NOTICE

This is the authors' version of a work that was accepted for publication in Cognitive Systems Research. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. The article was available on-line since: **5-OCTOBER-2011.** A definitive version is available in **Cognitive Systems Research, Volume 17-18 (2012), pp. 1-24. DOI: http://dx.doi.org/10.1016/j.cogsys.2011.09.001**.

A Computational Model of the Allocentric and Egocentric Spatial Memory by Means of Virtual Agents, or How Simple Virtual Agents Can Help to Build Complex Computational Models

Cyril Brom and Jan Vyhnánek

Faculty of Mathematics and Physics, Charles University in Prague

Jiří Lukavský

Institute of Psychology, Academy of Sciences of the Czech Republic

David Waller

Department of Psychology, Miami University

Rudolf Kadlec

Faculty of Mathematics and Physics, Charles University in Prague

Correspondence concerning this article should be addressed to Cyril Brom, Faculty of Mathematics and Physics, Charles University in Prague, Room 312, Malostranske Namesti 25, Prague, 11800, Czech Republic. Tel: (420) 221 914 216; Fax: (420) 221 914 281

Abstract

The ability to acquire, remember and use information about locations of objects in one's proximal surrounding is a fundamental aspect of human spatial cognition. In this paper, we present a computational model of this ability. The model provides a possible explanation of contradictory results from experimental psychology related to this ability, namely explanation of why some experiments have reproduced the so-called "disorientation effect" while others have failed to do so. Additionally, in contrast to other computational models of various aspects of spatial cognition, our model is integrated within an intelligent virtual agent. Thus, on a more general level, this paper also demonstrates that it is possible to use intelligent virtual agents as a powerful research tool in computational cognitive sciences.

Keywords: spatial cognition, paradigm of pointing, disorientation effect, intelligent virtual agent

1. Introduction

Computational approaches to cognitive science have become increasingly important in the past decade (Sun, 2008; Burgess, 2007). This is likely because, compared to verbally-based theories, *in silico* simulations enable a researcher to specify hypothetical mechanisms in precise detail, systematically explore the model and manipulate its parameters, and generate new predictions (Tyrrell, 1993; Burgess, 2007; Sun, 2008; Marsella & Gratch, 2009). At the same time, simulations can be both more complex and well-specified than analytical or numerical models (Kokko, 2007; ch. 8).

Intelligent virtual agents (IVAs) are pieces of software that are both autonomous and graphically *embodied* in a 2D or 3D virtual environment, capitalizing on the general agent metaphor used in software engineering and artificial intelligence (Wooldridge, 2002). IVAs are currently used in a large variety of applications, for instance, in educational simulations, virtual storytelling, cultural heritage, and computer games.

In general, research on IVAs attempts to make these entities more *believable*, that is, to increase the agents' ability to appear and behave in a lifelike manner. Believable IVAs enable users to suspend their disbelief by providing a convincing portrayal of the personality the user expects (Loyall, 1997). It is worth noting that for designers of many IVAs, the goal is to *imitate* a character's behavior, but not necessarily to develop a cognitively plausible model producing the behavior. Nonetheless, these agents can be used as tools for investigating plausible computational models of cognitive phenomena (Brom, Lukavsky, & Kadlec, 2010a) primarily because virtual worlds present convenient models of real worlds, and IVAs, owing to their modular architectures, can serve well as "vehicles" for carrying and testing the models. IVAs can generate input data for these models and allow output data to be meaningfully manifested. Compared to a robotic platform, which is sometimes used for

research in computational cognitive sciences (e.g. Krichmar et al., 2005), virtual reality is more technically accessible and allows for investigation of higher-level phenomena. The input representations can be more abstract and can enable researchers to ignore other difficult problems (e.g., robotic vision) (see Brom, Lukavsky, & Kadlec, 2010a for a longer discussion of positives and negatives of IVA platform vs. robotic platform).

However, to date, the use of cognitively plausible IVAs has been quite limited. In the past, some parts of IVA control mechanisms were psychologically inspired; most notably emotion models (e.g. Gebhard et al., 2008) and spatial memory models (e.g. Thomas & Donikian, 2006). Yet even these psychologically inspired models are not necessarily cognitively plausible when judged by standards of computational cognitive sciences. Indeed, most of psychologically inspired models for these agents enable a researcher neither to do comparative analysis against real world data nor to make predictions. Similarly, it has been demonstrated that IVAs can be controlled by Soar (Laird, 2000) as well as ACT-R (Best & Lebiere, 2006) cognitive architectures and that ACT-R can be integrated with Leabra, a neural architecture, and implemented in an IVA (Jilk, Lebiere, O'Reilly, & Anderson, 2008). However, even though these works presented valuable proof of concepts of connections of cognitive architectures with IVAs, they did not demonstrate fully the alleged advantages of IVAs for computational cognitive sciences.

This article has two goals. First, we will present a computational model of one fundamental aspect of human spatial cognition: the ability to acquire, remember and use locations of objects in one's proximal surrounding. In contrast to other computational models of various aspects of spatial cognition, our model is integrated within an IVA, and provides one possible explanation of contradictory results of experimental psychology related to this ability (Wang & Spelke, 2000; Holmes & Sholl, 2005; Waller & Hodgson, 2006). Second, and

more generally, this paper also demonstrates that it is possible to use IVAs as a powerful research tool in computational cognitive sciences.

1.1 Paradigm of pointing

Knowing locations of objects in one's environment is a critical component of many human behaviors. Accordingly, substantial effort has been devoted to understanding the processes and representations underpinning this ability. Most contemporary theories of human spatial knowledge posit (at least) two partly independent subsystems. First, a transient system is thought to update spatial relations as one moves through an environment. This system integrates multi-modal perceptual information and is generally thought to keep track of the locations of relatively few objects with respect to an egocentric frame of reference. A second system, based in long-term memory, codes locations of objects that are not perceptually available. This more enduring system of spatial knowledge may employ nonegocentric reference frames and has far greater capacity, although generally less precision, than the transient system (Burgess, 2006; Waller & Hodgson, 2006). The exact nature of these systems has been intensively investigated and debated in past (Easton & Sholl, 1995; Gallistel, 1990; Neisser, 1976; Wang & Spelke, 2002).

One experimental paradigm used to distinguish transient from enduring spatial knowledge involves pointing to remembered but unseen objects. In a typical experiment, a person learns the locations of several objects in a room-sized environment, e.g. Figures 1, 3. The person is subsequently asked to point to the remembered locations of these objects after the objects have been removed or occluded. Pointing while oriented to one's environment is generally assumed to tap into transient spatial knowledge, while pointing after being disoriented requires enduring spatial knowledge. Wang and Spelke (2000) documented a

reliable increase in the variability of a person's pointing errors as a result of disorientation (referred to here as the "disorientation effect") and argued that such an increase is not well-explained by theories that posit exclusive control of spatial behavior by an enduring system using nonegocenctric reference frames. Subsequent research (Waller & Hodgson, 2006) interpreted this increase in variable error as evidence for a switch from the relatively precise transient representation to the coarser enduring one.

Despite the apparent reliability of the disorientation effect in Wang and Spelke's (2000) work as well as Experiment 1 of Waller and Hodgson (2006), Holmes and Sholl (2005) (Experiments 3–7) repeatedly failed to replicate it. With the present research, we examine the possibility that these contradictory findings may have been a result of relatively subtle difference in experimental procedures. In so doing, we develop a computational model that enables us to replicate these experiments *in silico*, including the key difference in experimental procedures. This is possible because we model the subject's walking behavior using a walking IVA and we let the IVA walk in environments modeling the experimental rooms. The model enables us to fit precisely the data of Waller and Hodgson and to replicate qualitatively the key difference between results of the two experiments without altering any parameter settings. The model thus provides one possible explanation of the contradictory findings as well as generating new predictions that can, in future work, be tested empirically on human subjects.

1.2 Structure of the paper

The remainder of this paper begins with detailing the paradigm of pointing, explaining its importance, and outlining the findings of the two original studies we have modeled. Next, our computational model is introduced and underlying assumptions explained. We will also discuss how the model is integrated within cognitive architecture of our virtual agent. Then

the experimental method will be outlined and the results of the *in silico* study presented. The paper will conclude with a discussion of implications of our research for the spatial cognition studies as well as general computational psychology.

2. Paradigm of pointing and disorientation effect in detail

2.1 Motivation

As mentioned above, spatial knowledge is often conceptualized as comprising both transient, online, perceptual awareness of one's immediate surroundings as well as enduring, offline memorial representations of environmental structure. Understanding the nature of these two spatial systems and how they interact has been a central theme of much contemporary research in human spatial cognition.

Based partly on previous work with reorientation in children (Hermer & Spelke, 1996) and animals (Cheng & Gallistel, 1984), Wang and Spelke (2000) introduced a disorientation paradigm to examine adult human spatial knowledge. After learning the locations of six objects in their immediate environment, participants pointed to these objects while blindfolded, both before and after a long (e.g., one minute) rotation. Before rotation, pointing accuracy was high, and relatively unbiased, indicating that people were oriented to their environment. Wang and Spelke also calculated participants' "configuration error" as the variability of their errors. Configuration error before rotation was relatively low, indicative of a coherent knowledge structure on which online pointing was based.

After rotation, the uniform distribution of participants' pointing responses suggested that the rotation had disoriented the participants. More importantly, across all participants, configuration error while disoriented was significantly greater than when oriented. This effect persisted even after controlling for participants' variability in pointing to a single object – their so-called "pointing error." Wang and Spelke argued that if pointing had been based on

an enduring coherent "cognitive map" of an environment, configuration error would not have been affected by disorientation. In such a case, people would have presumably used the same mental representation before and after rotation and there should thus be no effect of disorientation. The significant increase in configuration error was therefore more consistent with the idea that people updated their directions to target objects in a piecemeal way while oriented, and that disorientation affected each direction estimate somewhat independently. These provocative results cast doubt on the ubiquity and importance of the concept of a "cognitive map" (a global and comprehensive internal spatial representation that does not code spatial relationships with respect to the observer) and notable attempts to replicate them were soon conducted by Holmes and Sholl (2005) and Waller and Hodgson (2006).

2.2 Holmes & Sholl procedure

In experiment 7 of Holmes and Sholl (2005), participants were tested in 4.90 m × 3.91 m room devoid of any furnishing other than 6 target objects as well as the tables and stool on which they were placed (Figure 1). During learning, participants walked around the room to familiarize themselves with the locations of the objects. Then they sat in a swivel chair in the middle of the room and they were tested in the three phases used by Wang and Spelke (2000). In the first phase they closed their eyes and pointed to each object location. If they made a mistake in this phase they were asked to study the room again. In the second phase (eyes-closed phase) the participants were blindfolded and were rotated approximately 45 degrees to either their left or right. In the third phase (disoriented phase) participants were slowly rotated so that they became disoriented. After they were stopped, they were told to imagine facing a direction of their choice. There were 24 trials for both the second and the third condition. Participants had to point at one object in a trial, four times to each of six objects chosen in a random order.

The critical results of this experiment are depicted on Figure 2, which illustrates a failure to replicate the disorientation effect originally reported by Wang and Spelke (2000, Exp. 1, 2).

--- Insert Figure 1 about here ---

--- Insert Figure 2 about here ---

2.3 Waller & Hodgson procedure

In another attempt to replicate Wang and Spelke's (2000) results, Waller and Hodgson (2006) asked participants to learn the locations of six objects in a 8.66 m × 8.40 m room. In contrast to the procedures of Holmes and Sholl (2005), Waller and Hodgson placed a 1.9 m × 1.9 m booth in the middle of the room, replicating closely the original conditions of the original Experiment 2 of Wang and Spelke (2000) (Figure 3). The procedure was the same as in Holmes & Sholl with the four small differences. First, participants sat on the chair placed in the booth. Second, in the eyes-closed phase blindfolded participants were rotated 40°. Third, in the disoriented phase, participants were asked to imagine facing an object of their choice. Fourth, in each phase, participants pointed twice to each of the six objects.

In this experiment, Waller and Hodgson replicated the disorientation effect of Wang and Spelke (2000). Figure 4 summarizes the data.

--- Insert Figure 3 about here ---

--- Insert Figure 4 about here ---

2.4 Our hypothesis

For both Holmes and Sholl's and Waller and Hodgson's experiments, we modeled the learning phases and the pointing phases (the eyes-closed phase and the disoriented phase). As detailed later, in our model, we assume the allocentric object-to-object relations are built only between objects that were perceived recently. These objects change while the agent is moving around the room. In the conditions of Holmes and Sholl the agent can perceive more objects at the same time than in the conditions of Waller and Hodgson because the booth is blocking the agent's view in the latter setting. Thus, our hypothesis is that the model will develop more accurate object-to-object spatial representation in the Holmes and Sholl conditions than in the Waller and Hodgson conditions. When confirmed, this discrepancy would help us to explain why Holmes and Sholl failed to replicate the disorientation effect. Although a similar idea has been expressed by Mou et al. (2006), here we aim at supporting it by *in silico* data. To examine our models' predictions with respect to object-to-object representations, we will also compare accuracy of developed object-to-object spatial representations in Waller and Hodgson's condition with the booth to the same condition *without* the booth, an opportunity we have due to the computational approach.

3. Computational model of disorientation paradigm

In this section, we introduce our computational model of disoriented paradigm (DP-model throughout). Because the model is integrated within an IVA, we will first introduce cognitive architecture of our IVA model. Then, we will outline assumptions behind the DP-model in detail. Finally, the DP-model will be introduced.

3.1 IVA architecture

The integration of the DP-model within an IVA is only possible because IVAs, including our own IVA, have elaborate and modular cognitive architectures. The IVA cognitive architectures are somewhat similar to general cognitive architectures, such as ACT-R, Soar or ICARUS (see Samsonovich (2010) for more of these architectures). However, in general, the IVA architectures depart from the general cognitive architectures in that the former tend to be special-purpose and application driven, while the latter predominantly serve general AI or neuro-/psychological goals.[1]

The architecture that drives our IVA is depicted on Figure 5. This figure depicts the IVA receiving inputs from the environment (ENV on Figure 5) via a threshold-based *attention filter*. The inputs fill up the *perceptual field* (PF) of the *short-term memory* (STM). The architecture has several long-term memory appendages, such as the long-term spatial memory (LTSM) and the long-term episodic memory (LTEM).[2]

The key feature of the architecture is hierarchical decomposition of an agent's behavior: IVA's behaviors are decomposed to sub-behaviors, which are further refined until some atomic actions are reached. This is similar to how representation of possible behavior is conceived by ICARUS cognitive architecture (Langley, Choi, & Rogers, 2009). In fact, our architecture distinguishes *tasks* from *goals* making the mechanism resembling also the Belief-Desire-Intention architecture (Bratman, 1987), but this is unimportant for present purposes (see Brom, Peskova, & Lukavsky (2007) for details). Note that hierarchical representations of behavior are popular in the field of IVAs and in videogames in particular (e.g., Isla, 2005; Champandard, 2008).

Every behavior may require several *resources*, i.e. objects, for execution. Behaviors to be pursued are selected within the *goal structure* and the *conflict resolution mechanism* based

on drives, external events, and a schedule. Currently pursued behaviors are represented within the *task field* (TF) of the STM. The *memory field* (MF) of the STM can hold temporarily information about an object recalled from a long-term memory. Every object is regarded as a tool for action, i.e. it is a set of "affordances" (Gibson, 1979), meaning it possesses pointers to behaviors it can be used for as a resource (dotted arrows on Figure 5). These pointers are perceived by IVAs when observing their environments. The architecture also features a simple valence-based *emotion module* and a simple *linguistic module* allowing the IVA to explain itself based on the content of its LTEM.

We have implemented several IVAs capitalizing on this architecture, most notably in the context of modeling various aspects of episodic memory for virtual characters (Brom et al., 2007; Brom et al., 2010a; Brom, Burkert, & Kadlec, 2010).[3] Individual parts of the architecture are implemented only when needed, e.g. only some of our IVAs feature emotions. Some of these IVAs inhabit 2D worlds while others 3D worlds of the game Unreal Tournament 2004 (Epic, 2004; Gemrot et al., 2009). The IVA developed for the DP-model uses a lightweight version of this architecture: it features neither emotions nor drives nor long-term episodic memory. In terminology of our architecture, the DM-model is an augmentation of the short-term memory and the long-term spatial memory.

--- Insert Figure 5 about here ---

3.2 The DP-model: overview and assumptions

Any researcher developing a complex computational model faces a problem of describing the model's properties in enough detail to enable *in silico* execution of the model. Here, we detail the assumptions we made when specifying the DP-model and overview the

model. We start with conceptual requirements and considerations, which we number with "G" prefix, and continue with technical considerations, numbered with "T" prefix. For brevity, we detail only the components of the model that are needed for explaining the results of Waller and Hodgson (2006; Exp. 1) and Holmes and Sholl (2005; Exp. 7). In Section 5, it will be shown that the model is actually more complex.

In general, the model is abstract; it operates with conceptual entities used for describing mental processes at the psychological level of abstraction. We make no links to neural substrate of human positional system in this paper. Additionally, we consider spatial relations only in the horizontal plane at eye height; in other words, objects' heights played no role in the experiments in question (other than the requirement that the objects could not occlude each other).

3.2.1 General requirements and conceptual considerations

G1. Egocentric and allocentric components. As already mentioned, dominant models of the human positional system posit (at least) two partly independent subsystems: one transient and the other more enduring (Burgess, 2006; Gallistel, 1990; Mou et al., 2004; Waller & Hodgson, 2006). Additionally, recent research suggests that both egocentric – subject to object – and allocentric – object to object – representations exist in parallel (see Burgess (2006) for a review). The idea of an egocentric module fits naturally with the idea of the transient system, while an allocentric module accords well with the notion of the enduring system (Burgess, 2006, pp. 555; but see also Waller & Hodgson, 2006). Thus, the extant literature supports two key subsystems for our model: an *egocentric* and an *allocentric*. While the representations held by the former one should be relatively short-lived, the representations of the latter should be more enduring. These two subsystems - transient egocentric and enduring allocentric will be featured by our model.

G2. Vectors, weights and errors. The natural assumption is to code egocentric information using self-to-object vectors while allocentric information using object-to-object vectors; in this way the information is to be coded in our model.

An important question is how to represent inaccurate information and how representational inaccuracies translate into output error (i.e. error in pointing). Although error in human pointing behavior can arise from a large number of underlying cognitive processes (e.g., storage, updating, maintenance, retrieval, etc.), for the sake of parsimony, our model will simplify these sources into only one kind of general memory error and one kind of motor error. Thus, we can represent all self-to-object and object-to-object vectors precisely, which is technically the most convenient solution, and generate these two errors during pointing tasks based on two sources of information. The first source is how well the vector was learnt during the learning phase; we will represent this information using a vector's weight (also called strength here). The second source is a noise from the motor and memory system.

To be more precise, the motor error will be generated using parameterized gaussian noise, which is an abstraction for the noise in the motor system. The motor error will not be influenced by the amount of learning.

The situation with the memory error is more complex. The memory error obviously should be influenced by both sources of information, not just by the noise in the memory system. We can assume that the longer a part of the objects' layout is visible the better it will be represented internally; the initial uncertainty in the internal representation of self-to-object and object-to-object relations will be gradually reduced. At the same time, we can assume that the longer a part of the layout will not be visible, the more will the respective internal relations become distorted, that is, the uncertainty will grow. This corresponds mainly to the noise accumulation due to updating of egocentric vectors towards non-visible objects and due

to general memory maintenance (at the neural level). We assume that the decrease and the increase of the accuracy are reverse processes. Thus, the weight is to correspond positively to the amount of learning the vector is exposed to during the learning phase and negatively to the amount of forgetting. The memory error will be then generated using gaussian noise the magnitude of which will be influenced by the respective weight. For convenience, we will use the scale 0 - 1 for weights (0 corresponds to a maximal uncertainty, 1 to a minimal uncertainty; later, we will see what exactly is meant by "minimal" and "maximal").

For brevity, we will often use terms *egocentric vectors* and *egocentric weights* for vectors and their weights from the egocentric subsystem, and *allocentric vectors* and *allocentric weights* for vectors and their weights from the allocentric subsystem.

G3. Relation between the subsystems. Although there are several ways how information can flow within the system, we assume that egocentric representation is updated based on information in the perceptual field, and the allocentric representation is updated based on the egocentric vectors, not the other way round. These assumptions are aligned with dominant theories of human information processing (e.g., Neisser, 1967; Marr, 1982), which generally consider the information used in perception and action to be egocentrically (e.g., retinotopically) organized, and for additional processing to be required in order to store non-egocentric information in long term memory. Thus, an egocentric weight is strengthened or weakened when a particular object is perceived or not perceived, respectively, i.e. based on whether the object is actually represented within the perceptual field. An allocentric vector between two objects is strengthened proportionally to the strengths of the two subject-to-object vectors corresponding to these two objects (i.e., let us assume we have an agent A and objects O_1 and O_2; the allocentric vector O_1–O_2 is updated based on the strengths of A–O_1 and A–O_2).

When a human moves, the coordinates of self-to-object vectors are updated. The exact nature of human updating is subject to intensive research (see, for example, Loomis, Klatzky, Golledge, & Philbeck, 1999). In the model, in accordance with Point G2, the self-to-object vector coordinates are updated precisely.

G4. Pointing. When pointing, we assume that the agent prefers to use the egocentric representation when it is available (Waller & Hodgson, 2006; Hodgson & Waller, 2006). Additionally, we assume that representations of self-to-object relations are disrupted after disorientation (e.g. Mou et al., 2006, pp. 1276). When this happens, the agent must use the often less precise object-to-object representation. Therefore, in our model the egocentric subsystem assists in the pointing task in the eyes-closed phase (i.e. before the disorientation) while the allocentric subsystem is used in the disoriented phase. We will call the former situation *egocentric pointing* and the latter *allocentric pointing*.[4]

G5. Speed of learning. Because the egocentric representation should be built (and forgotten) more rapidly than the allocentric one (Point G1), egocentric weights should increase and decrease more quickly than allocentric weights. But is this the only requirement on the speed of learning?

Psychologically speaking, human subjects will most likely know the objects' layout well at the end of the learning phase, but some self-to-object or object-to-object relations can be expected to be represented better than others. In terms of the model, this means that, at the end of the learning phase, the vectors should tend to be strong, but their weights should not be the same, and in particular, they should not be *all* very close to 1. With all weights very close to 1, a ceiling effect would be produced during generating memory error. We call this situation *full saturation*. Note a reverse problem, that is, with a floor effect, would appear when all weights are very small. Thus, we require the weights, when considered all together, to be

reasonable strong yet *diverse* at the end of the learning phase; some may approach the saturation (value 1), but some should be weaker. This holds for both egocentric and allocentric vectors.

How should a model achieve this? Let us start our thinking with allocentric vectors. In order to prevent full saturation of allocentric vectors, the *egocentric* vectors must be built and must decay *very* rapidly (strong egocentric weights basically mean that the allocentric weights are continuously being increased due to Point G3). Ideally, for the purpose of allocentric weights to be reasonably strong, but not fully saturated, at the end of the learning phase, egocentric weights should increase from 0 to 1, and vice versa, in about an order of magnitude shorter time period than the learning phase lasts. But is this possible?

Recall that egocentric weights should be also used for estimating errors during egocentric pointing (Points G2, G4). Should egocentric weights increase from 0 to 1, and vice versa, several times during the learning phase (due to the requirements on allocentric weights), two unwelcome things would happen at the end of the learning: a) several egocentric weights would most likely be close to 0, b) egocentric weights in general would not reflect the amount of learning/forgetting the vector is exposed to during the learning phase.

Thus, there are two contradictory requirements: on a rapid build-up and decay of the egocentric weights (due to the allocentric pointing) and on a slow build-up and decay of the egocentric weights (due to the egocentric pointing). To resolve this conflict, we define two kinds of egocentric weights; transient and enduring. The transient weights change more rapidly than the enduring weights and they are influenced by information in the perceptual field directly. The enduring weights are changed based on the transient weights and they assist

in the egocentric pointing. In addition, transient (but not enduring) egocentric weights are used for changing the allocentric weights.

Recall that the build-up of allocentric weights should be even slower than the build-up of enduring egocentric weights. Concerning the speed of allocentric weights' decay, because the forgetting in the allocentric representation is not crucial within the time scope we are investigating, we abstract from it: the allocentric weights will not decay in the model.

The requirements on the speed of learning of the three types of weights are summarized in Figure 6. For explanatory purposes, one can note a rough correspondence to Atkinson and Shiffrin's memory model (1968) (although this correspondence was not actually our guiding design principle). Note also that the "long-term" memory should be rather denoted as "intermediate" term memory given the time scope of our experiments.

--- Insert Figure 6 about here ---

3.2.2 Technical considerations

T1: Localization abilities and geometrical memory. Because the experiments that we model involve environments that are only room-sized, for simplicity, we endow our model with perfect knowledge of environmental topology, i.e. surrounding walls, and perfect ability to localize itself in the environment.

T2: Perception, attention. Our agent perceives all objects in its current visual field, which is 120° wide. The eye movements and foveation are not modeled. We do not limit the number of objects that can be represented in the perception field at one moment for there are

only up to 6 objects in the experiments we modeled. Additionally, we abstract from attention for its role has not been fully examined empirically and thus is not fully understood. Finally, because the experiments do not employ features of objects, the objects we model are state-less (nevertheless, the agent is able to distinguish one object from another).[5]

T3: Geometry of the perceived space. Although there is some evidence that the geometry of perceived proximal space is non-Euclidean (e.g. Cutting & Vishton, 1995), we make the simplifying assumption that both the peri-personal as well as extra-personal space have Euclidean geometry. Note that this consideration justifies the usage of gaussian-based noise for mimicking the error during pointing (Point G2).

T4: Grouping, alignment. For the purpose of this study we do not model high-order geometrical relationships between objects (e.g. gestalt laws—three objects being placed on the same line, objects forming two parallels etc.). Similarly, each object in our scenario is easily distinguishable from others and we do not implement any grouping. Because these principles were not explicitly treated in the experiments we modeled, we do not include them here.

T5: Intrinsic axis for the allocentric representation. In this paper, we assume that the 0° axis of the allocentric representation is the south-north axis of the rooms in which experiments were conducted (see Sec. 5 for more on this issue). In the egocentric representation, the natural 0° axis is represented by the agent's current heading.

T6. Steps. One agent's step approximates one human step or rotation by any angle (an arbitrary multiple of 15°).

3.3 The DP-model: details

Formally, the DP-model is a triple <E, A, P>, where E represents the egocentric component, A the allocentric one, and P is the perceptual field. The perceptual field is exactly what was described above: a set of internal representations of all objects that are currently perceived in the agent's visual field (Point T2 from Sec. 3.2).

In terms of our IVA architecture, the egocentric component augments the perceptual field and the allocentric extends the long-term spatial memory.

3.3.1 Egocentric component. The egocentric component E is a triple <H, ^{eg}V, ^{eg}C>, where H is an angle representing the agent's current heading (with respect to the south-north axis), ^{eg}V is a set of egocentric vectors (Point G2) and ^{eg}C is the egocentric updating configuration.

A vector $^{eg}v_A \in {}^{eg}$V represents a doubly weighted egocentric vector from the agent to the object A (Point G5) and it is a tuple <($^{eg}a_A$, $^{eg}b_A$), $^{eg\text{-}t}w_A$, $^{eg\text{-}t}d_A$, $^{eg\text{-}e}w_A$, $^{eg\text{-}e}d_A$ >, where:

· ($^{eg}a_A$, $^{eg}b_A$) are Cartesian coordinates (Point G2) of the vector between the agent and the object A with respect to the agent's actual heading H and position as shown in Figure 7;

· $^{eg\text{-}t}w_A$ is the vector's transient weight (or strength), a number from the interval <0,1>;

· $^{eg\text{-}t}d_A$ is a base for computing the transient weight, a number from interval <0, ^{eg}bound >. This variable is related to $^{eg\text{-}t}w_A$ via the standard logistic sigmoid function, as depicted on Figure 11.

· $^{eg\text{-}e}w_A$ is the vector's enduring weight, a number from the interval <0,1>;

· $^{eg\text{-}e}d_A$ is a base for computing the enduring weight, a number from interval <0, ^{eg}bound >. This variable is related to $^{eg\text{-}e}w_A$ via the standard logistic sigmoid.

The configuration ^{eg}C contains parameters $^{eg\text{-}t}inc$, $^{eg\text{-}t}dec$, $^{eg\text{-}e}inc$, $^{eg\text{-}e}dec$, $^{eg}bound$:

· $^{eg\text{-}t}inc$ is the speed of increasing the transient weights of the egocentric vectors;

· $^{eg\text{-}t}dec$ is the speed of decreasing the transient weights of the egocentric vectors;

· $^{eg\text{-}e}inc$ is the speed of increasing the enduring weights of the egocentric vectors;

· $^{eg\text{-}e}dec$ is the speed of decreasing the enduring weights of the egocentric vectors (Point G2, G3, G5);

· $^{eg}bound$ is an upper limit for the bases $^{eg\text{-}t}d_A$ and $^{eg\text{-}e}d_A$ for computing the vector strengths. Its purpose is to constrain the base of the egocentric vectors when an object is perceived for a long time. Such an over-learned egocentric vector would be then uneasily forgotten even a long time after the object was perceived (see Fig. 11).

These five parameters had been set before experiments started as described in Sec. 4.

--- Insert Figure 7 about here ---

3.3.2 Allocentric component. The allocentric component A is a tuple $< {}^{all}V, {}^{all}C >$, where ^{all}V represents a set of allocentric vectors (Point G2) and ^{all}C is the allocentric updating configuration.

A vector $^{all}v_{A,B}$, i.e. an allocentric vector from the object A to object B, is a triple $<({}^{all}a_{A,B}, {}^{all}b_{A,B}), {}^{all}w_{A,B}, {}^{all}d_{A,B}>$, where:

· ${}^{all}a_{A,B}$, ${}^{all}b_{A,B}$ are cartesian coordinates of the vector between objects A and B with respect to the main environment axis with direction 0° as shown if Figure 8 (Point T5);

· ${}^{all}w_{A,B}$ is the vector weight (or strength), a number from an interval <0,1>;

· ${}^{all}d_{A,B}$ is the base for computing the vector strength, a real positive number. This variable is related to ${}^{all}w_{A,B}$ via the standard logistic sigmoid function, as depicted on Figure 11. Note that ^{eg}bound is used as the upper limit for the base ${}^{all}d_{A,B}$ for computing the vector strengths.

The configuration ^{all}C contains only one parameter ^{all}inc which is the speed of increasing the weights of the allocentric vectors. As part of investigating our hypothesis, we examined how behavior of the model changes when the value of this parameter changes, as described in Sec. 4. Note that there is no ^{all}dec, an allocentric analogy to ${}^{eg\text{-}t}$dec and ${}^{eg\text{-}e}$dec because the allocentric weights do not decay (Point G5).

Note that vectors ${}^{all}v_{A,B}$ and ${}^{all}v_{B,A}$ are symmetric, meaning they have the same weights.

--- Insert Figure 8 about here ---

3.3.3 The initial state and update mechanisms. At the beginning of every experiment, all vector bases and weights are initialized to 0 for both the egocentric and the allocentric component. During the learning phase of every experiment, the IVA walks along a specified trajectory (how these trajectories were acquired is described in Sec. 4). The weights of the egocentric and allocentric vectors are updated each time the agent steps or turns.

Therefore, the final vector strengths depend on the path followed by the agent while exploring the room.

The perceptual field update. Every time step, we remove all objects that are no longer visible from the perceptual field P and insert all newly visible objects to P.

The egocentric component update. In every time step, every vector ${}^{eg}v_A \in {}^{eg}V$ is updated as follows:

(1) ${}^{eg\text{-}t}d_A(t+1) = {}^{eg\text{-}t}d_A(t) + {}^{eg\text{-}t}inc$, if object A was present in the perceptual field in time t;

${}^{eg\text{-}t}d_A(t+1) = {}^{eg\text{-}t}d_A(t) - {}^{eg\text{-}t}dec$, if object A was not present in the perceptual field in time t (Point G3);

with additional rule for bounding the vector base within predefined limits:

${}^{eg\text{-}t}d_A(t+1) = {}^{eg\text{-}t}d_A(t),$ if ${}^{eg\text{-}t}d_A(t+1) < 0;$

${}^{eg\text{-}t}d_A(t+1) = {}^{eg}bound,$ if ${}^{eg\text{-}t}d_A(t+1) > {}^{eg}bound;$

(2) ${}^{eg\text{-}t}w_A(t+1) = 1 / \{ 1+\exp[-({}^{eg\text{-}t}d_A(t+1) - {}^{eg}bound/2)] \};$

note that (2) is a shifted standard logistic sigmoid (Figure 11);

(3) ${}^{eg\text{-}e}d_A(t+1) = {}^{eg\text{-}e}d_A(t) + {}^{eg\text{-}t}w_A(t) . {}^{eg\text{-}e}inc - {}^{eg\text{-}e}dec;$

with additional rule for bounding the vector base within predefined limits:

${}^{eg\text{-}e}d_A(t+1) = {}^{eg\text{-}e}d_A(t),$ if ${}^{eg\text{-}e}d_A(t+1) < 0;$

${}^{eg\text{-}e}d_A(t+1) = {}^{eg}bound,$ if ${}^{eg\text{-}e}d_A(t+1) > {}^{eg}bound;$

(4) ${}^{eg\text{-}e}w_A(t+1) = 1 / \{ 1+\exp[-({}^{eg\text{-}e}d_A(t+1) - {}^{eg}bound/2)] \};$

(5) $({}^{eg}a_A, {}^{eg}b_A)$ is updated according to the agent's movement as depicted in Figures 9 and 10 (Points G2, T3).

--- Insert Figure 9 about here ---

--- Insert Figure 10 about here ---

--- Insert Figure 11 about here ---

The allocentric component update. In every time step, every vector ${}^{all}v_{A,B} \in {}^{all}V$ is updated as follows:

(6) $${}^{all}d_{A,B}(t+1) = {}^{all}d_{A,B}(t) + {}^{eg\text{-}t}w_A(t) \cdot {}^{eg\text{-}t}w_B(t) \cdot {}^{all}inc,$$

where ${}^{eg\text{-}t}w_A(t)$ and ${}^{eg\text{-}t}w_B(t)$ are transient weights of egocentric vectors to objects A and B in time t (Points G3, G5);

(7) $${}^{all}w_{A,B}(t+1) = 1 / \{ 1 + \exp [-({}^{all}d_{A,B}(t+1) - {}^{eg}bound/2)] \};$$

note that (7) is a shifted standard logistic sigmoid (Figure 11).

3.3.4 Action selection. The action selection mechanism of our agent enables it to: (a) follow a pre-specified trajectory in order to learn the spatial representations during the learning phase, and (b) perform one of the pointing tasks.

3.3.5 Pointing to objects. Recall that both egocentric and allocentric vectors are stored precisely and that the weight is an abstraction of representation precision (Point G2). In order to fit errors produced by real humans in the two pointing tasks (that is, in their eyes-closed phases and the disorientation phases), we need to devise a pointing mechanism that would "artificially" distort the precise representations based on the vector weights and

additional parameters. These parameters will than be subject to a data-fitting procedure. The point is that we need to fit data from both the experiments by the same setting of the parameters.

Because we have two representations used for the different phases of the pointing tasks, the enduring egocentric and allocentric (Point G4), we need two ways of distorting the precise representations, that is, two pointing mechanisms, one for each representation. Both mechanisms will operate in two steps: the first step models a memory error, and the second step models a motor error. First, when humans retrieve a spatial relation from memory, this relation may be inaccurate (recall that we are not concerned with when and how the memory error has been caused, we are only interested in replicating its magnitude here). Second, when an object's position (possibly inaccurate) is retrieved, humans may make another error during the pointing—the motor error—i.e. they may not point exactly where they intend to point.

Wang and Spelke (2000) measured quantities called the pointing and the configuration error (see Sec. 2.1). In our model the memory error corresponds mainly to the configuration error and the motor error corresponds to the pointing error: the higher the memory error the higher configuration error, and the higher motor error the higher pointing error. However, as Wang and Spelke noted, increases in pointing variability "within objects" (i.e., pointing error) affect measures of pointing variability "between objects" (i.e., configuration error). In the present context, this relationship implies that motor error also affects the configuration error. The configuration error in our model depends therefore on the memory error as well as on the motor error while the pointing error depends only on the motor error. The memory and the motor errors are independent.

In the eyes-closed phase (i.e. when the agent is *not* disoriented; see Sec. 2) the agent uses only the enduring egocentric representation (Points G4, G5). The pointing direction is generated from the enduring weight of the egocentric vector towards the object.

In the disoriented phase, the pointing is driven by the allocentric representation (Point G4). The egocentric representation is ignored in this phase. The agent is told which object it is facing. The imagined direction to the object that the agent has to point at, i.e. the memory error, is created by distorting vectors constituting a particular path towards this object in the allocentric representation (Figure 13). This path is chosen using a heuristic that attempts to make the pointing performance as precise as possible (i.e. the "shortest" path in terms of precision is found). Again, the vector weights influence the precision. The final pointing direction is deduced from the imagined direction towards the object with added motor error.

Formal description of pointing to objects. The pointing mechanism is a quadruple <M, E, D, C> where M is a an agent's model of spatial representation described above, E is the pointing mechanism used by the agent while the agent is oriented to the surrounding environment (eyes-closed phase, after rotation), D is the pointing mechanism used by the agent when it is disoriented (disoriented phase), and C is the configuration of pointing mechanisms E and D.

Configuration C contains four parameters determining:

- ${}^{eg}\sigma_{mem}$: a parameter responsible for the amount of the memory error while the agent is inferring the pointing direction from the egocentric representation (this parameter will be used by E);
- ${}^{eg}\sigma_{mot}$: a parameter responsible for the amount of the motor error in inferring the pointing direction from the egocentric representation (this parameter will be used by E);

· ${}^{all}\sigma_{mem}$: the first parameter responsible for the amount of the memory error in inferring the pointing direction from the allocentric representation (this parameter will be used by D);

· ${}^{all}\sigma_{mot}$: a parameters responsible for the amount of the motor error in inferring the pointing direction from the allocentric representation (this parameter will be used by D).

We note that all of these parameters relate only to the distortion of an accurate representation during the pointing; these parameters do not influence how the representation is built during learning. During learning, the topology of the room, the positions of objects, the agent's trajectories and parameters ^{all}inc, ${}^{eg\text{-}t}$inc, ${}^{eg\text{-}t}$dec, ${}^{eg\text{-}e}$inc, and ${}^{eg\text{-}e}$dec can influence how a representation is built.

These parameters will be subject to data-fitting procedure described in the next section.

Pointing mechanism E (eyes-closed phase). Let the agent be pointing to object A after a rotation. Let ${}^{eg}u_A$ denote the egocentric vector to the object A and α the accurate angle between the agent's current heading and the vector ${}^{eg}u_A$. The pointing direction is computed in two steps. First, imagined angle α' representing inaccurately retrieved pointing angle from the egocentric representation is generated with respect to normal distribution with the mean value α and the standard deviation ${}^{eg}\sigma_{mem}$ and further distorted depending on the enduring weight ${}^{eg\text{-}e}w_A$ of the vector ${}^{eg}u_A$ as described in the next paragraph. Let us denote the difference α - α' as ε. Note that ε actually corresponds only to that part of the total pointing direction's error that is caused by the noise in the processes of perception the object, storage of the memory trace of the object, maintenance of the trace and recall (but not execution of the motor gesture; see

Point G2). Thus, second, the final pointing direction is generated with respect to normal distribution with mean value α' and standard deviation ${}^{eg}\sigma_{mot}$.

How should the ${}^{eg\text{-}e}w_A$ influence ε? Because we are unaware of empirical results that would help us to answer this question with a reasonable degree of certainty, for parsimony, we used as simple mechanism as possible, which we derived based on the following considerations. First, when the weight is 0, the pointing should be entirely random; i.e., ε should be derived from the uniform distribution (-180, 180). Second, when the weight is 1, ε should be minimal, and this applies for subjects with *truly over-learned* representations. It seemed implausible to us that ε is zero even for such subjects. Instead, we assume that ε behaves like a normally distributed random variable and use parameter ${}^{eg}\sigma_{mem}$ to denote its standard deviation. Finally, for weights larger than 0 but smaller than 1, we use a linear combination between these two extremes as follows. Let *U* denotes the uniform distribution (-180, 180) and *N* denotes the normal distribution with the mean 0 and the standard deviation ${}^{eg}\sigma_{mem}$. The value of ε is generated by the *linear blend* function (Fig. 12):

$$\varepsilon({}^{eg\text{-}e}w_A) = \text{blend}({}^{eg\text{-}e}w_A) = (1 - {}^{eg\text{-}e}w_A) \,.\, U + {}^{eg\text{-}e}w_A \,.\, N \quad (8)$$

In the experiments we modeled, subjects should have studied the objects layout "as much as they can", but only over a relatively short period. Thus, the layout was most likely not truly over-learned by these human subjects, but we can assume that it was *nearly over-learned*, which we operationalize for the egocentric subsystem's purpose as follows:

(9) the average of the six enduring egocentric bases is between 0.8 and 0.95 at the end of a particular learning phase; all enduring egocentric bases ${}^{eg\text{-}e}d_A$ are larger than or equal to 5 at the end of a particular learning phase; at least 2 enduring egocentric bases ${}^{eg\text{-}e}d_A$ are larger than or equal to 9 at the end of a

particular learning phase; at most 2 enduring egocentric bases ${}^{eg\text{-}e}d_A$ are larger than or equal to 9.75 at the end of a particular learning phase.

Note that due to (8) and (9), we can fit parameter ${}^{eg}\sigma_{mem}$ even if we do not have data for subjects with truly over-learned objects' layouts in the settings of Waller & Hodgson and Holmes & Sholl. This will be detailed in Sec. 4.

--- Insert Figure 12 about here ---

Pointing mechanism D (disoriented phase). At the beginning of the disoriented phase, the agent "imagines" that it is heading towards one of the objects (chosen randomly). This corresponds to the experimental procedure of Waller and Hodgson (2006), see Sec. 2.3 and 4.4. Let the agent has imagined heading to object A while standing at position P, and let it point towards object B. The pointing direction is generated in three steps (Figure 13).

First, the self-to-object-A vector ${}^{eg}v'_A$ is inferred by the agent using the notion of its position in the allocentric map. Assuming the agent knows where it stands (Point T1) and that it is heading to object A (this accords with the method of the *in silico* experiment - see Sec. 4.4), the estimation of the angle (0°) is trivial and we only need to estimate the vector' length. The length is generated with respect to normal distribution with the mean given by the length of the real self-to-object-A vector ${}^{eg}v_A$. The standard deviation of the distribution is calculated as in the case of estimating the length of allocentric vectors described below. For this purpose we defined the vector ${}^{eg}v_A$ as fully learned with its weight ${}^{eg}w_A=1$. Then the estimated vector

${}^{eg}v'_A$ is aligned with the allocentric map, that is, it is rotated by the angle given by the angle between the agent's heading and the south-north axis of the environment (Point T5).

Second, a distorted allocentric vector ${}^{all}u_{A,B}$ from object A to B representing an inaccurately retrieved vector from the allocentric representation is calculated. This is done by vector addition of vectors $\{{}^{all}u'_{Ai,Aj} \mid i = 1..n-1; j = i+1 \}$ estimated from allocentric vectors $\{{}^{all}u_{Ai,Aj} \mid i = 1..n-1; j = i+1\}$. The vectors ${}^{all}u_{Ai,Aj}$ constitute the "best" path in the allocentric map of length n between objects A and B (going through objects A_i, where the object A_1 corresponds to A and A_n corresponds to B). The "best" path is selected as follows. All possible paths without cycles between objects A and B are ordered according to these criteria:

a) any path P_1 of length k+1 is better than path P_2 of length k if and only if the weight of the weakest vector of P_1 is twice as high as the weight of the weakest vector in path P_2;

b) if paths P_1 and P_2 has the same length, s_1 is the weakest vector of P_1 and s_2 is the weakest vector of P_2, than P_1 is better than P_2 if and only if s_1 is stronger than s_2.

The best path of these is chosen. The idea behind is that weak vectors and/or long paths lead to a greater memory error. The procedure above is a heuristic estimate of a path that might lead to smaller configuration error than other paths. Point (a) serves to estimate whether a longer path (counted in number of vectors) might be better than a shorter path and Point (b) compares two paths of an equal lengths. Obviously, these criteria are arbitrary. Note, however, that it is not possible to compute the shortest path using a standard algorithm for path searching, such as Dijkstra, for it is not clear what the distance between two objects should be. Setting the distance between two objects to 1-w, where w is the vector's weight, does not work very well. Such mechanism would propose that a path of the length 2

containing vectors weighting 0.99 and 0.01 is shorter than a path of the length 3 containing vectors weighting 0.6, 0.6, and 0.6.

The direction and the length of the vector ${}^{all}u'_{Ai,Aj}$ are estimated from the direction and the length of the retrieved (exact) vector ${}^{all}u_{Ai,Aj}$. Conceptually, the vector ${}^{all}u'_{Ai,Aj}$ is created by rotation of the vector ${}^{all}u_{Ai,Aj}$ and changing the length of the rotated vector (using two normally distributed random variables). This process seemed natural to us. Other option would be to add a normally distributed random variable to the two Cartesian components of the vector; this option seemed slightly less natural to us. Additionally, the former approach allows for different parameterization of estimating angular and distance errors, which is a feature that may be useful in future. Therefore, we had chosen the former approach.

More specifically, the direction of each vector ${}^{all}u'_{Ai,Aj}$ is estimated with respect to normal distribution with the mean given by the direction of the retrieved (exact) vector ${}^{all}u_{Ai,Aj}$ and standard deviation σ^{d}_{j}, where

$$\sigma^{d}_{j} = {}^{all}\sigma_{mem} \,.\, [1/({}^{all}w_{Ai,Aj})^{2}] \qquad (10)$$

The higher the weight is, the lower σ^{d}_{j} is. When the weight is equal to 1, σ^{d}_{j} is equal to ${}^{all}\sigma_{mem}$.

The length of each vector ${}^{all}u'_{Ai,Aj}$ is estimated with respect to normal distribution with the mean given by the length of the retrieved vector ${}^{all}u_{Ai,Aj}$ with standard deviation σ^{l}_{j}. The standard deviation σ^{l}_{j} is computed as follows. Let α denotes the angle defining a circular sector representing 95% confidence interval of the estimated *directions* as depicted on Figure 14. Let d is the shortest distance between the exact position of the object A_j and an edge of the circular sector given by α. Let now u denotes the length of vector ${}^{all}u_{Ai,Aj}$. The standard deviation σ^{l}_{j} is computed so that the range u±d is a 95% confidence interval of estimated

lengths. Because the distance d depends on u (in a specific way), the longer the original vector is, the more the estimated vector's length can differ from the real length.

At the end of the second step, the rotated egocentric vector ${}^{eg}v'_A$ calculated above in the first step is added to ${}^{all}u_{A,B}$, giving us a new (imagined) self-to-object-B vector ${}^{eg}v'_B$. In the third step, the final pointing direction is generated with respect to normal distribution with the mean given by the direction of ${}^{eg}v'_B$ and with the motor standard deviation ${}^{\text{all}}\sigma_{\text{mot}}$.

--- Insert Figure 13 about here ---

--- Insert Figure 14 about here ---

A note on the pointing mechanisms. In the *in silico* experiments, the agent actually points to every object twice. The step of modeling the memory error is conducted just once—we assume the inaccurate estimate from the memory is persistent for a while because the agent does not move. The step of modeling the motor memory is conducted each time the agent is pointing.

4. Experiments

This section describes replication of the experiments of Holmes and Sholl (2005; Exp. 7) and Waller and Hodgson (2006; Exp. 1) *in silico*. As explained in Sec. 2.4, in the conditions of Holmes and Sholl, the agent can perceive more objects at the same time than in the conditions of Waller and Hodgson because in the latter setting, the testing booth blocks the agent's view of much of the environment. Thus, **our hypothesis** is that the model will develop more accurate object-to-object spatial representation in the Holmes and Sholl's setting and

consequently, we will be able to replicate quantitatively the data of Waller and Hodgson yet qualitatively the data of Holmes and Sholl. On a more general level, the hypothesis postulates that the underlying cause is that the allocentric representation will develop more rapidly in the setting of Holmes and Sholl than in the setting of Waller and Hodgson and, in addition, the allocentric representation will develop more rapidly in the setting of Waller and Hodgson without the booth than in the setting of Waller and Hodgson with the booth.

Note that our hypothesis concerns mainly the allocentric representation. In fact, we have two main goals:

1) To find a particular parameters' setting for our model that would generate data replicating data from both the eyes-closed and disoriented condition of Waller & Hodgson quantitatively, and, at the same time, replicating the fact that the configuration error decreased in the disoriented condition of Holmes & Sholl comparing to their eyes-closed condition. Thus, we are interested in replicating the "reverse trend" in Holmes & Sholl conditions, not in exact replication of their results. There are differences e.g. between data from eyes-closed condition of Waller & Hodgson and eyes-closed condition of Holmes & Sholl (cf. Fig. 2 and 4) which are most likely caused by subtle differences in experimental procedures we do not model, e.g. differences in pointing devices or usage of different objects (with possibly different shapes and sizes). Additionally, we want to inspect an "artificial" condition of Waller & Hodgson without the booth. We are interested in whether the configuration error in the disoriented phase decreases in that "artificial" environment comparing to the disoriented condition of the original Waller & Hodgson.

2) To investigate how the model's behavior changes when values of its two particular parameters are being changed. These parameters are ${}^{all}inc$ and ${}^{all}\sigma_{mem}$ and their common denominator is that they are directly related to generating errors in pointing directions in the

disoriented phases of the experiments. Why are we interested in these two parameters only? As detailed next, the model has 10 parameters to be set, giving us extremely large 10-dimensional parameter-space. On the one hand, it is actually not very difficult to find one point in this parameter-space that would replicate the data as intended by Goal (1). It is more interesting whether behavior of the model would support our hypothesis within a broader range of parameter settings. On the other hand, it would be very hard if not impossible to investigate the whole parameter-space. Moreover, the preponderance of this space is likely to be psychologically meaningless. Thus, we will investigate its sub-space that is relevant for our main aim—how quickly the allocentric representation (but not the egocentric representation) develops under different experimental conditions (i.e., Waller & Hodgson with and without the booth, and Holmes & Sholl). For this aim, parameters ^{all}inc and ${}^{all}\sigma_{mem}$ are crucial. Note that by changing values of these parameters, we cannot expect to fit the real data precisely anymore, but we can observe general trends, such as "whether the allocentric representation develops more rapidly in one environment than in another", or "whether the configuration error is always higher for in one environment than in another."

Before the experiment starts, two issues have to be addressed. First, we must generate or acquire trajectories along which the IVA walks while learning the layout. Unfortunately, neither of the experiments reported the nature of their participants' trajectories. Second, we must determine the order in which we will fit the parameters and explain further constraints on the parameters. The model requires two sets of parameters to be set: a) configurations ^{eg}C and ^{all}C of the egocentric and the allocentric component, respectively (6 parameters), and b) configuration C of the pointing mechanism (4 parameters). The process of parameter setting

must be plausible, which requires subjecting the parameters and the order of their fitting to some constraints.

Our next steps will be as follows:

1) We will set the parameters for the speed of learning and forgetting of transient egocentric weights, ${}^{eg\text{-}t}$inc and ${}^{eg\text{-}t}$dec, respectively, based on an *a priori* assumption on behavior of these weights. These parameters should be set first because both allocentric subsystem and enduring part of the egocentric system uses transient egocentric weights as input. For similar reasons, we will also set ^{eg}bound. Sec. 4.1 details this step.

2) We will now generate trajectories along which the IVA will walk. This process is detailed in Sec. 4.2 where it is also explained that values of ${}^{eg\text{-}t}$inc, ${}^{eg\text{-}t}$dec and ^{eg}bound are already needed in this step.

3) We will then find parameters ${}^{eg\text{-}e}$inc and ${}^{eg\text{-}e}$dec based assumptions given in Def. (9) (Sec. 3.3.5) and using the trajectories generated in the previous step. Sec. 4.3 details this step. At that point, five parameters of the configuration ^{eg}C will be known. None of them will depend on the results from the real experiments.

4) We will then explain the procedure of the *in silico* experiments (Sec. 4.4), including fitting the parameter ${}^{all}\sigma_{mot}$ based on the Waller and Hodgson's data.

5) We will present results concerning Goal (2). That is, we will investigate behavior of the model in disoriented phases of the three environments when ^{all}inc and ${}^{all}\sigma_{mem}$ are altered but the other parameters kept fixed. Note we intentionally start with Goal (2) and proceed to Goal (1).

6) We will present results concerning Goal (1). In particular, we will find parameters of ^{all}inc, ${}^{all}\sigma_{mem}$, ${}^{eg}\sigma_{mem}$ and ${}^{eg}\sigma_{mot}$ based on the Waller & Hodgson's data, and use the model

with those parameter values to generate data for the replica of Holmes & Sholl and for the Waller & Hodgson's environment without the booth.

4.1 Setting the parameters of the transient part of the egocentric module ($^{eg\text{-}t}$inc, $^{eg\text{-}t}$dec, ^{eg}bound)

We considered it reasonable that the transient weight of an egocentric vector towards an object increases from 0 to 1 during a dozen of seconds when the agent is looking at the object, and decreases from 1 to 0 when the agent is not looking at the object. The latter period should be a bit longer. In both cases, these periods should be an order of magnitude shorter than the whole learning phase, which accords well with the required short-term memory characteristics of that part of the egocentric representation (Point G5 from Sec. 3.2).

Thus, we decided that increase of a transient egocentric vector *base*'s from 0 to ^{eg}bound would last five agent's steps while the decrease from ^{eg}bound to 0 would last 7 steps. We preferred to use the base instead of the weight in this definition due to the non-linear speed of strengthening and weakening of weights caused by the sigmoid function (Eq. 2). The chosen parameters are given in Tab. 1 and were held fixed throughout. The parameter ^{eg}bound was set to 10 in order to limit the maximal weight to approximately 0.995. Beyond the value of 10, the sigmoid function is nearly constant, which means that if the base is allowed to increase beyond 10, the weight almost would not change. However, for large values of the base, decaying of the weight would take too long given the duration of the experimental run.

4.2 Generating the trajectories

Because the trajectories of human subjects in the modeled experiments are unknown to us, we needed to generate trajectories along which the IVA will walk in the *in silico*

experiments artificially. To generate as plausible trajectories as possible, we used the following approach. One of the authors (J. V.) first generated 40 trajectories (20 for each environment) manually by walking the IVA around the room, as detailed below. Another author (D. W.) and four of his collaborators then rated these paths for plausibility. Each of the raters had previously served as an experimenter in a study on spatial representation and had direct experience with human participants learning room-sized layouts. From this set, ten of the most plausible trajectories were chosen for each environment for the *in silico* experiments (see Fig. 15, 16).

--- Insert Figure 15 about here ---

--- Insert Figure 16 about here ---

For generating each of the 40 trajectories, the experimenter (J. V.) chosen a different means of exploring the room (e.g. walking closely around the box clockwise, walking around the walls of the room in both directions etc.). These means corresponded to how the experimenter thought a human subject would behave. Additionally, we needed to solve the question how long each trajectory should have been? Recall that in the real experiments, subjects should have learnt the layout of the environment "as much as they could." We needed to find a similar "stop-rule" for the IVA. The rule we choose was to stop the IVA when the strength of allocentric vectors began to decelerate; i.e. they either approached saturation or remained weak. In other words, the bases corresponding to respective weights belonged to intervals in which the sigmoid function (Figure 11) was not steep.

The motivation for this rule was twofold. First, because of the sigmoid function used in learning and due to lack of allocentric forgetting, one of the following two things may happen with allocentric weights (in most cases). We may expect a weight to start to build up slowly, then increase relatively quickly and then slowly approach saturation. In this case, from a point, the weight does not change much, a "stop-rule". Or a weight may remain weak; in this case, it does not change much from the beginning.

Second, we had chosen the allocentric representation and not the enduring egocentric representation for the "stop-rule" because the former is more complex than the latter and it is much harder to build it up in a typical case. Recall that an allocentric vector between two objects increases only when transient weights of egocentric vectors towards both objects are strong, a case in which both *enduring* egocentric weights also increase (Eq. 3, 4). Because the allocentric representation is in our main focus, we needed the allocentric representations to be reasonably learnt at the end of the learning phase. This would not have been guaranteed when the enduring egocentric representation was used for the "stop-rule." But note also that allocentric representations should not be over-learned at the end of each trajectory (Fig. 17). All weights saturated for every trajectory would mean ceiling effect in the pointing task: the IVA would behave the same in the pointing task notwithstanding the particular learning trajectory.

--- Insert Figure 17 about here ---

For the purpose of the "stop-rule," four parameters must have been set: $^{eg\text{-}t}$inc, $^{eg\text{-}t}$dec, ^{eg}bound and ^{all}inc. The first three were already chosen (Sec. 4.1) but the last one was not. That parameter will be subject to further trialing. Thus, we could use only its "preliminary" value for the purpose of generating the trajectories; nevertheless, this value must be plausible. Our guiding idea was as follows. Imagine the IVA is looking at two objects without interruption. Provided the transient egocentric weights of vectors towards these two objects are 0 at the beginning, the base of an allocentric vector between these two objects should increase from 0 to ^{eg}bound in time period approximately three times longer than it takes the transient bases of the egocentric vectors towards these two objects to increase from 0 to ^{eg}bound. Based on this idea, the value 0.9 was chosen. Note that plausible values stemming from the fitting of the real data will be a bit lower but still similar. Note also that in practice, the build-up of the allocentric base takes usually longer, because the agent often does not look at two objects without interruption so long and the transient egocentric bases decreases quickly and must be re-built before the allocentric weight starts to increase notably again.

4.3 Setting the parameters of the enduring part of the egocentric module ($^{eg\text{-}e}$inc, $^{eg\text{-}e}$dec)

With the parameters $^{eg\text{-}t}$inc, $^{eg\text{-}t}$dec and ^{eg}bound from Sec. 4.1, we let the IVA walk along the 10 generated trajectories in the environment of Waller & Hodgson and set the parameters $^{eg\text{-}e}$inc, $^{eg\text{-}e}$dec and ^{eg}bound so that assumptions given in Def. (9) (Sec. 3.3.5) holds for the most of individual enduring egocentric representations as possible. The parameters chosen are given in Tab. 1. For such parameters, the assumptions given in Def. (9) holds for most (9/10) of individual representations. The remaining representation had one base 3.47 only. Consequently, the average of bases was 7.68 and only one base was larger than 9. However, this is still reasonably close to Def. (9).

We then checked how the weights of enduring egocentric representations were developed at the end of learning of the 10 trajectories with several slightly different settings of $^{eg\text{-}e}$inc, $^{eg\text{-}e}$dec. The model behaves coherently; mild changes of the parameter values result in mild changes of the weights.

--- Insert Table 1 about here ---

4.4 Procedure

The procedures for the simulation resembled those of the real experiments as much as possible, although small differences, e.g. in usage of different pointing devices (a PC USB Light Gun vs. a joystick), were not accounted for.

In the simulation of both experiments we tested 40 agents. Each agent followed one of the predefined paths; each path was used for four agents. During these walks the agents developed their spatial representations. We replaced the first phase of the original experiments, which was used to determine whether a participant has learned the layout enough, by the described methodology of generating paths (Sec. 4.2).

When modeling the second, eye-closed, phase we walked the agent along the defined trajectories at the pointing place (the middle of the chamber or the middle of the room, see Figure 1, 3) and let it point to every object twice. In this phase the agent used the egocentric representation for inferring the pointing directions.

When modeling the third, disoriented phase, the agent chose an imagined heading to a random object (uniform distribution), and then pointed to every object twice. The disorientation by rotation a human slowly was replaced by disabling the egocentric representation (Point G4 from Sec. 3.2): the agent used only the allocentric representation in this phase.

In the first experiment, we examined the influence of the parameters ${}^{all}inc$ and ${}^{all}\sigma_{mem}$ on configuration error in Experiment 1 of Waller and Hodgson (2006), Experiment 7 of Holmes and Sholl (2005), and Experiment 1 of Waller and Hodgson (2006) without the booth. First, we fitted the motor standard deviation ${}^{all}\sigma_{mot}$ in order to reach the pointing error as in the Experiment 1 of Waller and Hodgson (2006). Because the motor error is independent of memory representations (Point G2, Sec. 3.2), we basically found the value of ${}^{all}\sigma_{mot}$ so that the pointing error reported by Waller and Hodgson (Exp. 1; 2006) was reproduced as closely as possible. To this end, we let each of the 40 agents point to each of the six objects twice assuming that the vector towards the respective object retrieved from the memory is precise and the total error in pointing behavior is thus generated only through the motor error. The value chosen was ${}^{all}\sigma_{mot} = 15$. Next, we simulated for each experiment the learning and the disoriented phase 234 times with ${}^{eg\text{-}t}inc$, ${}^{eg\text{-}t}dec$, ${}^{eg\text{-}e}inc$, ${}^{eg\text{-}e}dec$, and ${}^{eg}bound$ given in Tab. 1, varying ${}^{all}inc$ and ${}^{all}\sigma_{mem}$. For ${}^{all}inc$, we iterated from 2 to 12 with the step 2, and for ${}^{all}\sigma_{mem}$, we iterated from 0.1 to 2 over 0.05, testing every combination of the two parameters and collecting the results. Note that we did not simulate the eyes-closed phase since we were varying only parameters related to allocentric pointing.

In the second experiment, we started with simulating the eyes-closed phase of Experiment 1 of Waller and Hodgson (2006) with ${}^{eg\text{-}t}inc$, ${}^{eg\text{-}t}dec$, ${}^{eg\text{-}e}inc$, ${}^{eg\text{-}e}dec$, and ${}^{eg}bound$ given in Tab. 1 in order to fit two parameters that influence configuration error in the eyes-

closed phase, i.e. ${}^{eg}\sigma_{mem}$, ${}^{eg}\sigma_{mot}$. These parameters were fitted manually so that the results of the configured model correspond to results of the real experiment as much as possible. The fitted values are ${}^{eg}\sigma_{mem}=8$ and ${}^{eg}\sigma_{mot}=13$ (note that ${}^{eg}\sigma_{mot}$ can be fitted without knowledge of the developed egocentric representations similarly to ${}^{all}\sigma_{mot}$). Then we picked several values of ${}^{all}\sigma_{mem}$ and ^{all}inc from our first experiment that led to a similar configuration error in the simulation of Experiment 1 of Waller and Hodgson (2006) as in the real experiment (using ${}^{all}\sigma_{mot} = 15$). Finally we ran both phases of Experiment 7 of Holmes and Sholl (2005) and Experiment 1 of Waller and Hodgson (2006) without the booth with the same parameters setting and collected the results.

4.5 Results

Figure 18 depicts the results of the **first experiment**. We clearly see that the configuration error controlled for the pointing variability was higher in the conditions of Waller and Hodgson than in the conditions of Holmes and Sholl when matched for corresponding ^{all}inc and ${}^{all}\sigma_{mem}$ (Wilcoxon signed rank test, $V=26928$, $p<0.001$).[6] Similarly, in the scenario with a booth, the mean configuration error controlled for the pointing variability was higher with the booth than without it (Wilcoxon signed rank test, $V = 27239$, $p<0.001$).

In general, for very low ^{all}inc values, a large configuration error was apparent in all conditions; these values are implausible since they result in underdeveloped representations. On the other hand, the configuration error approached an asymptote for larger values of ^{all}inc and did not decrease further; a ceiling effect and a sign of an over-learned representation. Arguably, the most meaningful values are between 0.5 and 1.

Our **second experiment** demonstrated that for some parameters ^{all}inc and ${}^{all}\sigma_{mem}$, first, the results fitted Waller and Hodgson (2006, Exp. 1) where the disorientation effect was

present, second, the “reverse” desorientation effect was observed in the settings of Holmes and Sholl (2005, Exp. 7) and, third, the configuration error in the desoriented phase decreased when the booth is removed in the condition of Waller and Hodgson (2006, Exp. 1).

The configuration and pointing errors in the eyes-closed phase of the replicas of the three experimental conditions are given in Table 2. The configuration and pointing errors in the disoriented phase of the replicas of the three experimental conditions are given in Table 3 for the three different values of ${}^{all}inc$ and ${}^{all}\sigma_{mem}$.

--- Insert Table 2 about here ---

--- Insert Table 3 about here ---

--- Insert Table 4 about here ---

The results of the replication of Waller and Hodgson (2006, Exp. 1) from both phases match the values observed in the real experiment (configuration error 15.96° and 20.12°, respectively, and pointing error 10.79° and 12.05°, respectively).

In replication of Holmes and Sholl we observe lower values than the values reported by Holmes and Sholl (2005, Exp. 7) for both phases; however, for the values of the parameters ${}^{all}inc$ and ${}^{all}\sigma_{mem}$ from Tab. 3, we observe a similar trend to lower errors in disorientation phase comparing to the eyes-closed phase, which is the key result of the second experiment. Additionally, corrected as well as uncorrected configuration errors in both phases of the Holmes and Sholl replica and the condition of Waller and Hodgson without the booth are lower then the respective configuration errors in the replica of the original Waller and

Hodgson with the booth, suggesting that the presence of booth influence the quality of the developed representations.

The results of t-tests for a decrease in the configuration errors when controlling for the pointing variability are shown in Table 4. This decrease is significant for small ${}^{all}\sigma_{mem}$.

For illustrative purposes, comparison of configuration errors in both phases for the two original experiments and the three replicas is given by Fig. 19 for ${}^{all}\sigma_{mem}=4$ and ${}^{all}inc=0.71$. Examples of allocentric representations for the trajectories depicted on Figures 15a and 16a when the allocentric component is configured with ${}^{all}inc=0.71$ are given on Figure 20.

--- Insert Figure 19 about here ---

--- Insert Figure 20 about here ---

4.6 Discussion

Results support our hypothesis. First, while the configuration error increases in the disoriented phase for replication of Waller and Hodgson, this error decreases for replication of Holmes and Sholl for the three particular parameter settings. Second, more importantly, the configuration error tends to be higher in general for the condition with the booth then for both conditions without the booth when matched for corresponding ${}^{all}inc$ and ${}^{all}\sigma_{mem}$. Figure 18 clearly shows that the presence of the booth results in slower decrease of the configuration error when ${}^{all}inc$ increases. Basically, the build-up of the allocentric map is slower with the booth. Importantly, because the model's parameters are the same in all the three experimental

conditions, the only plausible possibilities for explaining the difference are: (a) different learning trajectories and (b) the booth. Because the learning trajectories for the three simulations were generated by the same process, and in fact they were the same for the with-booth and without-booth simulations in the setting of Waller and Hodgson, it seems most likely that the primary reason for the difference in results is the presence of the testing booth and hence the visibility of the environment during learning. The following reasoning underpins this opinion.

In the model, the allocentric map is ultimately generated based on the co-occurrence of objects in the perceptual field. More objects in the perceptual field at the same time leads to better learning. Because there are presumably more objects seen at one time when there is no booth as opposed to when the booth is present, the resulting allocentric maps tend to be more complete in the replications of Holmes and Sholl but sparser in the replications of Waller and Hodgson (cf. Figures 20 left and right) (by the terms complete and sparse graphs we mean the terms of the graph theory in mathematics). To illustrate this further, we compared the average sum of weights over one allocentric map in Holmes and Sholl and Waller and Hodgson without the booth to the average sum of weights over one allocentric map in Waller and Hodgson with the boot for the model calibrated with parameters from Tab. 1 and the fitted parameters ${}^{all}\sigma_{mem}$=4 and ^{all}inc=0 .71. The weights are indeed higher in the former two settings than in the without-booth condition ($t(298) = 4.802$, $p<.001$ and $t(298) = 9.8205$, $p<.001$, respectively).

These considerations lead to experimentally testable prediction that, other variables held constant, trajectories along which a human can perceive more objects at the same time (on average) will lead to better allocentric pointing. Additionally, the model also predicts that a longer learning phase will lead to improved performance up to a point; however, it is likely

that additional parameters would have to be modeled in order to simulate the asymptotic learning that would be characteristic of people in this setting. Thus, in our opinion, empirically examining hypotheses about number of perceived objects will be more promising than examining the model's predictions about the path length or learning time.

Recall from Sec. 3.2 that the model makes several simplifying assumptions, e.g. the alignment effect, the effect of grouping, representing distances among objects and walls, looking around during learning, long-term forgetting etc. However, as long as these parameters influence the memory processes similarly in the two experiments we fitted, the different views produced by walking along different trajectories (thereby different objects seen) seem to remain the only cause influencing the speed of learning.

Finally, it is worthwhile to consider why the simulation did not replicate Holmes and Sholl's (2005) results exactly. First, note that there is a large difference in configuration errors in the eyes-closed phase of the original experiments (Figure 19). Arguably, this difference was caused due to subtle differences in the experimental procedures, such as different objects used or different pointing devices. We have intentionally avoided modeling these differences as they would require the creation of new methodological parameters that do not contribute to our understanding of spatial cognition. Second, in our model, they are the parameters ${}^{\text{eg-t}}\sigma_{\text{mem}}$, ${}^{\text{eg-t}}\sigma_{\text{mot}}$ that are used for generating the configuration error in the eyes-closed phase and the error is reproduced by adding a gaussian-based noise to the precise egocentric representation, a noise that grows with lower egocentric weights. Because these two parameters are the same for Waller and Hodgson's and Holmes and Sholl's conditions, because there the same number of objects in both the experiments, and because there are presumably more objects seen at one time when there is no booth as opposed to when the booth is present, which means that more enduring egocentric vectors can be strengthened at

the same time, it is really no surprise that the magnitude of the error in the eyes-closed phase is similar in both the *in silico* replicates (in fact, it is even lower for Holmes and Sholl's condition comparing to the Waller and Hodgson's condition with the booth). What is important though is how quickly the allocentric representations are built up and whether the configuration error decreases or increases in the disorientation phase compared to the eyes-closed phase. Concerning the first point, the booth clearly slows down the development of the representations, and concerning the second point, we have got the same trend as in the original experiments. Fitting the exact numbers with new parameters invented in an *ad hoc* manner would be trivial but bring nothing new.

5. General discussion and future directions

In this paper, we have presented a phenomenological computational model of allocentric and egocentric memory for locations of objects in a human's proximal surrounding. The model has been integrated within a simple intelligent virtual agent, calibrated using data of one psychological experiment that employed the paradigm of pointing, and used to produce data of another experiment employing the same paradigm. The two real experiments produced contradictory findings concerning the so-called disorientation effect, and we were able to qualitatively replicate these contradictory findings *in silico* with equally calibrated model. Because the model was internally the same in both the replicates, the difference must have been caused by external conditions. Our results and the analysis support the idea that the notable difference in the experimental procedures of the two real experiments is the agent's access to visual information during learning, which differed between the two experiments because of an occluding booth. This idea is further supported by results of another *in silico* experiment where we manipulated directly the presence of the booth in one of the experimental environments.

These findings have several important implications for current theories of human spatial representation. First and most generally, the present results lend further support to the existence (and sufficiency) of two distinct psychological systems of spatial representation – one that codes transient spatial information primarily through egocentric relationships, and another that codes enduring spatial information through allocentric relationships. These two systems are posited by many – but not all – theories of spatial representation (see for example Burgess, 2006; Mou et al., 2004; Easton & Sholl, 1995). Theories that do not incorporate an allocentric component (e.g., Wang and Spelke, 2000) may be relatively hard-pressed to account both for the results of Holmes and Sholl (2005) and for the discrepancy between these results and those of Waller and Hodgson (2006) – two sets of findings that the present work was able to fit. Second, and more incisively, the present work suggests specific functional ways in which the two systems of spatial representation interact with each other, trade information, and switch control over behavior. For example, it is significant that the present model requires spatial information in the allocentric system to have been first processed by the egocentric system, rather than being picked-up directly from the perceptual field. It is also significant that the presence of disorientation in our virtual agents determines which system exerts control over pointing. These and other properties of our model suggest specific aspects of psychological functioning that enable researchers to parsimoniously conceptualize the spatial aspects of human cognition. Finally, by successfully instantiating plausible versions of these two systems of spatial representation in a computational model, we have enabled rapid testing of theoretical predictions, as well as providing an extensible platform that can incorporate other behavioral phenomena in spatial learning and cognition and thus facilitate a more rapid development and refinement of theory.

The results of this paper can be also understood from the methodological perspective. We have demonstrated that an intelligent virtual agent can be employed in computational cognitive sciences. Without an IVA acting in space, i.e. walking along different trajectories, we would not have been able to reproduce the real-world data: the IVA's embodiment added the vital detail to the process of data reproduction.

Of course, it can be argued that our IVA is rather simple and that we have not employed many modules of our IVA's architecture (Figure 5). In other work, our IVA and its memory module are more complex than described in this paper. For instance, we have a complex application that can model arbitrary 2D virtual worlds of the size of a flat (Figure 21). Additionally, in our model, more allocentric maps can exist in parallel, each having a different intrinsic axis. This feature of the model reflects the on-going debate about intrinsic axes of the environment (e.g. Mou et al., 2004; Burgess, 2006; Kelly & McNamara, 2008). Moreover, the model also features other kind of vectors: those representing distances between objects and walls. The spatial memory model has another component estimating positions of objects in dynamic environments (think of a pen that can be moved by another agent; Brom et al., 2009), though this has been implemented independently. Finally, our action selection can be more complex than described in this paper, in particular, the agent can do various tasks, move objects, wait and inspect its surrounding.

--- Insert Figure 21 about here ---

Why have we removed all of these features? First, for the sake of parsimony; it is clear that they were not necessary to replicate the effects that we modeled from the literature. Second, in order to simulate these additional features realistically, we would have needed

sound empirical data in order to fix extra parameters, and these data are mostly missing. For instance, during the learning phase, when walking along a trajectory, the IVA can wait for a while or turn around inspecting the objects' layout.

More generally, it is worth noting that one advantage of IVAs is that they force researchers to think in terms of space, actions and events, which leads to thinking about complex models with many features, which in turn helps to identify gaps in knowledge. One can design and implement a complex model using an IVA and then iteratively set some of the parameters, identify some gaps, run new real-world experiments to get new data, refine the model, set new parameters, identify new gaps, etc. In this paper, we have made the first step, showing that this way exists. Future and ongoing work can refine these initial steps.

For researchers willing to follow the similar path, we provide here a summary of main lessons learnt concerning utilizing IVAs for computational cognitive modeling:

1. IVAs work well in relatively complex settings employing space. They can also work well with settings featuring multiple events and action possibilities, but the space (and embodiment) is critical.

2. Input representations should be relatively abstract and high-level (contrary to usage of robotic artifacts).

3. It is generally better to start with a simpler model with a small number of parameters and gradually proceed to more complex models with more parameters. As with other scientific endeavors, Occam's razor should be one's guiding principle.

Taken together, these points argue for employing IVAs in modeling of high-level cognitive processes, such as episodic, spatial or semantic memory (but not low-level procedural memory or iconic memory), or socio-affective processes, human-level complexity

decision making and action selection that involves space (e.g. one need not an IVA for modeling decision making in Wisconsin card sorting tasks). In our opinion, IVAs can become a testing tool complementary to robotic platform. While robots serve well for investigating relatively low-level perceptual and motor processes due to their physical embodiment, IVAs are useful for implementing models of high-level cognitive abilities due to complex environments they can act in and because virtual environments enable by-passing problems stemming from modeling perception and motor behavior at a low level (see Brom, Lukavsky and Kadlec, 2010 for more on this point).

Finally, two things deserve a comment. First, it is worth reflecting on whether computational modeling in cognitive sciences can contribute back to the field of intelligent virtual agents. For instance, can virtual characters from computer games behave more believably (Loyal, 1997) and/or have better cognitive abilities (Doherty & O'Riordan, 2008) when having more plausible computational models? In our opinion, it is disputable whether IVAs would benefit directly from computational cognitive psychology models. Perhaps more plausible is the idea that researchers will be inspired by psychological models and produce their own kind of models that, although simplified from the psychological perspective, will work reasonably well in a broad range of settings (see Brom & Lukavský, 2009 for more on this point). For instance, the subfield of videogame artificial intelligence studying spatial abilities of videogames characters has a long tradition and went so far in a different way than psychology. Most notably, in videogames, researchers and developers invested in computational efficiency and optimality issues (e.g. Axelrod, 2008). It is more plausible that these people will augment their models with some ideas stemming in psychology and sociology (e.g. Pedica & Vilhjamsson, 2009) than that they will start to use psychological

models such as the model presented in this paper. In our opinion, the subfields where the interaction may be most intensive in the next decade are emotion modeling, modeling of social relations and high-level action selection and planning.

Second, can IVAs help to bridge the gap between multiple levels of analysis, e.g. between psychological and neurobiological levels of abstraction? Here, we are mildly optimistic because IVAs' environments are well suited for presenting inputs at various levels of abstraction (which is harder to achieve with robotic artifacts). For instance, an IVA can perceive a 3D virtual reality both in terms of image-based input as well as abstractly (e.g. it can perceive the distance between two objects directly). So far, it has been demonstrated that it is possible to connect within an IVA an abstract cognitive architecture, namely ACT-R, with a neural architecture, namely Leabra (Jilk et al., 2008). In fact, one of our work in progress is to connect our platform Pogamut for fast developing of 3D IVAs (Gemrot et al., 2009) with a simulator Emergent implementing the Leabra (Aisa, O'Reilly, & Mingus, 2007). This would enable researchers to develop IVAs working with image-based input representations.

References

Aisa, B., O'Reilly, R., & Mingus, B. (2007). Emergent Neural Network Simulation System [WWW page], http://grey.colorado.edu/emergent/index.php/Main_Page [29.8.2011].

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 2, 89-195.

Axelrod, R. (2008). Navigation Graph Generation in Highly Dynamic Worlds. In S. Rabin (Ed.), *AI Game Programming Wisdom 4* (pp. 125 – 141). Course Technology, Cengage Learning.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Psychology Press.

Best, B., & Lebiere, C. (2006). Cognitive Agents Interacting in Real and Virtual Worlds. In Sun R. (Ed.), *Cognition and Multi-Agent Interaction* (pp. 187 – 218). Cambridge University Press.

Bratman, M. E. (1987). *Intention, plans, and practical reason*. Harvard University Press.

Brom, C., Pešková, K., & Lukavský, J. (2007). What does your actor remember. Towards characters with a full episodic memory. *Proc. of 4th International Conference on Virtual Storytelling*, LNCS 4871, 89-101.

Brom, C., & Lukavský, J (2009). Towards Virtual Characters with a Full Episodic Memory II: The Episodic Memory Strikes Back. *Proc. Empathic Agents, AAMAS workshop*, 1-9.

Brom, C., Korenko, T., & Lukavský, J. (2009). How Do Place and Objects Combine? "What-Where" Memory for Human-Like Agents. *Proc. 9th Intelligent Virtual Agents,* LNAI 5773, 42-48.

Brom, C., Lukavský, J., & Kadlec, R. (2010a). Episodic memory for human-line agents and human-like agents for episodic memory. *International Journal of Machine Consciousness, 2*(2), 227 – 244.

Brom, C., Burkert, O., & Kadlec, R.(2010b). Timing in Episodic Memory for Virtual Characters. *Proc. of the IEEE 2010 Symposium on Computational Intelligence and Games (CIG2010)*, 305-312.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends Cogn. Sci., 10* (12), 551-557.

Burgess, N. (2007). Computational models of the spatial and mnemonic function of the hippocampus. In Anderson, P., Morris, R., Amaral, D., Bliss, R., & O'Keefe, J. (Eds.), *The Hippocampus Book* (pp. 715-750). Oxford University Press.

Champandard, A. (2008). Getting started with decision making and control systems. In S. Rabin (Ed.), *AI Game Programming Wisdom 4* (pp. 257-264). Course Technology, Cengage Learning.

Cheng, K., & Gallistel, C. R. (1984). Testing the geometric power of an animal's spatial representation. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 409 – 423). Hillsdale, NJ: Erlbaum.

Cutting J. E., & Vishton P. M. (1995). Perceiving Layout and Knowing Distances. The Integration, Relative Potency, and Contextual Use of Different Information about Depth. In *Perception of Space and Motion* (pp. 69 –117). Academic Press.

Doherty, D., & O'Riordan C. (2008). Toward More Humanlike NPCs for First-/Third-Person Shooter Games. In S. Rabin (Ed.), *AI Game Programming Wisdom 4* (pp. 499-512). Course Technology, Cengage Learning.

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 483 – 500.

Epic Epic Games (2004). UnrealTournament 2004 [WWW page]. http://www.unrealtournament.com [29.8.2011].

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gebhard, P., Schröder, M., Charfuelan, M., Endres, C., Kipp, M., Pammi, S., Rumpler, M., & Türk, O. (2008). IDEAS4Games: Building Expressive Virtual Characters for Computer Games. *Proc. of the 8th International Conference on Intelligent Virtual Agents (IVA'08)*, LNCS 5208, 426-440.

Gemrot, J., Kadlec, R., Bída, M., Burkert, O., Píbil, R., Havlicek, J., Zemčák, T., Šimlovič, J., Vansa, R., Štolba, M., Plch, T., & Brom, C. (2009). Pogamut 3 Can Assist Developers in Building AI (Not Only) for Their Videogame Agents. *Agents for Games and Simulations*, LNCS 5920, 1-15.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Muffin.

Hermer, L., & Spelke E.j (1996). Modularity and development: the case of spatial reorientation. *Cognition, 61*, 195 – 232.

Hodgson, E., Waller, D. (2006). Lack of size effects in spatial updating: Evidence for offline updating. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 854-866.

Holmes, M. C. and Sholl, M. J. (2005). Allocentric coding of object-to-object relations in overlearned and novel environments. *J. Exp. Psychol.: General*, *119***,** 1069-1087.

Isla, D. (2005). Handling complexity in Halo 2 [WWW page], Gamasutra Online, http://www.gamasutra.com/view/feature/2250/gdc_2005_proceeding_handling_.php [29.8.2011]

Jilk, D.J., Lebiere, C., O'Reilly, R.C., & Anderson, J.R. (2008). SAL: An explicitly pluralistic cognitive architecture. *Journal of Experimental and Theoretical Artificial Intelligence, 20*, 197-218.

Kelly, J. W., & McNamara, T. P. (2008). Spatial Memory and Spatial Orientation. *Spatial Cognition VI*, LNAI 5248, 22-38.

Kim Y., Hill R. W. Jr., & Traum D. R. (2005). A Computational Model of Dynamic Perceptual Attention for Virtual Humans. *Proc. on Conference on Behavior Representation in Modeling and Simulation (BRIMS)*, Universal City, CA.

Kokko, H. (2007). *Modelling for Field Biologists and Other Interesting People.* Cambridge University Press.

Krichmar, J. L., Seth, A. K., Douglas, A., N., Fleischer, J. G., & Edelman, G. M. (2005). Spatial Navigation and Causal Analysis in a Brain-Based Device Modelling Cortical-Hippocampal Interactions. *Neuroinformatics,* 3(3), 197-221.

Laird, J. E. (2000). It Knows What You're Going to Do: Adding Anticipation to a Quakebot. *AAAI 2000 Spring Symposium Series: Artificial Intelligence and Interactive Entertainment*, AAAI Technical Report SS-00-02.

Langley, P., Choi, D., Rogers, S. (2009). Acquisition of hierarchical reactive skills in a unified cognitive architecture. *Cognitive Systems Research*, *10*, 316-332.

Leonard T.: Building an AI Sensory System (2003). Examining The Design of Thief: The Dark Project [WWW page]. Gamasutra Online. http://www.gamasutra.com/view/feature/2888/building_an_ai_sensory_system_.php [29.8. 2011].

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.) *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 125 – 151). Baltimore: Johns Hopkins University Press.

Loyall, A. B. (1997). *Believable Agents: Building Interactive Personalities.* Ph.D. diss. Computer Science Dept., Carnegie Mellon University.

Marr, D. (1982) *Vision*. San Francisco: W. H. Freeman

Marsella, S. C., Gratch, J. EMA (2009). A process model of appraisal dynamics. *Cognitive Systems Research*, *10* (1), 70 – 90.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 142 – 157.

Mou, W., McNamara, W. P., Rump, B., & Xiao, C. (2006). Roles of Egocentric and Allocentric Spatial Representations in Locomotion and Reorientation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32* (6), 1274-1290.

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice Hall.

Neisser, U. (1976). *Cognition & Reality*. San Francisco: W. H. Freeman.

Pedica, C., & Vilhjalmsson, H.H. (2009). Spontaneous Avatar Behavior for Human Territoriality. *Proc. 9th IVA*, LNAI vol. 5773, 344-357.

Samsonovich, A. (2010). Toward a unified catalog of implemented cognitive architectures (review). *Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the BICA Society. Frontiers in Artificial Intelligence and Applications*, vol. 221, 195-244.

Sun, R. (2008). Introduction to Computational Cognitive Modeling. In Sun R. (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 3-21). Cambridge University Press.

Thomas, R., & Donikian, S. (2006). A spatial cognitive map and a human-like memory model dedicated to pedestrian navigation in virtual urban environments. *Proc. Spatial Cognition V*, LNCS 4387, 421-436.

Tyrrell T. (1993). *Computational Mechanisms for Action Selection*. PhD thesis, Centre for Cognitive Science, University of Edinburgh.

Waller, D. and Hodgson, E. (2006). Transient and Enduring Spatial Representations Under Disorientation and Self-Rotation. *Jn. Exp. Psychol.: Learning, Mem., Cogn., 32* (4), 867-882.

Wang, R., Spelke, E. (2000). Updating egocentric representations in human navigation. *Cognition*, *77*, 215-250.

Wang, R. F., & Spelke, E. (2002). Human spatial representation: insight from animals, *Trends Cogn. Sci.*, *6* (9), 376-382.

Wooldridge M. (2002). An Introduction to MultiAgent Systems. Chichester: John Wiley & Sons.

Author Note

Cyril Brom, Jan Vyhnánek and Rudolf Kadlec, Faculty of Mathematics and Physics, Charles University in Prague.

Jiří Lukavský, Institute of Psychology, Academy of Sciences of the Czech Republic.

David Waller, Department of Psychology, Miami University.

C. Brom is partly supported by the research project P103/10/1287 (GACR). R. Kadlec and J. Vyhnánek are partly supported by the student grant GA UK 21809 and by the grant 201/09/H057.

Correspondence concerning this article should be addressed to Cyril Brom, Faculty of Mathematics and Physics, Charles University in Prague, Room 312, Malostranske Namesti 25, Prague, 11800, Czech Republic. Tel: (420) 221 914 216; Fax: (420) 221 914 281

Footnotes

Footnote 1:

It has been demonstrated that some general cognitive architectures can serve for controlling IVAs, e.g., (Jilk et al., 2008; Laird, 2000). However, special-purpose IVA architectures, including architectures employed in 3D videogames, often serve better for that purpose. On the *other* hand, the neuro-/psychological plausibility of IVA architectures is only limited. Note also that there are large differences among general cognitive architectures as well as among IVA architectures. For instance, personal opinion of the authors of this paper is that Soar, implementing a computationally efficient rule-based system, is better suited for controlling IVAs than most other general cognitive architectures. On the other hand, ACT-R or LIDA are more plausible than Soar.

Footnote 2:

The closest psychological metaphor is the Atkinson & Shiffrin's (1968) model. Despite limitations of that model (e.g. Baddeley, Eysenck, & Anderson, 2009, pp. 42), it is the dominant model used in the field of virtual agents, most likely for its technical simplicity. Neither our architecture nor other IVA architectures, to our knowledge, feature elaborated short-term memory models that would include phonological loops, visuo-spatial sketch-pads etc.

Footnote 3:

Unlike computational cognitive sciences, research on episodic memory modeling in the context of virtual characters tend to focus on believable models, that is, models that

produce behavioral outcome appealing to target users. These models need not be psychologically plausible. In contrast, the model presented in this paper is intended to be plausible, but not necessarily producing believable outcome.

Footnote 4:

Note that pointing is intrinsically an egocentric gesture. The phrase “allocentric pointing” is merely a shortcut for “pointing on the basis of an allocentric representation”.

Footnote 5:

Note that IVA models of perception that distinguish between the central part of the visual field and the periphery exist (e.g. Leonard, 2003) as well as IVA models of attention (Kim et al., 2005). If an IVA application employs object features, they are represented symbolically. Similarly, our IVA architecture represents features of objects symbolically; nevertheless, the experiments do not employ features of objects; it is only important that the agent is able to distinguish one object from another.

Footnote 6:

Note that a part of configuration error can be explained on the basis of pointing error. Based on some statistical assumptions, Wang & Spelke (2000) proposed that such “corrected” (or “controlled for the pointing variability”) configuration error can be computed as configuration error minus the pointing error divided by square root of the number of trials. We are interested in this “corrected” configuration error rather than the absolute configuration error. See also (Mou, McNamara, Rump, and Xiao, 2006) for further discussion on this issue.

Figures

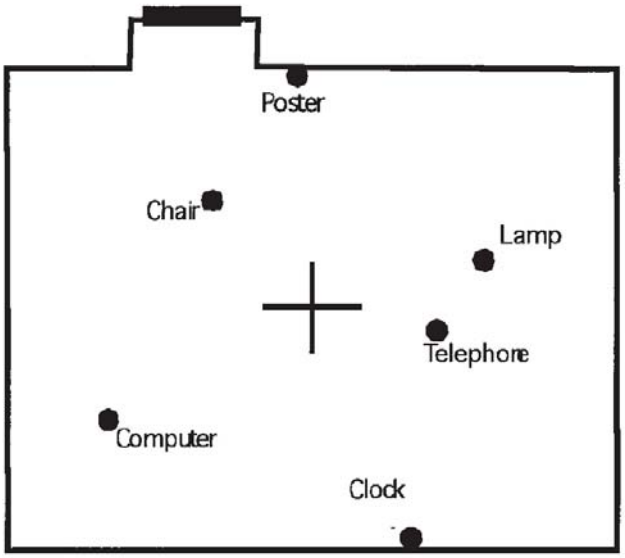


Figure 1. Diagram of the arrangement of objects in the experiment of Holmes and Sholl. Cross-hairs indicate participant's position when making pointing responses. Adopted from Holmes & Sholl (2005) with the publisher's permission.

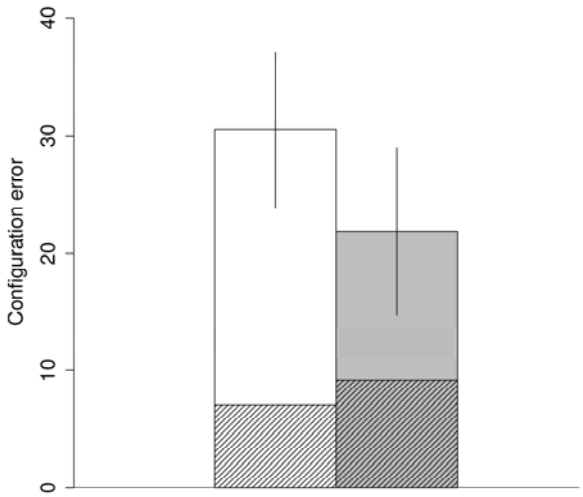


Figure 2. Results reported by Holmes & Sholl (2005; Exp. 7). White bars correspond to eyes-closed phase, grey bars to disoriented phase, lines represent 95% confidence intervals. The portions of the configuration error that can be accounted for by the respective pointing errors are shown as hatched regions (see Wang & Spelke, 2000, for statistical considerations). Note that the part of the configuration error that *cannot* be accounted for by the pointing error is *lower* for the disoriented phase, which means that the disorientation effect had not been replicated in this experiment.

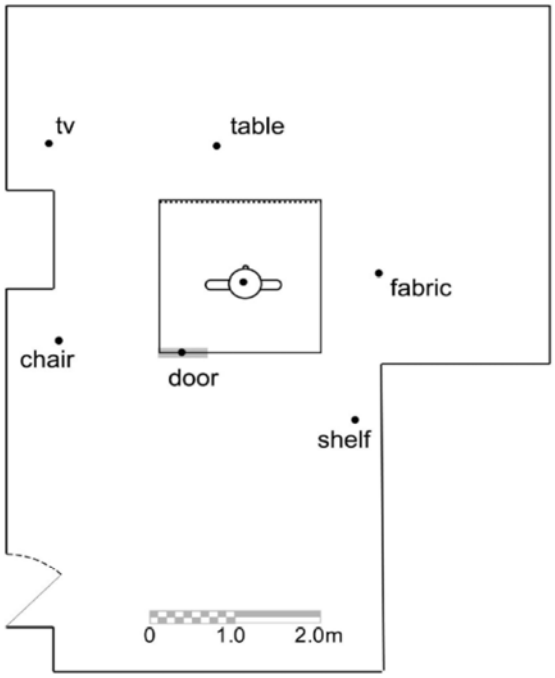


Figure 3. Diagram of the arrangement of objects in the experiment of Waller & Hodgson. When tested participants sat inside the booth. Adopted from Waller & Hodgson (2005) with the publisher's permission.

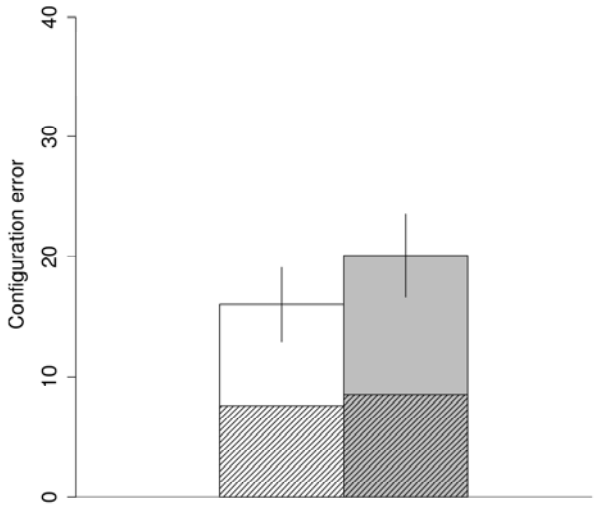


Figure 4. Results reported by Waller and Hodgson (2006; Exp. 1). White bars correspond to eyes-closed phase, grey bars to disoriented phase, lines represent 95% confidence intervals. The portions of the configuration error that can be accounted for by the respective pointing errors are shown as hatched regions. Note that the part of the configuration error that *cannot* be accounted for by the pointing error is *higher* for the disoriented phase—the replication of the disorientation effect originally found by Wang and Spelke (2000).

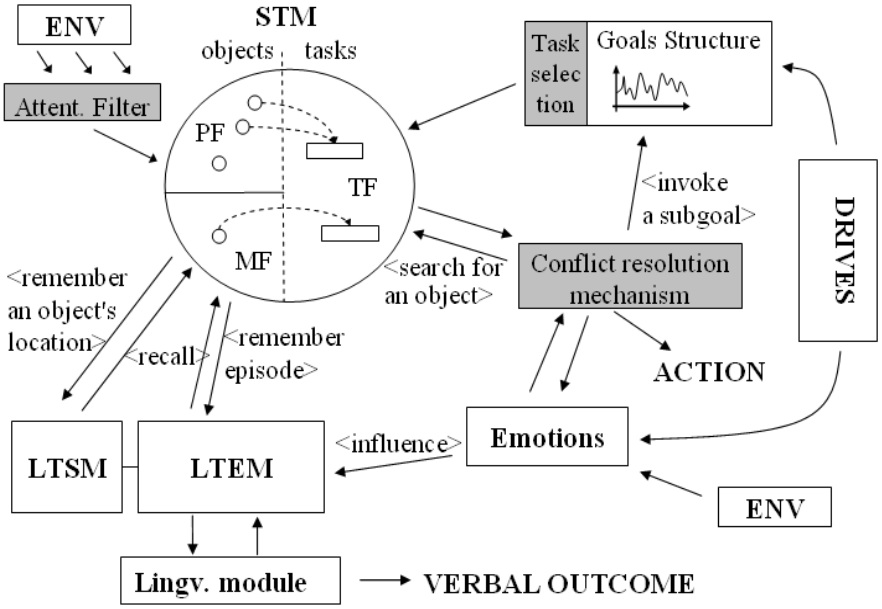


Fig 5. Our IVA's architecture. See the text for the further description. Note that for the experiments described here, the most relevant parts are the short-term memory, in particular its perceptual field, and the long-term spatial memory. The action selection part used here is fairly simple and the IVA features no LTEM, linguistic module, drives and emotions.

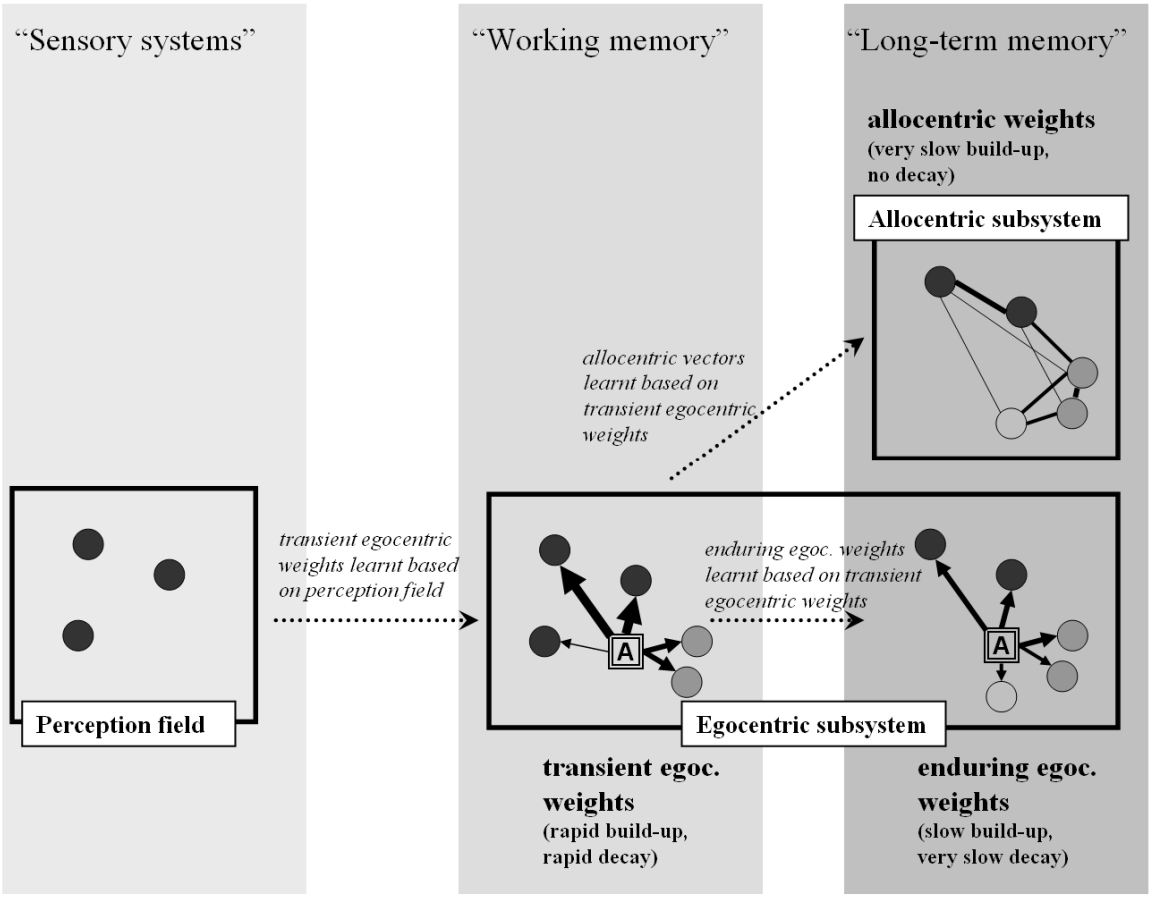


Figure 6. Information flow in the DP-model. There is a direct route from the perceptual field to the transient weights of the egocentric module, but neither to the allocentric module nor to the enduring weights of the egocentric module. Circles denote objects. Black circles denote objects just perceived, gray circles denote objects perceived a moment ago and white denotes unperceived objects memorized (only by enduring components). Lines and arrows denote vectors; the lines' strengths represent the vector weights. Every egocentric vector is depicted twice; once with a transient weight and once with an enduring weight. The square represents the agent. The rough correspondence to Atkinson and Shiffrin memory model (1968) is suggested by the labels at the upper part of the figure.

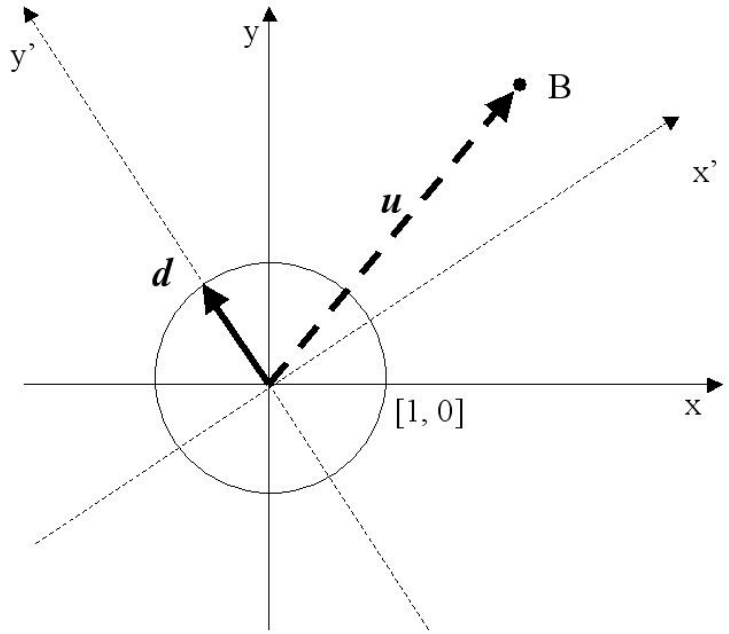


Figure 7. Representation of an egocentric vector *u* to object B when the agent is standing in the origin of coordinates and its heading corresponds to the direction of the vector *d* (this vector is depicted as unitary; in fact, its magnitude is not important). Axis *x* and *y* represent directions of walls of the room (i.e. the allocentric reference frame). Axis *x'* and *y'* represent the egocentric reference frame. Vector *u* would be (2, 2.5) with respect to the allocentric reference frame and (3, 1) with respect to the egocentric reference frame. In the egocentric representation, the egocentric reference frame is used.

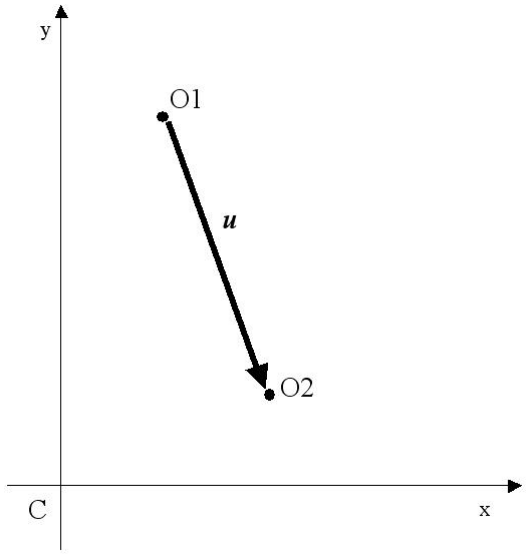


Figure 8. Representation of an allocentric vector from object O_1 to O_2. C is the position of left bottom corner of the room and axis *x* and *y* represent directions of walls. The absolute coordinates of O_1 are [1,4] and they are [2,1] for O_2. Thus coordinates of *u* are (1,-3). The agent's position and heading are not relevant to represent the relationship allocentrically.

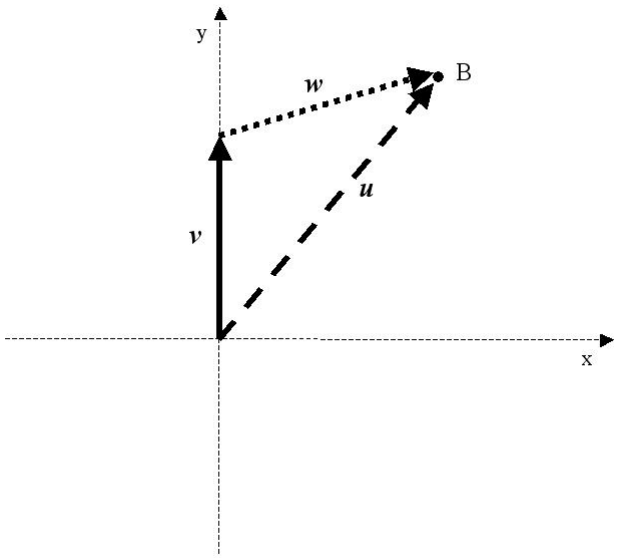


Figure 9. Updating of the egocentric vector *u* after the agent moves along the vector *v*. The coordinates of old egocentric vector *u* were (3, 3.5) while the new coordinates of *w* are (3, 1).

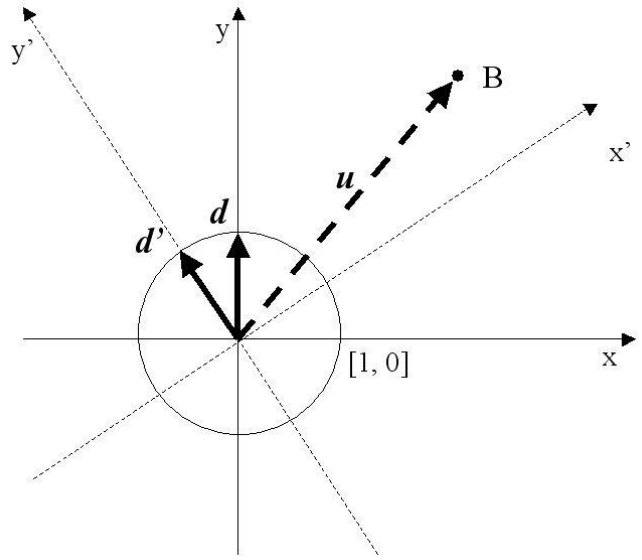


Figure 10. Updating of the egocentric vector *i* after a rotation. The vector *d* represents the agent's old heading, the vector *d'* the agent's new heading (both are depicted as unitary vectors). The coordinates of the vector *u* were (2, 2.5) with respect to the old heading and they have been updated to (3, 1) with respect to the new heading.

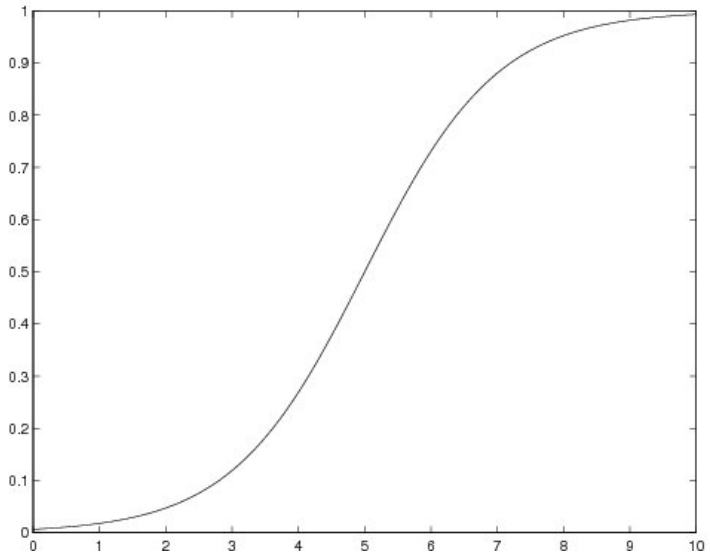


Fig 11. A logistic sigmoid used for computing a vector strength. The *x*-axis represents a vector base d and the y-axis represents the vector strength w. The scale of the *x*-axis depends on the [eg]bound parameter, which was set to 10 as detailed in Sec. 4.

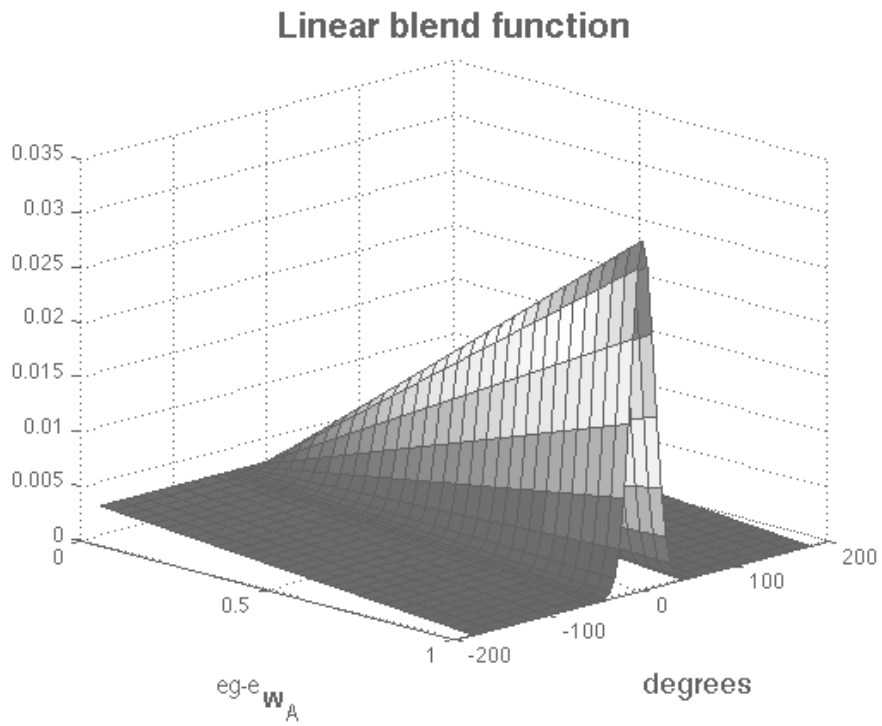


Fig 12. A linear blend function.

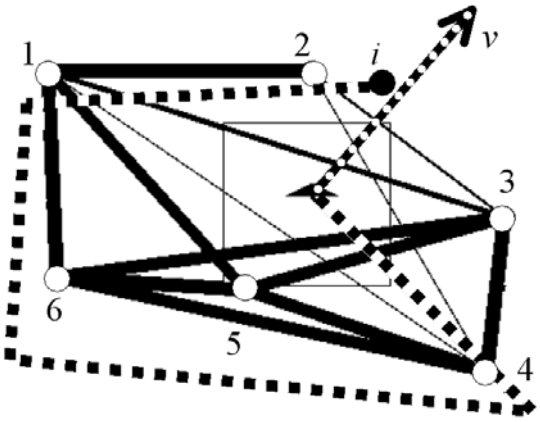


Fig 13. Estimation of the pointing direction to Object 2 (top middle on the figure) when the agent's imagined heading is to Object 4 (right bottom on the figure). The IVA stands in the middle of the booth. Lines between the objects (circles) represent allocentric vectors; their strengths denote the vector weights. The hashed line represents the "shortest" path from the agent towards the target object (the "shortest" is meant in terms of pointing precision, as detailed in the main text). The path goes over the objects 6 and 1. Note the distortion from the original representation. The image of the target object is denoted as *i*. However, because the model also imitates motor error, the agent does not point directly to the image of the target object, but a bit aside. The final pointing direction is depicted by the dotted line and marked as *v*.

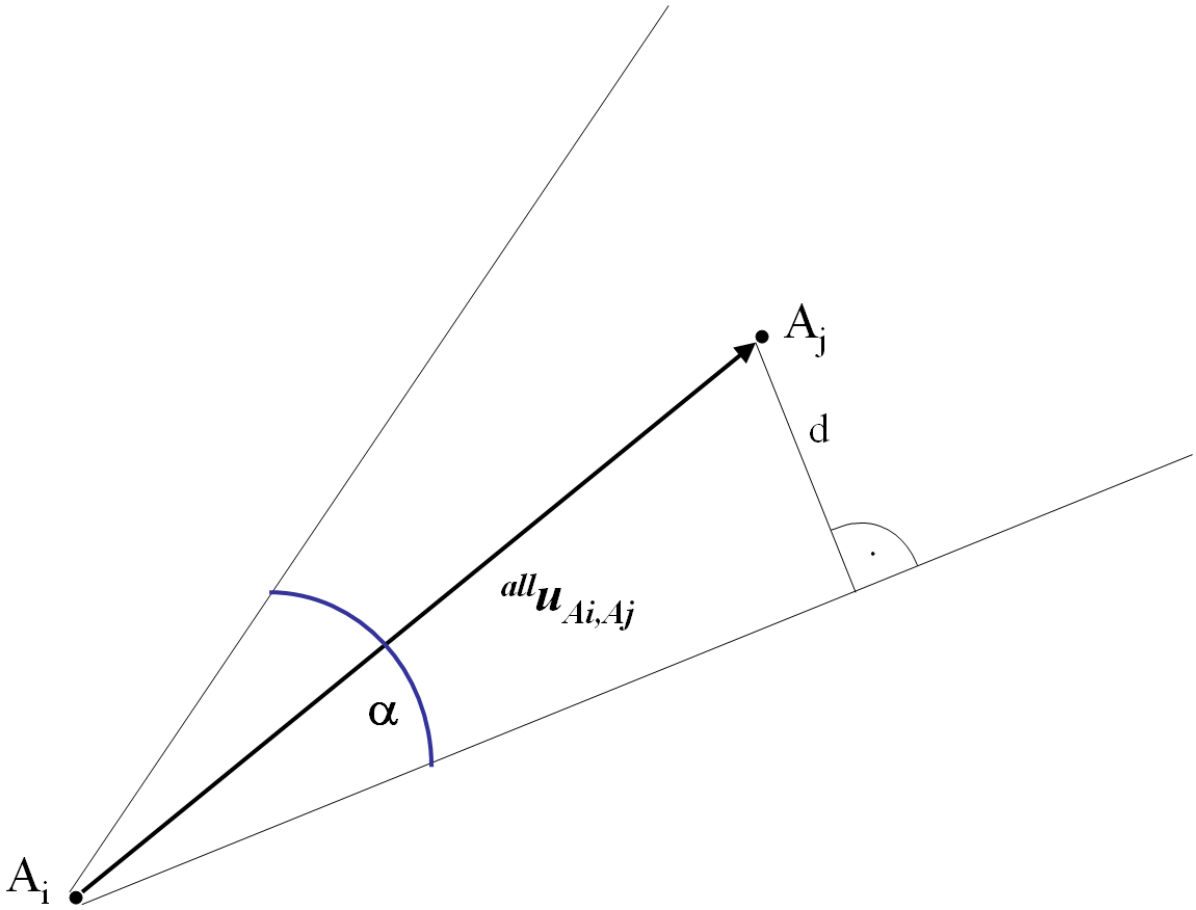
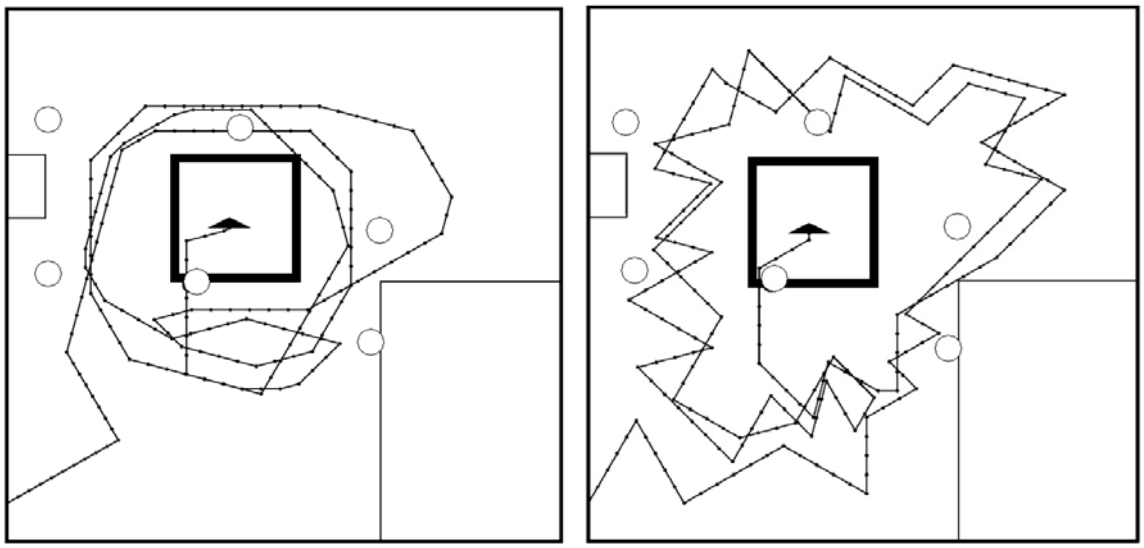
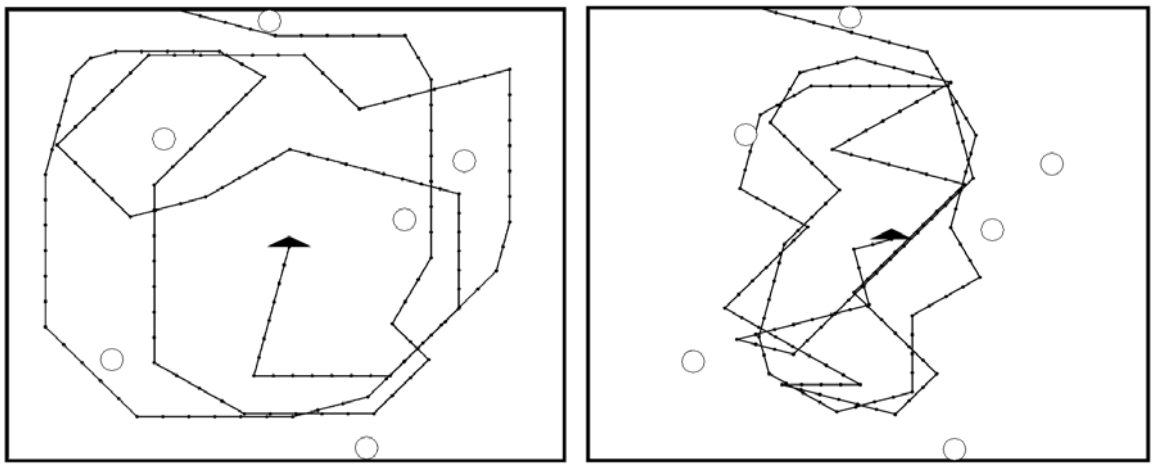


Fig. 14: Estimation of the vector from the object A_i to the object A_j. The angle α defines a circular sector representing 95% confidence interval of the estimated directions of the original vector ${}^{all}u_{Ai,Aj}$. The distance d is the shortest distance between the exact position of the object A_j and an edge of the circular sector.



(a) (b)

Figure 15. The most plausible path (a) and the least plausible path (b) in the conditions of Waller and Hodgson.



(a) (b)

Figure 16. The most plausible path (a) and the least plausible path (b) in condition of Holmes and Sholl.

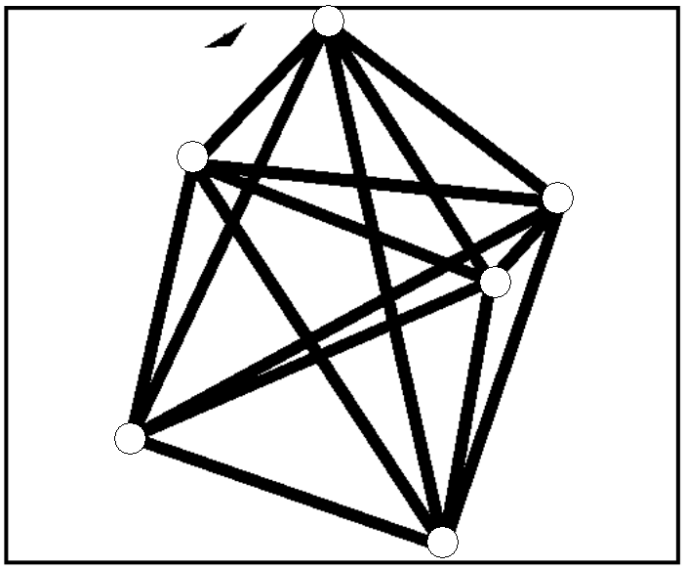


Figure 17. An example of an over-learnt allocentric map.

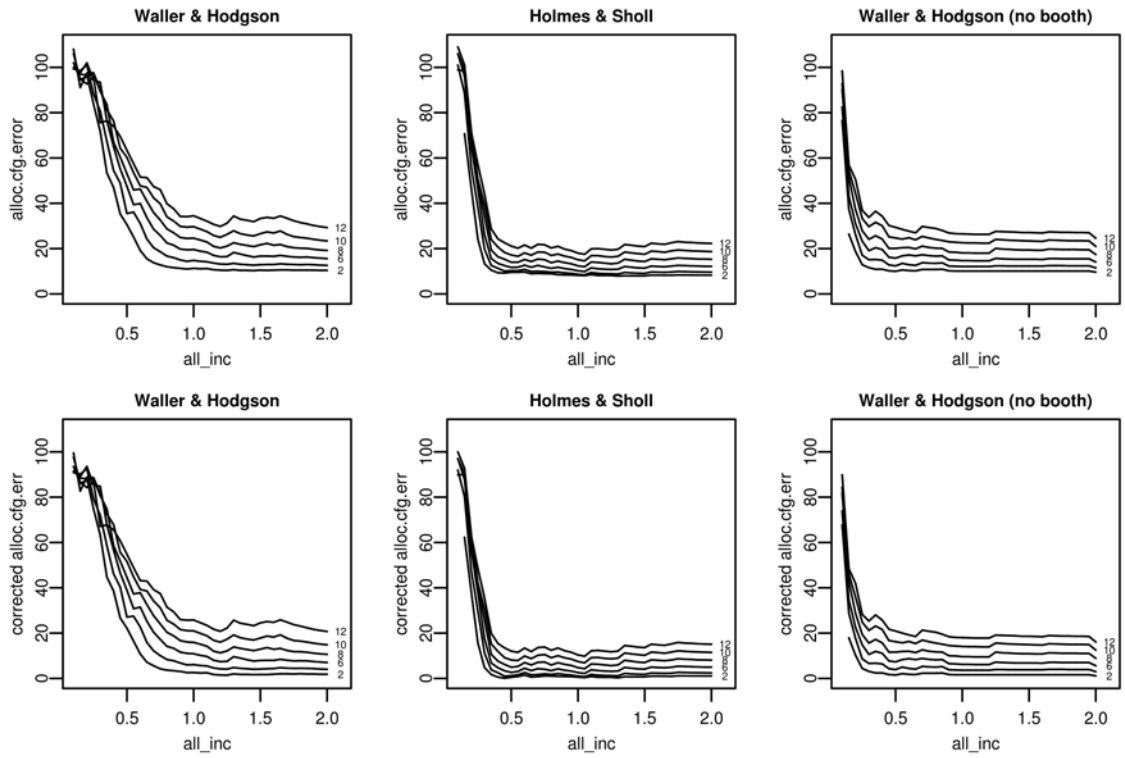


Figure 18. Configuration error (*y*-axis) as a function of ${}^{all}inc$ (*x*-axis) and ${}^{all}\sigma_{mem}$ (6 different lines corresponding to ${}^{all}\sigma_{mem}$ equal to 12, 10, 8, 6, 4, and 2; number 4 is not depicted on the figures due to space limitations). The upper three figures depict how configuration error changes in the three different environments. The lower three figures depict how configuration error controlled for pointing variability changes in the three different environments

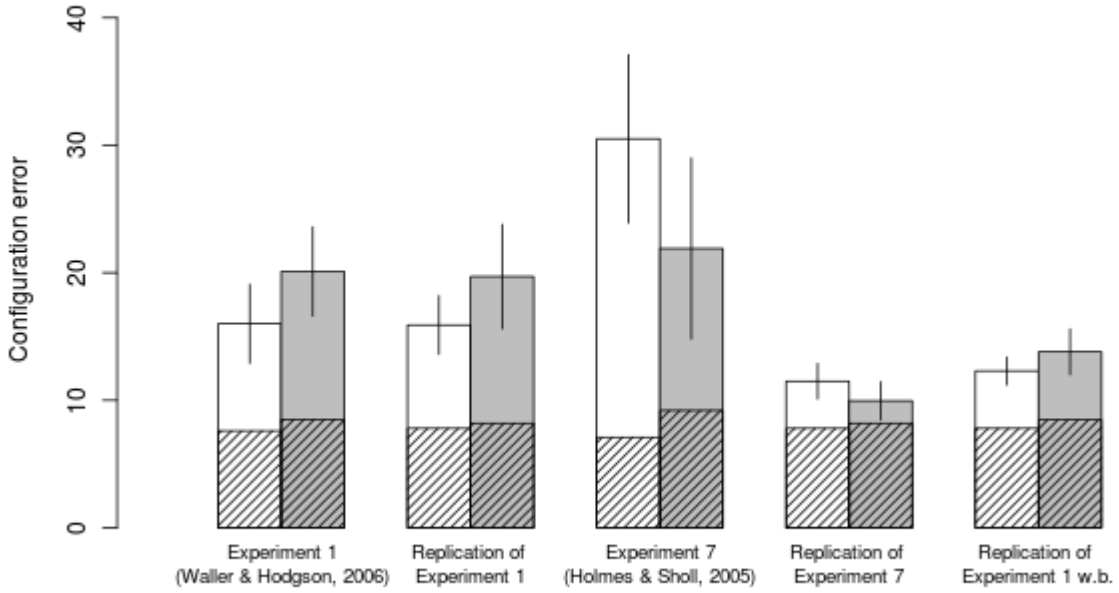


Figure 19. Configuration errors in original experiments and in our replication. White bars correspond to eyes-closed phase, grey bars to disoriented phase, lines represent 95% confidence intervals. The parts of the configuration errors that can be accounted for by the respective pointing errors are shown as hatched regions.

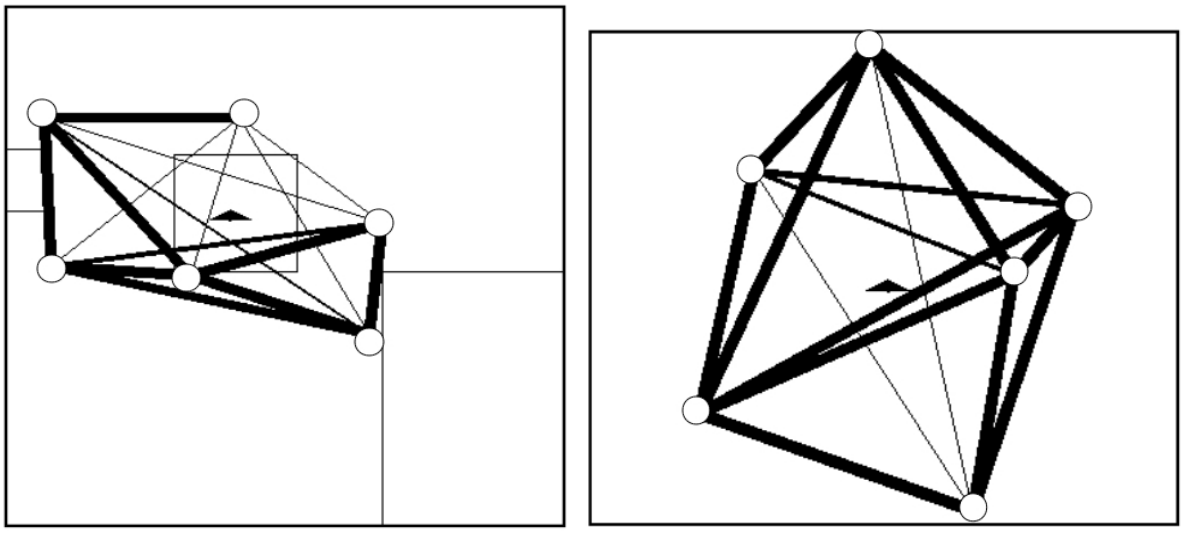


Figure 20. Examples of allocentric representations for the trajectories depicted on Figures 15a (left) and 16a (right). The IVA stands in the middle of the booth. Lines between the objects (circles) represent allocentric vectors; their strengths denote the vector weights. Generated for ${}^{all}inc = 0.71$ and ${}^{all}\sigma_{mem} = 4$.

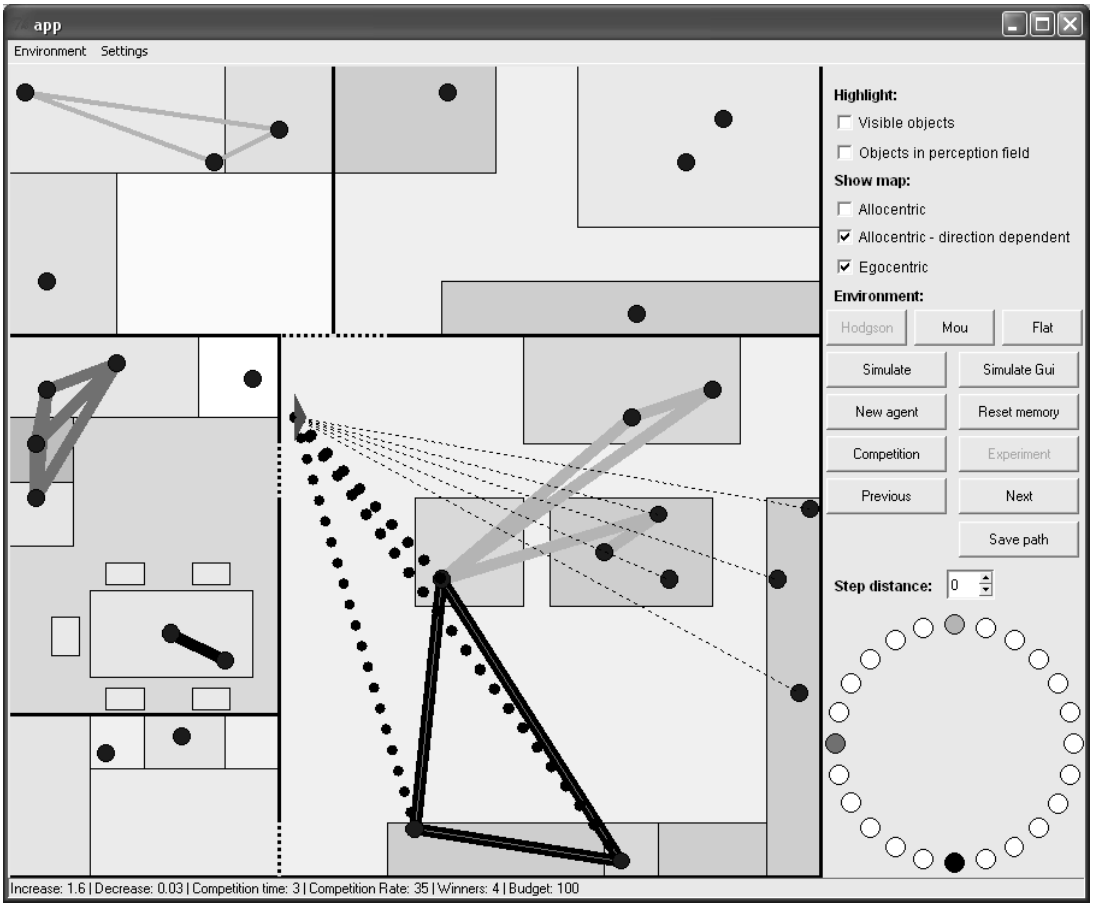


Figure 21. A screenshot showing our application for running experiments with the IVA with the spatial memory. There is a house with five rooms loaded as the virtual environment. This particular agent has the egocentric module as described in this paper, but direction dependent allocentric module; in particular, it builds three overlapping allocentric representations oriented along a different environmental axis (south - north, north - south, east - west) at the same time. Full lines represent the weighted allocentric vectors (gray scale denotes the intrinsic axis) and the dotted lines represent the weighted egocentric vectors. Doors between rooms are also depicted as dotted lines.

Tables

Table 1

The setting of the egocentric parameters for our experiments

$^{eg\text{-}t}inc$	$^{eg\text{-}t}dec$	$^{eg\text{-}e}inc$	$^{eg\text{-}e}dec$	$^{eg}bound$
2	1,5	1,5	0,05	10

Table 2

Configuration errors (CE) and pointing errors (PE) in the eyes-closed phase for simulations in the three experimental environments (mean and confidence intervals).

Waller & Hodgson		Holmes & Sholl		Waller & Hodgson (no booth)	
CE	PE	CE	PE	CE	PE
15.9° ± 2.3	11.0° ± 0.8	11.5° ± 1.4	11.0° ± 0.8	12.3° ± 1.1	11.0° ± 0.8

Table 3

Configuration errors (CE) and pointing errors (PE) in the disoriented phase for simulations in the three experimental environments for three different values of ^{all}inc *and* $^{all}\sigma_{mem}$*. (mean and confidence intervals).*

$^{all}\sigma_{mem}$	^{all}inc	Waller & Hodgson		Holmes & Sholl		Waller &Hodgson (no booth)	
		CE	PE	CE	PE	CE	PE
2	0.59	20.2° ± 5.5	11.9° ±1.2	9.67° ± 1.0	11.8° ± 1.4	10.4° ± 1.3	12.0° ± 1.3
4	0.71	19.7° ± 4.1	11.6° ± 1.3	9.96° ± 1.5	11.7° ± 1.3	13.8° ± 1.8	12.0° ± 1.3
5	0.78	20.1° ± 3.5	12.0° ± 1.4	10.4° ± 1.4	11.1° ± 1.3	15.1° ± 2.0	12.0° ± 1.3

Table 4

*Decrease in the configuration error in disoriented phase of replication of Holmes and Sholl (2005, Exp. 7) and disoriented phase of Waller and Hodgson (2006, Exp. 1) without the booth. Changes in configuration error (CE) are significant if they differed using one-sample t-test from predicted change in configuration error estimated from the mean pointing error (using the procedure controlling for the pointing variability from Waller & Hodgson, 2006 and Wang & Spelke, 2000). Statistical significant values are marked with ** (p<0.01) and * (p<0.05), while + denotes a trend (p<0.1).*

		H & S eyes-closed vs. disoriented			**W & H (no booth, disoriented) vs. W & H (with booth, disoriented)**		
${}^{all}\sigma_{mem}$	${}^{all}inc$	CE change	Predicted CE change	*t*	CE change	Predicted CE change	*t*
2	0.59	-1.922	0.575	-2.762 **	-9.145	0.084	-3.104 **
4	0.71	-1.631	0.496	-2.241 *	-5.213	0.266	-2.214 *
5	0.78	-1.113	0.027	-1.198 (n.s.)	-4.256	-0.005	-1.855 +