the most efficient way of producing a common system of representations referring to perceived and to-be-produced events is by creating codes with distal reference. I would rather say that this is, again by definition, the only way of implementing any system of representations referring to events, whether or not they are coded within a common medium.

Remember that "event" means here "event in the outer world." On that ground, the creation of event codes implies the creation of distal codes because representations that refer to events in the environment are, necessarily, distal representations. Hence, creating representations of distal entities is not a more or less efficient way of implementing representations of events. It is, rather, a trivial consequence of the way in which the terms "distal representations" and "representations of events" are defined. Claiming that it is the most efficient way is like claiming that the most efficient way for a married man to become a bachelor is by getting unmarried. If this were the most efficient way, there would be other, less efficient, ways. But there are none. Therefore, it is true that producing event codes requires creating distal codes – but note, first, that this is not an empirical but a conceptual truth; and, second, it has nothing to do with the common representational system for perception and action advocated by Hommel et al.

There is no Authors' Response to this continuing commentary.

Reference

Hommel, B., Müsseler, J., Aschersleben, G. & Prinz, W. (2001) The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences* 24(5):849–78. [MS]

Commentary on **Zenon W. Pylyshyn (2002). Mental imagery? In search of a theory. BBS 25(2):157–182.**

Abstract of the original article: It is generally accepted that there is something special about reasoning by using mental images. The question of how it is special, however, has never been satisfactorily spelled out, despite more than thirty years of research in the post-behaviorist tradition. This article considers some of the general motivation for the assumption that entertaining mental images involves inspecting a picture-like object. It sets out a distinction between phenomena attributable to the nature of mind – what is called the cognitive architecture, and ones that are attributable to tacit knowledge used to simulate what would happen in a visual situation. With this distinction in mind the paper then considers in detail the widely held assumption that in some important sense images are *spatially displayed* or are *depictive,* and that examining images uses the same mechanisms that are deployed in visual perception. I argue that the assumption of the spatial or depictive nature of images is only explanatory if taken literally, as a claim about how images are physically instantiated in the brain, and that the literal view fails for a number of empirical reasons – for example, because of the cognitive penetrability of the phenomena cited in its favor. Similarly, while it is arguably the case that imagery and vision involve some of the same mechanisms, this tells us very little about the nature of mental imagery and does not support claims about the pictorial nature of mental images. Finally, I consider whether recent neuroscience evidence clarifies the debate over the nature of mental images. I claim that when such questions as whether images are depictive or spatial are formulated more clearly, the evidence does not provide support for the picture-theory over a symbol structure theory of mental imagery. Even if all the empirical claims were true, they do not warrant the conclusion that many people have drawn from them; that mental images are depictive or are displayed in some (possibly cortical) space. Such a conclusion is incompatible with what is known about how images function in thought. We are then left with the provisional counterintuitive conclusion that the available evidence does not support rejection of what I call the "null hypothesis"; namely, that reasoning with mental images involves the same form of representation and the same processes as that of reasoning in general, except that the content or subject matter of thoughts experienced as images includes information about how things would look.

Spatial inference: No difference between mental images and mental models

Markus Knauff[a,b] and Christoph Schlieder[c]

[a]*Max-Planck-Institute for Biological Cybernetics, D-72076 Tübingen, Germany, and* [b]*Center for Cognitive Science, University of Freiburg, D-79098 Freiburg, Germany;* [c]*Faculty of Information Systems and Applied Computer Sciences, University of Bamberg, D-96045 Bamberg, Germany.*
markus.knauff@tuebingen.mpg.de
christoph.schlieder@wiai.uni-bamberg.de

Abstract: In contrast to Pylyshyn's view, there is no such thing as "reasoning in general." Different types of reasoning tasks are solved with different reasoning strategies. A more specific null hypothesis is that *spatial inference* with mental *images* involves the same representational formalism as that of spatial inference with mental *models.* There is no evidence that this hypothesis must be rejected.

Pylyshyn's line of argumentation starts with a "null hypothesis" and ends, after a scholarly review of empirical work, with the conclusion that the null hypothesis cannot be refuted. His hypothesis reads: "reasoning with mental images involves the same form of representation and the same processes as that of reasoning in general" (Pylyshyn 2002, p. 157). Unfortunately, the way in which the hypothesis is stated prevents it from being refuted. It has the form *X involves the same Y as* Z, with an *X* whose ontological status is disputed and a non-existing Z. There is no such thing as "reasoning in general" because different types of reasoning tasks are solved with different reasoning strategies. Consider, for instance, the following inference problem:

The hammer is to the right of the pliers
The screwdriver is to the left of the pliers
The wrench is in front of the screwdriver
The saw is in front of the hammer
Which spatial relation holds between the wrench and the saw?

In principle, such tasks could be solved by a reasoning strategy that consists of applying inference rules as suggested by Rips (1994). But this account has an important consequence: the premises must be kept separately in mind throughout the whole reasoning process. No integrated representation of the problem information is constructed. Empirical evidence suggests that a different type of strategy is adopted by most subjects: An integrated representation, or mental model, is constructed in working memory and constitutes the information on which the reasoning process operates (see, e.g., Johnson-Laird & Byrne 1991). A mental model for a spatial inference task encodes a specific spatial configuration satisfying all the premises; for example:

screwdriver	pliers	hammer
wrench		saw

What is specific (and spatial) in this integrated representation? First, we need to clarify what is meant by properties of a representation. Discussing the functional space, Pylyshyn argues that a matrix considered as a data structure without the computational processes operating on it does not embody any intrinsic representational constraints. Although this is a common view in cognitive science (Anderson 1978), it is surprising from a computational perspective, because representation without processing is a void concept. All representational constraints on images are extrinsic in Pylyshyn's sense because they are resulting from the interaction between data structures and processing (Schlieder 1998). Therefore, it seems that the distinction between intrinsic and extrinsic constraints of data structures should be abandoned. In contrast, properties (e.g., intrinsic vs. extrinsic) of representational formalisms should be considered that encompass both data and processes. In the case of spatial inferences, the representational formalism consists of a data structure, the mental model, and processes that construct, inspect, and modify this model.

Does mental imagery play a role in this representational system as reports of subjective experiences suggest (cf. Kosslyn 1994, p. 324)? Or does spatial inference, as mental model theory suggests, rely on a representational formalism that is spatial in the sense that it represents the relative position of objects to other objects, but not necessarily visual features such as shape, size, distance, and so on? In response to these questions, we propose to reformulate Pylyshyn's null hypothesis: Spatial inference with mental *images* involves the same form of representation and the same processes as spatial inference with mental *models*. Obviously, this hypothesis is more specific as it refers not to reasoning in general but to a particular reasoning strategy (X = spatial inference with images; Y = representational formalism; Z = spatial inference with mental models). At the same time, it is more amenable to empirical testing because it asks for the involvement of cognitive systems devoted either to visual or to spatial processing. There are several groups of experiments that help to answer our questions. For instance, if reasoning involves visual images that differ from models, then the "imageability" of the materials should affect reasoning performance. Johnson-Laird et al. (1989) failed to find any differences between problems that were easy or difficult to visualize. And Knauff and Johnson-Laird (2002) found that if the content of an inference problem helps to construct a spatial mental model, then comprehension and reasoning proceed smoothly. But, if the content gives rise to visual images, reasoning is impeded by irrelevant visual details. Klauer et al. (1997) showed that spatial inference interferes with the preoccupation of visual and spatial working memory, whereas phonological secondary tasks do not impair reasoning performance (cf. also Vandierendonck & de Vooght 1997). However, if the visual and the spatial components in the secondary tasks are varied separately, processes devoted to purely visual features of objects turn out to be irrelevant (Knauff et al. 2004). Experiments using functional brain imaging measured cortical activity during spatial inference in parietal areas but not in early visual areas (e.g., Goel & Dolan 2001; Knauff et al. 2002). The parietal cortex is thought to play a key role in spatial processing and working memory (e.g., Baker et al. 1996) and in the integration of sensory information from all modalities into egocentric spatial representations (Burgess et al. 2001; Xing & Andersen 2000).

The details of the reported studies are not as important as the general points: (1) In the case of spatial inferences, it seems reasonable that mental models together with the processes that construct, inspect, and modify these models encode the spatial properties described in the premises of the inference problem. Hence, the representational formalism (i.e., representations and processes) is intrinsically spatial. (2) The abstract spatial nature of the mental models corresponds to activity in areas of the brain that are involved in the processing of spatial information from different modalities. (3) The mental models do not necessarily represent visual features (shape, size, etc.) and the subjective experience of visual imagery is probably an irrelevant side effect. Overall, based on this evidence we cannot reject the specific null hypothesis that spatial inference with mental *images* involves the same representational formalism as that of spatial inference with mental *models*. Whether this conclusion generalizes to all sorts of reasoning, as claimed by Pylyshyn, definitely needs further investigation.

ACKNOWLEDGMENTS
The first author is supported by a Heisenberg Award from the Deutsche Forschungsgemeinschaft (DFG; German National Research Foundation). The research was also supported by grants from the DFG under contract number Kn465/2-4 and in the Transregional Collaborative Research Center Spatial Cognition (SFB/TR 8). The authors greatly appreciate the helpful comments of Christian Ruff and Gerhard Strube on an earlier draft of this work.

Author's Response

From reifying mental pictures to reifying spatial models

Zenon W. Pylyshyn
Rutgers Center for Cognitive Science, Rutgers University at New Brunswick, Piscataway, NJ 08854-8020. **zenon@ruccs.rutgers.edu**

Abstract: Assuming that the vehicle of imaginal thought is a spatial model may not be quite as egregious an error as assuming it is a two-dimensional picture, but it represents no less a reification error. Because the model is not a literal physical layout, one is still owed an explanation of *why* spatial properties hold in the model – whether because of architectural constraints or by stipulation. The difference is like the difference between explaining behavior from a principle and predicting it by looking it up in a list. In the latter case no purpose is being served by calling it a mental model.

There is much in the **Knauff & Schlieder** commentary with which I agree. For example, I agree that spatial representations are not images in the sense of being perceivable pictures and I agree that representations are to be viewed in conjunction with the processes that operate over them. What I disagree with is the conclusion these commentators draw from their premises. In particular they appear to make exactly the same reification error that pervades the entire imagery literature, only they reify spatial layouts rather than pictorial objects.

Interestingly, the example they offer is the very one I have discussed elsewhere (Pylyshyn 2003, pp. 451–55). It concerns the location of a collection of objects. Discussion of such examples invariably involves exhibiting a table on which the objects are laid out in a fixed pattern. **Knauff & Schlieder** refer to this as an "integrated representation . . . constructed in working memory" and claim that it is spatial "in the sense that it represents the relative position of objects to other objects." They clearly do not wish to assume there is a table in the head, so it is a puzzle why the "integrated" characterization is not equally true of a representation such as the one they give in the problem statement, augmented by some preparatory inferences such as those concerning the relative locations of pliers and hammer, screwdriver and saw, and so on – exactly the sorts of inferences that have been shown to occur in the course of understanding a problem statement (e.g., Bransford & Johnson 1973). Why is this not an "integrated representation"?

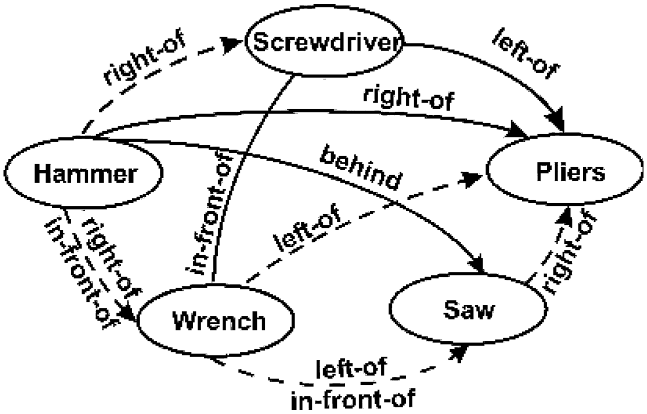

Figure R1 (Pylyshyn). A network implementation of the mental model example cited by Knauff & Schlieder. Solid lines are part of the problem statement. Dotted lines indicate inferred relations inserted in the course of understanding the problem statement.

When it is expressed as a network such as in Figure R1, is that integrated enough? It allows some inferences to be "read off" the data-structure quickly, just as in the mental model.

It is clear that **Knauff & Schlieder** do not want such a data structure to qualify as a "mental model" although they may be hard put to say why. In fact it is at this point that talk of mental models invariably lapses into the same reification error as does talk of mental pictures: It assumes that the objects in the mental model conform to physical principles such as rigidity, group movement, invariance of relative locations with changes to some parts of the model, and so on, and that objects stay put and are available for inspection as attention is applied to it. The "frame problem" in Artificial Intelligence (Pylyshyn 1987) arises precisely because one is not entitled to make such assumptions in general. To the extent that any such assumptions hold in a particular problem, they have to be explicitly stated. If they hold for every problem in a certain general domain (e.g., all problems involving spatial layouts), there is some prospect that they might be built in to the computational architecture. But spatial properties of the sort that are appealed to in the example are clearly contingent – the problem could have been one in which the objects were floating in a turbulent liquid or attached to balloons. Though this would still be a problem involving spatial locations, nobody would reason about it by examining the sort of spatial model that Knauff & Schlieder advocate. Why not? Clearly because in this case the assumptions about rigid configuration would not hold. But this just shows that the constraints are part of the particular problem statement (they are cognitively penetrable) and therefore need to be expressed in "frame axioms" rather than in the fixed architecture.

This has nothing to do with whether the properties are in the data structures or in the algorithm. Rather, it has to do with whether the properties are being modeled as part of the mental architecture or as part of the problem understanding. It does no good to point to the computer model and say that the array *together with its access process* is functionally like a table, because it could be functionally like anything one chooses since the function it can compute is unconstrained (beyond being a recursive function). An explanation has to have fewer degrees of freedom than the data it explains – it has to say why *this* pattern rather than *that* pattern occurs when either could have been programmed.

In order for the program to serve an explanatory role, the theorist must say how it maps onto the mental process. In particular, the theorist has to make a commitment as to whether the array-plus-algorithm is presented as a model of the cognitive process or as an emulation of the medium of representation (i.e., the mental architecture; see Pylyshyn 1984). To abandon the distinction between architecture and process is to abandon the distinction between explaining and describing (or asserting). It is the difference between deriving a prediction from an assumption about the nature of the computational architecture and reading it off a list in which one has stored what will happen under various circumstances. Both may be equally predictive, but the list is not explanatory because it is unconstrained.

References

The letter "r" before author's initials stand for Author's Response to CC references

Anderson, J. (1978) Arguments concerning representations for mental imagery. *Psychological Review* 85: 249–77. [MK]

Baker, S. C., Frith, C. D., Frackowiak, R. S. J. & Dolan, R. J. (1996) Active representation of shape and spatial location in man. *Cerebral Cortex* 6:612–19. [MK]

Bransford, J. D. & Johnson, M. K. (1973). Considerations of some problems of comprehension: In: *Visual information processing*, ed. W. Chase. Academic Press. [rZWP]

Burgess, N., Maguire, E. A., Spiers, H. J. & O'Keefe, J. (2001) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *NeuroImage* 14:439–53. [MK]

Goel, V. & Dolan, R. J. (2001) Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia* 39:901–909. [MK]

Johnson-Laird, P. N. & Byrne, R. M. J. (1991) *Deduction*. Erlbaum. [MK]

Johnson-Laird, P. N., Byrne, R. M. J. & Tabossi, P. (1989) Reasoning by model: The case of multiple quantifiers. *Psychological Review* 96:658–73. [MK]

Klauer, K. C., Stegmaier, R. & Meiser, T. (1997) Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning* 3:9–47. [MK]

Knauff, M. & Johnson-Laird, P. N. (2002) Visual imagery can impede reasoning. *Memory and Cognition* 30:363–71. [MK]

Knauff, M., Jola, C., Strube, G., Rauh, R. & Schlieder, C. (2004) The psychological validity of qualitative spatial reasoning in one dimension. *Spatial Cognition and Computation* 4: 167–88. [MK]

Knauff, M., Mulack, T., Kassubek, J., Salih, H. R. & Greenlee, M. W. (2002) Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research* 13:203–12. [MK]

Kosslyn, S. M. (1994) *Image and brain*. MIT Press. [MK]

Pylyshyn, Z. W. (1984) *Computation and cognition: Toward a foundation for cognitive science*. MIT Press. [rZWP]

(Ed.) (1987) *The robot's dilemma: The frame problem in artificial intelligence*. Ablex. [rZWP]

(2003) Seeing and visualizing: It's not what you think. MIT Press/Bradford Books. [rZWP]

Rips, L. J. (1994) *The psychology of proof*. MIT Press. [MK]

Schlieder, C. (1998) Diagrammatic transformation processes on two-dimensional relational maps. *Journal of Visual Languages and Computing* 9:45–59. [MK]

Vandierendonck, A. & de Vooght, G. (1997) Working memory constraints on linear reasoning with temporal and spatial contents. *Quarterly Journal of Experimental Psychology* 50A:803–20. [MK]

Xing, J. & Andersen, R. A. (2000) Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *Journal of Cognitive Neuroscience* 12:601–14. [MK]