Spatial Principles in Control of Focus in Reasoning with Mental Representations, Images, and Diagrams

Dominik Engel, Sven Bertel, and Thomas Barkowsky

SFB/TR 8 Spatial Cognition, Universität Bremen, Germany
{dengel, bertel, barkowsky}@sfbtr8.uni-bremen.de

Abstract. The effective control of attentional focus is an essential requirement in mental reasoning based on mental models and mental images, as well as in the interaction with external diagrams. In this paper, we argue for spatial organization principles common to various mental subsystems that entail a non-centralistic control of focus. We give a brief overview of mental spatial reasoning and present a review of psychological findings related to cognitive control. We review existing modeling approaches that realize control of focus in imagery, scene recognition, and mental animation. Based on these foundations, we identify basic spatial organizing principles that are shared by the diverse subsystems collaborating in mental spatial reasoning. We discuss the implications of these principles in the framework of a computational modeling approach and give an outline of the conception of control of focus in our computational architecture *Casimir*.

1 Introduction

This paper addresses issues of computational modeling of mental image-based reasoning with spatial configurations. The work is based on the assumption that representational and procedural aspects of cognitive systems come together, and are but two sides of a single coin. As a matter of fact, all computational attempts to model cognitive phenomena are based on the hypothesis that computational metaphors such as data and storage, information and processing, function and system provide adequate concepts for understanding and describing cognitive phenomena. Different approaches differ in which of the metaphors they follow and to what extent; yet, the basic dual abstraction into representations and processes persists.

For the computer scientist, it can hardly come as a surprise that representations and processes should be related (neither should it surprise the psychologist, cf. Palmer, 1978); they only make sense when seen as pairs. On a representation theoretic level, representations in fact set the standards for processes, and vice versa. On a practical level, however, it can be worthwhile to focus on properties of one *given* the other. This is especially true where our knowledge of a system is incomplete. Such is often the case in the modeling of cognitive phenomena, as the level of detail of what is postulated for mental representations frequently differs from that of the corresponding processes. The investigation into visual mental imagery provides a prime example of a debate that for a long time has been guided in particular by considerations of various representational formats.

We argue that spatial properties in mental knowledge representations influence mechanisms in the mental processes which operate on such representations. Second, we propose that there exist basic spatial organizing principles which are common to different types of mental representations and processes, and that it is in part based on those principles that the different types are related or interact. Specifically, we will address the role that mechanisms resulting from spatial or spatio-analogical structures play in the control of focus in reasoning with mental models, visual mental images, and external diagrams. To this end, the next section gives a minimalist review of mental spatial reasoning, followed by a section providing a synopsis and discussion of selected psychological findings on attentional control. Section 4 presents an overview of existing approaches that give a range of functional accounts of the processes involved. Subsequently in Section 5, we will identify basic spatial organizing principles and mechanisms that are common across different mental representations and processes. In Section 6, implications of these principles and mechanisms for the development of a specific computational model of mental image-based reasoning will be discussed. Section 7 concludes the paper and gives an outline of future work.

2   Mental Spatial Reasoning

This section gives an overview of the topic of mental spatial representations, reasoning with mental images, and the role of externalization and diagrams from the perspective of spatial reasoning.

2.1   Mental Representations of Space and Visual Mental Images

As we move through the world, a variety of sensory inputs are continually presented to the brain. Motor and sensory information are used to construct mental representations of the space in which we move. Studies on deficits following from parietal lobe lesions suggest that multiple mental representations are constructed (cf. Colby & Duhamel, 1991), for instance differing in the frame of reference, but that some of these representations are truly intermodal (e.g. in that they receive input through visual as well as somatosensory streams; cf. Duhamel et al., 1998).

As a bottom line, mental representations of the surrounding space are in many ways analogous to the space. Sometimes the analogies lie in distinct spatial properties (e.g. topological ones, Sereno et al., 2001), in accessibility to manipulation (e.g. Shepard & Metzler, 1971, for mental rotation of images, or Hegarty, 1992, for mental animation of mechanical systems), or in properties with respect to reasoning (Kosslyn et al., 1978, for image scanning; Moyer, 1973, for size judgments). The construction of mental images and mental image-based reasoning can be seen as rather extreme cases in which a plethora of analogies between mental and external representations can be drawn. These analogies have led to vivid debates about the actual representational format of the mental representations (cf. Tye, 1991; Kosslyn & Thompson, 2003; Pylyshyn, 2003).

Much is known about human mental conceptions of large-scale spaces; there exist a range of theories that detail their properties and structure in relation to developmental issues, to learning, to different classes of space (e.g. to geographic or environmental spaces), or to a variety of spatial reasoning tasks. Over the past decades, the increase in knowledge along various lines of research has been reflected by the introduction of metaphors such as cognitive maps (Tolman, 1948), spatial images (Lynch, 1960), cognitive atlases (Kuipers, 1982), geographic information systems (Hirtle, 1998), or cognitive collages (Tversky, 1993). None of these metaphors must be taken literally, since, for example, mental knowledge about geographic spaces is found to be frequently distorted, fragmentary, incomplete (cf. Montello, 1992; Tversky, 1993), or potentially contradictory. In addition, mental representations of spatial knowledge are often organized hierarchically (cf. Stevens & Coupe, 1978) or chunked together to more complex structures (e.g. as visual chunks; cf. Kosslyn & Pomerantz, 1977).

Sometimes, spatial information is associated with mental representations of non-spatial information (i.e. in the form of spatial tags attached to semantic representations). A use of locational indexing procedures has been revealed in memory tasks for which location should be irrelevant (Richardson & Spivey, 2000), again confirming the elemental role of external (spatial) structures for inner mechanisms of an embodied cognitive system (cf. Wilson, 2002; Lockhead & Pomerantz, 1991). Functionally, using spatial indexes can be computationally efficient as they relieve working memory load (Ballard et al., 1997).

2.2 Externalization and the Spatial Properties of Diagrams

The perceptual and cognitive advantages of external diagrams over sentential representation have been frequently stressed, both with respect to representational and procedural properties; consequent to these are computational advantages. It is because of the way locality and indexing are realized in diagrams that, for instance, information which is needed at the same time is displayed in groups, and that correspondences between diagram parts are established without the need to introduce explicit labels (Larkin & Simon, 1987). Also, the drawing of direct "perceptual inferences" is permitted by a close coupling between (bottom-up) processes in visual perception and (top-down) processes in mental imagery (e.g. Kosslyn & Sussman, 1995). Further computational advantages arise from the specificity of information required by the diagrammatic representation format (Stenning & Oberlander, 1995); in this respect, diagrams have similar advantages over propositional formats as mental images have over isolated knowledge fragments. The use of diagrams supports processes of creativity and reflection (Goldschmidt, 1995), fosters synchronized communication (Healey et al., 2002), and introduces structure (Purcell & Gero, 1998) into problem solving processes.

With respect to the spatial properties and effects discussed in this paper, we argue that diagrams as accessed through visual perception can be compared to mental representations of spatial knowledge, such as in long-term memory (LTM), in working memory, and in particular in visual mental images. The near relationship between images and diagrams is further supported by findings of similar patterns of

eye movements in visual perception of a diagram and consequent re-instantiations as a mental image, as well as of similar functional roles of such patterns (Laeng & Teodorescu, 2002). Accordingly, spatial effects of scanning, sequentialization, neighborhoods or locality, and of grouping in reasoning with external diagrams can be related to those in reasoning with mental representations of space.

3   Control of Mental Visuo-Spatial Processes

After having reviewed some issues of mental reasoning about space, in this section we discuss a selection of psychological findings relating to cognitive control of the processes involved. In cognitive psychology, the issue of control has traditionally been discussed in the context of the research on *attention*. Within the multitude of contributions from cognitive science that have been made to the topic of attention, we discuss three issues that are fundamental to the phenomena we are interested in.

3.1   Control of Visual Attention and Selectivity

Traditional research on attentional control centered around the *selection* of information from perceptual data. Broadbent's (1958) filter theory suggests a bottleneck in the processing system located at the transition from a parallel, high-capacity system to a linear, limited-capacity system. Much research followed this tradition (e.g. Treisman, 1964; Moray, 1973) and focused on the question where this bottleneck is located. Johnston and Dark (1986) categorize the theories in Broadbent's tradition as *cause* theories. For cause theories, they distinguish between two domains of processing, *Domain A* and *Domain B*, which they abstract from various labels used in causal theories about attention: "nonconscious and conscious, automatic and controlled, peripheral and central, intraperceptual and extraperceptual, preattentive and attentive, and passive and active". What they subsume under *Domain A* then is the "large-capacity, non-conscious, and passive system that is responsible for encoding environmental stimuli", whereas *Domain B* is characterized as "a relatively small-capacity, conscious, and active system that is responsible for controlling various forms of information processing including selective attention. […] Domain B is … an attentional mechanism or director, a cause of selective processing" that selects portions of the data offered by *Domain A*.

   The problem they identify for causal theories is that "Domain B has all the characteristics of a processing homunculus" and thus the question how an individual pays attention has to be asked how this "attentional director" pays attention, resulting in an infinite regress (Johnston & Dark, 1986). More recently, Allport (1993) criticizes the set of assumptions that the causal theories often take for granted; some of these assumptions seem at least questionable from the perspective of more recent neuropsychological and neurophysiological findings. In many causal theories, it is assumed that mental processing is taken as a "linearly ordered, unidirectional sequence from sensory input to overt output" (Allport, 1993). Furthermore, processing of nonsemantic, especially spatial, features is assumed to be done prior to semantic categorization.

If these assumptions do not hold, however, then the discussion of early vs. late selection loses its basis. Regarding the location of attentional selection, Allport points out that the mutually exclusive distinction between early and late selection presumes a single locus of attentional selection, and thus also a unitary computational process. The most problematic assumption is the postulation of such a unitary central system that is the sole component responsible for attentional control. Monsell and Driver (2000) address the same problem and remark that even with advances in neuropsychology the problem of postulating a homunculus in attentional control has not been sufficiently resolved.

A possible answer to the problem of cause theories are approaches that view attention as "the consequence of natural priming effects" (Johnston & Dark, 1986). Johnston and Dark fail to report what exactly these effects are. However, the idea of attention emerging from a set of effects is reflected in more recent psychological theories (cf. Monsell & Driver, 2000; Posner, 1993), and it is supported by findings that suggest that control of attention involves distributed systems, both on the functional and on the neural level (Nobre et al., 2004; Ishai et al., 2000; Hommel et al., 2004; Allport, 1993; Posner, 1993).

In the light of these results, another traditional assumption on attention is to be questioned: the view of attention as a kind of limited resource that forms a bottleneck in the processing system and induces selectivity in information processing. In evolutionary terms, the idea that processes developed *on* the available resources is more sensible than the idea that processes developed independently of the (limited) resources, and then had to get by with these resources through selectivity. While the idea of selectivity remains important, recent evidence points out the perspective of attention as a set of distributed control mechanisms that are not only concerned with selection but also cover management, scheduling, and communication tasks (cf. Kieras et al., 2000). Together with the idea of attentional control as an emergent, distributed phenomenon, these theories have interesting implications for a computational model.

3.2   Control of Focus in Mental Imagery and Visual Perception

Traditionally, research on attention has been concerned with external sources, i.e. perception of objects and events in the extrapersonal world. Only recently, attention in the domain of internal mental representations is moving into the focus of research programs (e.g. Nobre et al., 2004; Griffin & Nobre, 2003). Nobre and colleagues (2004), for example, performed an array of brain imaging studies that try to link the orienting of spatial attention to extrapersonal objects and events with the orienting of spatial attention on internal representations held in working memory (WM). They report evidence that may indicate that "orienting of attention in the perceptual and working memory domains share common substrates". They also present data that suggest "that shifting and zooming the spatial focus of attention in the absence versus presence of a memory context recruits highly overlapping but not coextensive systems". In their brief discussion of the relevance of their experiments to

mental imagery, Nobre and colleagues point out the link between mental imagery and internal orienting of attention, based on work by Ishai and Sagi (1995) and Ishai and coworkers (2000).

It seems plausible that some of the mechanism at work in visual attention can also be postulated for attention in mental imagery. Not only is there evidence suggesting strong overlap between the two systems, both on the functional and on the neural level (cf. Michelon & Zacks, 2003; Ishai & Sagi, 1995) for higher level cognitive processing, but there are also findings that support the view that imagery and perception systems are interfaced at an early processing stage (Ishai & Sagi, 1997). In this view, mechanisms of attentional control induced by mental imagery and visual perception are likely to exhibit a similar coupling.

3.3 Space as Structure in the Control of Focus

As Allport (1993) points out, "spatial location is far from being 'a simple characteristic' so far as its coding in the brain is concerned". Space is fundamental in providing structure during the control of focus. This has been shown, for example, in studies with disorders of spatial attention, which also affect internal, non-sensory, mental representation (cf. Bartolomeo & Chokron, 2002), for instance during mental imagery. There is abundant research on the topic of selection based on spatial attention. Posner (1984) suggests three internal mental operations in the covert orienting of spatial attention: disengagement of attention from its current focus, moving attention to the target focus, and engaging attention to the new target. In the context of mental imagery, Hazlett and Woldorff (2004) subdivide the shifting of spatial attention further into separable stages of planning and executing. As Eimer and colleagues (2003) argue, functional imaging studies have shown that the processes involved in control of spatial attention are also distributed among different attentional networks, possibly employing different spatial reference frames.

In Kosslyn's (1994) model of mental imagery, the issues of space as structure and control of focus come together in the formation of *multipart images* (i.e. images that consist of more than one component). According to Kosslyn's theory, mental images can be formed in a piecemeal manner with several activations from long-term memory and subsequent visualization of the activated information in the visual buffer. The locations of the visualization of the image parts are determined by the position of the attention window. In order to construct multipart images, the window of attention is moved to the right positions by the attention shifting system, based on spatial information from long-term memory. The same system is also responsible for moving the window of attention to the sequence of scanning positions during image inspection.

Noton and Stark (1971) introduce the spatial notion of the *scanpath* as the sequence of eye movements with which a subject scans a picture and organizes its features in a linear order. They suggest that the scanpath, i.e. the sequence of eye movements, could be part of the mental representation of a visual scene (cf. Stark & Choi, 1996).

Brandt and Stark (1997) show that the scanpath during mental imagery reflects the content of the imagined scene. Laeng and Teodorescu (2002) present evidence that suggests that eye scanpaths during mental imagery play a functional role. They argue that scanpaths might provide a spatial index to the parts of a mental image, a position that is embraced by Mast and Kosslyn (2002). In this view, the scanpath can be conceived of as a part of mental representations that is abstracted from the actual eye movements, and that plays a functional role in general relative spatial indexing. The frames of spatial reference in this respect still remain subject to research (cf. Eimer et al., 2003).

3.4 Summary

In summary, the psychological evidence discussed above supports the view of attentional control as a set of distributed processes rather than just as a resource. Control of attentional focus is a cognitive effect realized by distributed executive processes that are based on fundamental mechanisms. In the case of spatial cognition, there is evidence that spatial structures and representations lead to spatial mechanisms in the control of focus.

4 Existing Approaches

The idea that specific spatial knowledge representations are used to control focus-related processes in the mind has been used with respect to several partial aspects in cognitive modeling approaches so far. Examples are structural encodings that are used to control attention shifts in mental image construction, representations that guide eye movements employed for detecting salient features in visual scene analysis, or control of attention based on functional implications in understanding diagrams that convey dynamic operations.

4.1 Control of Focus in Mental Image Processing

In the construction of visual mental images, control of attention is an essential requirement to combine several image parts in a meaningful manner. Mental images are formed from well-structured components that are retrieved from long-term memory. In *multi-part images*, control of focus is used on the visual buffer representation structure to detect where a new image part has to be integrated into a partially constructed image. In the implementation of his early mental imagery model, Kosslyn (1980) proposes a *cathode-ray tube metaphor* to illustrate the successive, focus-oriented construction of the mental image on the *surface representation* (the visual buffer structure that holds the image proper). The content to be visualized on the surface representation is retrieved from the *deep representation*, which conceptually corresponds to human long-term memory.

In the deep representation structure, there are two types of representations: *literal* or *perceptual* representations that encode what entities look like (i.e. shape representations); and *discursive* representations that propositionally encode compositional in-

formation related to mental images (i.e. general knowledge about part-whole relationships, rough size information, or information about object categories). The literal image representation comes in two types: as *skeletal encodings* and as *individual encodings*. Skeletal encodings contain the overall shape and structure of an image, whereas individual encodings provide further detail. When a complex image is to be constructed in the visual buffer, a rough overall image based on a skeletal image is generated first. The encoding of the skeletal image contains propositional descriptions of details that may further specify the image. With respect to the control of focus, the positions of the detailing image parts are given as sequences of descriptions of abstract search procedures that are performed on the image in the visual buffer until the proper location is detected.

The mechanisms realized in this implementation have also been proven to be plausible through the later, neuropsychologically motivated work by Kosslyn (1994). According to his later conception, in a first processing step a global image is retrieved from associative memory, which afterwards is further detailed. Subsequent parts are arranged according to spatial descriptions retrieved from associative memory (the mechanisms employed here are the same that are also used in top-down hypothesis testing in object recognition during visual object recognition). The spatial descriptions retrieved are used directly to position the attention window on the visual buffer.

In a more technical application inspired by Kosslyn's mental imagery research, similar strategies are used in the *computational imagery* system by Glasgow and Papadias (1992): propositional descriptions of spatial relationships are used to generate a spatio-analogical representation on a symbolic array structure. Also, the same propositional descriptions are used to control a region of attention on the array structure that, for instance, can be used to reason about spatial motion planning.

4.2   Control of Eye Movements in Visual Scene Analysis

It has been shown that the mental imagery mechanisms described above can be regarded as being functional in vision processes, for instance, when a complex object or an entire scene has to be inspected and analyzed (cf. Kosslyn & Sussman, 1995). In a technical field of application, these ideas have been adopted in the *active perception* paradigm in computer vision (see Aloimonos, 1993, for a review).

A cognitively motivated computational system that integrates spatial representations of objects and visual scenes with sensorimotor representations has been developed by Schill and co-workers (1998, 2001). These integrated knowledge representations are used directly for the control of saccadic eye movements during visual scene analysis tasks. The basic representation scheme used in this approach is a feature triple of the form: *current sensory feature – eye movement – target sensory feature*.

The basic sensorimotor features (e.g. polygon patterns representing the characteristics of a polygon's vertices) are organized in hierarchical structures that represent entire visual scenes. The overall architecture of the system consists of a visual preprocessing component that extracts salient two-dimensional features, and a reasoning component that operates on the basis of belief measures according to Shafer (1976). The system adapts itself to possible scenes using a supervised learning component

that learns to relate scene concepts and sensorimotor features for subsequent processing. At each processing step the system analyzes the current visual feature (i.e. the feature that currently is in the focus); based on this feature and knowledge gained from already performed eye movements, the system determines a new target feature (i.e. its relative location) that promises to provide the maximum information gain towards the goal of identifying the scene under consideration. The corresponding shift of focus (modeling eye movement in natural cognitive systems) is performed and, the whole process starts over again. This cycle is repeated until a certain threshold of belief is reached and the scene is claimed to be identified.

4.3   Control of Attention in Diagram Understanding

Aiming at describing and modeling the mental processes involved in understanding external diagrammatic depictions is related to both fields discussed above, and thus it exhibits specific requirements with respect to control of focus: on the one hand, diagrams are conveyed through external media and are therefore related to object and scene recognition; on the other hand they are combined from interpreted and well-organized components that mentally need be dealt with in a specific manner, thus requiring extensive use of corresponding mental imagery processes.

Diagram understanding is especially demanding when the objective of the diagram is to convey dynamic behavior, for instance of a chain of events or of some mechanical device. Research on *mental animation* (e.g. Hegarty, 1992) aims at explaining how people manage to infer dynamic motion from static diagrammatic depictions. Since mental animation relies on complex reasoning processes in working memory, mental capacity limits make an efficient allocation of mental resources necessary. The prevailing mental strategy employed is by decomposing the task and by solving partial problems subsequently (*piecemeal* strategy). This decomposition is done according to the causal chain of events that characterizes the system's behavior.

Through dual task experiments, it has been shown that mental animation is highly related to visual mental processes: there are strong interferences between visuo-spatial working memory and mental animation, whereas verbal tasks do not interfere with animation tasks (Sims & Hegarty, 1997). Moreover, there is an immediate correlation between eye fixations of specific image components and the entities in the causal chain that need to be analyzed to investigate specific motion properties in the diagram.

Hegarty (1992) presents a production system that models mental animation of pulley systems that account for a strong successiveness in the reasoning chain (i.e. motion of a specific part can only be inferred based on immediately neighboring parts) as well as for a limited working memory capacity with respect to the number of image parts that can be simultaneously animated. So, in this model, the focus of the reasoning process successively moves from one entity in the chain to the next one, which propagates the inferred direction of motion through the pulley system. Being a production system, however, the model does not account for the analogical aspects of mental animation, i.e. for mental animation being performed in visuo-spatial working memory.

5   Spatial Representations Lead to Spatial Mechanisms

In the last two sections, we have presented an overview of some existing approaches and psychological models to get hands on a number of phenomena of attentional control and focus. In the following, we will identify basic spatial organizing principles and mechanisms across different mental representations and the processes associated with these representations.

In an editorial, Hommel et al. (2004) make the case for a highly distributed view of executive control of human behavior; with respect to the processes involved, they conclude that "most if not all of these processes may turn out to be disappointingly common and it may be their concert that creates the emergent property of being 'executive'". With respect to attentional control and focus in mental imagery, the rationale for the current paper is to motivate that basic spatial organizing principles and basic spatial processing mechanisms can account for a good part of those processes – at least from a computer science point of view. The argument for which we will try to provide grounds in the following is that (a) spatial properties in mental knowledge representations lead to spatial mechanisms in the mental processes which operate on such representations, and that (b) there exist basic spatial organizing principles and mechanisms which are common to different types of mental representations and processes, and which play an important role for relating representations of the different types. The interplay of spatial mechanisms resulting from spatial properties of the underlying representations and the way in which mechanisms in one domain can trigger mechanism in another domain can form the basis for an abstracted, functional model view of control of focus. This model view may avoid the problem of causal theories and may fulfill the requirements put forward by recent psychological evidence as discussed in Section 3.

In its postulation of basic (spatial) organization schemes and mechanisms common to different types of mental representations, our approach finds an ally in Cowan's embedded-process model of working memory (Cowan, 1988, 1999). His model posits a set of basic mechanisms of activation and attention which are universal to all or most working memory components. It is the comparable functional nature of mental procedures across different memory subsystems that should be considered rather than their respective and potentially differing implementations. Cowan states: "there is no single separate theoretical entity that I would call working memory; that is a practical, task-oriented label. What are potentially more meaningful in a theoretical sense are the basic mechanisms proposed to underlie this complex system, including activation of memory contents of an attentional process, and the contextual organization of memory" (Cowan, 1999, p.88).

In the following, we will specifically discuss spatial knowledge held in three mental faculties: representations in long-term memory, (spatial) mental models in working memory, and visual mental images. In addition, we will consider depictions in external diagrams as a fourth kind of representation. Figure 1 illustrates these four representational domains of mental spatial reasoning along with the connections we propose on the basis of spatial mechanisms and organizing principles.

5.1   Origins of Spatial Properties in Mental Representations of Space

There are two ways in which we can consider spatial properties of mental representations: First, there exist properties that are induced by the spatial structure inherent to the entities that are represented. Second, there exist spatial properties that are induced by the organization of the mental representation structures. Richardson and Spivey (2000), for example, present evidence that "spatial indexes are being employed by the cognitive system, even in a memory task where location is irrelevant".

An example for the first type are mereological properties of the represented entities that are reflected in mental representation by contiguity, leading to subsequent activation. In spatial configuration tasks we find another example in the way in which spatial ordering information is reflected in mental models as proposed by Johnson-Laird (1983). In the case of external diagrams that show causal connections, for example the working of a system of pulleys, the spatial organization of the external diagram is reflected in the sequence of reasoning steps (Hegarty, 2000).
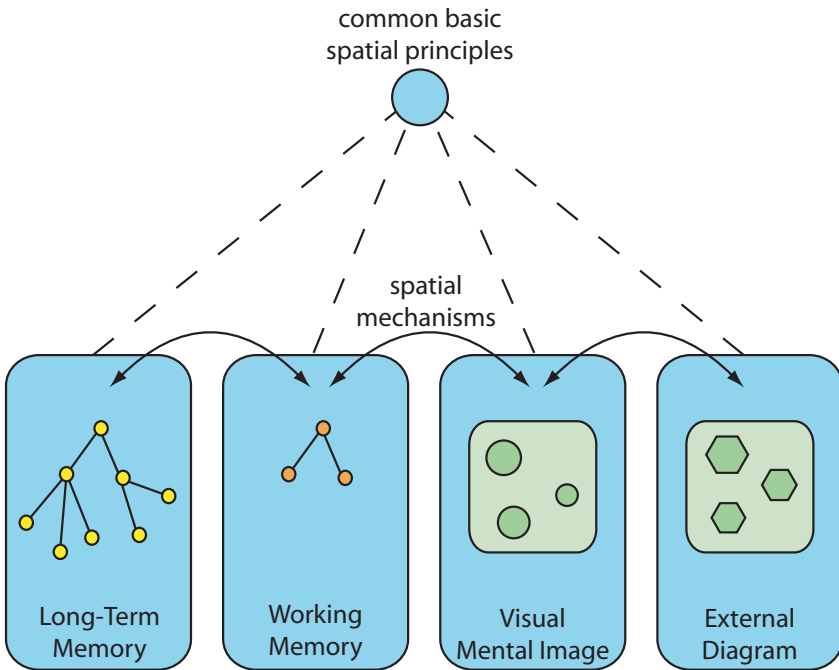
Fig. 1. Domains of spatial mechanisms and principles

Mental images provide good examples for the second type, i.e. representations in which spatial properties are induced by the organization of the representation structure: their construction involves an instantiation of spatial knowledge that exists in a non-visual mental format into a visuo-analogical form. Representational properties of this visuo-analogical form are largely determined by the representation structure (the

cognitive substrate, if you will) that holds the mental image. Along with an increase in specificity that necessarily goes along with the instantiation process (i.e. through graphical constraining, Scaife & Rogers, 1996), it is the specific representation format that allows to read off novel bits of information. In this property, mental images exhibit many similarities to external diagrams: through analogical representation and knowledge contained implicitly they facilitate reasoning (Shimojima, 1996). With respect to attentional focus during mental image inspection, the quasi-pictorial structure of the image has a strong influence on focus shifts. Spatial properties also get induced by spreading activation in mental representations that integrate individual knowledge fragments. Examples are integrated semantic structures in long-term memory as well as mental models in working memory. Effects of hierarchical organization of spatial knowledge (such as observed in the classical experiments of Stevens and Coupe, 1978) might be explained by activation spreading in semantic nets.

5.2   Spatial Principles Resulting from Effects of Spatial Structures

We have discussed two types of origins for spatial structures in mental representations, and we have pointed out that spatial properties are more than just another characteristic for a range of representations. This is true for visual mental images and external diagrams, but also for non-visual mental representations attributed to long-term memory and working memory. This subsection addresses the role that spatial effects and mechanisms, resulting from this spatial structure, play for attentional control and focus in reasoning with working memory representations, visual mental images, and spatial knowledge in long-term memory. The number of discussed effects have to be seen as an exemplary selection, rather than an exhaustive account. We thus focus on principles that are representative for our claims and that span all four representational domains (Fig. 1). In the following, we concentrate on the issues of *grouping* and *chunking*, *zooming* and *scanning,* and *sequentialization*.

5.2.1   Grouping and Chunking

Organizing pieces of information into larger, meaningful units is a universal cognitive principle. It can be found in all four domains of interest. If, throughout different representational formats, one conceptualizes spatial knowledge in terms of individual knowledge fragments, a number of different organizational schemes can be found. In the context of spatial reasoning, the different schemes find a unifying characteristic in that they all relate to the spatial properties of the representations involved.

In long-term memory, knowledge may be organized semantically (e.g. such as described by metaphors of semantic networks) or by temporal coherence in terms of episodes. Models of activation spreading on networks are frequently used to describe aspects of activation flow through those structures (e.g. Kokinov, 1997), and connection strengths between individual pairs of fragments in the structure may vary significantly. Individual fragments are largely described by propositional metaphors; however, certain information of shapes and configuration, certainly with respect to human faces, may be encoded otherwise (McDermott et al., 1999). The interlinking of the fragments frequently is conceptualized in the term of networks, possibly exhibiting a hierarchical structure (Stevens & Coupe, 1978). While long-term memory remains the

domain that is dominated least by spatial mechanisms, the characteristics of these network conceptualizations are at the transition from the purely propositional to the spatial, paving the way from semantically encoded knowledge to spatio-analogical representations in spatial reasoning.

For spatial reasoning in working memory, knowledge fragments may be organized in terms of mental models which fulfill roles as problem instantiations, mental theories, or are employed in mental simulations. The number of fragments integrated into one model is usually small, especially when compared to the structures that exist in long-term memory from which knowledge fragments in WM are typically derived. However, the degree of integration between fragments is high in mental models. With respect to computational characterizations of structures of spatial knowledge in WM, annotated graphs have often been suggested (e.g. by Barkowsky, 2002), where spatial entities serve as graph nodes and are connected by (mostly binary) topological, directional, distance, or other relations.

Visual mental images, on the other hand, are mental constructions in working memory that adhere to a specific representational format. The prevalent imagery paradigm stresses the relationship between mental images as representations and the visual buffer as the images' main representational substrate (Kosslyn, 1994; Kosslyn & Thompson, 2003). Mental images are constructed from mental models, and the individual actual organization of image parts into substructures largely matches that found in the corresponding mental model. Early work by Kosslyn and Pomerantz (1977) reported chunking in mental images, and the chunks reported seem to relate to a basic spatial structure underlying the mental model, as well (e.g. in the schematized figure of a human body, visual parts that make up an arm are chunked together). Content in mental images is frequently described to be interpreted (Logie, 2001), since, for most of its parts, it stems from semantic structures in LTM. The assumption of underlying spatial and semantic organization is in line with findings that structural reinterpretations of figures which are easy in perception are often hard in imagery (Chambers & Reisberg, 1992; Verstijnen et al., 1998): images already possess such a spatial and semantic grouping, where visual percepts yet have to be assigned one.

When visual mental images are externalized into diagrams, the groupings present in the image are preserved in the external diagram. Conversely, grouping effects during inspection, i.e. internalization of external diagrams, are reflected in the mental image formed, for instance according to Gestalt laws (Rock & Palmer, 1990).

Grouping, then, is one of the effects resulting from spatial organization principles. Although the actual mechanisms are different in each of the representational domains, they are connected on a functional level by their spatial grounding, and thus contribute to cognitive control during spatial reasoning: groupings are not only reflected in the various representations, they also entail one another on the basis of a common principle.

5.2.2 Zooming and Scanning

Inspection of visual mental images provides a good illustration of how the assumption of a common basic spatial organization underlying spatial knowledge in LTM representations, mental models, and mental images can help explain mechanisms of

retrieval or maintenance. Spatial operations during inspection of a visual mental image, like *zooming*, i.e. focussing on a detail of an image, or *scanning*, i.e. bringing another part of the image into focus, have direct effect on the image representation. Zooming might be conceptualized in terms of concentration whereas scanning might be conceptualized in terms of translation. They both lead to a change of focus on the image, or, in Kosslyn's (1994) terms, to a shift of the attention window. As with grouping, the effects are not limited to the single domain of mental imagery, but have direct influence in control in the other domains. For example, a scanning operation on a mental image may trigger retrieval processes that lead to activation of associated long-term memory content that thus becomes part of working memory.

In information-rich LTM structures, zooming can be conceptualized in terms of different extents of activation across structural parts: On a coarser level of inspection, only some distributed parts of the representation are fairly activated. Zooming in on one specific part leads to higher activation of its neighborhood along the structural connections, thus making neighboring components part of working memory (cf. the mechanisms of activation and attention suggested by Cowan, 1999).

The scanpath, as discussed in Section 3.3, is an example for how LTM content can influence operations in mental imagery: during the construction of multipart images, the scanpath is retrieved from long-term memory and enters working memory along with the stored knowledge fragments. During visualization, the spatial information of the scanpath directs the construction of multipart images, and thus has direct influence of the control of focus. The same can be suggested for the externalization processes of the content of the visual mental image into an external diagram: the scanpath might provide the spatial index during the piecemeal inspection and externalization. This feature brings us to the principle of sequentialization.

5.2.3 Sequentialization

We have identified the control of focus as a phenomenon resulting from distributed processes. In order to jointly achieve computational processing goals in reasoning with mental images, these processes need to work on a succession of representations. We identify the principle of sequentialization, i.e. the ordering of salient features in a complex structure, as a mechanism making the emergence of goal-directed behavior possible from distributed processes.

Sequentializations work in – and across – all four representational domains as introduced above. We have already identified one of the prime structures involved in sequentialization: the scanpath. During the inspection of external diagrams, eye movements, i.e. overt shifts of attention, organize the salient features of the perceived input into a meaningful sequence. As discussed above, this is the basis for the model of saccadic eye control, devised by Schill et al. (2001). According to the evidence presented in Section 3, this external scanpath can be conceived of as being transformed into a sequence of relative spatial indices that can be stored along with the perceived information in long-term memory. The scanpath in this apprehension is an abstracted spatial index. In this view, it can be attributed as a sequencing mechanism to all of the four representational domains of spatial mental reasoning. As with the other mechanisms, the actual implementations might differ in the domains, but the principle stays the same.

During spatial reasoning with mental images, spatial indexing plays a crucial role. It structures the parts of the working memory mental model to be visualized as a mental image in the spatio-analogical representation format. The sequence also controls the succession of activations from long-term memory for the generation of multipart images and the position for new parts of a multipart image to be integrated in an existing image.

During image inspection, the scanpath influences the chain of positions for the attention window, that in turn determine the succession of readings from the visual mental image. As an aside, the spatial sequentialization of the scan points of visual mental image could also be held partially accountable for the fact that mental images are harder to reinterpret than visual percepts (Verstijnen et al., 1998): the sequentialization of the scanpath provides a commitment to an interpretation that is not easily overruled. Only by re-externalization and re-inspection of the content can there be new sequentializations and thus new interpretations.

Finally, sequentialization is a necessary principle in the externalization of mental images. While many processes are engaged in the inspection of the mental image, the transformation to motor actions, and the feedback from eye movements and other sensory input in a highly parallel manner, the sequentialization of spatial indices provides the thread that gives structure to the adequate collaboration between the processes involved.

5.3 Summary

At the examples of grouping, scanning, and sequentialization, we have shown that on an abstracted level, there exist basic principles in reasoning with mental images that stem from spatial properties inherent in the representations involved. These principles cross the different representational domains of mental spatial reasoning, in each of which their actual implementation might differ. Through this universality, they guide the control of focus and aid goal-directed behavior emerging from autonomously working processes.

6 Implications for a Computational Model

In this section, we show how the insight of the previous sections can be reflected in a computational model. We discuss the implications of spatial principles on the one hand and distributed, non-central control mechanisms on the other, in the control of focus with respect to the *Casimir* model (computational architecture, specification and implementation of mental-image based reasoning). In the present discussion of cognitive control, we will concentrate on aspects of Casimir's underlying conceptual model.

6.1 Overview of the Conceptual Model

Figure 2 shows an exemplary processing cycle of imagery-based reasoning in the conceptual model of *Casimir*. We will use this high-level flow to discuss the workings of five major subsystems: long-term memory activation, image construction, image

inspection, memory update, and diagram inspection. As can be seen, these subsystems span all the four representational domains discussed in the previous section.

6.1.1 Long-Term Memory Activation

Long-term memory representations are accessed based on the representation of a (propositionally stated) spatial problem. Such a problem could be, for example, to decide upon the relative spatial orientation of two geographic locations. The problem representation could contain a query regarding the spatial relations that holds between two entities. The access process activates spatial knowledge fragments, i.e. structures that bring n entities into an n-ary spatial relation in long-term memory. Through this process, the knowledge fragments are transferred to working memory where they are integrated into a mental model representation.
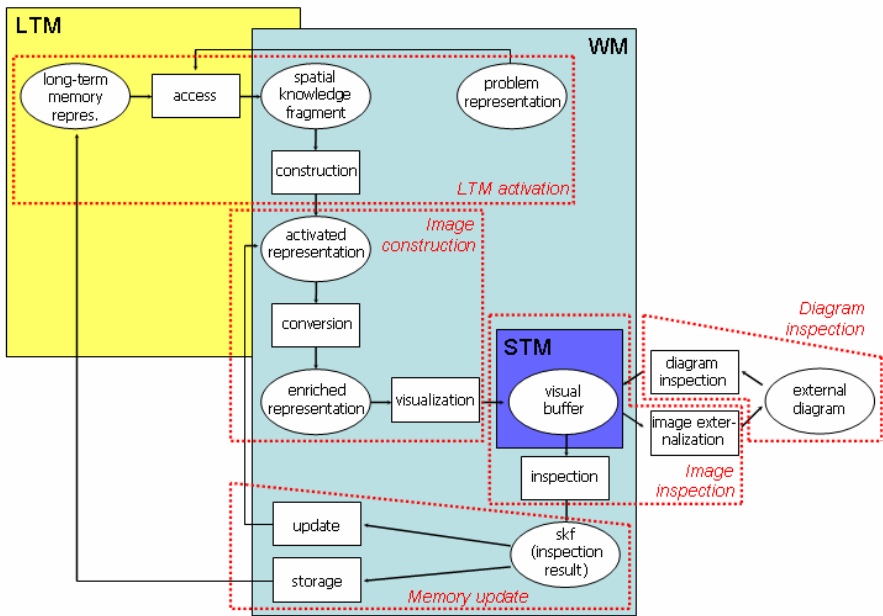
Fig. 2. Exemplary processing cycle and division into subsystems in the conceptual model of *Casimir*

6.1.2 Image Construction

The knowledge contained in this initial representation is usually underdetermined; it might be incomplete or too coarse to be visualized in a mental image. The *conversion* process in working memory enriches the representation by filling gaps in specificity with default assumptions, assigning ontological types and completing relations where necessary. The enriched mental model representation is the basis for the visualization process that activates parts of the visual buffer to produce a spatio-analogical representation, i.e. a mental image.

6.1.3   Image Inspection

The mental image held in the visual buffer is subject to inspection. The processes involved in inspection employ scanning and zooming on the mental image in order to read off previously undetected knowledge that can aid in reasoning and solving the problem at hand. The inspection processes thus may produce new spatial knowledge fragments, which can either serve as input to the memory update system for storage or further scrutiny during a subsequent processing cycle, or which might be passed on to the externalization processes, leading to the construction of an external diagram.

6.1.4   Memory Update

Regarding the further use of spatial knowledge fragments passed on by the inspection system, the memory update system has a twofold responsibility: On the one hand, these fragments can serve as an input to memory update processes in working and long-term memory, where long-term memory update is conceptualized as storage. The inspection result may also be passed to further cognitive functions, for example it may be verbalized or result in a motor action being taken.

6.1.5   (External) Diagram Inspection

The processes and structures responsible for interpreting perceptual input from external diagrams are grouped in the external diagram inspection subsystem. This system can receive perceptual input as well as input from working memory systems, reflecting the overlap between bottom-up perceptual and top-down imagery processes discussed in Section 3.

6.2   Control of Focus in *Casimir*

In line with the psychological findings we presented in Section 3, the flow of control in our model is conceived of as a distributed phenomenon of the interplay between autonomous components. In this collaboration, each of the components aims at achieving a local goal specific to that component, thus furthering the convergence of the system towards a global goal that is derived from the initial problem representation. The same is true on a macro-level for the individual subsystems: each of the systems triggers actions that refine a representation to match a local goal, for example the image construction subsystem works towards a representation that is suitable for visualization in the visual buffer. The global flow of control in *Casimir* thus results from the hierarchical composition of local goals on different granularities. The structure of the representations involved plays an important role in this process, as do unifying spatial principles. *Grouping*, for example, is a mechanism that is employed to transcend hierarchies of organization, whereas *sequentialization* is used to implement causal relationships on the same level of hierarchy; both principles can serve to provide action triggers that spread control of focus over subsystems and representational domains. Thus, on the finest granularity level, local goals are causal relationships, i.e., they are triggered from other goals on the basis of the structure of the representations they work with. On coarser granularity levels, the goals gradually assume a more global character, up to the goal of solving the initial problem.

The spatial principles are used to connect different subsystems and domains, both vertically and horizontally in terms of organizational hierarchy. A shift of focus in one domain can entail a shift of focus in a neighboring domain. This is best illustrated with a metaphoric conception of the transcendence of representations between the representational domains discussed above. Metaphorically, the representational domains can be conceived of as rotating disks on which the mental representations lie (Fig. 3). Consider the example of working memory content being visualized in a mental image. The trigger to activate visualization can be seen as a rotation of the "mental image disk", as illustrated by the rightmost arrow in Fig. 3. The "rotation" of more content into the mental image domain inevitable leads to other parts of working memory to come into focus and to the activation of associated long-term memory content, illustrated by the rotation of the leftmost arrow that is triggered by the rotation of the middle disk. While the "rotation" of the disks, i.e. the actual mechanisms, might differ in implementation in the different domains, there are principles that work across the domains and facilitate more global goals.
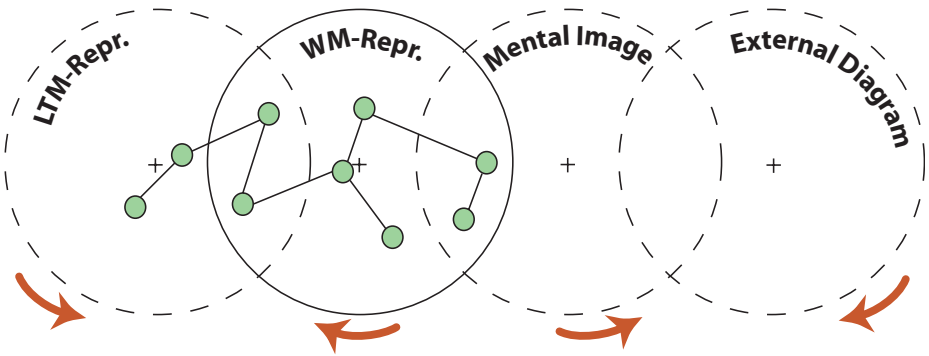
Fig. 3. *Rotating disks* metaphor for computational control mechanisms

As an example, consider zooming on a mental image as illustrated in Fig. 4. Based on activated long-term memory, i.e., content that is integrated into a working memory mental model ①, a first mental image is visualized in the visual buffer ②. According to the goals of local inspection processes, a part of the mental image is subjected to zooming ③. This triggers a shift of focus in the mental model representation ④ and leads to further retrieval from long-term memory ⑤, to provide a more detailed model ⑥. The changed focus in the mental model leads to an adaptation of focus in the mental image, which provides the frame for a detailed visualization ⑦. As this rough sketch shows, changes in the flow of control in any one of the representational domains entail a change in focus in the other domains and may trigger actions in other domains that serve to achieve local goals, for instance acquiring more detail for the entity in focus, regardless of the implementation of the parallel processes of the next lower level.

The aim to achieve locally defined goals and the connection of representations over different representational domains result in dynamic collaboration between neighboring system components. This can facilitate the construction of collaborative networks in which components join processes, capacities, and resources on the basis of the connecting principles. A possible result could, for example, be that *access*, *construction*, and *conversion* processes are joined in a special-purpose network to collaborate in solving a simple spatial configuration problem with only little help from other components.
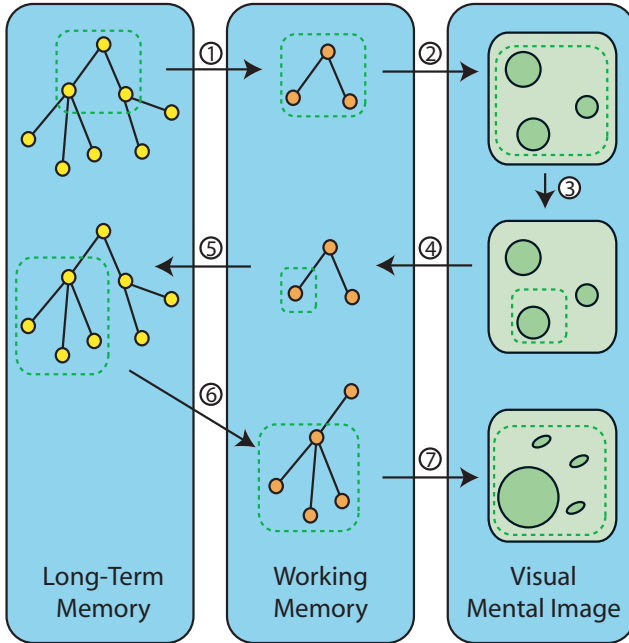

Fig. 4. The effect of zooming in different domains

7   Conclusion

We have reviewed psychological evidence relating to cognitive control and mental spatial reasoning with representations in four domains: long-term memory and working memory representations, visual mental images, and external diagrams. After having discussed the origins of spatial structures in all four of these domains, we have presented arguments that basic spatial organizational principles exert direct influence on the control of focus during mental spatial reasoning. We have outlined possible implications for a specific computational model.

All of the above has been conducted on a relatively high level of abstraction, with the aim of pointing to the existence of unifying principles across representational domains that have an influence on cognitive control. In further work, we will address

these principles in more detail and look into the involved mechanisms in each of the domains. As regards the application in a computational model, we are currently refining the metaphors we presented in this paper and extract implications and structures for the implementation of *Casimir*.

With respect to the application perspective in an assistance scenario, for example in spatial configuration tasks, the presented work provides the foundation for more research into principles effecting cognitive control for providing users with adequate assistance with regard to (a) representations in external diagrams, (b) process flows, and (c) cognitive load. Sequentialization operations, for example, might be applied to external representations for highlighting important features (cf. the notion of *aspectualization* as introduced in Bertel et al., 2004). The insights of connecting principles across different domains of representations regarding shifts of focus could be reflected in the design of work flows and process flows; thus, navigation through dynamic configurations may be facilitated. Finally, the insights into mechanisms of cognitive control of focus in the *Casimir* model may serve to give a rough estimate of the cognitive load (induced by mechanisms associated with cognitive control) that a specific spatial task imposes on a human reasoner.

Acknowledgments

References

Allport, A. (1993). Attention and control: have we been asking the wrong questions? A critical review of twenty-five years. In Meyer, D. E. & Kornblum, S. M. (Eds.), *Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, Vol. XIV of *Attention and Performance*. Cambridge, MA: MIT Press.

Aloimonos, Y. (Ed.) (1993). *Active Perception*. Mahwah, NJ: Lawrence Erlbaum.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20* (4), 723–767.

Barkowsky, T. (2002). *Mental representation and processing of geographic knowledge – A computational approach*. Berlin: Springer.

Bartolomeo, P., & Chokron, S. (2002). Orienting of attention in left unilateral neglect. *Neuroscience and Biobehavioral Reviews*, *26*, 217–234.

Bertel, S., Freksa, C., & Vrachliotis, G. (2004). Aspectualize and conquer in architectural design. In J. S. Gero, B. Tversky, & T. Knight (Eds.), *Visual and Spatial Reasoning in Design III* (pp. 255–279). Key Centre of Design Computing and Cognition, University of Sydney.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*, 27–38.

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon.

Chambers, D., & Reisberg, D. (1992). What an image depicts depends on what an image means. *Cognitive Psychology*, *24*, 145–174.

Colby, C. L., & Duhamel, J. R. (1991). Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia, 29*, 517–537.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104,* 163–191.

Cowan, N. (1999). An embedded-process model of working memory. In P. Shah, & A. Miyake (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). New York: Cambridge University Press.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: Convergent visual and somatic response properties. *Journal of Neurophysiology, 79*, 126–136.

Eimer, M., Forster, B., & van Velzen, J. (2003). Anterior and posterior attentional control systems use different spatial reference frames: ERP evidence from covert tactile-spatial orienting. *Psychophysiology*, *40*, 924–933.

Glasgow, J., & Papadias, D. (1992). Computational imagery. *Cognitive Science*, *16*, 355–394.

Goldschmidt, G. (1995). The designer as a team of one. *Design Studies*, *16*, 189–209.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15* (8), 1176–1194.

Hazlett, C. J., & Woldorff, M. G. (2004). Mechanisms of moving the mind's eye: Planning and execution of spatial shifts of attention. *Journal of Cognitive Neuroscience*, *16* (5), 742–750.

Healey, P. G. T., Swoboda, N., Umata, I., & Katagiri, Y. (2002). Graphical representation in graphical dialogue. *International Journal of Human-Computer Studies*, *57*, 375–395.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18* (5), 1084–1102.

Hegarty, M. (2000). Capacity limits in diagrammatic reasoning. In M. Anderson, P. Cheng., & V. Haarslev (Eds.), *Theory and application of diagrams* (pp. 194–206). Berlin: Springer.

Hirtle, S. C. (1998). The cognitive atlas: using GIS as a metaphor for memory. In M. Egenhofer & R. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 267–276). Oxford University Press.

Hommel, B., Daum, I., & Kluwe, R. H. (2004). Exorcizing the homunculus, phase two: Editors' introduction. *Acta Psychologica*, *115*, 99–104.

Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, *268*, 1772–1774.

Ishai, A., & Sagi, D. (1997). Visual imagery facilitates visual perception: Psychophysical evidence. *Journal of Cognitive Neuroscience*, *9* (4), 476–489.

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, *28*, 979–990.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, *37*, 43–75.

Kieras, D. E., Meyer, D. E., Ballas, J. A., & Lauber, E. J. (2000). Modern computational perspectives on executive mental processes and cognitive control: Where to from here? In S. Monsell, & J. Driver, (Eds.), *Control of Cognitive Processes*, Vol. XVIII of *Attention and Performance* (pp. 681–712). Cambridge, MA: MIT Press.

Kokinov, B. (1997). Micro-level hybridization in the cognitive architecture dual. In R. Sun, & F. Alexander (Eds.), *Connectionist-symbolic integrations: From unified to hybrid architectures*. Hilsdale, NJ: Lawrence Erlbaum Associates.

Kosslyn, S. M., & Pomerantz, J. P. (1977). Imagery, propositions and the form of internal representations. *Cognitive Science*, *9*, 52–76.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 47–60.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1994). *Image and brain – The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Kosslyn, S. M., & Sussman, A. L. (1995). Roles of imagery in perception: Or, there is no such thing as immaculate perception. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1035–1042). Cambridge, MA: MIT Press.

Kosslyn, S. M., & Thompson, W. M. (2003). When is early visual cortex activated during visual mental imagery?. *Psychological Bulletin*, *129* (5), 723–746.

Kuipers, B. (1982). The 'map in the head' metaphor. *Environment and Behavior, 14* (2), 202–220.

Laeng, B., & Teodorescu, D. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scence. *Cognitive Science*, *26*, 207–231.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.

Lockhead, G. R., & Pomerantz, J. R. (1991). *The perception of structure.* Washington, D.C.: American Psychological Association.

Logie, R. H. (2001). Working memory: A mental space for design and discovery. In J. S. Gero, B. Tversky, & T. Purcell (Eds.), *Visual and spatial reasoning in design II* (pp. 223–235). University of Sydney, Australia: Key Centre of Design Computing and Cognition.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Mast, F. W., & Kosslyn, S. M. (2002). Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition*, *86*, 57–70.

McDermott, K. B., Buckner, R. L., Peterson, S. E., Kelley, W. M., & Sanders, A. L. (1999). Set- and codespecific activation in frontal cortex: an fMRI study of encoding and retrieval of faces and words. *Journal of Cognitive Neuroscience, 11* (6): 631-40.

Michelon, P., & Zacks, J. M. (2003). What is primed in priming from imagery? *Psychological Research*, *67*, 71–79.

Monsell, S., & Driver, J. (2000). Banishing the control homunculus. In S. Monsell, & J. Driver (Eds.), *Control of Cognitive Processes*, Vol. XVIII of *Attention and Performance* (pp. 3–32). Cambridge, MA: MIT Press.

Montello, D. R. (1992). The geometry of environmental knowledge. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 136–152). Berlin: Springer.

Moray, R. S. (1973). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *9*, 56–90.

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, *8*, 228–246.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. *Journal of Cognitive Neuroscience*, *16* (3), 363–373.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American*, *224*(6).

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259–303). Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*, *4*, 1863–1874.

Posner, M. I. (1993). Attention before and during the decade of the brain. In D. E. Meyer, & S. M. Kornblum (Eds.), *Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, Vol. XIV of *Attention and Performance*. Cambridge, MA: MIT Press.

Purcell, A. T, & Gero, J. S. (1998). Drawings and the design process. *Design Studies*, *19*, 389–430.

Pylyshyn, Z. W. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences, 7*, 113–118.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: looking at things that aren't there anymore. *Cognition*, *76* , 269–295.

Rock, I., & Palmer, S. (1990). The legacy of Gestalt Psychology. *Scientific American*, *263* (6), 84–90.

Scaife, M., & Rogers, Y. (1996) External cognition: How do graphical representations work?, *International Journal of Human-Computer Studies*, *45*, 185–213.

Schill, K., Umkehrer, E., Beinlich, S., Zetzsche, C., Deubel, H., & Pöppel, E. (1998). A hybrid system for the scene analysis with saccadic eye movements: Learning of feature relations. *Perception, 27*, 6.

Schill, K., Umkehrer, E., Beinlich, S., Krieger, G., & Zetzsche, C. (2001). Scene analysis with saccadic eye movements: top-down and bottom-up modeling. *Journal of Electronic Imaging, Special Issue on Human Vision and Electronic Imaging, 10* (1), 152–160.

Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, *294*, 1350–1354.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Shimojima, A. (1996). Operational constraints in diagrammatic reasoning. In Allwein, & Barwise (Eds), *Logical reasoning with diagrams* (pp. 27–48). Oxford University Press.

Sims, V. K., & Hegarty, M. (1997). Mental animation in the visual-spatial sketchpad: Evidence from dual-task studies. *Memory & Cognition, 25*, 321–332.

Stark, L. W., & Choi, Y. S. (1996). Experimental metaphysics: The scanpath as an epistemological mechanism. In W. H. Zangemeister, H. S. Stiehl, & C. Freksa (Eds.), *Visual Attention and Cognition*. Amsterdam: Elsevier.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, *19*, 97–140.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422–437.

Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review, 55* (4), 189–208.

Treisman, A. M. (1964). Monitoring and storage of irrelevant messages and selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*, 449–459.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. Frank, & I. Campari (Eds.), *Spatial information theory – A theoretical basis for GIS* (pp. 14–24). Berlin: Springer.

Tye, M. (1991). *The imagery debate*. Cambridge, MA: MIT Press.

Verstijnen, I. M., van Leeuwen, C., Goldschmidt, G., Hamel, R., & Hennessey, J. M. (1998). Creative discovery in imagery and perception: Combining is relatively easy, restructuring takes a sketch. *Acta Psychologica, 99*, 177–200.

Wilson, M. (2002). Six views of embodied cognition, *Psychonomic Bulletin and Review*, *9*, 625–636.