expected outcome to check the actual outcome against. If they match, everything is fine, and the reality check process will probably not come to consciousness or require any attention. If not, this might require some attention or action, that is, we might for example want to look again to make sure that everything is okay or allocate some cognitive resources to resolve the mismatch or act appropriately. An example might illustrate this.

If we walk and slip on ice our reflexes correct the mismatch between expected spatial behavior (walking) and the real danger of falling. In this situation the currently required position of our limbs would disagree with the motor commands generated for walking and the *reality check* (implemented in this example as hardwired) would correct for the occurring error. Therefore we propose: *reality check* ⇒ *spatial perception.*

This double-checking is the obvious connection to *spatial perception*. One rather far-fetched hypothesis would be to propose: *spatial perception* ⇒ *reality check*, implying that we can only perceive *if* we expect and maybe even *what* we expect. Naturally, this cannot be sufficient to explain perception, but it sheds a new light on change blindness results - even considerable changes in our surround go unnoticed if we do not expect them to occur.

Spatial Learning [PROCESS] If the *reality check* encounters an unexpected event, there might be something we could learn from this discrepancy. Since the organism cannot predict everything right from the start, its internal prediction model needs to be developed though learning. As we are concerned here with *spatial* behavior only, we would like to constrain ourselves here to *spatial learning*. *Spatial learning* can be seen as the process of building up and modifying spatial knowledge, that is, the process which operates over time on the *allocentric spatial memory* (see below). We hypothesize the logical connection to be: *spatial learning* ⇒ *expectation & imagination*. Many learning algorithms as understood in the neurosciences require an error signal, which can be defined as the difference between stimulus and prediction. Hence, without *expectation & imagination* we would not be able to perform *spatial learning*.

Allocentric Spatial Memory [DATA] Through *spatial learning*, we can acquire *allocentric spatial memory*, e.g., spatial memory in the form of a "cognitive map" allowing for novel shortcuts (see, e.g. Poucet, 1993; Tolman, 1948; Trullier, Wiener, Berthoz, & Meyer, 1997). Therefore, *spatial learning* can be seen as an ongoing process operating on the knowledge stored in *allocentric spatial memory*. We would like to state that learning and memory are tightly coupled, require one another and thus cannot be strictly separated. We express this as a direct coupling (equivalence on the logical, but of course not on a functional level) between *spatial learning* and *allocentric spatial memory*.

Object/Landmark Memory [DATA] Having described the dynamic, path integration-based left branch of the framework, we will now discuss the more static, landmark-based right branch. *Object/landmark memory*, which is the most basic data structure in our framework, contains knowledge about objects and landmarks *without* their spatial context or relationships. This is the data structure needed for, e.g., object recognition (see below). We do not assume any preferred storage format, but presume that we cannot built up any knowledge of spatially extended objects or landmarks without some kind of *spatial perception* (*object and landmark memory* ⇒ *spatial perception*).

Identification [Process] Having the ability to store knowledge about objects and landmarks, it makes sense to demand some recognition process which can identify objects, in order to label them as individuals and potentially recognized them later. This *identification* process can be seen as the "what path" in the perception model by Milner & Goodale (1995). The logical relation here is as follows: *identification* $\Rightarrow$ *object and landmark memory*. In other words, if one cannot remember any objects, it should not be possible to recognize and to identify them later.

Localization [Process] As soon as we perceive anything spatially, we can localize it even without necessarily being able to identify it. That is, the *localization* process does not assume any attribution of identity. One could compare this to the "where path" in Milner and Goodale's model of perception (Milner & Goodale, 1995). The logical relation between these two terms is: *localization* $\Rightarrow$ *spatial perception*. In other words, without any *spatial perception* we could have no *localization* process (i.e., $\neg$ *spatial perception* $\Rightarrow$ $\neg$ *localization*).

Instantaneous Spatial Updating [Process] In order to convincingly explain recent results from spatial updating experiments by Riecke, von der Heyde, & Bülthoff (2002a) in the context of this framework, we need to refine our concept of spatial updating. That is, we would like to extend the prevailing definition of spatial updating by distinguishing between the classical *continuous spatial updating* known from the blindfolded spatial updating literature and the hereby introduced "*instantaneous spatial updating*".

Spatial updating in general can be thought of as the spatial transformation process operating on the egocentric mental spatial representation. In this manner, *continuous spatial updating* is the process of continuously and incrementally (smoothly) transforming our egocentric reference frame, where as *instantaneous spatial updating* is the immediate, and if need be discontinuous ("jump"- or "teleport"-like) process. Where as the continuous process might have some limitations in terms of transformation speed (e.g., a limited mental rotation speed), the instantaneous one probably does not.

As continuous spatial updating alone is based on path integration and leads to exponentially increasing alignment errors over time, it seems sensible to propose a second process that can re-anchor the potentially misaligned mental reference frame to the physical surround. We would like to introduce the term *instantaneous spatial updating* to refer to this process. To give an example, imagine the following: You are at home at night when the main fuse blows. You will have to walk around in darkness until you manage to find the fuse box or some light source. When walking around in complete darkness, we become increasingly uncertain about our current ego-position. That is, we still have some intuitive feeling of where we are, but we would not bet much on the exact location. The situation changes as soon as we can perceive the location of a known landmark. This instantaneous position fixing could occur via different sensory modalities: Auditorily, for example the phone could be ringing. Haptically, we might touch or run into the kitchen table. Visually, somebody else might already have replaced the fuse, or lightning might have lit the room for a fraction of a second. That is, any clearly identifiable spatial cue (landmark) could re-anchor our mental reference frame instantaneously, without much cognitive effort or time needed. This process of re-aligning or re-anchoring the mental reference frame to the surround is what we refer to as *instantaneous spatial updating*.

When locomoting under full-cue conditions, this instantaneous spatial updating probably occurs at any instance in time and is thus indistinguishable from continuous spatial updating, as both processes are in close agreement and complement each other. Moreover, they can be considered as a mutual back-up system for the case that one of them fails or does not receive sufficient information.

In sum, *instantaneous spatial updating* refers to the process of re-aligning or re-anchoring the mental spatial reference frame to the surround using position-fixing via landmarks (*instantaneous spatial updating* ⇒ *egocentric reference frame*). This process can be triggered by, for example, haptic, auditory, and, probably most frequently, visual landmarks. *Instantaneous spatial updating* is thus critically depending both on the *localization* and *identification* process: *Instantaneous spatial updating* ⇒ *localization* process means that it would not make sense to re-anchor the mental reference frame if we were not sure about the exact coordinates to use. Moreover, *instantaneous spatial updating* ⇒ *identification* means that it would not make sense to re-anchor the mental reference frame if we could not recognize anything familiar that told us where we were. Furthermore, we propose that *spatial presence & immersion* is a necessary prerequisite for automatically triggering *instantaneous spatial updating*, just as it was for *continuous spatial updating* (*instantaneous spatial updating* ⇒ *spatial presence & immersion*).

Piloting [PROCESS] Position- or recognition-based navigation (also called *piloting*) uses exteroceptive information to determine one's current position and orientation. Such information sources include visible, audible or otherwise localizable and identifiable reference points, so-called landmarks (i.e., distinct, stationary, and salient objects or cues). This implies *piloting* ⇒ *localization* and *piloting* ⇒ *identification*. Many studies have demonstrated the usage and usability of different types of landmarks for navigation purposes, (see Golledge (1999), Hunt & Waller (1999) for an extensive review). Piloting allows for correction of errors in perceived position and orientation through reference points (position fixing) and is thus well-suited for large-scale navigation. Piloting mechanisms often used include scene matching or recognition-triggered responses. Note that no aligned egocentric reference frame is needed for piloting. Furthermore, note that no higher cognitive processes are needed for piloting, as even simple robots can use for example snapshot-based piloting for navigation (Franz, Schölkopf, Mallot, & Bülthoff, 1998).

Spatial Orientation [GOAL/DESIRED PROPERTY] The main overall goal of the organism is in this context, as stated above, proper *spatial orientation*, which is essentially the ability to easily find one's way around.

Spatial Behavior [PROCESS] Last but not least, we seem to have all basic ingredients to define *spatial behavior* as behavior performed in space and time and at the same time relying on spatial knowledge about the world.

First of all, it seems plausible to assume *spatial behavior* ⇒ *spatial learning*: Without learning spatial knowledge, we would not be able to adapt to new situations and find our way around in a novel or changing environment. That is, we propose that *spatial learning* is required for the **adaptability** of *spatial behavior*.

As spatial behavior (especially in animals) is typically quick and intuitive, many of the required computational processes need to be largely automated. Hence, we propose that automatic spatial updating is a necessary prerequisite for **quick** and **intuitive** *spatial behavior*. Hence, we propose that quick and intuitive *spatial behavior* ⇒ *continuous spatial updating* or *spatial behavior* ⇒ *instantaneous spatial updating*. That is, quick and intuitive *spatial behavior* should not be possible without either *continuous spatial updating* or *instantaneous spatial updating* or both being operational[2]. For the

[2] As both continuous and instantaneous spatial updating logically imply *spatial presence & immersion*, we hereby

consistency of this model, we would like to exclude for the time being behavior that can be modeled by simple direct coupling of perception and action, without any spatial knowledge (e.g., Braitenberg vehicles (Braitenberg, 1984)). Instead, we do limit our view of *spatial behavior* being depended on good *spatial orientation*. Without *spatial orientation* we are not able to perform the required *spatial behavior* (*spatial behavior* ⇒ *spatial orientation*). Consequently, *spatial behavior* can be used to measure and evaluate the successful *spatial orientation* in psychological experiments.

Obviously enough, spatial behavior should be most accurate and precise if we can recognize and localize unique reference points. As *instantaneous spatial updating* as well as *piloting* are the two processes relying on the *localization* and *identification* of such landmarks, we propose that at least one of them has to work for us to have accurate and precise spatial behavior. Hence, we propose that **accurate** and **precise** *spatial behavior* ⇒ *instantaneous spatial updating* or *spatial behavior* ⇒ *piloting*.

Having identified specific items that are required for different aspects of *spatial behavior* (accurate & precise behavior, adaptability, and quick & intuitive behavior), we are enabled to analyze spatial or experimental situations accordingly: If the observed spatial behavior is for example accurate and precise, but response times are long and participants report not having much of an intuitive spatial orientation, we could conclude that *piloting* (the landmark-based static right branch of the framework) is intact, whereas *continuous spatial updating* as well as *instantaneous spatial updating* are probably largely impaired. This might in turn, for example, be due to the lack of convincing *spatial presence & immersion.*

Conversely, if the observed spatial behavior is quick & intuitive but lacks accuracy and precision, we would argue that automatic *continuous spatial updating* was working, but neither *instantaneous spatial updating* nor *piloting* were intact. Thus, the central and left dynamic part seems to be intact, where as the landmarks-based right branch is not. Examples for this case include blindfolded walking, getting lost in deep forest, and of course visually induced vection in an optic drum. Note that sensory cues that might allow for *continuous spatial updating* include vestibular cues (accelerations), proprioceptive cues (e.g., from walking), but also visual or auditory from optic or acoustic flow, respectively.

2.1 Where does cognition fit into the model?

So far have attempted to lay out a consistent framework based on functional and logical connections between related items. The contribution of higher cognitive processes or strategies has not so far been taken into consideration. Moreover, especially the lower part of the framework seems to be largely beyond conscious control: For example, even if we might consciously decide to do so, it is virtually impossible to influence *identification* (not recognize your friend's face) or *ego-motion perception* (consciously elicit the convincing sensation of ego-motion).

So where does cognition fit into this model? By its very nature, cognition is flexible and versatile and consequently cannot simply be represented as one box logically dependent on other boxes. Rather, cognition might be considered as an optional process that can be resorted to if the partly automated framework fails or does not allow for the desired spatial behavior. That is, we have conscious access

indirectly claim that *spatial behavior* ⇒ *spatial presence & immersion*. In other words, when we do not feel ourselves at a specific location and orientation, we cannot interact with the world in a natural and effortless manner. Hence, we proposed indirectly that *spatial presence & immersion* are required for quick and intuitive *spatial behavior*.

to for example the lower items of the framework (*motion perception*, *localization*, and *identification*), even though we cannot consciously control them. That is, we can for example consciously question *motion perception* to cognitively derive the simulated displacement, even though we might not perceive any ego-motion. We are, however, unable to use this abstract knowledge about the simulated turning angle to intentionally evoke the percept of convincing ego-motion. That is, the lower items in the framework can be queried, but are nevertheless to a large degree cognitively impenetrable.

Cognition [PROCESS] Ultimately, this leads to a forth connection to *spatial behavior*: Cognition can be used to develop for example novel strategies to solve a complex navigation problem, or to use mental spatial reasoning to derive the desired spatial behavior. Hence, *cognition* can be considered a necessary condition for *spatial behavior* based on non-automated **abstract strategies** and **mental spatial reasoning**. This can be represented in the framework as *spatial behavior* $\Rightarrow$ *cognition*. Due to the inherent flexibility of *cognition*, however, there are no other fixed links to *cognition*. Rather, cognition can be used to flexibly query the desired information from most or maybe even all of the other items of the framework. Hence, if we observe spatial behavior that is neither quick & intuitive nor very accurate & precise, we could argue that the behavior might have been based on abstract cognitive strategies. As mental geometric reasoning can lead to quite accurate and precise spatial behavior (see, e.g., Riecke, van Veen, & Bülthoff (2002)), we propose *cognition* as a third possibility for achieving accurate and precise *spatial behavior* (apart from *instantaneous spatial updating* and *piloting*): **Accurate** and **precise** *spatial behavior* $\Rightarrow$ *cognition*.

2.2 Ways to measure spatial presence and immersion

Until very recently, quantifying presence and immersion has been typically attempted using highly subjective and introspective methods like questionnaires (Hendrix & Barfield, 1996a, 1996b; Ijsselsteijn, de Ridder, Freeman, Avons, & Bouwhuis, 2001; Lessiter, Freeman, Keogh, & Davidoff, 2001; Schloerb, 1995; Schubert, Friedmann, & Regenbrecht, 2001; Witmer & Singer, 1998). These methods were an important first step towards understanding the nature and relevance of presence and immersion for many applications. In the following, we would like to sketch novel quantification methods that rely not on introspection but rather on psychophysical measures, thus complementing the existing methodologies and allowing for more robust and reliable measures.

Having embedded *spatial presence & immersion* into a functional framework allows us to devise new quantification methods by either measuring all necessary prerequisites or, even more elegantly, measuring the sufficient conditions. As we have seen in the previous section, spatial presence is embedded into a collection of processes with useful and testable properties. We found three sufficient but not necessary prerequisites of spatial presence: *continuous spatial updating*, *instantaneous spatial updating*, and *obligatory behavior*. In addition, we have one necessary, but not sufficient, prerequisite (*consistency check*). Having laid out the logical framework, we can now use this prerequisite to measure presence: The mismatch between the primary egocentric reference frame and other potentially conflicting reference frames becomes a proposed measure for spatial presence. The actual things to measure are the reference frames from different modalities and the potential mismatch between them by appropriate psychophysical methods.

Furthermore, certain spatial behaviors seem impossible without sufficient *spatial presence & immersion*. Measuring the functioning of *obligatory behavior* is a potential and currently discussed method to quantify *spatial presence & immersion*. In the same line of reasoning, effortless *continuous* or *instantaneous spatial updating* cannot occur without sufficient *spatial presence & immersion*. Following the logical chain further up in our model, we see that spatial updating (continuous or instantaneous) is a necessary prerequisite for quick and intuitive *spatial behavior*. Conversely, the observation of such quick and intuitive *spatial behavior* implies automatic spatial updating and consequently also *spatial presence & immersion*. Those examples represent indirect measures of spatial presence that can readily lead to novel experiments complementing current presence research (see, e.g., Riecke, von der Heyde, & Bülthoff (2001)). We hope that this framework will help to structure the way we think about spatial presence and ways to quantify it.

2.3 Further hypotheses about functional relations

So far we tried to sketch a clear chain of logical connections which can be summarized as *spatial behavior* $\Rightarrow$ *spatial perception*, which is plausible per se. In addition to some assumptions we had to make in laying out our string of arguments, we would now like to introduce two hypothetical additional loops.

Inspired by recent experimental evidence (Riecke, von der Heyde, & Bülthoff, 2002b), we propose that *spatial presence & immersion*, *continuity*, *ego-motion perception* and an *egocentric reference frame* **together** are sufficient to enable proper *continuous spatial updating* (*spatial presence & immersion* $\wedge$ *continuity* $\wedge$ *ego-motion perception* $\wedge$ *egocentric reference frame* $\Rightarrow$ *continuous spatial updating*). In other word, continuous spatial updating works if all four prerequisites are true. Conversely, if we observe impaired *continuous spatial updating*, then we can conclude that at least one of the prerequisites is violated. ($A \wedge B \wedge C \wedge D \Rightarrow E$ is equivalent to $\neg E \Rightarrow \neg A \vee \neg B \vee \neg C \vee \neg D$).

Taken together with the previously established logical connections (*continuous spatial updating* $\Rightarrow$ *spatial presence & immersion* $\Rightarrow$ *consistency check* $\Rightarrow$ *egocentric reference frame)* $\wedge$ *(continuous spatial updating* $\Rightarrow$ *continuity* $\Rightarrow$ *egocentric reference frame)* $\wedge$ *(continuous spatial updating* $\Rightarrow$ *ego-motion perception*), we can furthermore conclude the following: if any of the four prerequisites is violated, *continuous spatial updating* would be rendered impossible or at least largely impaired ($\neg A \vee \neg B \vee \neg C \vee \neg D \Rightarrow \neg E$). Together with the above argument, this leads to the following equivalence: $\neg E \iff \neg A \vee \neg B \vee \neg C \vee \neg D$, which is the same as saying that $E \iff A \wedge B \wedge C \wedge D$. In other words, this means that instead of measuring *continuous spatial updating*, we can measure *consistency check* $\wedge$ *spatial presence & immersion* $\wedge$ *egocentric reference frame* $\wedge$ *ego-motion perception*.

Furthermore, as *spatial presence & immersion* implies both *consistency check* and *egocentric reference frame*, we can as well state that measuring *continuous spatial updating* is equal to measuring *spatial presence & immersion* $\wedge$ *ego-motion perception*. This opens up many interesting experimental investigations. For example, *spatial presence & immersion* can be quantified by measuring *continuous spatial updating* and *ego-motion perception* and vice versa.

A very similar second loop is located in the static right part of the framework. Based on experimental evidence by Riecke, von der Heyde, & Bülthoff (2002c), we propose that *spatial presence & immersion* $\wedge$ *egocentric reference frame* $\wedge$ *localization* $\wedge$ *identification* $\Rightarrow$ *instantaneous spatial updating*. Following the same reasoning as before, this opens up the possibility to measure *instantaneous spatial updating* instead of *spatial presence & immersion* $\wedge$ *localization* $\wedge$ *identification*. Even more pragmatically, one could use standard psychophysics to measure the latter two of the conditions (*localization* $\wedge$ *identification*) as well as the new method of quantifying *instantaneous spatial updating* (Riecke et al., 2002c) in order to quantify *spatial presence & immersion* in quasi-static situations.

3 Discussion

In summary, we attempted to devise a preliminary model relating issues related to spatial orientation, spatial updating, and spatial presence on a logical and functional level. Being a preliminary working model, we do not yet have a complete set of experimental evidence to support all of our assumptions and hypotheses. There is, however, evidence from psychophysical experiments to support the logical chain from *spatial perception* to *localization* and *identification* and also towards *motion perception*. Some of the experiments described above also showed that certain spatial behavior is possible (quick pointing to currently invisible landmarks) which would suggest either the path via *continuous spatial updating* or *instantaneous spatial updating*. Most likely, both possibilities involve *spatial presence & immersion* for the efficient and effortless solution of the task.

We are currently starting "POEMS" (Perceptually Oriented Ego-Motion Simulation), a new EU-collaboration in the context of the FET presence initiative. This project is closely linked to the framework presented and will employ multiple sensory modalities to investigate convincing ego-motion sensation and spatial presence. We will combine and correlate traditional presence measures like questionnaires and physiological measures with novel methods of quantifying spatial presence and immersion derived from the framework presented. To give an example, our framework would predict high correlation between the different presence measures if and only if one coherent and consistent egocentric reference frame exists. Other projects in the FET Presence initiative will look for correlations of obligatory behavior with presence (for example fear of heights).

So far, we have not attempted to relate each item in the framework to a corresponding information flow. Many of the proposed connections may indeed be closely linked to corresponding processing steps and neural connections in the human brain. Most of the boxes might also be considered as being localized in specific brain regions. There is for example a large body of literature indicating that the hippocampus is critically involved in path integration as well as landmark-based navigation and cognitive maps in animals including humans (Berthoz, 1997; Maguire, Frith, Burgess, Donnett, & O'Keefe, 1998b; Maguire, Burgess, Donnett, Frackowiak, Frith, & O'Keefe, 1998a; McNaughton, Barnes, Gerrard, Gothard, Jung, Knierim, Kudrimoti, Qin, Skaggs, Suster, & Weaver, 1996; Mittelstaedt, 2000; O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978; Poucet, 1993; Samsonovich & McNaughton, 1997). Trying to associate the individual boxes and logical connections of the current framework with corresponding neural substrate would be a challenging as well as promising endeavor. Obviously enough, however, it goes well beyond the scope of this paper.

4 Conclusions

To sum up, we embedded current terminology from the field of spatial presence and immersion in a functional and logical framework. This framework covers aspects ranging from spatial perception over allocentric and egocentric spatial memory up to spatial behavior. Finally, we used this framework to generate hypotheses which can guide future research and can be experimentally tested.

References

Berthoz, A. (1997). Parietal and hippocampal contribution to topokinetic and topographic memory. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.*, **352**(1360), 1437–1448.

Braitenberg, V. (1984). *Vehicles*. Cambridge, MA: MIT Press.

Dichgans, J., & Brandt, T. (1978). Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.), *Perception*, Vol. VIII of *Handbook of Sensory Physiology* (pp. 756–804). Berlin Heidelberg: Springer.

Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. (1998). Where did I take that snapshot? Scene-based homing by image matching. *Biol. Cybern.*, **79**(3), 191–202.

Golledge, R. G. (Ed.). (1999). *Wayfinding Behavior: Cognitive mapping and other spatial processes*. Baltimore: Johns Hopkins.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments*, **5**(3), 274–289.

Hendrix, C., & Barfield, W. (1996b). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments*, **5**(3), 290–301.

Hunt, E., & Waller, D. (1999). Orientation and Wayfinding: A Review. [Online] Available: http://depts.washington.edu/huntlab/vr/pubs/huntreview.pdf.

Ijsselsteijn, W., de Ridder, H., Freeman, J., Avons, S. E., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments*, **10**(3), 298–311.

Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1997). Encoding spatial reresentations through nonvisually guided locomotion: Test of human path integration. In D. Medin (Ed.), *The psychology of learning and motivation*, Vol. 37 (pp. 41–84). San Diego, CA: Acad. Press.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, **10**(3), 282–297.

Mach, E. (1922). *Die Analyse der Empfindungen [The analysis of sensations]*. Jena: Gustav Fischer.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998a). Knowing where and getting there: a human navigation network.. *Science*, **280**(1), 921–924.

Maguire, E. A., Frith, C. D., Burgess, N., Donnett, J. G., & O'Keefe, J. (1998b). Knowing where things are: Parahippocampal involvement in encoding object locations in virtual large-scale space.. *Journal of Cognitive Neuroscience*, **10**(1), 61–76.

Maurer, R., & Séguinot, V. (1995). What is modelling for?: A critical review of the models of path integration. *Journal of Theoretical Biology*, **175**(4), 457–475.

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., Kudrimoti, H., Qin, Y., Skaggs, W. E., Suster, M., & Weaver, K. L. (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *Journal of Experimental Biology*, **199**, 173–185.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Mittelstaedt, H. (2000). Triple-loop model of path control by head direction and place cells. *Biol. Cybern.*, **83**(3), 261–270.

Mittelstaedt, H., & Mittelstaedt, M.-L. (1982). Homing by path integration. In F. Papi & H. Wallraff (Eds.), *Avian navigation* (pp. 290–297). Berlin: Springer.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as spatial map. Preliminary evidence from unit activity in the freely moving rat. *Brain Research*, **34**, 171–175.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, England: Clarendon.

Poucet, B. (1993). Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms. *Psychological-Review*, **100**(2), 163–182.

Riecke, B. E., van Veen, H. A. H. C., & Bülthoff, H. H. (2002). Visual Homing Is Possible Without Landmarks: A Path Integration Study in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, **11**(5).

Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2001). How Real is Virtual Reality Really? Comparing Spatial Updating using Pointing Tasks in Real and Virtual Environments. *Journal of Vision*, **1**(3), 321a. http://www.journalofvision.org/1/3/321/.

Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2002a). Spatial updating experiments in Virtual Reality: What makes the world turn around in our head?. In H. H. Bülthoff, K. G. Gegenfurtner, H. A. Mallot, & R. Ulrich (Eds.), *Beiträge zur 5. Tübinger Wahrnehmungskonferenz*, p. 162. Knirsch Verlag, Kirchentellinsfurt, Germany. Available: http://www.kyb.tuebingen.mpg.de/publication.html?publ=632.

Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2002b). Spatial updating in virtual environments: What are vestibular cues good for?. In K. Nakayama et al. (Ed.), *VisionScienceS02 (in press)*, Talk presented at the VisionScienceS Meeting, Sarasota, Florida, United States. Available: http://www.kyb.tuebingen.mpg.de/publication.html?publ=628.

Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2002c). Teleporting works - Spatial updating experiments in Virtual Tübingen. In *OPAM (accepted talk)* Kansas City, United States. Available: http://www.kyb.tuebingen.mpg.de/publication.html?searchstring=Teleporting+works.

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, **17**(15), 5900–5920.

Schloerb, D. W. (1995). A Quantitative Measure of Telepresence. *Presence: Teleoperators and Virtual Environments*, **4**(1), 64–81.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, **10**(3), 266–281.

Tolman, E. C. (1948). Cognitive maps in Rats and Men. *Psychological Review*, **55**, 189–208.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J. A. (1997). Biologically based artificial navigation systems: Review and prospects. *Prog. Neurobiol.*, **51**(5), 483–544.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, **7**(3), 225–240.