Psychological Review

A Theory and a Computational Model of Spatial Reasoning With Preferred Mental Models

Marco Ragni and Markus Knauff

Online First Publication, June 10, 2013. doi: 10.1037/a0032460

Psychological Review

0033-295X/13/$12.00 DOI: 10.1037/a0032460

A Theory and a Computational Model of Spatial Reasoning With Preferred Mental Models

Marco Ragni
University of Freiburg

Markus Knauff
University of Giessen

Inferences about spatial arrangements and relations like "The Porsche is parked to the left of the Dodge and the Ferrari is parked to the right of the Dodge, thus, the Porsche is parked to the left of the Ferrari," are ubiquitous. However, spatial descriptions are often interpretable in many different ways and compatible with several alternative mental models. This article suggests that individuals tackle such indeterminate multiple-model problems by constructing a single, simple, and typical mental model but neglect other possible models. The model that first comes to reasoners' minds is the *preferred mental model*. It helps save cognitive resources but also leads to reasoning errors and illusory inferences. The article presents a *preferred model theory* and an instantiation of this theory in the form of a computational model, *preferred inferences in reasoning with spatial mental models* (PRISM). PRISM can be used to simulate and explain how preferred models are constructed, inspected, and varied in a spatial array that functions as if it were a spatial working memory. A spatial focus inserts tokens into the array, inspects the array to find new spatial relations, and relocates tokens in the array to generate alternative models of the problem description, if necessary. The article also introduces a general measure of difficulty based on the number of necessary focus operations (rather than the number of models). A comparison with results from psychological experiments shows that the theory can explain preferences, errors, and the difficulty of spatial reasoning problems.

Keywords: thinking, reasoning, spatial cognition, preferred mental models, computational modeling

Supplemental materials: http://dx.doi.org/10.1037/a0032460.supp

Spatial cognition is a particularly active field of research in psychology. One reason is that the space around our bodies is inextricably connected to perception and action (Allen, 2004; Burgess, 2008; Denis & Loomis, 2007). Another reason is that we constantly have to solve spatial problems, plan, and make decisions in space. People, objects, regions, events, and all kinds of entities are situated somewhere in the physical world and are spatially related to one another. Many spatial relations, though, are not obvious or explicitly given, and we are forced to infer where entities are located in relation to each other. Imagine, for instance, that we tell you that Hamburg is to the north of Berlin and Berlin is to the north of Dresden. You may not know exactly where in Germany the cities are located. Nevertheless, you should not have a problem inferring that Hamburg must be to the north of Dresden. The inference is simple for most people, but psychologists have identified several factors that make a spatial reasoning problem difficult to solve. Examples of such factors include the influence of information presentation order (Ehrlich & Johnson-Laird, 1982, to which we return later) and the use of nontransitive relations such as "next to," "overlap," or "contact" instead of transitive relations like "to the north of," "left of," "right of," "in front of," "behind," and so on (Goodwin & Johnson-Laird, 2008; Knauff & Ragni, 2011). Also complicating the matter are the effect of complex n-place relations such as "in between" or "equidistant" instead of binary relations (Goodwin & Johnson-Laird, 2005; Jahn, Knauff, & Johnson-Laird, 2007; Ragni, Fangmeier, Webber, & Knauff, 2006), the imageability of spatial relations (Knauff & Johnson-Laird, 2002; Knauff & May, 2006), and, most importantly, the contrast between determinate and indeterminate spatial reasoning problems (Byrne & Johnson-Laird, 1989; Carreiras & Santamaría, 1997; Johnson-Laird & Byrne, 1991; Roberts, 2000; Schaeken, Girotto, & Johnson-Laird, 1998). To understand the differences between *determinate* and *indeterminate* problems—which is a

Marco Ragni, Center for Cognitive Science, University of Freiburg, Freiburg, Germany; Markus Knauff, Department of Psychology, University of Giessen, Giessen, Germany.

The experiments were first presented at the 28th Annual Conference of the Cognitive Science Society in Vancouver, British Columbia, Canada, and in summarized form at the International Conference on Spatial Cognition in Bremen, Germany, 2006 (Ragni, Fangmeier, Webber, & Knauff, 2006, 2007). Marco Ragni's research was supported by the German Research Foundation (DFG) in the Transregional Collaborative Research Center, SFB/TR 8 within projects R8-[CSPACE], and the strategic project ACTIVATIONSPACE. Markus Knauff's research was supported by the DFG under contract numbers KN 465/6-2 and KN 465/11-1 and by grants from the DFG within the priority program "New Frameworks of Rationality" (SPP 1516). Our thanks go to Thomas Fangmeier for analyzing the second experiment, to Felix Steffenhagen for programming parts of the PRISM model, and to Sven Brüssow for helpful discussions. We also thank Rebecca Albrecht for drawing portions of our figures and Stephanie Schwenke for proofreading and editing the manuscript. Finally, we would like to thank Dario Salvucci for valuable comments and advice.

Correspondence concerning this article should be addressed to Marco Ragni, Center for Cognitive Science, University of Freiburg, Friedrichstraße 50, 79098 Freiburg, Germany. E-mail: ragni@cognition.uni-freiburg.de

central issue of this article—try to solve the following two problems in your head rather than with a drawing:

1: The Ferrari is parked to the left of the Porsche.

The Beetle is parked to the right of the Porsche.

The Beetle is parked to the left of the Hummer.

The Hummer is parked to the left of the Dodge.

Is the Porsche (necessarily) parked to the left of the Dodge?

2: The Ferrari is parked to the left of the Porsche.

The Beetle is parked to the right of the Porsche.

The Porsche is parked to the left of the Hummer.

The Hummer is parked to the left of the Dodge.

Is the Porsche (necessarily) parked to the left of the Dodge?

Neither problem (with the premises above the line and the to-be-validated conclusion below the line) is easy to solve. However, for most people, the second problem is even harder to solve than the first, although just one word has been changed from the first problem ("Porsche" instead of "Beetle" in the third premise). So, what causes the difference in difficulty? Byrne and Johnson-Laird (1989) initially argued that the main difference between the two problems is that Problem 1 allows only one valid interpretation, while the premises of Problem 2 can be interpreted in more than one way. In other words, Problem 1 agrees with just *one mental model,* whereas Problem 2 concurs with *multiple models*. This is easy to realize when we try to construct an arrangement of the vehicles along a horizontal line. For Problem 1, we obtain the arrangement

Ferrari	Porsche	Beetle	Hummer	Dodge

In contrast, Problem 2 leads to three possible arrangements of the cars:

Ferrari	Porsche	Beetle	Hummer	Dodge
Ferrari	Porsche	Hummer	Beetle	Dodge
Ferrari	Porsche	Hummer	Dodge	Beetle

In both problems, the valid conclusion is, "Yes, the Porsche is parked to the left of the Dodge." But, in Problem 1, this inference concurs with one spatial arrangement, whereas we must consider three arrangements in Problem 2. Note that, by definition, the conclusion is only logically valid if it holds in all possible interpretations of the premises. Thus, if you try to solve Problem 2 solely in your mind, you have to mentally walk through all possible models, one by one, to check whether the conclusion "The Porsche is parked to the left of the Dodge" is indeed true in all models. This is very difficult for most of us, and not surprisingly, several studies have shown that problems of Type 2 (which are also called multiple-model problems) are much harder to solve than problems of Type 1 (which are also called single-model problems; e.g., Boudreau & Pigeau, 2001; Carreiras & Santamaría, 1997; Roberts, 2000; Schaeken et al., 1998; Schaeken & Johnson-Laird, 2000; Schaeken, Johnson-Laird, & d'Ydewalle, 1996).

Spatial relational inferences are so fundamental that one might think that the underlying cognitive processes and the differences in reasoning difficulty are easy to explain. However, this is not the case. On the one hand, most psychologists agree that the *theory of mental models* is currently the best framework for understanding human spatial reasoning (Oberauer, Weidenfeld, & Hörnig, 2006; Schaeken, Van der Henst, & Schroyens, 2007; Vandierendonck, Dierckx, & DeVooght, 2004). The conjecture that people reason by constructing and manipulating mental models of the spatial situation described in premises is supported by a large number of behavioral findings (Goodwin & Johnson-Laird, 2005, 2008; Jahn et al., 2007; Johnson-Laird, 2006; Knauff, 2007, 1999) and neuroimaging results (Fangmeier & Knauff, 2009; Fangmeier, Knauff, Ruff, & Sloutsky, 2006; Goel & Dolan, 2001; Goel, Stollstorff, Nakic, Knutson, & Grafman, 2009; Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003; Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002; Prado, Van der Henst, & Noveck, 2011; Prado, Chadha, & Booth, 2011; Ruff et al., 2003).

On the other hand, the model theory makes quite general assumptions about what makes some reasoning problem difficult to solve. For the model theory, the difficulty of an inference depends on the number of models it calls for (e.g., Johnson-Laird, 2006; Johnson-Laird & Byrne, 1991). The more models a person must bear in mind to solve a reasoning problem, the higher the risk that the person does not consider all models. Thus, for proponents of the model theory, the number of mental models required to solve a reasoning problem provides a measure of the difficulty of reasoning problems. But does this mean that we really know what happens in the human mind when people reason spatially? And is the number of possible models really an adequate measure for reasoning difficulty? What is still missing is a detailed theory of human spatial reasoning that explains how mental models are constructed, inspected, and manipulated in working memory and how the number of necessary models affects reasoning difficulty. This theory must also account for empirical findings indicating that individuals prefer to mentally create some models but have great difficulty considering other models that also satisfy the premises (Jahn et al., 2007; Knauff, Rauh, & Schlieder, 1995; Rauh et al., 2005). For instance, in the indeterminate cars problem above, most people construct just the first model and fail to consider the other two models, although they are also consistent with the premises, as we show later. Can we predict such preferences? Is the number of models an adequate measure for reasoning difficulty if people do not consider all possible models? How are preferred mental models constructed? Do people really consider alternative interpretations of the problem description? What makes the preferred models so special?

Our goal in this article is to present a comprehensive theory of human spatial reasoning with preferred mental models and to show how this theory can be implemented in a psychologically realistic computational model of how humans think spatially. In the first part of the article, we present our *preferred model theory* reflecting our main assumption that people usually construct just a single, simple, and typical model but fail to consider other models in which the premises hold. In the second part of the article, we present the computational model *preferred inferences in reasoning with spatial mental models* (PRISM), which is a specific instantiation of the preferred model theory. Here, we describe the overall architecture of PRISM, the *construction* of a preferred model from premises, the *inspection* of this model that leads to a putative conclusion, and the model *variation* in which the preferred model is varied to obtain alternative models, if necessary. We also present some processing examples in which we describe, step by step, how PRISM works on different sorts of reasoning problems. The third part of the article describes the consequences of the theory and then assesses the theory and the computational model in light of the empirical evidence, including two experiments from our laboratory. We furthermore argue that the number of possible models might not be an adequate measure for the difficulty of reasoning problems. Finally, the article discusses some corollaries and consequences from the theory. We compare our theory to other computer models of human reasoning, show that our theory is not limited to reasoning with spatial relations, and consider some limitations and open questions of our theory. We conclude that our theory has important implications for a general theory of human thought.

Core Ideas of the Preferred Models Theory of Spatial Reasoning

Cognitive psychologists have investigated reasoning about spatial relations for many years but have disagreed about the underlying mental processes (Breslow, 1981; Evans, Newstead, & Byrne, 1993; Roberts 1993). Today, however, the vast majority of researchers consider the model theory to be the empirically best supported theory of human spatial inference (Goodwin & Johnson-Laird, 2005; Knauff, 2009; Vandierendonck, 1996; Vandierendonck et al., 2004; for an exception, see Van der Henst, 2002). According to Johnson-Laird (1983, 2006, 2010) and Johnson-Laird and Byrne (1991) a mental model is an *integrated representation* of the information presented in the reasoning problem. That is, the diverse pieces of information from the premises are not kept as separate entities in the reasoner's mind. Rather, they are merged into a single representation that reflects the information given in the problem description. In other words, a mental model is a mental representation of objects and relations (structure) that constitutes a model (in the usual logical sense) of the premises given in the reasoning task. Or as Johnson-Laird (1998) put it, "the parts of the model correspond to the relevant parts of what it represents, and the structural relations between the parts of the model are analogous to the structural relations in the world" (p. 447). According to the model theory, people translate a perceived or imagined situation into such a mental model and use this representation to solve associated inference problems (Johnson-Laird, 1983, 2001, 2006, 2010; Johnson-Laird & Byrne, 1991).

A crucial assumption of the model theory is that reasoning is a process in which, first, unified mental models of the given premises are generated and then, due to the fact that this information can be ambiguous, alternative models of the premises are sequentially generated and inspected. This process can be broken down into three separate phases, which Johnson-Laird and Byrne (1991) called the comprehension, description, and validation phases. In the preferred model theory, we favor the terms *model construction, model inspection,* and *model variation phases* because these terms better characterize what actually happens in the phases. In the *model construction phase,* reasoners use their general knowledge and knowledge about the semantics of spatial expressions to construct an internal model of the state of affairs that the premises describe. This is the stage of the reasoning process in which the given premises are integrated into a unified mental model. According to the theory, only this mental model needs to be kept in memory, that is, the premises may be forgotten (Mani & Johnson-Laird, 1982). In the *model inspection phase,* a parsimonious description of the mental model is constructed, including a preliminary conclusion. In other words, the mental model is inspected to find relations not explicitly given. This phase was called the description phase by Johnson-Laird and Byrne because they conceived the preliminary conclusion as a kind of description of the model: "This description should assert something new that is not explicitly stated in the premises" (Johnson-Laird & Byrne, 1991, p. 35). In the *model variation phase*—according to the classical mental model theory—people try to find alternative models of the premises in which the conclusion is false. If they cannot find such a model, the conclusion must be true. If they find a contradiction, they return to the first stage and so on until all possible models are generated and tested (Johnson-Laird & Byrne, 1991). According to this view, the variation phase is an iteration of the first two phases in which alternative models are constructed and inspected in turn.

We now present our *preferred model theory* and describe where our account differs from the standard mental model theory. The differences mainly regard the model construction and model variation phases. First, we believe that the model theory needs a more detailed concept of what actually happens in the model construction phase. Notably, many questions are still open for indeterminate problems with an ambiguous set of premises. For instance, how do people keep track of all possible models for indeterminate problems? In their early work, Byrne and Johnson-Laird (1989) speculated that people try to consider all models of the premises of multimodel problems but then fail to maintain these models in working memory. Another option is that people represent just a single model but also symbolically represent (by means of annotations) that certain relations are ambiguous (Vandierendonck, De Vooght, Desimpelaere, & Dierckx, 1999). A third alternative is that people simply pay no attention to the information from irrelevant premises. For instance, in the problem above, people would not represent the first premise because the relation between the Ferrari and the Porsche is irrelevant for the evaluation of the conclusion (Schaeken et al., 1998; Schaeken & Johnson-Laird, 2000). Several ideas have been developed as to how people deal with the ambiguity of spatial descriptions and reduce the cognitive complexity of indeterminate problems. In a review study, Vandierendonck et al. (2004) showed that these approaches range from the construction of several fully elaborated models, through the use of partial models that capture the uncertain element of the spatial arrangement, to the construction of a single model with annotations coding the ambiguity.

Our suggestion in this article is even more radical. The crucial question from our point of view is whether people actually recognize all of the many different interpretations that agree with a set of ambiguous premises. We do not think so. This poses a new question: Do all possible models of indeterminate premises have the same chance of being considered by the reasoner? Again, we do not think so. A crucial assumption within our preferred model theory is that, in most situations, people only construct a single, simplified, and typical mental model and ignore all others. People are almost blind to alternative interpretations of the premises. We might only construct further models if the reasoning problem clearly requires us to consider alternatives. In other words, we do not think that the construction of an initial model is a stochastic process that produces one model this time and another the next time. In the preferred model theory, the construction of the initial model is, in principle, a deterministic process that always produces the same model for the same premises. We assume that this preferred model is the same for most people and that such preferred models bias people in a predictable way (Knauff, 2013; Rauh et al., 2005). The preferred model is favored over others because it is easier to construct in spatial working memory. In the next stage of the reasoning process, this preferred model is inspected to find new spatial relations that are not explicitly specified in the premises. This is done by a spatial focus that can be conceived as attention shift mechanism, to which we return later.

The second difference to the standard model theory is that we suggest a major revision of the model theory's assumptions about the third phase of inference. Johnson-Laird called this phase the model validation phase because in this phase a person generates and tests alternative models of the premises to check whether an alternative model of the premises exists in which the conclusion is false. From a logical point of view, this is essential because formally a conclusion is logically valid if and only if the relation holds in all thinkable models of the premises (Barwise, 1982; Russell & Norvig, 2010). This is also why, according to the standard model theory, the number of models required for an inference affects its difficulty: The risk of an invalid inference increases if the number of models becomes too great. However, we do not believe that people normally reason this way. To the contrary, if an individual is confronted with a determinate problem, we assume that there is no validation of the initial model. According to the preferred model theory, variation is only required for indeterminate problems and even then only if the reasoning problem clearly requires the individual to consider alternatives, for instance, if the person is explicitly asked for all possible models. We further assume that the sequence of alternative models is not random. Rather, people are heavily biased toward models that are similar to the preferred model. The reason is that the generation of alternative models follows the principle of minimal changes, which is a core principle in the belief revision literature (Gärdenfors, 1988, 1990; Harman, 1986; Knauff, Bucher, Krumnack, & Nejasmic, 2013). In the preferred model theory, minimal change means that all possible models are sorted by similarity to the preferred model and that other models can only by obtained by local transformations of the preferred model. We model this procedure by using a neighborhood graph, to which we return later, when we describe how the model variation phase is implemented in PRISM.

Before we present our computational model PRISM, which is an instantiation of the preferred model theory, we summarize the main assumptions of our preferred model theory. These are as follows:

1. When individuals are confronted with indeterminate reasoning problems, they are likely to construct just a single, simple, and typical model, even when a description is compatible with several alternative models. This model that first comes to the reasoner's mind is the preferred mental model.

2. Preferred mental models of spatial descriptions are those constructed according to the principle that new objects are added to a model without disturbing the arrangement of those tokens already represented in the model.

3. Reasoning with indeterminate premises is biased toward preferred mental models. Thus, inferences about relations conforming to a preferred model are easier than inferences about relations that hold only in alternative models.

4. The difficulty of an inference does not depend on the number of logically possible models but on the difficulty of mentally constructing preferred and alternative mental models of the circumstances the premises describe.

5. People search for alternative interpretations of the premises only if this is explicitly required. If a search for alternative models is required, it always starts with the preferred model. Alternative models are constructed by local transformations, and the process follows the principle of minimal changes.

6. Alternative models that require a longer sequence of local transformations are more likely to be neglected than models that are only minor variations of the preferred model. Therefore, the danger of missing a particular alternative model increases with its distance from the preferred model.

7. Logical errors and illusory inferences result from omitting models in which the conclusion from the preferred model does not hold.

We now present our computational model PRISM that realizes these assumptions. PRISM is a successor to our previous *Spatial Reasoning by Models* (SRM) model (Ragni & Knauff, 2008; Ragni, Knauff, & Nebel, 2005) and continues the emphasis of preferred and neglected mental models for explaining human spatial reasoning, including invalid inferences and reasoning difficulties. PRISM is completely implemented in a program that is written in Python and can be downloaded from http://imodspace.iig.uni-freiburg.de/prism

The Architecture of PRISM

PRISM is a symbolic cognitive architecture in which tokens and the spatial relations among them are represented in a spatial array that can be inspected and manipulated by a spatial focus. PRISM reasons with binary spatial expressions such as "left of," "right of," "in front of," "behind," and so on and generates preferred models, alternative models (if necessary), and logically valid and invalid conclusions. The problems can take one of two forms: In a *generation problem,* the input is a set of spatial premises, and PRISM must generate a conclusion as output. In a *verification problem* the premises are given, and PRISM receives a query about the logical validity of a to-be-verified conclusion. In such problems, the output is a judgment of "valid" or "invalid." To realize such relational inferences, the architecture of the PRISM consists of *five components*.

The first component is the *input mechanism* that encodes the relations in the premises as a triplet (X, r, Y) in which

X is the referent,

r is a binary spatial relation, and

Y is the relatum.

The referent X is the *located object* (LO), and the relatum Y is the *reference object* (RO). The distinction is standard in psycholinguistics and spatial language research, where researchers agree on the assumption that a spatial locational description refers to the position of an object relative to another object or area (Hayward & Tarr, 1995; Herskovits, 1986; Jackendoff & Landau, 1991; Miller & Johnson-Laird, 1976; Talmy, 1983; Tenbrink, Andonova, & Coventry, 2011). Following this distinction, in PRISM, the RO of a premise is usually inserted into the model first, followed by the LO. One exception is the first premise. Here, we assume that individuals prefer to change the roles of RO and LO in favor of an incremental model construction, as shown by Oberauer and Wilhelm (2000). PRISM does not account for the problems related to the ambiguity of spatial relations (Gapp, 1997; Hayward & Tarr, 1995; Knauff, 1999; Vorwerg & Rickheit, 1998). We simply assume that "left" means that the LO is to the left of the RO and along exactly the same line in a spatial array (see below). It can be adjacent to the RO, or there can be other cells (empty or filled) in between. The relation "in front of" means that the LO is in a cell in front of the RO and along exactly the same line. It can be adjacent to the RO, or there can be other cells in between. "Right" and "behind" are defined accordingly. We have not implemented a language-understanding device because our goal with PRISM is to simulate the actual reasoning process rather than language-specific processes involved in reading the premises, which in principle can be presented as sentences on the screen, or verbally, or in any other format, for instance, as pictures on the screen or as arrangement of "real" physical objects, and so on. In fact, several different devices would be necessary to model this input, and these devices could certainly have effects on the inference. However, in developing the PRISM model, we are interested not in this type of effect but in phenomena that we can clearly trace back to the actual reasoning processes. In adopting this approach, we follow most reasoning researchers' opinion that a theory of reasoning starts after the input device and before the output device comes into play (Braine & O'Brien, 1998; Evans et al., 1993; Goodwin & Johnson-Laird, 2005; Johnson-Laird, 1983; Rips, 1994).

The second component of PRISM concerns the *set of tokens.* The tokens can represent any kind of physical entities—people, natural objects, human-made objects, geographic regions, events, and so on. They are usually the arguments (terms) of the reasoning problems. However, PRISM does not make any specific assumptions about how these entities are represented. The reason is that a variety of evidence suggests that the human brain/mind processes object properties and location information separately. The object properties system presumably employs position-invariant information, whereas the second system is responsible for representing and processing spatial information. Empirical evidence supporting this view comes from almost all areas of the cognitive sciences, ranging from low-level perception, through working memory and long-term memory, up to the task of expressing spatial experience through language (Baddeley, 1986, 1990; Landau & Jackendoff, 1993; Logie, 1995; Ungerleider & Mishkin, 1982; for a summary, see Knauff, 2013). Computational analysis has revealed that splitting processing into separate systems for identifying and locating objects leads to better performance than a single system processing both object properties and spatial information (Rueckl et al., 1989). Research from our group has revealed that the distinction between object properties and spatial processing also resolves many inconsistencies in the previous reasoning literature (Knauff, 2009; Knauff et al., 2003). In particular, we were able to show that considering irrelevant objects' properties can even impede individuals' performance in spatial reasoning (Knauff & Johnson-Laird, 2002; Knauff & May, 2006). For PRISM, this distinction means that the tokens that must be placed in spatial working memory are usually abstract symbolic representations linked to knowledge in long-term memory that provides in-depth information about these entities (Hollingworth, 2004; Tversky & Hemenway, 1983). One advantage of not representing detailed object information is that models in PRISM are not confused with visual images. We have argued for a careful distinction between spatial models and pictorial mental images in many other publications (Knauff, 2009, 2013). Another benefit of not representing objects' properties is that PRISM can solve problems on all spatial scales, ranging from small-scale space that we can reach with our hands; through vista-space, the space we can apprehend from one place without necessary locomotion (Montello, 1993); to large-scale space, for instance, geographic regions or areas. All these entities are treated equivalently in PRISM.

The third component of PRISM is the *spatial array* that functions like a spatial working memory. The background is that most memory researchers accept the assumption of a specialized working memory subcomponent involved in the representation and processing of spatial information from different input modalities such as vision, touch, hearing, or language (Logie, 1995). According to this view, the spatial array in PRISM is a supramodal structure, rather than a modality-specific system such as the visual buffer in Kosslyn's theory of visual mental imagery (Kosslyn, 1994; Kosslyn, Ganis, & Thompson, 2006). The spatial array is the essential construct for the representation the mental model, and in principle, the properties of models originate from the properties of the array in which they occur. Formally, the spatial arrays can be considered indexes that spatially connect objects to create a scene or, in the present context, to allow inferences (Papadias & Sellis, 1992, 1994). In PRISM, it is realized as a two-dimensional grid structure, in which the tokens from the premises can be inserted, moved, and removed by the fourth component, the focus operator.

The *spatial focus,* the fourth component of PRISM, operates on the spatial array. The spatial focus can be thought of as an attention shift mechanism that can place tokens into a model or inspect the model to discover new information. Thus, the focus constitutes the central operating device of PRISM and is therefore the second component from which a mental model derives its properties. We assume that many experimental findings can be explained by means of the working principles of this focus. In particular, an important question that arises here is how alternative models can be constructed in the variation phase. We return to this point below.

The fifth component of PRISM is the *control process,* which is responsible for controlling the focus and all other components. It is responsible for two functions; one of these can be defined exactly, while the second—less important—still lacks some empirical evidence. The well-defined function of the control process is how it controls token insertion, model inspection, and model variation. This component, together with the focus and the spatial array, mirrors the core assumptions of the preferred model theory and its implementation in PRISM. The control process is involved in all operations that lead to the preferred model and the order in which alternative models are constructed. One of the main questions in this context is where a token is inserted in the array if another token already occupies this position. The second, currently less well-defined function of the control process is to add symbolic annotations to a model if it detects indeterminacy in the premises. Vandierendonck et al. (2004) proposed such annotations to deal with indeterminate problems where part of the premise information must be maintained for the construction of alternative models (Mani & Johnson-Laird, 1982). Currently, we merely presume that the control process is able to add such annotations to a model, without specific assumptions about how this is actually realized in the human cognitive system.

The Operations of PRISM

In our description of the preferred model theory, we introduced the three phases of a reasoning process, which we referred to as model construction, model inspection, and model variation. We now describe how these phases are implemented in PRISM. We start with a description of how PRISM works with determinate problems and then describe the more complex case of indeterminate problems.

Model Construction

In PRISM, the construction of a mental model starts with the first premise and an empty spatial array. As previously described, we assume that the input mechanisms process the presented premises and generate an output that defines the RO, the LO, and the spatial relation between these tokens. So imagine, for instance, that the input mechanisms received the following two-dimensional determinate problem and translated it into a readable format:

3: The Porsche is to the right of the Ferrari.

The Beetle is to the left of the Ferrari.

The Dodge is in front of the Beetle.

The Volvo is in front of the Ferrari.

Which relation holds between the Dodge and the Volvo?

At the beginning of the reasoning process, PRISM's focus is at the position (0, 0), and there are four possible directions in which the focus can be moved: right, left, forward, backward. Additionally, a no-move operation is possible. The movement and the operations of the focus depend on the types of the different premises. PRISM distinguishes four types of premises:

Type 1—Initial premise: This is the first premise of the reasoning problem. This premise is the starting point of the model construction process.

Type 2—One-new-token premise: This sort of premises consists of two tokens, of which one has already appeared in a preceding premise and one is a new token from the present premise. The token that appeared in a previous premise has already been inserted into the spatial array; only the new token must be placed into the array (e.g., the Beetle in the second premise in Problem 3). To account for the difference between determinate and indeterminate problems, PRISM makes a case distinction between Type 2d and Type 2i premises. In Type 2d premises (which stands for determinate), the new token can be inserted at exactly one position, whereas, in Type 2i (for indeterminate), there is more than one possible position that agrees with the semantics of the spatial relation in the premise. Preferred models come into play when PRISM processes Type 2i premises, and we return to this point later.

Type 3—Two-new-tokens premise: In such premises, two new tokens appear, that is, none of the tokens in the present premise were mentioned in a previous premise. An example is the second premise of a discontinuous premise order (e.g., C r_3 D; A r_1 B; B r_2 C, with r_n for the spatial relations), in which the first and second premises have no objects in common and thus cannot be immediately integrated into one model.

Type 4—Connecting-submodels premise: These premises are those in which a token appears that connects two partial models. This is the case when the third premise of a discontinuous premise order must be processed (such as the B r_2 C in premises of the form C r_3 D; A r_1 B; B r_2 C).

PRISM now works on the premises in four steps:

1. Initially PRISM receives a premise of Type 1.

2. PRISM inserts the first token of the first premise in cell (0, 0). Then, it uses this token as RO and adds the second token to the next adjacent cell according to the spatial relation.

3. The "parser" reads the next premise.

4. PRISM decides on the type of premise:

—If the premise is of Type 2, the focus moves to the RO, and from there, it inserts the LO into the next cell according to the relation. If a token is already present in the desired cell—as is the case in indeterminate Type 2i premises—the simulation moves back to the RO and makes an annotation (which we describe later). Then, it moves to the next free cell according to the relation and inserts the token into the next free position (according to the relation to the RO).

—If the premise is of Type 3, a new spatial array is generated, and both tokens are inserted in the manner of premises of Type 1 (see Step 2).

—If the premise is of Type 4, the focus groups one model and inserts it into the other model (Bara, Bucciarelli, & Lombardo, 2001).

Figure 1 provides a rough illustration of these processes, and the algorithm is presented as a flow chart in Figure 2. For Problem 3, for instance, the resulting model in spatial working memory has the following form:

Beetle	Ferrari	Porsche
Dodge	<u>Volvo</u>	

In the illustration of the model the underlined token is the final position of the focus after premise processing is finished. The focus now remains at this position, which is also the starting point of the model inspection process. Now, let us consider the following indeterminate reasoning problem:

4: The Porsche is to the right of the Ferrari.

The Beetle is to the left of the Porsche.

The Dodge is in front of the Beetle.

The Volvo is in front of the Porsche.

When PRISM reads the first premise of Problem 4, you place the Porsche and the Ferrari in the array. But then you receive the next

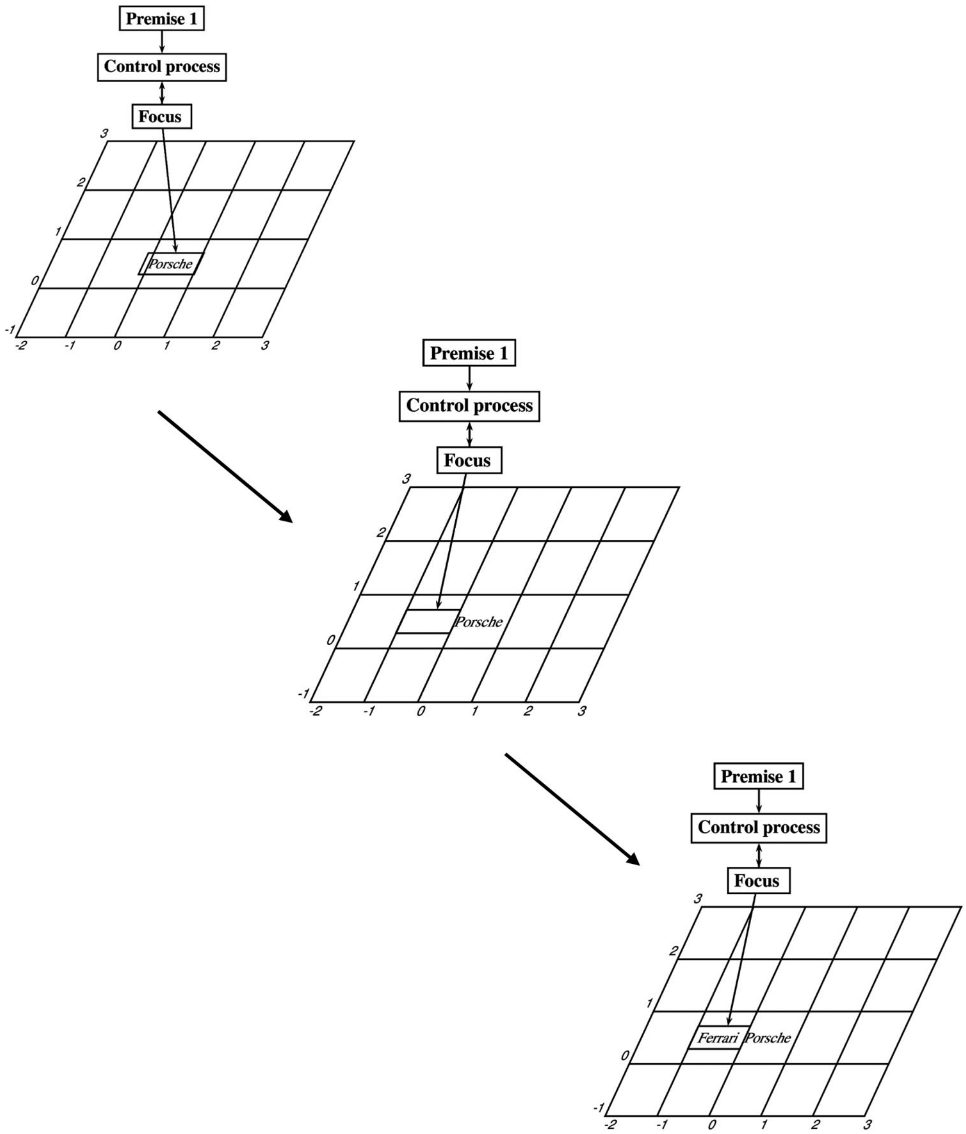

Figure 1. PRISM processing the premise “The Porsche is to the right of the Ferrari.” The control process of PRISM inserts the tokens successively by focus operations. The control process for the model construction phase is illustrated in Figure 2. PRISM = preferred inferences in reasoning with spatial mental models.

premise, which is a Type 2i premise, and realize that you have a problem: The cell directly to the left of the Porsche, where you would normally place the Beetle, is already occupied by the Ferrari. You now have two options, which we term the *fff-strategy* (first free fit) and the *ff-strategy* (first fit) of model construction: In the fff-strategy, the system detects that the first possible cell in the array is already occupied by a token and therefore moves farther to the left and inserts the Beetle in the next free cell that agrees with the meaning of “left of.” This is called the *fff-strategy* because the focus inserts a token at the first free position that fits with the premise. The model resulting from the application of the fff-strategy is as follows:

Beetle	Ferrari	Porsche
Dodge		Volvo

Another option would be to apply the ff-strategy. Using this strategy, the system would squeeze the Beetle in between the Ferrari and the Porsche, which would basically mean that you have to temporarily keep the Beetle in memory, pick up the Ferrari again, and move it one cell to the right. The focus then moves back and inserts the Beetle (whose relation to the Ferrari is buffered in an annotation) into the next cell directly to the left of the Porsche. With the ff-strategy, the focus is forced to insert the token at the first position—the first cell that fulfills the premise—even though doing so means that other objects must be relocated. The model resulting from the application of the ff-strategy is as follows:

Ferrari	Beetle	Porsche
	Dodge	Volvo

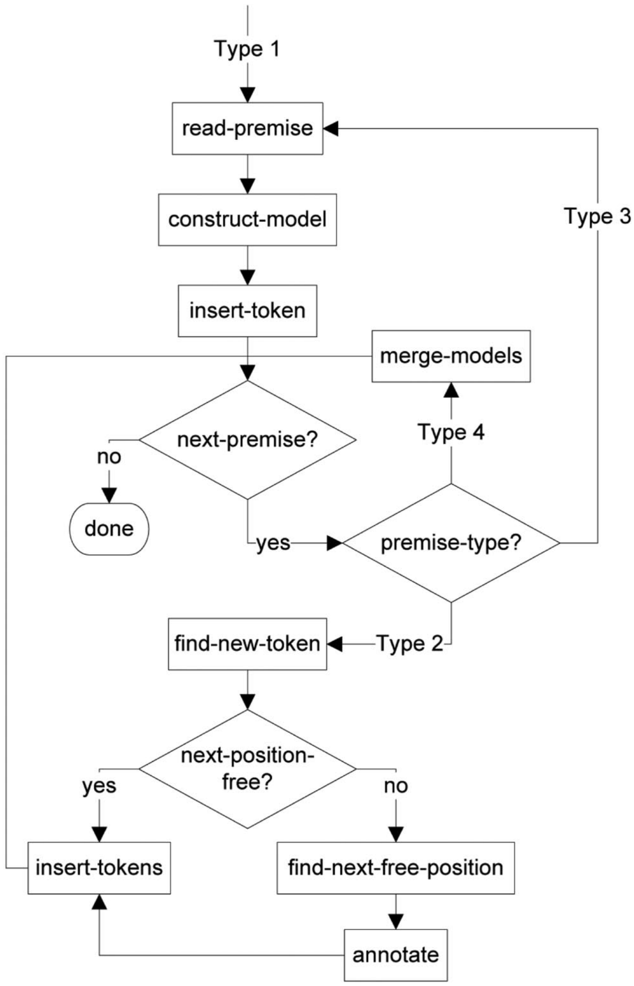

Figure 2. The model construction phase in PRISM. PRISM inserts tokens into mental models according to the different types of premises: A Type 1 premise is the initial premise to be processed. This leads to a construction of a model and the insertion of the tokens. A Type 2 premise introduces one new token, which is related to a token already in the model. Here, indeterminacy can occur. A Type 3 premise introduces two new tokens not in the model. This leads to the same construction as a Type 1 premise. Finally, a Type 4 premise connects two separate models. PRISM = preferred inferences in reasoning with spatial mental models.

In the experiments, which we describe later, we empirically tested the two token insertion strategies (ff- and fff-strategies) against each other and found strong support for the fff-strategy. This result is the main reason why the first model is the preferred model, whereas the second model is only very rarely constructed by human reasoners. The power of PRISM is that it explains the difference between the different constructed models and the insertion strategy. The next step in the inference process is the model inspection phase.

Model Inspection

The model inspection process always inspects a single model in the array, and we do not, therefore, have to distinguish between determinate and indeterminate problems. However, we must distinguish between generation problems (PRISM has to generate a conclusion) and *verification* problems (PRISM must decide whether a to-be-verified conclusion is logically valid or invalid). Take, for instance, the determinate Problem 3 above. After model construction, PRISM's focus is still at the last position of the model construction (Volvo), which is now the starting point for the focus that inspects the model to find the relation that holds between the two tokens mentioned in the question. Therefore, the focus starts at the Volvo (RO) and then inspects the model to find the Dodge (LO). If the LO is found in the scan direction, the relation between the two tokens is known (the meaning is provided by an external mental lexicon); otherwise, the system has to change direction first and then find the second token. Another problem, however, arises if PRISM must verify a presented conclusion, that is, if the question in the previous example is replaced by a conclusion that PRISM must check for its validity. In Problem 4, this would mean that, for instance, PRISM has to check whether the conclusion "The Volvo is to the right of the Dodge" is valid. As the Dodge is found with a leftward scan starting from the Volvo, PRISM generates "valid conclusion" as output.

Model Variation

The model variation phase is a crucial part of our theory. Johnson-Laird and Byrne (1991) stated that "only in the third stage is an essential deductive work carried out; the first two stages are merely normal processes of comprehension and description" (p. 36). We have shown that this is not entirely true because the creation of preferred models in the construction phase has an enormous effect on the further process of inference. On the other hand, Johnson-Laird and Byrne were correct in that only in this phase is actual logical work carried out because in this phase an individual (and PRISM) has to make a decision about whether or not a conclusion follows from the premises. Indeed, this process is crucial in the entire stream of thought and lies at the heart of human reasoning. It is also the essential phase in which alternative models come into play, if this is required by the task at hand. However, our first assumption about the model variation phase is that it only rarely takes place. This is a major departure from the standard model theory, which assumes that model validation always happens because people search for counterexamples to verify a putative conclusion. In PRISM, in contrast, no search for counterexamples exists, and therefore, the difficulty of an inference does not depend on the number of alternative models. We postulate a "blindness for multiple models," which means that people are almost blind to the existence of alternative models and basically treat multiple-model problems as though there were a single possible model. Our second conjecture is that model variation—not validation—is required only for indeterminate problems and only if the task at hand clearly requires the consideration of alternative models. This is the case if, for instance, a reasoner is asked to verify a relation that does not hold in the preferred model or when the task explicitly requires generating all possible models. In both cases, the search for alternative models in PRISM is *not* an iteration of the first two inference phases in which alternative models are generated and inspected in turn (Johnson-Laird & Byrne, 1991). Instead, in PRISM, model variation is the third phase of an inference with its own characteristics and processing principles (see Figure 3).

The variation process starts from the preferred model and then successively generates alternative models by applying minimal changes to the preferred model. This leads step by step to alternative models. It is important to see that this procedure has the consequence that alternative models that require more alterations to the preferred model are more difficult to create than models that involve only minor revisions of the preferred model.

Now, we introduce another important concept of our theory: *the neighborhood graph.* How do we obtain alternative models? In graph theory, a neighborhood graph is a directed graph consisting of vertices and edges (Diestel, 2012). In a graph, k-nearest neighbors are two vertices that are connected by exactly k edges. A special case of a k-nearest neighbor in the graph is the direct neighbor ($k = 1$), where two vertices are directly connected by a single edge. Freksa (1991) introduced the neighborhood graph as formalism to model spatial (and temporal) inferences in artificial intelligence, notably in the area of qualitative reasoning, which is a field that seeks to develop computer programs that reason with imprecise, incomplete commonsense knowledge (e.g., Cohn & Hazarika, 2001; Cohn & Renz, 2007). In Freksa's *conceptual neighborhood graph,* the vertices represent spatial relations, while the edges connect the relations with the fewest differences. Since some relations are closer to each other in the graph (connected by fewer edges than other relations), similarity between relations can be determined (Freksa, 1991, 1992). In the following, we suggest using a neighborhood graph to determine the similarity between different models of a set of premises and, thus, the sequence in which models and alternative models are constructed. The idea is that the vertices in the neighborhood graph represent the models, and the edges connect the model with the fewest differences. Since some models are connected by fewer edges with other models, the similarity between models can be determined by the shortest path in this neighborhood graph. If a one-step transformation from one model to another model exists, then two models are called 1-nearest neighbors. In general, if two models can be connected by a minimal path of length $k,$ then we call these two models k-neighbors. To demonstrate the main idea of this account, we again use a one-dimensional problem, but PRISM works with two-dimensional problems in the same way. Consider the following problem:

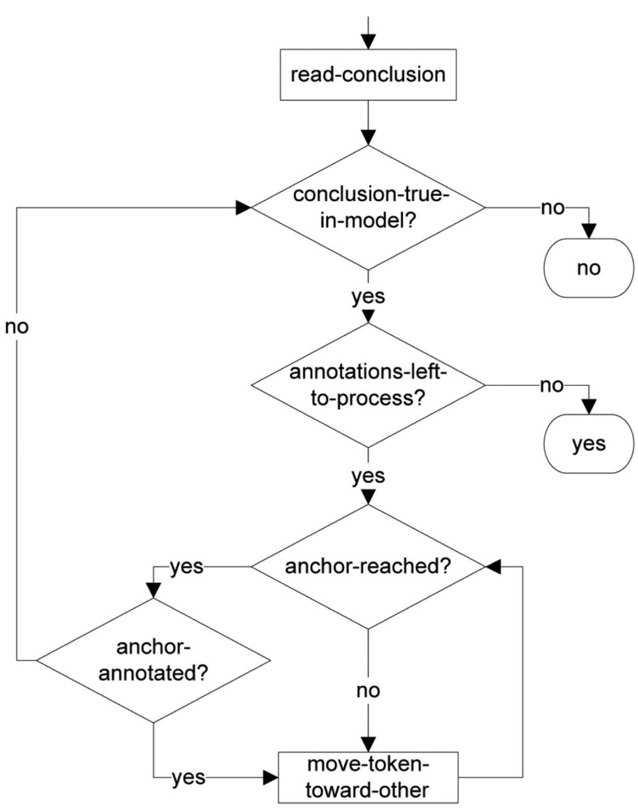

Figure 3. The model variation phase of PRISM. PRISM performs this process only if required by the task. If so, PRISM checks if the conclusion already holds in the model at hand. If not, then it checks if there are tokens, which were annotated in the model construction phase. The annotations comprise information about so-called anchor tokens. These anchor tokens from the annotation are constraints for the token to be varied. PRISM = preferred inferences in reasoning with spatial mental models.

5: The Ferrari is parked to the left of the Porsche.

The Beetle is parked to the right of the Porsche.

The Porsche is parked to the left of the Hummer.

The Hummer is parked to the left of the Dodge.

Which relation holds between the Porsche and the Dodge?

The situation is slightly different from Problem 2 because we now have to generate a relation that holds between the Porsche and the Dodge rather than verify a presented relation (we return to the differences between verification and generation problems later). You can easily see that Problem 5 (again) leads to the following three possible models:

Ferrari	Porsche	Beetle	Hummer	Dodge
Ferrari	Porsche	Hummer	Beetle	Dodge
Ferrari	Porsche	Hummer	Dodge	Beetle

Does each of these models have the same chance of being constructed? We do not think so. Instead, we predict that the first model will be the preferred model, which we explain later in terms of our difficulty measure. The second model is harder to construct than the preferred model, and the third in turn is harder to construct than the second one. The reason for this hierarchy is that the first nonpreferred model requires only one *swap* operation from the preferred model (changing the positions of Beetle and Hummer), whereas the second nonpreferred model necessitates two such swap operations from the preferred model. From this follows the neighborhood graph

$$\mathbf{FPBHD} \rightarrow \mathbf{FPHBD} \rightarrow \mathbf{FPHDB}.$$

By applying formal methods from mathematics, in Ragni and Wölfl (2005), we were able to show that there is definitely no continuous transformation from the preferred model to the third model without generating the second model. Moreover, the neighborhood graph also follows directly from the principle of minimal changes, which is an important concept in the belief revision literature and says that a belief set (or set of models) before and after the belief revision should be as similar as possible (Gärdenfors, 1984; Harman, 1986). In the case of model revision, this principle forces as much information as possible to be preserved by the change and implies a minimal transformation distance between models, that is, the shortest path in the neighborhood graph between two models. In our theory and PRISM, this transformation distance is essential to explaining human reasoning difficulty: If an alternative model has a high transformation distance from the preferred model, it is difficult to construct and is therefore more likely to be neglected.

Processing Examples

Several examples of how PRISM works with the determinate and indeterminate generation and verification problems now follow. The processing of the determinate generation Problem 3 is simple: PRISM receives the information from the first premise, "The Porsche is to the right of the Ferrari," and the focus takes the Porsche as the RO (because it is mentioned first) and inserts it into the spatial array. Next, the focus moves to the left and inserts the Ferrari into the next cell according to the relation. Then, PRISM processes the second premise, which is "The Beetle is to the left of the Ferrari." Because the focus is still on the Ferrari, the model inserts the Beetle into the next free cell to the left of the Ferrari. The third premise is "The Dodge is in front of the Beetle"; now, the focus changes direction and moves to the front cell and inserts the Dodge. Then, PRISM reads the fourth premise, "The Volvo is in front of the Ferrari." The focus moves back to the Ferrari and after that moves one step in front of the Ferrari and inserts the Volvo. Note that this is a determinate problem and thus there is no model variation phase. The result is the following model:

Beetle	Ferrari	Porsche
Dodge	<u>Volvo</u>	

The next step is the inspection phase. PRISM must now find out which relation holds between the Dodge and the Volvo. Note that the focus is still at the marked cell; PRISM checks if this cell contains a token mentioned in the conclusion, which is the case in the example (Volvo). Now the focus moves from the Volvo to the Dodge and checks in which direction it has moved. Thus, PRISM generates the valid conclusion "The Dodge is to the left of the Volvo." The process is more complex for indeterminate premises, such as the following:

6: The Porsche is to the right of the Ferrari.

The Beetle is to the left of the Porsche.

The Dodge is in front of the Beetle.

The Volvo is in front of the Porsche.

PRISM now works in the following manner: It reads the first premise and starts constructing the model in the same way as described for the determinate problem. When PRISM then processes the second premise, which is a Type 2i premise, it moves to the RO (Porsche) and from there to the left. Then, the focus detects that the Ferrari is already in the cell. Thus, the focus makes an annotation at the Porsche (saying that the Beetle is to the left of the Porsche) and then—by applying the fff-strategy—moves further to the next cell to the left of the Ferrari, where it inserts the Beetle. The resulting model is the preferred model:

Beetle	Ferrari	Porsche
Dodge		Volvo

In a generation problem, PRISM has to answer a query like "Which relation holds between the Dodge and the Volvo?" As in the previous example, the focus moves from the Volvo (which was inserted last) to the Dodge and, based on the direction it moved, decides that the valid conclusion is "The Dodge is to the left of the Volvo," which is indeed logically valid.

Now, imagine that PRISM receives the query "Which relation holds between the Beetle and the Ferrari?" The first premise of Example 6 is processed as previously described. PRISM moves the focus to the RO, the Ferrari, and because it began at the Volvo (the last focus position) and only passes the Porsche, it continues with a leftward scan—because the focus follows the principle of minimal direction changes—and finds the Beetle in the next cell in the left-hand side cell of the array. From this, PRISM concludes that "The Beetle is to the left of the Ferrari." However, this is a logically invalid inference because there is an alternative model of the premises in which the Beetle is to the right of the Ferrari. This model is the following:

Ferrari	Beetle	Porsche
	Dodge	Volvo

Because PRISM does not normally search for alternative models, it overlooks this model, although it would have falsified the inference "The Beetle is to the left of the Ferrari." However, because PRISM does not recognize this model, it produces a logically invalid illusory inference.

The only types of problems in which PRISM generates alternative models are (a) generation problems that explicitly ask for all possible models and (b) verification problems in which the program must decide on a relation that does not hold in the preferred model. In the first case, PRISM starts with the preferred model and then, in principle, constructs all possible models by minimal changes. The sequence of constructed models follows the neighborhood graph, and the risk of neglecting a model increases as a function of its distance from the preferred model. The longer the path through the neighborhood graph, the higher the risk that the model is neglected. We do not describe the entire process here but return to the implications in the next section. The second case in which PRISM generates alternative models (in which the program must decide on a relation that does not hold in the preferred model) is given, for example, if PRISM has to verify whether "The Beetle is to the left of the Ferrari" holds in the previous example. Now, PRISM generates the preferred model and puts an annotation at the Beetle saying that this token could also be placed somewhere else to the left of the Porsche. Then the following happens: The normal inspection process notices that the relation at hand does not hold in this preferred model but also that there is an annotation at the Beetle. If this annotation is forgotten, not readable, or not available for any other reason, PRISM generates an incorrect response because it neglects the alternative model (we do not yet have clear assumptions on this point). If, however, the annotation is processed by the focus, PRISM is signalized that an alternative model might exist and thus starts to search for other models in which the relation at hand holds.

PRISM does this by applying minimal changes to the preferred model. Thus, the focus moves to the Beetle with the annotation, picks up this token, and moves it to the other possible position in the array. Now, the inspection process continues as usual. The result of this inspection is that PRISM recognizes that "The Beetle is to the right of the Ferrari" holds in the alternative model but not in the preferred model. That means that the relation is possible but not logically necessary. This example is simple because there is just one alternative to the preferred model. In problems with more than one alternative model, the sequence of constructed models can again be determined by the neighborhood graph, as we show later in our psychological experiments.

Consequences of the Theory

The previous parts of this article have presented the preferred model theory, how this theory is implemented in our computer simulation program PRISM, and how PRISM deals with different kinds of problems. An important consequence of the preferred model theory is that we carefully discriminate between different sorts of reasoning problems. We distinguish between determinate (single-model) problems and indeterminate (multiple-model) problems and between verification and generation problems. In a verification problem, a set of premises and a conclusion are presented, and individuals have to decide whether or not the conclusion follows from the premises. In generation problems, the premises are presented, and individuals are asked to generate valid conclusions. For indeterminate problems, we further distinguish between one-model-generation problems and multiple-model-generation problems. In one-model-generation problems, a person is instructed to generate a single model from the set of all possible models. In multiple-model-generation problems, we ask a person for all possible models that agree with the premise. For all these different sorts of problems, our theory and PRISM make clear predictions as to the responses of individuals, and we also show whether these responses are logically correct or not. In Table 1, the consequences of our theory for all types of problems are presented. In the first column on the left, we verbally describe each problem-type. The next three columns stand for the three processes of model construction, model inspection, and model variation. A "+" indicates that we assume that this process is mentally carried out by human reasoners when they are confronted with the respective sort of problem. A "−" indicates that we believe that this process is usually *not* carried out by logically untrained people. The fifth column shows how we believe a reasoner responds when he or she is asked to verify or generate a conclusion (we come back to this in a moment). The sixth column shows whether the inference is valid from a logical point of view, and the last column assesses the individual's response as logically correct or incorrect.

Table 1
Different Forms of Relational Reasoning Problems, the Reasoning Phases, the Predicted Responses, Logical Validity, and the Correctness of the Individual's Response

Type	Description	Model construction	Model inspection	Model variation	Predicted response	Logical validity of the inference	Correctness of the individual's response
			Determinate problems (single-model problems)				
I	Verification of a relation that holds in the model, as valid conclusion	+	+	−	Yes	Valid	Correct
II	Verification of a relation that does not hold in the model, as valid conclusion	+	+	?	No	Invalid	Correct
III	Generation of a relation that is a valid conclusion	+	+	−	Conclusion generated	Valid	Correct
			Indeterminate problems (multiple-model problems)				
IV	Verification of a relation that holds in all possible models, as valid conclusion	+	+	−	Yes (because it holds in the PMM)	Valid	Correct
V	Verification of a relation that holds in the PMM, as valid conclusion	+	+	−	Yes (because it holds in the PMM)	Invalid	Incorrect
VI	Verification of a relations that holds in an AMM, but not in the PMM, as valid conclusion	+	+	+ (maybe −) (depends on the distance to the PMM)	Yes or maybe no (depends on similarity to PMM)	Valid	Incorrect
VII	Generation of all (or some) valid relations that hold in the PMM and AMM	+	+	+ (in the order from the neighborhood graph)	Valid relations generated	Valid	Correct
VIII	Generation of a specific common relation in PMM and AMM	+	+	+	"No valid conclusion possible"	Valid	Correct

Note. A "+" indicates that we assume this process is mentally carried out by human reasoners when they are confronted with the respective sort of problem. A "−" indicates that we believe that this process is usually not carried out by logically untrained people. A "?" indicates a problem where we are not entirely sure. PMM = preferred mental models; AMM = alternative mental model (nonpreferred mental model).

There are several important things to note in the table. First, it is clear that all types of problems need a model construction process and that this model must be inspected to solve the task at hand. However, the column Model variation has only a few "+"s, which reflects our opinion that model variation is only rarely carried out by human reasoners. Usually, people are satisfied with a single model, which is the only model in determinate problems and the preferred model in indeterminate problems. In principle, we assume that people never try to generate alternative models in determinate problems. This is a major difference to the standard model theory, in which people try to validate their putative conclusion by ruling out that other models exist in which this conclusion does not hold. In problems where a valid conclusion for a determinate problem is presented, we do not expect that people search for counterexamples (Problem I in Table 1). A problem where we are not entirely sure yet is marked with the "?" (Problem II). What happens if people have to evaluate a relation that does not hold in the model? Do they try to see whether a model exists in which this relation holds? Logically, of course, this is senseless. However, it is possible that some people think that a relation is valid if they can find just one model in which the relation holds. Thus, some people might vary the model they have in mind, but others will not. Moreover, the variance might result not only from individual strategies but also from the exact wording of the problem. Unfortunately, reasoning researchers use many different instructions and often express tasks in different ways. They use terms like "logically follows," "necessarily follows," "thus, it follows," and "therefore . . .." We do not go into detail here, but we are convinced that such differences have an immense effect on how people reason. The differences in wording might be responsible for many inconsistencies in the literature (Van der Henst & Schaeken, 2005) and might also affect whether people think that the experimenter asks for a possible or a necessary model, conclusion, or relation. However, in the current version, PRISM performs logically correctly in this regard, and thus, no model variation takes place for these sorts of problems. In the third type of determinate problem, participants have to generate the only possible conclusion (Problem III). Often such problems are relatively easy to solve, but discontinuous premise orders (see below), too many premises, complex relations (see below), or other factors can make even these problems difficult to solve.

The four sorts of indeterminate problems are even more interesting. In some problems, people are instructed to say whether or not a relation is a logically valid conclusion (Problem IV). The problem is indeterminate, and thus, following the norms of formal logic, it is necessary to search all possible interpretations of the premises to find a conclusion that holds in all models. According to our theory, however, people do not reason this way but nevertheless give a correct answer because they just consider the preferred model, but the relation in this preferred model holds in all models. They do not think like a logician, but fortunately, the response is correct. In the second sort of indeterminate problem (Problem V), the situation is different, and we consider this one of the theoretically most interesting cases. Here, a preferred model leads people to a logically incorrect decision: The presented relation holds in the preferred model but not in the alternative models. However, the preferred model is so powerful that people are not aware that other models might exist. The bias toward the preferred model produces an illusory inference because the person treats an inference from the preferred model as sufficient for a logically valid inference. The third sort of indeterminate problem (Problem VI) is also interesting because here the neighborhood graph and the principle of minimal change come into play. In fact, this is one of the rare cases where we think that some individuals enter into a model variation phase to find out whether a model exists in which the presented relation holds. Again, this is logical nonsense, but we cannot rule out that people try to solve the problem this way. If so, we make another prediction. The risk of a logically incorrect response is a function of the position of the model in the neighborhood graph. The further the alternative model is from the preferred model, the lower the probability that the person takes the model into account. Since considering a relation from the alternative model as valid is a logical mistake, fewer errors are made for more distant models (in the neighborhood graph) than for models that are only minor revisions of the preferred model (nearer to the preferred model in the neighborhood graph). The next sort of indeterminate problem (Problem VII) is very straightforward. Our assumption is that people conduct a model variation process starting with the preferred model. In the next step, they generate the next neighbor of the preferred model by applying minimal changes. The next model is a minor alteration of the first alternative model and so on. Alternative models that require a longer sequence of local transformations are more likely to be neglected than models that are only minor variations of the preferred model. Thus, the hazard of missing a particular alternative model increases with its distance from the preferred model. The last sort of problem is special (Problem VIII); we just added this sort of problem to the table because such problems were used in the influential experiments on indeterminacy by Byrne and Johnson-Laird (1989). In all other problems (Problems I–VII), the participants of experiments had to generate a conclusion or they were instructed to decide whether a given relation is a valid conclusion or not (Does it follow that . . .; yes or no?). In problems of Type VIII, in contrast, Byrne and Johnson-Laird asked people to generate a conclusion (a specific relation between two tokens) that is common in all models; they also offered "nothing follows" as a third alternative to answer (the only logically correct answer). In the next section, we refer to these sorts of problems presented in Table 1 and assess how well our theory and PRISM can deal with these problems.

Indeterminacy and the Difficulty of Reasoning Problems

The previous sections have demonstrated the great differences between diverse sorts of reasoning problems and have showed that PRISM is able to solve all of these problems. We now return to the difference between determinate and indeterminate problems and show how our theory diverges from the standard model theory. The main point upon which we disagree is the connection between indeterminacy and the difficulty of an inference. In the standard model theory, the difficulty of a reasoning problem depends on the number of models it calls for (Johnson-Laird, 1983, 2001, 2006; Johnson-Laird & Byrne, 1991). The more models an individual must consider, the higher the load on working memory. Thus, for proponents of the model theory, the number of mental models required to solve a reasoning problem provides a measure of the difficulty of reasoning problems. However, in the following, we challenge this hypothesis and propose an alternative account in its stead. Our main idea is to use the *number of necessary focus operations* as a measure for the difficulty of an inference rather than the number of possible models. This measure is more specific than the number of models and also more precise as it permits exact predictions to be made for many different classes of reasoning problems.

The basis of our suggestion is that our theory and PRISM carefully distinguish between the representational and processing assumptions embedded in the program, that each process in PRISM is realized by a distinct procedure, and that PRISM specifies which particular operation is used to manipulate a specific kind of mental representation. One advantage of such a computational model is that it allows precise predictions that can be tested with humans. Another advantage of PRISM is that the system permits the number of operations that must be performed by the system to solve a certain task to be counted. A difficulty measure based on this number of operations agrees with common approaches in computational complexity theory, a branch of the theory of computation concerned with the classification of computational problems according to their intrinsic difficulty (Papadimitriou, 1994). In the following, we report the original indeterminacy effect and then our explanation for the difference between determinate and indeterminate problems. Then, we show that our difficulty measure also predicts many other findings from the psychology of reasoning.

As previously stated, the *indeterminacy effect* was first reported by Byrne and Johnson-Laird (1989). We have mentioned the effect several times in this article but now come back to the precise structure of the problems. In fact, Byrne and Johnson-Laird compared not only the two sorts of premise sets (determinate and indeterminate) presented so far but the following three kinds of problems:

7a: Type 1:
A is to the right of B.
C is to the left of B.
D is in front of C.
E is in front of B.

Hence, D is to the left of E.

7b: Type 2:
B is on the right of A.
C is to the left of B.
D is in front of C.
E is in front of B.

Hence, D is to the left of E.

7c: Type 3:
B is to the right of A.
C is to the left of B.
D is in front of C.
E is in front of A.

What is the relation between D and E?

The Type 1 and Type 2 problems (7a and 7b) are the sorts of problems that we have mainly explored in the present article. Byrne and Johnson-Laird (1989) and other psychologists reported that problems of Type 1 are easier to solve than Type 2 problems (e.g., Carreiras & Santamaría, 1997; Roberts, 2000; Schaeken et al., 1998; Van der Henst, 2002). The standard explanation is that Type 1 problems support one model but Type 2 problems agree with two models. The models are different but lead to the same conclusion, "D is to the left of E." In this article, we have argued that one of the two models is preferred (the fff-model). However, we have paid less attention to Type 3 problems, in which the situation is different (Problem 7c). These are the last class of problems (Problem VIII) in Table 1. If you think about Problem 7c, you will see that the only correct answer to the question is that no valid conclusion exists. This is because you can construct multiple models of the premises, but in these models, there is no common relation between D and E. Thus, the only logically correct response is that nothing follows from the premises. Byrne and Johnson-Laird and others showed that these kinds of problems (e.g., Problem 7c) are the most difficult; they are, probably not surprisingly, more difficult than Type 1 problems (e.g., Problem 7a). But they are also more difficult than Type 2 problems (e.g., Problem 7b), although both types of problems (e.g., Problem 7a and Problem 7b) require two models. However, when we started our research on the indeterminacy effect and its connection to preferred mental models, we speculated that the number of possible models might not be the actual cause for the findings reported by Byrne and Johnson-Laird and other groups (e.g., Vandierendonck et al., 2004).

One reason for our skepticism is that Byrne and Johnson-Laird (1989) studied problems with one model and problems with two models. The two-model problems they called multiple-model problems, although just two models might be not enough to explore the connection between the number of possible models and the difficulty of reasoning problems. In our previous research, we studied problems with up to 13 models (with a set of relations from AI) and did not find a correlation between the number of possible models and reasoning difficulty (Knauff, Rauh, Schlieder, & Strube, 1998b). There was a difference when we pooled all problems with multiple models (3 models, 5 models, 9 models, and 13 models) together and compared these multimodel problems with the single-model problems. But there was no correlation between participants' reasoning performance and the exact number of possible models (Rauh et al., 2005). A second reason for our skepticism is that we think that preferred mental models normally prevent people from considering all possible models. Most of the time, people just reason with the preferred model but ignore other possible models. The third reason for our doubt is that the search for counterexamples in the standard model theory seems to us too much driven by the analogy to formal logic. For a logician, a conclusion is logically valid if and only if it is consistent with all possible models of the premises. But do logically untrained people use this definition of logical validity? There is indeed empirical evidence showing that people consider a conclusion logically valid if it holds in just one model (Rauh et al., 2005). From a psychological point of view, the search for counterexamples seems necessary only for very specific tasks, but not as a general rule for how people reason (see Table 1). However, if people do not search for counterexamples, in other words, if they do not try to generate and test all possible models, then the assumption about the connection between reasoning difficulty and the number of models is also questionable.

Here is our alternative explanation for the differences in difficulty: First, in PRISM, the solution for Type 1 and 2 problems (e.g., Problem 7a and Problem 7b) only relies on two processes: model construction and model inspection. There is no model variation because the task is to verify a valid conclusion, which, according to our theory, does not require considering alternative models (see Problems I and IV in Table 1). Second, both processes (model construction and inspection) comprise several more elementary operations of the focus, which inserts, moves, focuses, refocuses, and groups tokens in the array and moves groups of tokens. Our suggestion is to count the *number of focus operations* and to test whether this measure allows us to explain differences in reasoning difficulty. Do difficult problems require more focus operations than easier problems? Table 2 illustrates the necessary focus operations and the accumulated number of focus operations for the three sorts of problems (determinate, indeterminate with valid conclusion, indeterminate with invalid conclusion). The first column presents the premise that must be processed, the second column shows the model in the array in which the underlined token is the current position of the focus, the third column is a verbal description of the focus operation, and the fourth column shows the accumulated number of focus operations. As one can see in Table 2, PRISM processes the three types of problems differently. The determinate problem needs 12 focus operations, the indeterminate problems with a valid conclusion need 15 focus operations, and the indeterminate problems with an invalid conclusion require 21 focus operations. Note that we treat all operations of the focus equivalently as if they all have the same costs in terms of working memory capacity and processing costs. In principle, we could introduce several free parameters and values for the different operations, but we decided not do this in the current version of PRISM. We return to this point in the General Discussion.

An Assessment of the Preferred Models Theory

In the previous section, we have argued that the number of necessary focus operations is an adequate measure for the difficulty of reasoning problems. We have argued that this measure is even better than the less precise measure based on the number of possible models. The number of possible models might be a more intuitive measure, and we are aware that psychological theories occasionally benefit when they concur with our intuitions. It makes the theory easier to explain, and even experts in the field are more willing to agree with theories that are in accord with their intuition. However, in the following, we demonstrate that our new measure also permits us to predict many other well-known phenomena from the psychology of reasoning.

Take, for instance, the famous *premise order effect,* which was first reported in Ehrlich and Johnson-Laird (1982; see also Knauff, Rauh, Schlieder, & Strube, 1998a). The authors gave participants reasoning problems in which the premises were presented in three forms:

8a: Continuous premise order: A r_1 B, B r_2 C, C r_3 D

8b: Semicontinuous premise order: B r_2 C, C r_3 D, A r_1 B

8c: Discontinuous premise order: C r_3 D, A r_1 B, B r_2 C

The participants had to infer the conclusion A r_4 D. The authors reported that continuous order (37% error; Problem 8a) was easier than discontinuous order (60% error; Problem 8c) and that there was no significant difference between continuous and semicontinuous problems (39% error; Problem 8b; Ehrlich & Johnson-Laird, 1982). The model theory explains this premise order effect as being due to the difficulty of integrating the information from the premises. In the continuous (Problem 8a) and semicontinuous (Problem 8b) orders, it is possible to integrate the information from the first two premises into one model at the outset, whereas, when they are presented in a discontinuous order (Problem 8c), subjects must wait for the third premise in order to integrate the information from the premises into a unified representation. Similar results were reported by Carreiras and Santamaría (1997). Table 3 shows the number of focus operations necessary to solve problems in continuous, semicontinuous, and discontinuous premise order. The new measure for reasoning difficultly also clearly explains the premise order effect: The continuous problem requires nine focus operations, the semicontinuous problem requires 10 focus operations, and the discontinuous problem requires 12 focus operations. Note particularly that the processing of the last premises in the discontinuous problem requires significantly more focus operations. This agrees with the results from Knauff et al. (1998a) that are presented in Table 4. As visible in the table, we could show that the processing of the third premise in discontinuous problems (Problem 8c) takes significantly longer than the processing of third premise in continuous (Problem 8a) and semicontinuous (Problem 8b) problems. In the last column of Table 4, the corresponding number of focus operations is presented.

Another phenomenon in the domain of relational reasoning is the *figural effect*. The effect was reported in Johnson-Laird and Bara (1984) and originates from the area of syllogistic inference (see Khemlani & Johnson-Laird, 2012, for a comparison of psychological theories). The authors gave participants relational reasoning problems of the following two forms:

9a: figure 1: A is related B; B is related to C

9b: figure 2: B is related A; C is related to B

The participants had to draw a conclusion about A and C in their own words. Johnson-Laird and Bara (1984) reported that the participants showed an overall bias to generate more A–C conclusions than C–A conclusions. Moreover, this effect was modulated by the figure of the problem: In problems of figure 1 (Problem 9a), participants generated more conclusions in the form A–C than the other correct conclusion C–A. In contrast, they generated more conclusions in the form C–A for problems of figure 2 (Problem 9b). Johnson-Laird and Bara explained this figural effect by assuming that the integration of the premises in working memory is more difficult in figure 2 problems because of the need to bring the B term into the middle. According to this view, the construction of a mental model is easier for premises that have the repeated term as the first term in the next premise because, in this case, the information of the given premises can be integrated immediately and no cognitive resources are needed for mental operations that bring the middle term into the middle. We can explain this result with our new measure. However, we actually doubt that this effect is also transferable to the domain of reasoning with spatial relations. Note that Johnson-Laird and Bara explored reasoning with relations, but the relations were not spatial relations. Moreover, the relation "is related to" serves as its own converse, which is different in the spatial domain (where the converse of a relation is not the same relation, e.g., "left of," "right of"). In our laboratory we therefore conducted an experiment with spatial relations in which the participants had to solve spatial reasoning

Table 2
The Processing Steps of PRISM for the Three Types of Problems Introduced by Byrne and Johnson-Laird (1989)

Premise	Mental model	Operation of the focus	Accumulated number of focus operations
		Determinate problem with valid conclusion (Type 1 problem)	
Model construction			
A is right of B.	A	Insertion of object A	1
	_ A	Moves to the left	2
	$\underline{B}$ A	Insertion of object B	3
C is left of B.	_ B A	Moves one step left	4
	$\underline{C}$ B A	Inserts object C	5
D is in front of C.	C B A _	Move one step in front of object C	6
	C B A $\underline{D}$	Insert object D	7
E is in front of B.	$\underline{C}$ B A D	Search object B	8
	C $\underline{B}$ A D	Move one step in front of object B	9
	C B A D _	Move one step in front of object B	10
	C B A D $\underline{E}$	Insert object E	11
Model inspection			
D left of E? Yes or No?	C B A D $\underline{E}$	Focus is at E after model construction	11
	C B A $\underline{D}$ E	Move left (according to the relation in the conclusion)	12
PRISM output		YES, D is to the left of E. (which is logically correct)	12
		Indeterminate problem with valid conclusion (Type 2 problem)	
Model construction			
B is right of A.	$\underline{B}$	Insertion of object B	1
	_ B	Move one step left	2
	$\underline{A}$ B	Insertion of object A	3
C is left of B.	A $\underline{B}$	Search reference object B	4
	$\underline{A}$ B	Move one step left (ff-principle, not empty)	5
	_ A B	Move one step further to the left	6
	$\underline{C_B}$ A B	Insert object C (fff-principle) with annotation B (somewhere left of B)	7
D is in front of C.	C_B A B _	Move one step in front of object C	8
	C_B A B $\underline{D}$	Insert object D	9
E is in front of B.	$\underline{C_B}$ A B D	Search object B	10
	C_B $\underline{A}$ B D	Search object B	11
	C_B A $\underline{B}$ D	Move one step in front of object B	12
	C_B A B D _	Move one step in front of object B	13
	C_B A B D $\underline{E}$	Insert object E— preferred model	14
Model inspection			
D left of E? Yes or No?	C_B A B D $\underline{E}$	Focus is still at E after model construction	14
	C_B A B $\underline{D}$ E	Move left (according to the relation in the conclusion)	15
PRISM output		YES, D is to the left of E. (which is logically correct)	15
		Indeterminate problem with no valid conclusion (Type 3 problem)	
Model construction			
B is right of A.	$\underline{B}$	Insertion of object B	1
	_ B	Move one step left	2
	$\underline{A}$ B	Insertion of object A	3

(*table continues*)

Table 2 (*continued*)

Premise	Mental model	Operation of the focus	Accumulated number of focus operations
C is left of B.	A $\underline{B}$	Search reference object B	4
	$\underline{A}$ B	Move one step left	5
	_ A B	Move one step left	6
	$\underline{C_B}$ A B	Insert object C (fff-principle) with annotation B (somewhere left of B)	7
D is in front of C.	C_B A B _	Move one step in front of object C	8
	C_B A B $\underline{D}$	Insert object D	9
E is in front of A.	$\underline{C_B}$ A B D	Search object A	10
	C_B $\underline{A}$ B D	Search object A (found)	11
	C_B A B D _	Move one step in front of object A	12
	C_B A B D $\underline{E}$	Insert object E	13
Model inspection			
What is the relation between D and E?	C_B A B D $\underline{E}$	Focus is still at E after model construction	
	C_B A B $\underline{D}$ E	Move left to search for D. <u>Tentative conclusion: D left of E.</u> This conclusion is usually accepted which is logically incorrect. However, PRISM can also search for alternative models.	14
Model validation			
	C_B A B $\underline{D}$ E	Focus is still at D after model inspection	
Model validation starts because Annotation A says that C can also be directly left of B, thus to the right of A.	$\underline{C_B}$ A B D E	Search C, because C is RO for D	15
	$\underline{\mathbf{C_B}}$ A B **D** E	Group column (C D)	17
	C_B $\underline{A}$ B D E	Move one step right	19
	C_B $\underline{\mathbf{A}}$ B D **E**	Group column (AE)	20
	A C_B B E D	Exchange the grouped columns (AE) and (CD). <u>Second conclusion: D right of E</u>	21
	PRISM compares "D left of E" and "D right of E" (not part of PRISM)!		
PRISM output: No valid conclusion.			21

Note. A determinate problem with a valid conclusion (Type 1 problem) requires more focus operations than an indeterminate problem with a valid conclusion (Type 2 problem), and both in turn require more focus operations than inferences with aan indeterminate problem with no valid conclusion (Type 3 problem). This explains differences in reasoning difficulty from human experiments. PRISM = preferred inferences in reasoning with spatial mental models.

problems in four different figures (Knauff, Rauh, Schlieder, & Strube, 1998a):

10a:	figure 1:	A is to the left of B; B is to the left of C
10b:	figure 2:	B is to the right of A; B is to the left of C
10c:	figure 3:	A is to the left of B; C is to the right of B
10d:	figure 4:	B is to the right of A; C is to the right of B

All four figures (abbreviated as A–B, B–C; B–A, B–C; A–B, C–B; B–A, C–B) describe the same spatial arrangement. In contrast to Johnson-Laird and Bara's experiment, the participants here generated more C–A conclusions than A–C conclusions overall. However, we also found that this effect was modulated by the figure of the premises. In particular, the conclusion C–A was used even more often for the premise orders A–B, B–C and B–A, B–C (Problems 10a and 10b) than for the orders A–B, C–B and B–A, C–B (Problems 10c and 10d). Note that the figural effect is not a difference in difficulty but rather demonstrates a preference for a specific term order in the conclusion. Nevertheless, we can explain the effect with our *number of operations* measure. The explanation is so obvious that it is not necessary to present the whole procedure step by step. As we have described in the previous sections, an important rule in PRISM is that after the construction of the model, the focus remains at the last inserted token, in this case, C. As this last position is also the starting point of the inspection process, it is reasonable that people prefer to generate conclusions that begin with C. The reason is that an A–C conclusion would require shifting the focus back to token A before beginning the inspection process, which would require an additional

Table 3
Premise Order Effect From Knauff, Rauh, Schlieder, and Strube (1998a)

Premise order	Premise 1	Premise 2	Premise 3	Total number of focus operations
Continuous	13,0 (3)	11,2 (3)	10,9 (3)	9
Semicontinuous	13,6 (3)	11,0 (3)	14,4 (4)	10
Discontinuous	12,4 (3)	13,9 (3)	19,5 (6)	12

Note. Average premise reading time in seconds and the associated number of focus operations in PRISM (in parentheses). The last column shows the total number of focus operations as measure for the difficulty of the inference. The number of required focus operations explains the differences in premise processing times from the experiment. PRISM = preferred inferences in reasoning with spatial mental models.

focus operation. In contrast, C–A conclusions do not demand a relocation of the focus (Knauff et al., 1998a).

A relatively new finding is the *effect of relational complexity* reported in Goodwin and Johnson-Laird (2005). The background is the theory of relational complexity introduced by Halford, Wilson, and Phillips (1998; see also Halford, Wilson, & Phillips, 2010). Goodwin and Johnson-Laird presented problems with three premises about the relative starting positions of runners in a race to participants. The participants had to answer a query about the exact starting positions of two out of four runners on five lanes. Of course, because four runners A, B, C, D are on five lanes, one lane remains empty. The problems were as follows:

11a: Lower relational complexity:

A is left of C and B is left of A.
B is left of C and D is left of B.
A is further away from C than B is from A.

Who is closer to the empty lane, B or A?

11b: Higher relational complexity:

A is left of C and B is left of A.
B is left of C and D is left of B.
B is further away from C than D is from A.

Who is closer to the empty lane, B or A?

Note that the third premise of Problem 11a involves three runners, while in Problem 11b, it involves four runners. Accordingly, the third premise of Problem 11a involves a ternary relation (lower relational complexity), while the corresponding premise of Problem 11b involves a quaternary relation (higher relational complexity). However, both problems yield the same model:

D B A _ C

In this model, the _ denotes the empty lane, and the answer to the question in both problems is that A is closer to the empty lane than B is. However, Goodwin and Johnson-Laird reported that participants gave more correct answers to the low relational complexity problems (90 % correct; Problem 11a) than to the high relational complexity problems (77 % correct; Problem 11b). The participants also responded faster to lower relational complexity problems (15 s) than to the higher relational complexity problems (23 s). The experiment supports the theory of relational complexity (Halford et al., 1998) and shows that the difficulty in reasoning with relations is affected by the complexity of the relations in the premises. Related results on reasoning with *n*-place relations were reported in Jahn et al. (2007). Can our *number of focus operations* measure account for the relational complexity effect? Yes, it can. As visible in Table 5, the processing of the problem with lower relational complexity (11a) requires 22 focus operations, whereas the higher relational complexity problem (11b) requires 24 operations of the focus.

We have now shown that the number of necessary focus operations explains the difficulty of reasoning problems. We have further showed that the number of focus operations also explains other effects from the psychology of reasoning. The measure explains the premise order effect, the figural effect, and the effect of relational complexity. These phenomena do not only refer to reasoning difficulty but also represent other well-known empirical results from the psychology of reasoning. Figure 4 summarizes all four phenomena discussed in this and the previous sections.

Experiments

So far, we have demonstrated that our theory can explain why people prefer to construct some models but often struggle to consider alternative models. We have also provided a detailed explanation as to how people alter preferred models in order to obtain alternative interpretations of the premises. And we have demonstrated that our theory can account for many classic empirical results from the literature. The PRISM system demonstrates that it is capable of simulating the empirical findings, including reasoning difficulty, errors, and preferences. However, our theory not only explains many phenomena from the literature and our laboratory, it also leads to new predictions that can be tested in psychological experiments. We now report two experiments that tested the following predictions:

1. People construct just a single preferred model and neglect other possible models.

2. The preferred models are those models that are constructed according to the fff-strategy.

3. Models that require a longer sequence of local transformations are more likely to be neglected than models that are only minor variations of the preferred model.

All three predictions are a direct consequence of our PRISM model. The first two predictions are related to the model construction process. We tested them in Experiment 1. The third prediction is concerned with the model variation process, and we tested the hypothesis in Experiment 2. In Experiment 1, we used a generation paradigm and, in Experiment 2, a verification paradigm. The problems in Experiment 1 correspond to problems of Type VII in Table 1. The problems in Experiment 2 correspond to problems of Types IV, V, and VI in Table 1.

Experiment 1: Preferred Mental Models

In this experiment, we tested whether individuals construct preferred models based on the fff-strategy and whether models that

Table 4
Processing Steps in PRISM for Continuous, Semicontinuous, and Discontinuous Premise Order

Premise	Mental model		Operation of the focus	Accumulated number of focus operations
			Continuous premise order	
A is left of B.	A		Insertion of A	1
	A _		Moves to the right	2
	A B		Insertion of object B	3
B is left of C.	A B		Reads reference object B of the premise	4
	A B _		Moves one step right	5
	A B C		Inserts object C	6
C is left of D.	A B C		Reads reference object C of the premise	7
	A B C _		Moves one step right	8
	A B C D		Inserts object D	9
			Semicontinuous premise order	
B is left of C.	B		Insertion of B	1
	B _		Moves to the right	2
	B C		Insertion of object C	3
C is left of D.	B C		Reads reference object C of the premise	4
	B C _		Moves one step right	5
	B C D		Inserts object D	6
A is left of B.	B C D		Searches reference object B of the premise	7
	B C D		Reads reference object B of the premise	8
	_ B C D		Moves one step to the left	9
	A B C D		Inserts object A	10
			Discontinuous premise order	
A is left of B.	A		Insertion of A	1
	A _		Moves to the right	2
	A B		Insertion of object B	3
C is left of D.	A B	_	Generation of new array with a focus	4
	A B	C	Focus 2 inserts object C	5
	A B	C _	Focus 2 moves right	6
	A B	C D	Focus 2 inserts object D	7
B is left of C.	A B	C D	Focus 2 searches reference object C	8
	A B	C D	Focus 2 groups submodel CD	10
	A B C D		Grouped model CD inserted in first model	12

Note. PRISM needs more focus operations for continuous premise order than does semicontinuous order, and both require less focus operations than does discontinuous premise order. This explains differences in premise processing from human experiments. Note that PRISM processes differ for the continuous and semicontinuous order only in the number of refocus operations; the discontinuous problems require the generation of an alternative submodel, which is then integrated in the first model. This integration process requires grouping the objects of the second submodel (in this case, the objects C and D) and integrating them in the first model. The number of focus operations in problems with discontinuous premise order exceeds the number of focus operations in continuous premise order by three and in semicontinuous premise order by two focus operations. PRISM = preferred inferences in reasoning with spatial mental models.

result from other construction strategies are more difficult to build. The fff-strategy, which we defend, would mean that reasoners prefer to sacrifice adjacency in favor of outside insertion. The ff-strategy, in contrast, would mean that people prefer to relocate a token in favor of immediate insertion of the token at hand.

Method

We tested 20 students (9 female, $M = 23.4$ years) from Princeton University (Princeton, New Jersey). They acted as their own controls and evaluated two determinate, two indeterminate one-dimensional, and two indeterminate two-dimensional problems. The problems are illustrated in Table 6.

Each type of problem was presented twice to the participants, making a total of 12 problems. The problems were presented in a randomized order, and the tokens were fruits (lemon, orange, kiwi, peach, mango, and apple), which were also randomly assigned to the problems. The participants were tested individually using a computer that administered the experiment. The presentation of the premises was self-paced and followed the separate-stage paradigm introduced by Potts and Scholz (1975). Each premise was presented sequentially on the computer screen and remained there until the presentation of the last premise was finished. After the final button press, all premises were removed, and the participant was asked to draw the model on a sheet of paper. Participants were free to draw more than one model if they noticed that this was possible. However, they were neither instructed to draw more than one model nor told that in some problems more than one model was possible.

Table 5
The Processing of the Relational Complexity Problems From Goodwin and Johnson-Laird (2005) in PRISM

Premise	Mental model	Operation of the focus	Accumulated number of focus operations
		Construction process for the first two premises	
A is left of C.	$\underline{A}$	Insertion of A	1
	A _	Moves to the right	2
	A $\underline{C}$	Insertion of object C	3
B is left of A.	$\underline{A}$ C	Search reference object A	4
	_ A C	Move one step left	5
	$\underline{B}$ A C	Inserts object B	6
B is left of C.	B $\underline{A}$ C	Checks if B is left of C by searching object C	7
	B A $\underline{C}$	Found object C	8
D is left of B.	B $\underline{A}$ C	Searches reference object B	9
	$\underline{B}$ A C	Found object B	10
	_ B A C	Moves one step left	11
	$\underline{D}$ B A C	Inserts object D	12
		Low-complexity problem	
A is further away from C than B is from A.	D $\underline{B}$ A C	Searches object A	13
	D B $\underline{A}$ C	Searches object A	14
	D B A $\underline{C}$	Searches object C/counts distance	15
	D B $\underline{A}$ C	Searches object A	16
	D $\underline{B}$ A C	Searches object B/counts distance	17
	D B $\underline{A}$ C	Searches object A	18
	D B A $\underline{C}$	Searches object C/counts distance	19
	D B A _	Deletes object C	20
	D B A □ _	Moves one step right	21
	D B A □ $\underline{C}$	Inserts object C	22
		High-complexity problem	
B is further away from C than D is from A.	D $\underline{B}$ A C	Searches object B	13
	D B $\underline{A}$ C	Searches object C/counts distance	14
	D B A $\underline{C}$	Found object C/Distance 2	15
	D B $\underline{A}$ C	Searches object A, D	16
	D $\underline{B}$ A C	Searches object D/counts distance	17
	$\underline{D}$ B A C	Found object D/Distance 2	18
	D $\underline{B}$ A C	Searches object B	19
	D B $\underline{A}$ C	Searches object C/counts distance	20
	D B A $\underline{C}$	Found object C/Distance 2	21
	D B A _	Deletes object C	22
	D B A □ _	Moves one step right	23
	D B A □ $\underline{C}$	Inserts object C	24

Note. High-complexity problems require more focus operations than low-complexity problems. This explains the findings from the human experiments. The first two premises are constructed by PRISM as before; for the low-complexity case, 22 operations are necessary. The third premise, "A is further away from C than B is from A," requires counting the distances between positions and moving objects accordingly. For the high-complexity case, 24 operations are necessary. The higher relational difficulty depends on the necessity to compare the distance of four diverse objects. The square indicates an empty lane. PRISM = preferred inferences in reasoning with spatial mental models.

Results and Discussion

The participants performed well in both determinate (83% correct drawings) and indeterminate problems (92% correct drawings; $p < .01$). At first glance, the higher rate of correct drawings for indeterminate problems seems to conflict with the indeterminacy effect (indeterminate problems are more difficult). The effect is, however, understandable when we remember that we are talking about an experiment in a generation paradigm. In this paradigm, the probability of finding a valid model for a single-model problem just by chance is in fact lower than for multiple-model problems. This explains why our participants drew more correct models for indeterminate problems than for determinate tasks. However, the analysis of indeterminate problems is more important. Although the premises were consistent with more than one model, in 90% of the problems, the participants drew just one model. They ignored that there were alternative interpretations of the premises. Only two participants (10%) produced more than one model. This result agrees with our first prediction that people construct just a single model and neglect other possible models. We then, in a second step, further analyzed the correct drawings and found that the participants drew significantly more models according to the fff-strategy, whereas they only rarely applied the ff-strategy (fff-models: 78%, ff-models: 22%; $p < .01$). This result supports our prediction that preferred models exist and that these preferred models are those models constructed according to the fff-strategy. Our corollary from this experiment is that if the first possible position in the model that fulfills the spatial relation of a premise is already occupied by

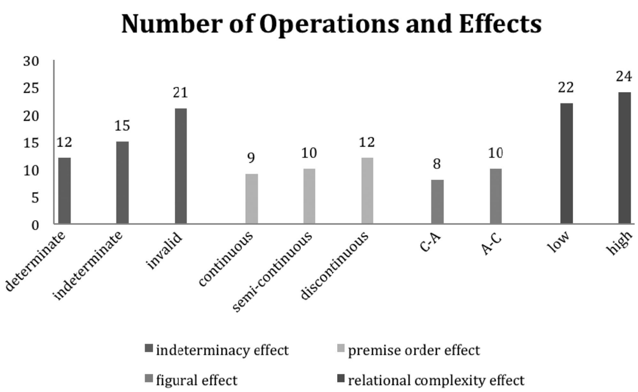

Figure 4. The number of necessary focus operations in PRISM explains the difficultly of spatial reasoning problems. This difficulty measure can account for a number of classical findings from the reasoning literature: the indeterminacy effect (Byrne & Johnson-Laird, 1989), the figural effect (Johnson-Laird & Bara, 1984; Knauff, Rauh, Schlieder, & Strube, 1998a), the premise order effect (Ehrlich & Johnson-Laird; 1982; Knauff et al. 1998a), and the relational complexity effect (Goodwin & Johnson-Laird, 2005). The measure is a better predictor for reasoning difficulty than the classical difference between single-model problems and multiple-model problems. PRISM = preferred inferences in reasoning with spatial mental models.

another token, human reasoners prefer to sacrifice adjacency (ff-strategy) in favor of outside insertion (fff-strategy; Jahn et al., 2007). In other words, reasoners avoid relocating an object that is already in the model to free up the first possible position. Instead, they place the object in question at the end of the line, where the relation is also fulfilled. PRISM predicts exactly this behavior.

Experiment 2: Alternative Mental Models

In this experiment, we tested whether alternative models are built from scratch or by local transformations of the preferred model. We investigated this during the variation phase of the inference process. The number of correct inferences and the solution times for different presented models were analyzed.

Method

We tested 25 students (11 female, $M = 23.9$ years) from the University of Freiburg (Freiburg, Germany). They had to solve 24 indeterminate spatial reasoning problems with four premises. Half of the problems were valid, the other half invalid. Results from invalid problems are not reported in this article. Three different types of possible models were offered: fff-models, ff-models, and mix-models, which were a combination of the ff- and fff-strategies. Each premise consisted of two tokens (from a set of five fruits, apple, pear, peach, kiwi, or mango) and one spatial relation (e.g., "left of"). For each problem, one of six models (three valid) was offered. The "valid" models could be constructed using the fff-strategy (preferred model), the ff-strategy, or a mixture of both strategies. The models constructed according to the ff-strategy and the mix-strategy were the alternative mental models. The participants read the four premises self-paced in a sequential order. Each premise was displayed in the middle of the screen and disappeared before the next was presented. After premise presentation, a model was shown, and the participants had to decide whether this model was a possible model of the premises.

Results and Discussion

Our participants identified models constructed according to the fff-strategy significantly more often (correctness: 92%; response time = 3.797 ms) than other models that were also consistent with the premises but followed the mix- (correctness = 81%; response time = 4.359 ms) or ff-strategy (correctness = 44%; response time = 6.410 ms). In fact, only 8% of the preferred models (fff-strategy) were erroneously rejected, whereas more than 50% of the models constructed according to the ff-strategy were erroneously rejected. This clearly indicates that it is much easier to identify the preferred fff-model than the other models (ff- and mix-model). The same trend is also visible in the response times. Here, the acceptance of fff-models was significantly faster than with the other models. Both trends were statistically significant (error rate: preferred fff-model < mix-strategy model < ff-strategy model; Page's L test: $L = 3.55$, $p < .001$; response time: preferred fff-model < mix-strategy < ff-strategy; Page's L test: $L = 234.5$, $p < .01$). Note that the participants performed at chance level for ff-models.

The experiment shows, first, that the fff-models were indeed preferred and easier to identify. However, another essential finding is that mixed models were more often identified than ff-models. This is important because it indicates that reasoners have difficulty using the ff-strategy although it means applying only a single strategy, which one might think should be much easier than changing strategies within a problem. But if we assume that the participants who constructed the preferred model followed the fff-strategy, it is easier to swap only the fourth token with the third token to transfer the preferred model into a model that is constructed according to the mix-strategy. The variation of the preferred model, which is constructed in accordance with the ff-strategy, requires more effort because two tokens must be swapped and more relations have to be taken into account. This supports our third prediction that alternative models are generated by applying minimal changes to the preferred model. The mix-model is indeed the next neighbor to the preferred model, while the ff-model is further away in the neighborhood graph (cf. the neighborhood graph above). This is exactly what PRISM predicts.

General Discussion

We had two goals in the research reported in this article. The first goal was to advance a comprehensive theory of human spatial reasoning. The theory we suggest relies on mental models, but in this article, we have also showed where the standard model falls short of explaining some phenomena of human spatial thinking. We have therefore presented our preferred model theory, which explains how humans deal with the ambiguity of spatial descriptions, why people prefer to construct some models while neglecting others, and how this leads to reasoning errors and logically invalid inferences. The second goal of this article has been to show

Table 6
Problems Presented in Experiment 1

Problem	Possible models						Construction strategy
[a] A is to the left of B.	**(1)**	**E**		**D**			fff (preferred model)
C is to the right of A.		**A**	**B**	**C**			
D is behind C.	(2)	E	D				ff (alternative model)
E is behind A.		A	C	B			
[b] A is to the left of B.	**(1)**	**E**					fff (preferred model)
C is to the right of A.		**A**	**B**	**C**			
D is behind C.				**D**			
E is behind A.	(2)	E					ff (alternative model)
		A	C	B			
			D				
[c] B is to the right of A.	**(1)**	**A**	**B**	**C**	**D**	**E**	fff (preferred model)
C is to the right of B.	(2)	A	B	D	C	E	ff (alternative model)
D is to the right of B.	(3)	A	B	C	E	D	mix (alternative model)
E is to the right of C.							
[d] B is to the right of A.	**(1)**	**A**	**B**	**C**	**D**	**E**	fff (preferred model)
C is to the right of B.	(2)	A	B	C	E	D	mix (alternative model)
D is to the right of B.	(3)	A	B	D	E	C	mix (alternative model)
E is to the right of B.	(4)	A	B	D	C	E	mix (alternative model)
	(5)	A	B	E	C	D	mix (alternative model)
	(6)	A	B	E	D	C	ff (alternative model)
[e] A is to the left of B.		E		D			determinate
C is to the right of B.		A	B	C			
D is behind C.							
E is behind A.							
[f] B is to the right of A.		A	B	C	D	E	determinate
C is to the right of B.							
D is to the right of C.							
E is to the right of D.							

Note. The left column contains the premise structure of the different problems, the middle column contains the possible models, and the last column contains additional comments for each possible model. The problems [a] and [b] are indeterminate two-dimensional with two possible models, the problems [c] and [d] are one-dimensional with three or six possible models, and the problems [e] and [f] are determinate one- and two-dimensional with only one possible model. The model indicated in bold letters is the hypothesized preferred model for each problem constructed in line with the fff-principle. ff = first fit; fff = first free fit.

how our theory can be translated into a psychologically realistic computational model of human spatial reasoning. We have argued that the key to modeling human spatial reasoning is to provide an architecture that preserves the spatial relation between objects from the problem description and where each postulated mental process is realized by a distinct procedure. The PRISM system provides an existence proof that such a system is possible and can simulate human spatial reasoning performance, including reasoning difficulty, errors, and preferred and neglected models. We have also showed that PRISM leads to new predictions, and we have reported results from two experiments that agree with these predictions. In the remaining parts of the article, we, first, compare PRISM to other computer models of human reasoning. Second, we show that our theory is not limited to reasoning with spatial relations but is a more universal theory of reasoning with all kinds of relations, even if they are not spatial in nature. Finally, we discuss some limitations of PRISM and some open problems with our theory. We conclude that our theory has many implications for a more general theory of human thinking.

A Comparison of PRISM With Other Computational Models

PRISM is not the first attempt to simulate human spatial reasoning. John Hummel and coworkers developed the LISA system, which is not specifically concerned with spatial relations but is a more general model of human relational inferences (Hummel & Holyoak, 2003, 2005). LISA solves relational inferences and learns from analogical mapping (which we do not describe here). However, Hummel and Holyoak (2001) developed a mental array module (MAM) that maps pairwise relations between objects onto locations in a spatial array. The system can compute transitive inferences and is able to simulate some empirical findings, including the relational complexity effect and the effect of markedness as suggested by Clark (1969). An umarked, positive form of an adjective such as "taller" is more easily (and quickly) processed than its marked counterpart "smaller." Hummel and Holyoak (2001) simulated related findings from Sternberg (1980) in the MAM of LISA and showed how the meanings of comparatives such as "greater" and "lesser" can be defined in terms of the operation of the spatial module itself. In PRISM, we did not try to simulate markedness, as the effect seems be at least partially caused by linguistic factors, which are not the focus of PRISM. Moreover, the effect of markedness for spatial expressions such as "left of," "right of," "above," "below," and so on is not clear and might be marginal (Baguley & Payne, 2000). Other effects from the domain of spatial reasoning, such as the indeterminacy effect, figural effect, and premise order effect, have not been reconstructed in LISA. However, the principle difference between PRISM and LISA is that LISA is a connectionist model with some symbolic ingredients, whereas PRISM is a strictly symbolic account. Certainly, both connectionist and symbol approaches have their strengths and weaknesses, but many authors have claimed that symbolic accounts must make the representational and processing assumptions explicit, while connectionist models implicitly build them into patterns of connections, rules for how weights change, and so on (e.g., Anderson, 1993). Another difference between PRISM and LISA is that the MAM processes metrical distance information (Hummel & Holyoak, 2001). The representation in PRISM is much more parsimonious as it requires only qualitative information about the relationship between tokens in a symbolic spatial array, without needing to incorporate metrical distance information. Nevertheless, PRISM is adequate to reconstruct many experimental findings from the literature, including our own studies. In particular, it adequately rebuilds preferred models and offers a comprehensive account for model variation by means of minimal changes. For a further comparison, it would be important to see how LISA can deal with preferred and alternative mental models in reasoning with spatial relations.

There are also some strictly symbolic computational implementations of human spatial inference that have commonalities with PRISM. The classical program by Johnson-Laird and Byrne (1991) is able to parse relational premises and to insert objects into an array. Yet the program does not explain how the insertion and focus operations work, and it is also not clear on how the system would deal with empirical model preferences. Another model is the UNICORE system of Bara and colleagues (2001). UNICORE also relies on the placement of tokens in a spatial array and a scanning process that realizes the inference. UNICORE does not, however, provide a detailed reconstruction of each reasoning step but basically consists of one large algorithm that cannot be split into its individual subprocesses. Other computational models have been developed by Baguley and Payne (2000) and by our colleague Christoph Schlieder. Schlieder modeled the effects of preferred models within a system that uses a focus to place and manipulate objects in a linear order on a left-to-right axis. Schlieder also suggested the principle of minimal direction change and that the model inspection process starts at the last inserted token in a model, which we also use in PRISM. However, his system is confined to a single spatial dimension and deals with rather arbitrary spatial relations from artificial intelligence research (Schlieder & Berendt, 1998).

A remarkable computer implementation of spatial reasoning was developed by Goodwin and Johnson-Laird (2005). In their article, the authors also raised the question of how people deal with ambiguous multiple-model problems and referred to our earlier findings on preferred mental models (Jahn et al., 2007; Knauff et al., 1995). Thus, the authors also suggested that "individuals tend to construct only a single mental model of a set of relations, to construct the simplest possible model, and to use their knowledge to yield a typical model" (Goodwin & Johnson-Laird, 2005, p. 478). However, they did not offer an explanation as to why a specific model is preferred over others and also defended the opinion that human reasoners systematically search for counterexamples to falsify a putative conclusion. In this respect, the system by Goodwin and Johnson-Laird leaves open many questions that we can answer with PRISM. In another respect, however, Goodwin and Johnson-Laird's system has features that PRISM does not yet offer. Particularly, Goodwin and Johnson-Laird's system is interesting because the first stage of the program yields a representation of the meaning of a spatial description. In an interpretative process, each word refers to a lexical entry specifying the meaning of the word. The program uses corresponding semantic rules for the location of the objects in a three-dimensional array of cells and is also able "to combine the meanings of words and phrases according to the grammatical relations among them" (Goodwin & Johnson-Laird, 2005, p. 477). This leads to compositional semantics as suggested by Montague (1974). As stated at the beginning of this article, PRISM does not include such a parser that understands the meaning of spatial relational expression because our aim is to reconstruct actual reasoning processes rather than the understanding of the premises. This makes PRISM highly flexible because, in principle, the premise input can be in any format (sentences, pictures, arrangements of real physical objects, etc.). Moreover, this approach agrees with most reasoning researchers' view that a reasoning theory starts after the processing of a verbal input and before the reasoning processes' result must be translated into a verbal output (Braine & O'Brien, 1998; Evans et al., 1993; Goodwin & Johnson-Laird, 2005; Johnson-Laird, 1983; Rips, 1994). However, apart from that, it is probably desirable to complement PRISM with a parser that uses semantic rules to understand the spatial expressions from the premises. One of our next tasks will be to explore how the approach of Goodwin and Johnson-Laird can be combined with PRISM.

Reasoning With Other Sorts of Relations

Thus far, we have only discussed reasoning with spatial relations. However, we now show that our theory and PRISM can also deal with all kinds of nonspatial relations, such as earlier–later, hotter–colder, better–worse, smarter–dumber, is a relative of, is more popular than, is more tasty than, and so on. Take, for example, the following two problems about German chancellors:

12: Willy Brandt was more popular than Gerhard Schröder was.
Gerhard Schröder was more popular than Angela Merkel is.

13: Gerhard Schröder was more popular than Angela Merkel is.
Willy Brandt was more popular than Angela Merkel is.

Who was the least popular? Problem 12 is a determinate problem, and the answer is that Ms. Merkel is the most unpopular. You can easily see that if you map the nonspatial relation of more or less popularity on a spatial relation and represent the relation in the spatial model:

Brandt	Schröder	Merkel

You can see that PRISM can cope with this problem in the same way it does with spatial relations. Problem 13 is an indeterminate problem with two models:

Brandt	Schröder	Merkel
Schröder	Brandt	Merkel

PRISM predicts that you construct the first model because after the first premise the focus is at the token "Merkel." Since this is the starting point of the processing of the second premise, the focus starts from there and detects that the next cell already is filled with "Schröder." Following the fff-strategy, the focus thus moves one step further to the left in the array and inserts "Brandt" into the model. Again, the least popular is Merkel, but the most popular is Brandt because this results from the preferred model. Notice that the second dimension in the array can be used to represent that two tokens have the same value in a certain dimension, for example, that they are equally popular or old. For instance, imagine PRISM receives the following premises:

14: Ann is older than Brenda.

Brenda is older than Cathy.

David is as old as Ann.

Eric is as old as Cathy.

Try by yourself to find out who is oldest. The important thing to notice it that our account can deal with nonspatial relations in the same way that it solves spatial problems. Many studies have shown that the mapping of all kinds of nonspatial relations on spatial relations is deeply rooted in human cognition (Gattis & Holyoak, 1996; Tversky, Kugelmass, & Winter, 1991). Many of these findings found their way into the theory of relational complexity (Halford et al., 1998, 2010) and were corroborated by functional brain-imaging experiments (e.g., Knowlton & Holyoak, 2009; Waltz et al., 1999). This aspect of our theory is psychologically appealing because, in our daily life, relations and inferences about them give rise to different actions and result in certain preferences. People form beliefs and desires based on relations and develop attitudes toward sets of objects or events, typically reflected in an implicit or explicit decision or choice (Lichtenstein & Slovic, 2006). One of our next projects will be to explore how preferred models might affect such decisions, attitudes, and prejudices in our everyday life.

Limitations, Extensions, and Future Work

Our current theory also has a few limitations and leads to some new questions. Let us first consider some of the limitations. Perhaps the most serious limitation of the current PRISM model is that, like virtually all computational simulations of human reasoning, it does not tell us how reasoning interacts with knowledge and beliefs. From a pure reasoning perspective, this might be feasible, but of course, focusing on isolated reasoning processes does not mirror how humans reason in everyday life. Many studies have shown that what we know, or believe we know, affects how we think. The famous *belief bias,* which has been observed in almost all domains of reasoning, is just one example of how strongly people's thinking is affected by their beliefs (overview in Evans, 1990). Belief biases have also been explored in the domain of spatial reasoning, notably in the area of reasoning with geographic information. In the famous Reno–San Diego effect, for instance, American university students inferred from the fact that Reno is in Nevada and Nevada is generally east of California that Reno is east of San Diego (which is incorrect; Stevens & Coupe, 1978). Tversky (1981) reported many similar distortions in judged spatial relations in cognitive maps. In our laboratory, we conducted experiments in which we explored how preferred mental models are modified by the background knowledge and personal opinions of reasoners. The participants' task was to draw spatial inferences about the contamination of sectors of a river with chemicals and the resulting environmental hazard. If two sectors with different chemicals overlapped, it resulted in environmental danger. One group of participants was "ecologists" (environmental perspective); the other half were "owners of the chemical firm" (economical perspective). A third group was "neutral" people. We could, first, clearly identify the preferred mental models of the neutral people. The second finding, however, was that ecologists had a strong tendency to draw conclusions that would force a stoppage of the further inlet of chemicals, whereas the economists most often saw no reasons for stopping the inlet. Interestingly, the decisions were made faster if they were supported by the preferred model in the neutral perspective, but the reasoner needed more time if she or he had to vary the preferred model to make an inference matching his or her own interests (Kuβ, 1998; Rauh, 2000). Such effects are highly interesting from a basic and from an applied psychological perspective but cannot be modeled in our current PRISM model.

A related limitation of PRISM is that the model has no link to a mental lexicon in which, for instance, the semantics of spatial prepositions are represented. This was discussed in the previous section in which we presented the computer simulation suggested by Goodwin and Johnson-Laird (2005). For PRISM, this limitation has some drawbacks. For instance, PRISM currently uses a rather narrow deictic interpretation of spatial prepositions and does not tolerate a slightly misaligned placement of tokens. The deictic meaning of spatial propositions depends on the observers' position and ignores that the spatial arrangements look different depending on a person's perspective. The deictic meaning also disregards that many objects have an intrinsic orientation. The expression "in front of a car," for instance, for many people means that an object is somewhere in front of the windshield of the car even if the observer is facing one side of the car (for an overview, see Zlatev, 2007). PRISM also does not account for the fuzziness and vagueness of spatial expressions. Hayward and Tarr (1995), as well as many other groups, presented spatial prepositions to people together with diagrams showing arrangements of objects. Next, they asked participants to rate the goodness of a spatial preposition as a description of the spatial relationship between the objects in the diagram. It is surprising how imprecise the usage of such spatial prepositions is. People are highly flexible in using prepositions, and their semantics are far from clear (Gapp, 1997; Hayward & Tarr, 1995; Knauff, 1999; Vorwerg & Rickheit, 1998). The current version of PRISM cannot account for the fuzziness and ambiguity of spatial relations in natural language.

A further limitation of PRISM is that while the system has some free parameters, all of these are currently set on a fixed default value. In the current phase of our project, this is certainly a great advantage rather than a disadvantage because we cannot use free parameters for data fitting. In the future, however, it might be reasonable to set some of the parameters to specific values. For example, we presently treat all operations of the focus equivalently, as though the diverse operations had the same costs in terms of working memory capacity and processing costs and speed. We currently also set a parameter for the decay of information in spatial working memory to 0, which basically means that PRISM can maintain a model in the spatial array forever. The capacity of spatial working memory is currently also not limited. We consider the fact that PRISM currently does not have one free parameter to play with to fit the empirical data beneficial to our theory. However, further experimental findings will probably help us to set some parameters based on empirical grounds.

There are also a few technical questions regarding the sorts of spatial relations PRISM can deal with. In particular, PRISM is limited to reasoning with relations in the dimensions from left to right and from front to back; it does not deal with vertical relations. However, we do not think that this is a principal problem for our theory. For instance, Goodwin and Johnson-Laird (2005) developed a simulation that works in all three spatial dimensions. We instead decided to confine PRISM to two-dimensional reasoning because we want to demonstrate how PRISM works with the classical indeterminate problems from the psychology of reasoning (which are two-dimensional). These problems are complex enough to display the main ideas of our theory. Technically, it is easy to extend PRISM to the third dimension. However, we are not sure whether this would lead to additional insights that are psychologically interesting.

PRISM is also incapable of reasoning with some relations in the horizontal plane. On the one hand, PRISM can reason with topological relations such as "in between" or "next to." This is important, as we were able to show in earlier experiments that people also construct preferred mental models when reasoning with topological relations (Knauff & Ragni, 2011). PRISM can also reason with qualitative (but not metrical) distances like "further away" or "closer by," as we showed in the context of relational complexity. Thus, PRISM can reason with qualitative ordinal, topological, and distance relations. On the other hand, PRISM struggles with inferences in the diagonal of the spatial array (reasoning with "diagonally behind' is not possible) and also cannot handle the spatial relation "in the same place." That PRISM does not allow two (or more) tokens to be in the same cell of the array was an important conceptual decision during the design of PRISM, which we think is sensible. In the real world, solid objects cannot usually occupy the some place at the same time. Liquids or container-like objects might be exceptions, but to deal with such entities would make PRISM unjustifiably complex and would have too far-reaching consequences for our model construction process. Participants in our experiments also do not usually locate two objects at the same place.

A final question we want to discuss is whether the standard model theory actually assumes that people consider all possible models and always search for counterexamples to falsify a putative conclusion. The first thing to say is that Polk and Newell (1995), in their theory of verbal reasoning, also eschewed the use of counterexamples. The authors developed a verbal model theory and also argued, as we do, that falsification and search for counterexamples might be less important than the standard model theory suggests. Polk and Newell's account prompted various mental model theorists, on the one hand, to play down the idea of an automatic search for counterexamples and alternative models and, on the other hand, to find evidence for the use of counterexamples in certain reasoning tasks (Neth & Johnson-Laird, 1999; Vandierendonck et al., 2004). In recent work from our group, we developed a computational model of spatial reasoning as verbal reasoning that also works without falsifications and counterexamples (Krumnack, Bucher, Nejasmic, Nebel, & Knauff, 2011).

Another point is that the classical model theory also makes attempts to abandon the assumption that people consider all possible models. In Knauff et al. (1995), we showed for the first time that reasoners have strong preferences in combining spatial premises in just a single model—the preferred mental model. Later, Vandierendonck et al. (1999) argued that reasoners could try to keep track of all possible interpretations in a single model but represent symbolically that the relation between some tokens is indeterminate. Schaeken and Johnson-Laird (2000) developed a system for reasoning with temporal relations that reduces the number of models by ignoring all premises that are irrelevant for an inference. Goodwin and Johnson-Laird (2005) also argued that people might develop different strategies to deal with indeterminate premises. Their article referred to a set of principles that determine a preferred model, which it derived from Jahn, Johnson-Laird, and Knauff (2004). The authors also considered the possibility that people overlook counterexamples and wrote that their "program constructs multiple models of indeterminacies, but only to a limited number" (Goodwin & Johnson-Laird, 2005, p. 479). So, on the one hand, it should be clear that we consider the theory of mental models to be the most successful theory of human spatial reasoning. On the other hand, there are also genuine differences between the standard model theory and the approach we have presented in this article. An important difference between the two accounts is that we are able to predict and explain why some models are preferred over others, while other models are difficult to grasp for most people. The preferred models are those constructed following the principle that new entities are added to a model without relocating those entities already represented in the model. Another novelty of our theory is that we use a neighborhood graph to model how alternative models are constructed by local transformations. In this article, we have shown that alternative models that necessitate a longer sequence of local transformations are more likely to be overlooked than models that are only minor variations of the preferred model. According to our theory, that is the main cause for reasoning errors and illusory inferences. The errors people make are not random but depend on the distance of the necessary model from the preferred model in the neighborhood graph. The risk of missing a particular alternative model rises with its distance from the preferred model.

Conclusions

In this article, we have showed that human spatial reasoning is heavily biased toward preferred models and that these preferences result in erroneous conclusions and illusory inferences. Models should not be confused with solutions, though they often give rise to solutions. Thus, preferred models have an instant effect on how people solve problems. Preferred models bias people toward certain solutions and make them almost blind to alternative ways of solving a problem. And the more revisions of an initial model a person must mentally perform, the higher the risk that he or she will neglect new solutions to a problem. We are probably able to grasp readily that a problem has an alternative solution that is similar to our preferred way of solving it. But it seems to be difficult to revise our preferred solution so radically that we consider an entirely new solution. We have shown this in the spatial domain, but we think that preferred and neglected mental models my also play an important role in other domains of thinking, reasoning, and problem solving. They can help us to think efficiently, but the preferred models can also bias our thinking toward too simple solutions for complex problems.

References

Allen, G. (2004). *Human spatial memory: Remembering where*. Mahwah, NJ: Erlbaum.

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1986). *Working memory*. New York, NY: Oxford University Press.

Baddeley, A. D. (1990). *Human memory*. Hove, England: Erlbaum.

Baguley, T. S., & Payne, S. J. (2000). Long-term memory for spatial and temporal mental models includes construction processes and model structure. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53*A, 479–512.

Bara, B., Bucciarelli, M., & Lombardo, V. (2001). Model theory of deduction: A unified computational approach. *Cognitive Science, 25,* 839–901. doi:10.1207/s15516709cog2506_1

Barwise, J. (1982). *Handbook of mathematical logic*. Amsterdam, the Netherlands: North-Holland.

Boudreau, G., & Pigeau, R. (2001). The mental representation and processes of spatial deductive reasoning with diagrams and sentences. *International Journal of Psychology, 36,* 42–52. doi:10.1080/00207590042000001

Braine, M. D. S., & O'Brien, D. P. (1998). *Mental logic*. Mahwah, NJ: Erlbaum.

Breslow, L. (1981). Reevaluation of the literature on the development of transitive inferences. *Psychological Bulletin, 89,* 325–351. doi:10.1037/0033-2909.89.2.325

Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences, 1124,* 77–97. doi:10.1196/annals.1440.002

Byrne, R., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language, 28,* 564–575. doi:10.1016/0749-596X(89)90013-2

Carreiras, M., & Santamaría, C. (1997). Reasoning about relations: Spatial and nonspatial problems. *Thinking & Reasoning, 3,* 191–208. doi:10.1080/135467897394347

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76,* 387–404. doi:10.1037/h0027578

Cohn, A. G., & Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae, 46,* 1–29.

Cohn, A. G., & Renz, J. (2007). Qualitative spatial representation and reasoning. In F. van Harmelen, V. Lifschitz, & B. Porter (Eds.), *Handbook of knowledge representation* (pp. 551–596). Amsterdam, the Netherlands: Elsevier.

Denis, M., & Loomis, J. (2007). Perspectives on human spatial cognition: Memory, navigation, and environmental learning. *Psychological Research, 71,* 235–239. doi:10.1007/s00426-006-0079-x

Diestel, R. (2012). *Graph theory*. Heidelberg, Germany: Springer.

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior, 21,* 296–306. doi:10.1016/S0022-5371(82)90626-0

Evans, J. S. B. T. (1990). *Bias in human reasoning: Causes and consequences*. Hove, England: Psychology Press.

Evans, J. S. B. T., Newstead, S. E., & Byrne, R. (1993). *Human reasoning: The psychology of deduction*. London, England: Erlbaum.

Fangmeier, T., & Knauff, M. (2009). Neural correlates of acoustic reasoning. *Brain Research, 1249,* 181–190. doi:10.1016/j.brainres.2008.10.025

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. M. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience, 18,* 320–334. doi:10.1162/jocn.2006.18.3.320

Freksa, C. (1991). Qualitative and spatial reasoning. In D. M. Mark & F. U. Andrew (Eds.), *Cognitive and linguistic aspects of geographic space* (pp. 361–372). London, England: Kluwer. doi:10.1007/978-94-011-2606-9_20

Freksa, C. (1992). Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54,* 199–227. doi:10.1016/0004-3702(92)90090-K

Gapp, K. P. (1997). *Objektlokalisation: Ein System zur sprachlichen Raumbeschreibung* [Object localization: A system for verbal spatial descriptions]. Wiesbaden, Germany: Deutscher Universitätsverlag. doi:10.1007/978-3-663-01531-4

Gärdenfors, P. (1984). Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy, 62,* 136–157. doi:10.1080/00048408412341331

Gärdenfors, P. (1988). *Knowledge in flux*. Cambridge, MA: MIT Press.

Gärdenfors, P. (1990). The dynamics of belief systems: Foundations vs. coherence theories. *Revue Internationale de Philosophie, 172,* 24–46.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 231–239. doi:10.1037/0278-7393.22.1.231

Goel, V., & Dolan, R. J. (2001). Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia, 39,* 901–909. doi:10.1016/S0028-3932(01)00024-0

Goel, V., Stollstorff, M., Nakic, M., Knutson, K., & Grafman, J. (2009). A role for right ventral lateral prefrontal cortex in reasoning about indeterminate relations. *Neuropsychologia, 47,* 2790–2797. doi:10.1016/j.neuropsychologia.2009.06.002

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review, 112,* 468–493. doi:10.1037/0033-295X.112.2.468

Goodwin, G. P., & Johnson-Laird, P. N. (2008). Transitive and pseudo-transitive inferences. *Cognition, 108,* 320–352. doi:10.1016/j.cognition.2008.02.010

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21,* 803–831. doi:10.1017/S0140525X98001769

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14,* 497–505. doi:10.1016/j.tics.2010.08.005

Harman, G. (1986). *Change in view: Principles of reasoning*. Cambridge, MA: MIT Press.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition, 55,* 39–84. doi:10.1016/0010-0277(94)00643-Y

Herskovits, A. (1986). *Language and spatial cognition*. Cambridge, England: Cambridge University Press.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 519–537. doi:10.1037/0096-1523.30.3.519

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In M. L. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 279–305). Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220–264. doi:10.1037/0033-295X.110.2.220

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neutrally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Psychological Science, 14,* 153–157. doi:10.1111/j.0963-7214.2005.00350.x

Jackendoff, R., & Landau, B. (1991). Spatial language and spatial cognition. In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman* (pp. 145–169). Hillsdale, NJ: Erlbaum.

Jahn, G., Johnson-Laird, P. N., & Knauff, M. (2004). Reasoning about consistency with spatial mental models: Hidden and obvious indeterminacy in spatial descriptions. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial cognition IV: Reasoning, action, interaction* (pp. 165–180). Berlin, Germany: Springer.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition, 35,* 2075–2087. doi:10.3758/BF03192939

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 441–467). San Diego, CA: Academic Press. doi:10.1016/B978-012301160-2/50016-2

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5,* 434–442. doi:10.1016/S1364-6613(00)01751-4

Johnson-Laird, P. N. (2006). *How we reason*. New York, NY: Oxford University Press.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *PNAS: Proceedings of the National Academy of Sciences, USA*. Advance online publication. Retrieved from http://www.pnas.org/content/early/2010/10/11/1012933107.abstract

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16,* 1–61. doi:10.1016/0010-0277(84)90035-0

Johnson-Laird, P. N., & Byrne, R. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin, 138,* 427–457. doi:10.1037/a0026841

Knauff, M. (1999). The cognitive adequacy of Allen's interval calculus for qualitative spatial representation and reasoning. *Spatial Cognition and Computation, 1,* 261–290. doi:10.1023/A:1010097601575

Knauff, M. (2007). How our brains reason logically. *Topio, 26,* 19–36. doi:10.1007/s11245-006-9002-8

Knauff, M. (2009). A neuro-cognitive theory of deductive relational reasoning with mental models and visual images. *Spatial Cognition and Computation, 9,* 109–137. doi:10.1080/13875860902887605

Knauff, M. (2013). *Space to reason*. Cambridge, MA: MIT Press.

Knauff, M., Bucher, L., Krumnack, A., & Nejasmic, J. (2013). Spatial belief revision. *Journal of Cognitive Psychology, 25,* 147–156. doi:10.1080/20445911.2102.751910

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity. *Journal of Cognitive Neuroscience, 15,* 559–573. doi:10.1162/089892903321662949

Knauff, M., & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition, 30,* 363–371. doi:10.3758/BF03194937

Knauff, M., & May, E. (2006). Mental imagery, reasoning, and blindness. *Quarterly Journal of Experimental Psychology, 59,* 161–177. doi:10.1080/17470210500149992

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,* 203–212. doi:10.1016/S0926-6410(01)00116-1

Knauff, M., & Ragni, M. (2011). Cross-cultural preferences in spatial reasoning. *Journal of Cognition and Culture, 11,* 1–21. doi:10.1163/156853711X568662

Knauff, M., Rauh, R., & Schlieder, C. (1995). Preferred mental models in qualitative spatial reasoning: A cognitive assessment of Allen's calculus. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 200–205). Mahwah, NJ: Erlbaum.

Knauff, M., Rauh, R., Schlieder, C., & Strube, G. (1998a). Continuity effect and figural bias in spatial relational inference. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 573–578). Mahwah, NJ: Erlbaum.

Knauff, M., Rauh, R., Schlieder, C., & Strube, G. (1998b). Mental models in spatial reasoning. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition: An interdisciplinary approach to representation and processing of spatial knowledge* (pp. 267–291). Berlin, Germany: Springer.

Knowlton, B. J., & Holyoak, K. J. (2009). Prefrontal Substrate of Human Relational Reasoning. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences,* pp. 1005–1017. Cambridge, MA: MIT Press.

Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2006). Mental imagery and the human brain. In Q. Jing, M. R. Rosenzweig, G. d'Ydewalle, H. Zhang, H.-C. Chen, & K. Zhang (Eds.), Progress in psychological science around the world*1. Neural, cognitive, and developmental issues* (pp. 195–209). Hove, England: Psychology Press.

Krumnack, A., Bucher, L., Nejasmic, J., Nebel, B., & Knauff, M. (2011). A model for relational reasoning as verbal reasoning. *Cognitive Systems Research, 12,* 377–392. doi:10.1016/j.cogsys.2010.11.001

Kuß, T. (1998). *Der Einfluß der domänenspezifischen Interessenlage auf die Konstruktion präferierter mentaler Modelle beim räumlich-relationalen Schließen* [The influence of social perspective on the construction of preferred mental models in spatial relational inference] (Unpublished diploma thesis). Albert-Ludwigs-Universität Freiburg, Psychologisches Institut, Freiburg, Germany.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16,* 217–265. doi:10.1017/S0140525X00029733

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511618031

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, England: Erlbaum.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10,* 181–187. doi:10.3758/BF03209220

Miller, G., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, England: Cambridge University Press.

Montague, R. (1974). *Formal philosophy: Selected papers*. New Haven, CT: Yale University Press.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 312–321). Berlin, Germany: Springer-Verlag.

Neth, H., & Johnson-Laird, P. N. (1999). The search for counterexamples in human reasoning. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society* (p. 806). Mahwah, NJ: Erlbaum.

Oberauer, K., Weidenfeld, A., & Hörnig, R. (2006). Working memory capacity and the construction of spatial mental models in comprehension and deductive reasoning. *Quarterly Journal of Experimental Psychology, 59,* 426–447. doi:10.1080/17470210500151717

Oberauer, K., & Wilhelm, O. (2000). Effects of directionality in deductive reasoning: I. The comprehension of single relational premises. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1702–1712. doi:10.1037/0278-7393.26.6.1702

Papadias, D., & Sellis, T. (1992). Spatial reasoning using symbolic arrays. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *International Conference GIS: From space to territory—Theories and methods of spatio-temporal reasoning in geographic space* (pp. 153–161). Berlin, Germany: Springer.

Papadias, D., & Sellis, T. (1994). Qualitative representation of spatial knowledge in two-dimensional space. *VLDB Journal, 3,* 479–516. doi:10.1007/BF01231605

Papadimitriou, C. H. (1994). *Computational complexity*. Reading, MA: Addison-Wesley.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review, 102,* 533–566.

Potts, G. R., & Scholz, K. W. (1975). The internal representation of a three-term series problem. *Journal of Verbal Learning and Verbal Behavior, 14,* 439–452. doi:10.1016/S0022-5371(75)80023-5

Prado, J., Chadha, A., & Booth, J. R. (2011). The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. *Journal of Cognitive Neuroscience, 23,* 3483–3497. doi:10.1162/jocn_a_00063

Ragni, M., Fangmeier, T., Webber, L., & Knauff, M. (2006). Complexity in spatial reasoning. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1986–1991). Mahwah, NJ: Erlbaum.

Ragni, M., Fangmeier, T., Webber, L., & Knauff, M. (2007). Preferred mental models: How and why they are so important in human reasoning with spatial relations. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial Cognition V: Reasoning, action, interaction* (pp. 175–190). Berlin, Germany: Springer. doi:10.1007/978-3-540-75666-8_11

Ragni, M., & Knauff, M. (2008). Deductive spatial reasoning: A computational and cognitive perspective. *Künstliche Intelligenz, 1,* 13–17.

Ragni, M., Knauff, M., & Nebel, B. (2005). A computational model for spatial reasoning with mental models. In B. Bara, B. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1064–1070). Mahwah, NJ: Erlbaum.

Ragni, M., & Wölfl, S. (2005). Temporalizing spatial calculi: On generalized neighborhood graphs. In U. Furbach (Ed.), *KI 2005: Advances in artificial intelligence—Proceedings of the 28th Annual German Conference on AI* (pp. 64–78). Berlin, Germany: Springer.

Rauh, R. (2000). Strategies of constructing preferred mental models in spatial relational inference. In W. Schaeken, G. De, Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 177–190). Mahwah, NJ: Erlbaum.

Rauh, R., Hagen, C., Knauff, M., Kuß, T., Schlieder, C., & Strube, G. (2005). Preferred and alternative mental models in spatial reasoning. *Spatial Cognition and Computation, 5,* 239–269.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Roberts, M. J. (1993). Human reasoning: Deduction rules or mental models, or both? *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 46,* 569–589. doi:10.1080/14640749308401028

Roberts, M. J. (2000). Strategies in relational inference. *Thinking and Reasoning, 6,* 1–26.

Rueckl, J. G., Cave, K. R., & Kosslyn, S. M. (1989). Why are "what" and "where" processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience, 1,* 171–186.

Ruff, C. C., Knauff, M., Fangmeier, T., & Spreer, J. (2003). Reasoning and working memory: Common and distinct neuronal processes. *Neuropsychologia, 41,* 1241–1253. doi:10.1016/S0028-3932(03)00016-2

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Schaeken, W., Girotto, V., & Johnson-Laird, P. N. (1998). The effect of irrelevant premise on temporal and spatial reasoning. *Kognitionswissenchaft, 7,* 27–32. doi:10.1007/s001970050047

Schaeken, W., & Johnson-Laird, P. N. (2000). Strategies in temporal reasoning. *Thinking & Reasoning, 6,* 193–219. doi:10.1080/13546780050114500

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition, 60,* 205–234. doi:10.1016/0010-0277(96)00708-1

Schaeken, W., Van der Henst, J.-B., & Schroyens, W. (2007). The mental models theory of relational reasoning: Premises' relevance, conclusions' phrasing, and cognitive economy. In W. Schaeken, A. Vandierendonck, W. Schroyens, & G. d'Ydewalle (Eds.), *The mental models theory of reasoning: Refinements and extensions* (pp. 129–149). Mahwah, NJ: Erlbaum.

Schlieder, C., & Berendt, B. (1998). Mental model construction in spatial reasoning: A comparison of two computational theories. In U. Schmid & J. F. Krems (Eds.), *Mind modelling: A cognitive science approach to reasoning, learning and discovery* (pp. 133–162). Lengerich, Germany: Pabst.

Sternberg, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109,* 119–159. doi:10.1037/0096-3445.109.2.119

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10,* 422–437. doi:10.1016/0010-0285(78)90006-3

Talmy, L. (1983). How language structures space. In H. L. J. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and applications* (pp. 225–282). New York, NY: Plenum Press. doi:10.1007/978-1-4615-9325-6_11

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

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology, 13,* 407–433. doi:10.1016/0010-0285(81)90016-5

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology, 15,* 121–149. doi:10.1016/0010-0285(83)90006-3

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23,* 515–557. doi:10.1016/0010-0285(91)90005-9

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–587). Cambridge, MA: MIT Press.

Van der Henst, J.-B. (2002). Mental model theory versus the inference rule approach in relational reasoning. *Thinking & Reasoning, 8,* 193–203. doi:10.1080/13546780244000024

Van der Henst, J.-B., & Schaeken, W. (2005). The wording of conclusions in relational reasoning. *Cognition, 97,* 1–22. doi:10.1016/j.cognition.2004.06.008

Vandierendonck, A. (1996). Evidence for mental-model-based reasoning: A comparison of reasoning with time and space concepts. *Thinking & Reasoning, 2,* 249–272. doi:10.1080/135467896394438

Vandierendonck, A., DeVooght, G., Desimpelaere, C., & Dierckx, V. (1999). Model construction and elaboration in spatial linear syllogisms.

In W. Schaeken, G. DeVooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 191–207). Mahwah, NJ: Erlbaum.

Vandierendonck, A., Dierckx, V., & DeVooght, G. (2004). Mental model construction in linear reasoning: Evidence for the construction of initial annotated models. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 57*A, 1369–1391.

Vorwerg, C., & Rickheit, G. (1998). Typicality effects in the categorization of spatial relations. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition* (pp. 203–222). Berlin, Germany: Springer. doi:10.1007/3-540-69342-4_10

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., Santos, M. de M., . . . Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10,* 119–125. doi:10.1111/1467-9280.00118

Zlatev, J. (2007). Spatial semantics. In D. Geeraerts & H. Cuyckens, (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 318–350). New York, NY: Oxford University Press.

Received November 7, 2011
Revision received February 15, 2013
Accepted February 16, 2013 ■