Kognitionswissenschaft (1998) 7: 33–38

Kognitions-
wissenschaft

Abstraction and Individual Style in Syllogistic and Heterogeneous Reasoning

Jon Oberlander

Human Communication Research Centre, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LW, UK
(e-mail: J.Oberlander@ed.ac.uk)

Abstraktionen und individuelle Stile beim syllogistischen und heterogenen Schließen

Zusammenfassung. Syllogistisches Schließen kann mit Hilfe mentaler Modelle modelliert werden. Heterogenes Schließen kombiniert propositionale und visuelle Information; in dem Logik-Lernsystem Hyperproof dienen bildliche Darstellungen als Modell (e) für propositionale Informationen. Gewisse Abstraktionen sind sowohl in der Notation der Theorie der mentalen Modelle als auch bei den von Hyperproof dargebotenen bildlichen Darstellungen möglich. Interindividuelle Unterschiede beim Denken wurden in beiden Bereichen gefunden. Wir schlagen daher vor, daß zumindest einige dieser Unterschiede auf die unterschiedlichen Fähigkeiten der Lernenden zurückzuführen sind, Abstraktionen zu bilden und zu manipulieren.

Summary. Syllogistic reasoning can be modelled with mental models. Heterogeneous reasoning brings together propositional and visual information, and in the Hyperproof logic-teaching system, the visual can act as a model for the propositional. Both mental models notation and Hyperproof's visual representations allow abstractions of a limited kind. Individual differences in reasoning performance have been found in both domains. We suggest that at least some of the differences can be attributed to students' differing ability to manipulate abstractions.

1 Introduction

The mental models approach to syllogistic reasoning is well established (Johnson-Laird 1983; Johnson-Laird & Byrne 1991). It proposes that subjects construct representations of finite situations and reason by checking and elaborating those representations. The Hyperproof logic-teaching system is a more recent development (Barwise & Etchemendy 1994). In it, students learn how to construct proofs in first-order logic by bringing together sentential and visual information. However, just as with mental models, this type of heterogeneous reasoning system allows subjects to construct representations of finite situations and to reason by checking and elaborating those representations. Here, the graphical acts as a model for the sentential.

Recent studies of subjects learning logic with Hyperproof have indicated the importance of individual differences in cognitive style (Stenning, Cox & Oberlander 1995; Oberlander et al. 1996). We maintain that these differences arise in part from subjects' varying ability to manipulate relatively abstract representations. Meanwhile, some existing studies of syllogistic reasoning have indicated that individual differences have an important impact not just on patterns of error, but also on the styles of subjects' proofs. We therefore suggest that abstraction ability may be related to at least some of the differences found in the domain of syllogistic reasoning.

2 The Importance of being abstract

The observation that graphical systems require certain classes of information to be specified goes back at least to Bishop Berkeley. A triangle drawn on a sheet of paper has to have a specific set of angles and line lengths and cannot, without further conventional interpretation, be a "generic" triangle. Stenning & Oberlander (1995) termed this property "specificity" and argued that it is useful because inference with specific representations can be very simple. They conceded, however, that actual graphical systems do allow abstractions to be expressed, and it is this that endows them with a usable level of expressive power. For example, they argued that a Euler circles notation could be used to model syllogistic reasoning so long as appropriate abstraction devices were incorporated into it. It follows that a key step in mastering any practical representational system – graphical or otherwise – is to learn which abstractions can be expressed and how.

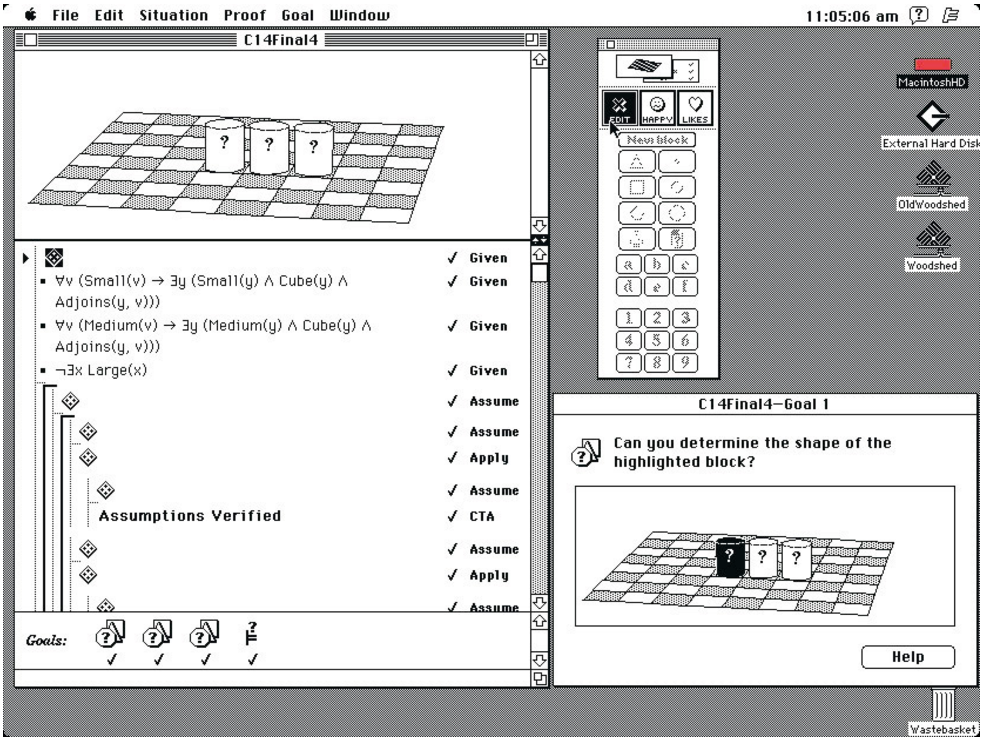

Fig. 1. The Hyperproof interface. The main window (*top left*) is divided into an upper graphical pane and a lower calculus pane. The tool palette is floating next to the main window, and other windows can pop up to reveal the set of goals which have been posed

2.1 Abstraction in mental models

It is sometimes claimed that mental models are unlike other representational systems (such as Euler circles, Venn diagrams and semantic networks) in that they are in some way "isomorphic" to the states of affairs they represent (see Johnson-Laird & Bara 1984, p. 24). If this were strictly true, then the notation would not sustain abstraction devices of the type we are concerned with. However, from Gurr (1998), it seems that mental models notation cannot be strictly isomorphic and that in satisfying looser conditions – such as homomorphism – it is not qualitatively different from the other representational systems. Moreover, inspection of the proposed notations for mental models reveals mechanisms designed for effective abstraction. In Johnson-Laird & Byrne (1991), for instance, the premise *All Xs are Ys* is represented as in (1):

(1) [X] Y
    [X] Y
    . . .

Unlike its predecessors (such as Johnson-Laird & Bara 1984), this notation provides an explicit representation of which types are fixed or necessary (via the brackets). In addition, the potential mutability of a given model is given prominent status, via the use of the ellipsis. The latter innovation is particularly important, since it emphasises the extent to which reasoning starts out from an *implicit* model and, by a process of "fleshing out", adds extra types of individual to it, gradually making it more *explicit*. Implicit models are relatively abstract (since they can be extended in various ways), whereas explicit models are relatively concrete (since they are compatible with fewer states of affairs). Reasoning reliably will require an ability to understand and manipulate abstract models, as well as the concrete ones which correspond to particular finite states of affairs.

2.2 Abstraction in hyperproof

Let us now turn to abstraction in Hyperproof. The system and its interface are described in detail elsewhere (see Barwise & Etchemendy 1994; Stenning, Cox & Oberlander 1995). Briefly, the system should be viewed as a proof-checking environment designed to support human theorem proving using heterogeneous information. As can be seen in Fig. 1, the interface contains two main window panes. One presents a diagrammatic view of a chess-board world containing geometric objects of various shapes and sizes. The other presents a list of sentences in predicate calculus, and control palettes are also available. These window panes are used in the construction and editing of proofs. Several types of goals can be proved, involving the shape, size, location, identity or sentential descriptions of objects; in each case, the goal can involve determining some property of an object or showing that a property *cannot* be determined from the given information. A number of rules are available for proof construction. Whereas some of these are traditional syntactic rules (such as &-elimination), others are "graphical", in the sense that they involve consulting or altering the situation visually depicted in the diagrammatic window. In addition, a number of rules check properties of a developing proof.

To construct a proof, students must often start out by either transferring information from the sentential premises (using the rule **Apply**) or by adding directly to the graphical situation (using **Assume**). They can move information from the graphical world back to the sentential component by using either **Observe** or **Inspect**. They can declare when all sentential and graphical assumptions are true in the current situation by using **CTA** (Check Truth of Assumptions), and equally, they can declare that a sentence is inconsistent with either another sentence or the current graphical situation by using **Close**. They can also declare that a set of graphical cases exhausts the possibilities via **Exhaust**.

Like mental models notation, Hyperproof supports abstraction. In the graphical world, concrete objects have one of three shapes and three sizes. However, the system also maintains a set of abstraction conventions for objects' spatial or visual attributes. As well as concrete depictions of objects, there are "graphical abstraction symbols", which leave attributes under-specified: a cylinder, for instance, depicts an object of unknown size. To use Hyperproof properly, students must master the use of these abstraction devices, as we indicate below.

2.3 Inference and abstraction

The process of adding information to an initially abstract graphical situation is illustrated in Fig. 2. This sequence of graphical situations occurs in an actual proof submitted by a student in answer to the question posed in Fig. 1. In this particular instance, each graphical situation is successively more concrete – and hence less abstract – than its predecessor. For questions such as these, students are pursuing a strategy remarkably similar to the process of "fleshing out" described in Johnson-Laird and Bryne's mental models theory.

To see that this is so, consider the sequence of mental models in Fig. 2. Building an initial implicit model corresponds to the use of **Apply** to build an abstract graphical world. Adding the second premise involves elaboration, via another **Apply**, of what is still an abstract graphic. The use of brackets is similar to the use of **Exhaust** declarations. The extraction of provisional conclusions corresponds to an **Inspect**, and attempted falsification, proceeding by making the model more explicit, is similar to an **Assume CTA** sequence. Reading off a weakened conclusion initiates a new cycle, as does **Inspect**. Hence it seems that, in this case, the representations and processes of mental models work by gradually making an implicit model explicit, while Hyperproof works by supporting proof by cases. Thus each allows us to prove conclusions by constructing a sequence of increasingly concrete cases and reading off the common conclusion.

However, both Hyperproof and mental models algorithms support episodes of reasoning rather more complex than those shown in Fig. 2. A step of a Hyperproof may involve a retreat to a more abstract situation, just as a step in mental models reasoning may plausibly make a model less explicit. Now, Johnson-Laird and Byrne divide reasoning into three stages: comprehension, description and validation. Following Rauh & Schlieder (1997), we may term these stages *model construction*, *model inspection* and *model variation* and apply them to Hyperproof reasoning. The current point can then be stated as follows: in principle, it is possible for the phase of model construction to issue in an abstract model. Thus what Knauff, Rauh & Schlieder (1995) term the *preferred* model would be abstract. The phase of model variation, which involves iterated cycles of further construction and inspection, will therefore lead to at least some cycles in which the concreteness of the current model is increased.

It remains an empirical matter as to whether – and for whom – the initially constructed model is actually

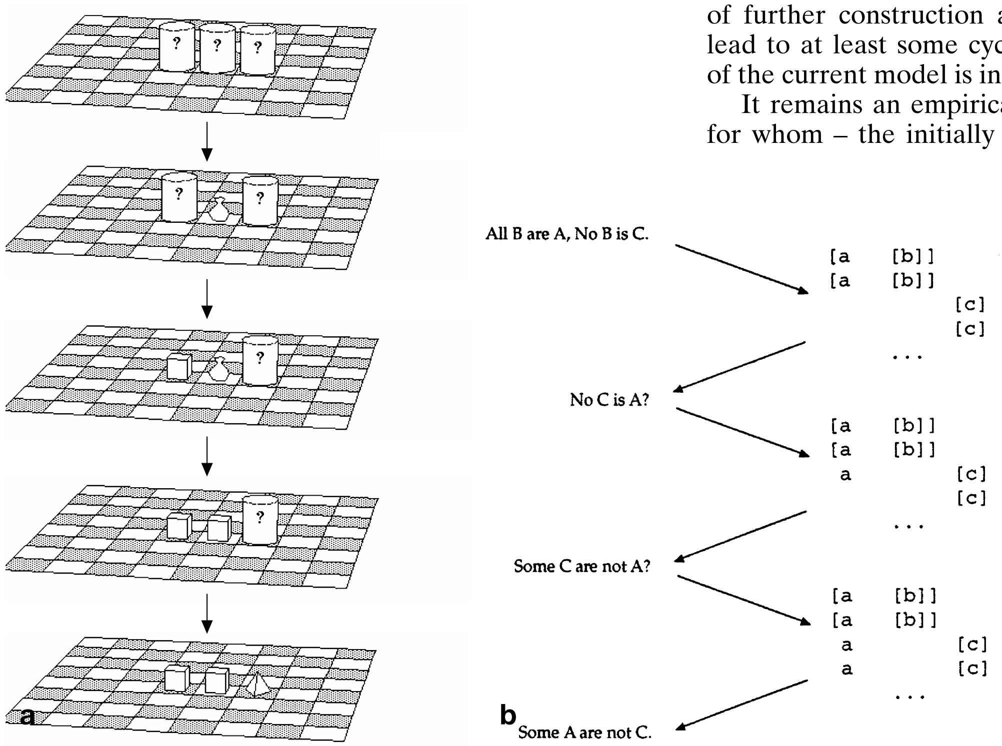

Fig. 2. Inference and concreteness in Hyperproof (*left*) and mental models (*right*). Here, each Hyperproof step builds a new, more concrete situation by applying sentential information or adding assumptions to the previous graphical situation. The sequence of mental models is from Johnson-Laird and Byrne (1991; p. 122). Each step adds a new entity type, making the initial implicit model increasingly explicit

abstract and implicit, rather than concrete and explicit. However, if there are subjects for whom the initial model is abstract, then competence requires the ability to pick and subsequently vary the "right" initial model. For them, the difficulty will reside in the amount of concreteness that must be added to whatever model is initially selected. It can therefore be predicted that performance using either Hyperproof or mental models will be influenced by subjects' differing abilities to deal with the abstractions allowed in the notations. Our hypothesis is that this abstraction ability will influence two things: firstly, patterns of error, a relatively familiar parameter in mental models research, and secondly, styles of inference, which have been less extensively studied. We now turn to evidence for style differences in Hyperproof.

3 Proof styles in hyperproof

We carried out a controlled field study of students learning first-order predicate calculus with or without Hyperproof. Details of the full study are given in Stenning, Cox & Oberlander (1995). For current purposes, what matters is that we can distinguish two subgroups of students according to their behaviour on a pre-test of their analytical reasoning ability, based on the Graduate Record Exam (GRE). The test contained two types of item, which we term *determinate* and *indeterminate*. The former are categorised as those items often best solved by constructing an external representation of some kind. Typically, they can be solved by constructing a single model of the premises and reasoning about that model. The latter are those items which require argument analysis and for which the construction of external representations is not usually recommended. Typically, they require the entertainment of multiple models of the premises for their solution. Scores on the determinate items in the GRE pre-test were used to perform a median split, dividing subjects into two subgroups. Det Hi subjects were those who performed well on determinate GRE items – those often best solved using an external representation. DetLo subjects were those who performed less well. Here we discuss the performance of two subgroups in the study: DetHi and DetLo subjects who followed a course in Hyperproof.

These subjects submitted solutions to computer-based exam problems. There were two questions in which the graphical pane contained fully concrete information, and two questions in which the graphical pane contained a considerable amount of abstractness. We call these question types *determinate* and *indeterminate*, respectively. It should be noted, of course, that, although we use the same terms again, this categorisation is not directly equivalent to the categorisation of GRE item types. We logged both the final proofs and the time course of the process by which they were constructed.

Substantial systematic differences in proof style arise between DetHi and DetLo subgroups. Some differences seem obvious to casual inspection, e.g. when one subject used many more steps than another for the same proof. At the group level, however, we found that significant differences were most pronounced on the indeterminate questions, which contained a high degree of initial graphical abstraction.

We generated representations of the concreteness of each of the graphical situations entertained in each proof and then graphed these to generate "proofograms", which chart the way in which concreteness varies through the course of a proof. The visual differences between proofograms proved to be striking and robust: one group is "spikey", and the other is "layered" (Oberlander et al. 1996). The visual grouping of proofograms suggests the existence of a *staging* phenomenon: DetHi's layered proofs introduce concreteness by stages, whereas DetLo introduce it more immediately. In terms of proof structure, DetHi tend to produce structured sets of cases, with superordinate cases involving graphical abstraction, whereas DetLo tend to produce sets of cases without such overt superordinate structure.

To probe this further, we considered the likelihood of one Hyperproof rule being used after another, performing an analysis of the corpus of final proofs following Dunning's (1993) Log-Likelihood method. This finds those pairs of elements (or "bigrams") which are reliably and significantly associated in a corpus of sequentially arranged tokens. It is usually applied to collections of natural language utterances, but is also suitable for the artificial proof constructs discussed here. Because of the proofogram results, the bigram analysis distinguishes uses of **Assume** which introduce full concreteness into a graphic (which we now term **Fullassume**) from those which leave the situation abstract (which we term **assume**). The corpus analysis again showed that differences between the subgroups were greatest on indeterminate questions. Firstly, DetHi subjects had seven bigrams significantly associated which DetLo did not, and DetLo had just two bigrams significantly associated which DetHi did not. Secondly, of the seven, five exploit **assume**. These are associated with proof construction moves which make a situation more concrete, but still leave it somewhat abstract. They are precisely the moves to be expected from reasoners exhibiting staging behaviour. What is striking – even seemingly paradoxical, in view of their original GRE performance – is that only the DetHi reasoners stage their concreteness. The others apparently move directly to fully concrete situations.

4 Proof styles in the syllogism

Thus subject subgroups using Hyperproof differ in their proof styles, and these differences, we argue, emerge when subjects have to manipulate graphical abstractness through rules such as assume. The question is now whether the staging phenomenon carries over to individual differences in syllogism performance. Obviously, subjects make errors on mentally performed syllogisms, and some subjects make more errors than others. However, the possibility that subgroups of subjects might exhibit distinctive patterns of proof style has remained relatively unexplored. Nonetheless, at least three sets of observations seem relevant.

4.1 Previous work on individual style differences

First, Sternberg & Weil (1980) examined the abilities and behaviour of subjects solving linear syllogisms. Four different strategies to solve this class of problems had been proposed: spatial, linguistic, mixed spatial/linguistic and algorithmic. Subjects were characterised in terms of their spatial and verbal abilities on the basis of their performance in two spatial and two verbal aptitude tests. They were instructed to solve linear syllogisms using one of three methods: through reading and answering, through visualisation and through the use of a simple algorithmic procedure. It was found that the effectiveness of a given strategy depends on the subjects' pattern of abilities: the mixed strategy draws on both verbal and spatial abilities; the linguistic strategy primarily on verbal abilities; the spatial strategy on spatial abilities; and the algorithmic strategy on both abilities, although to a lesser extent than the mixed strategy.

Secondly, Galotti, Baron & Sabini (1986) investigated whether reasoners differed systematically in the number of alternate models they constructed. Their "good" reasoners were the top third of the sample, with fewest errors, while the "poor" reasoners were the bottom third, with most errors. The groups differed most on items with no valid conclusions, and this is consistent with the fact that failure to generate a sufficient number of models would lead to error. Galotti et al. (1986, p. 23) suggest that poor reasoners "may be less self-critical in their search for alternative models, that is, less prone to seek models that are inconsistent with their initial conclusion."

Finally, Ford (1995) examined think-aloud protocols for a group of 20 subjects presented with the 27 syllogisms with valid conclusions and permitted the use of external representations. She found that eight subjects made extensive use of set diagrams, eight spoke in terms of algebraic substitutions, two showed evidence of mixed strategies, and two exhibited neither type of behaviour. Ford proposed two alternative algorithms, one couched in essentially spatial terms, and the other in terms of algebraic substitutions. Ford then argued that, as might be expected, subjects who use set diagrams exhibit patterns of error consistent with use of the spatial algorithm, while those who talk in terms of substitutions exhibit patterns consistent with the algebraic algorithm.

4.2 Predictions concerning abstraction and staging

We have argued that abstraction devices are important in both Hyperproof and in the mental models notation taken to account for much syllogistic data. In both areas, evidence has been found for systematic individual differences. On the one hand, our Hyperproof evidence suggests that subject groups' proof strategies hinge on their ability to build cases via available abstraction devices and that this ability is determined by their prior cognitive style. On the other hand, the existing evidence from studies of syllogistic reasoning – particularly Sternberg & Weil's and also, to a lesser extent, Ford's – suggests that subjects' visual and verbal abilities influence their effectiveness in pursuing different inference strategies and that this is therefore likely to influence which strategy they tend to adopt.

What remains to be seen, however, is whether the staging phenomenon carries over to differences in syllogistic reasoning. There are two main ways to bridge the gap. Firstly, we could try to relate our DetHi/DetLo dimension to psychometric measures of visual and spatial ability. The former is taken to be related to ability to solve one-model or diagrammable problems. However, it is not clear how this relates to spatial ability. The evidence to date suggests that the link is by no means direct and that other dimensions of variation (such as serialism versus holism) may be just as relevant. Work in progress is now pursuing this strategy: we are replicating the study of the Hyperproof group, with psychometric pre-tests added to the battery.

The second option is more speculative but, if successful, would underline the importance of abstraction ability. The rationale is as follows: DetHi students are those who successfully exploit staging, building models of increasing concreteness. DetLo are doing something else, a kind of purely concrete inference. Suppose these styles carry over directly to syllogistic reasoning. In that case, we should be able to find two distinctive groups of reasoner. One group will build relatively abstract models during the initial model construction phases, as sketched in Sect. 2.3. The other will prefer to construct fully concrete models immediately. Concomitantly, during the model variation phase, both groups will tend to vary their models by one or two attributes at a time. However, the first group will both increase and decrease their model's explicitness during reasoning episodes, while the second group will exhibit less variance overall in the levels of explicitness of their models.

5 Conclusions

Our own studies do not yet provide sufficient evidence to decide for or against abstraction-based differences in syllogistic reasoning. Similarly, neither Sternberg & Weil nor Ford will provide such evidence directly, since they focus only on whether or not spatial abilities are being used. Nonetheless, Galotti et al.'s remark is suggestive: recall that poor reasoners "may be less self-critical in their search for alternative models". It is therefore plausible to suggest that looking at individuals' variation processes in more detail may reveal distinctive proof styles. However, the most obvious next step would be to look for evidence that there is systematic variation in the relative explicitness of initial models selected by subjects and, if there is, to determine how it relates it to scores on GRE-type items.

In the meantime, however, we may agree with Roberts (1993) that individual differences raise real difficulties for any account of human reasoning that attempts to provide a single, fundamental mechanism and somehow distinguish it from "superficial" strategies that can be added on top of it. Thus the field is bound to be invigorated and enriched by meeting the challenge of individual differences.

This paper is based on joint work with Keith Stenning and Richard Cox. Special thanks go to John Etchemendy, Dave Barker-Plummer, Mark Greaves and Padraic Monaghan.

References

Barwise, J. & Etchemendy, J. (1994) *Hyperproof*. CSLI Lecture Notes. Chicago: Chicago University Press.

Dunning, T. (1993) Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* **19,** 61–74.

Ford, M. (1995) Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition* **54**, 1–71.

Galotti, K. M., Baron, J. & Sabini, J. P. (1986) Individual differences in syllogistic reasoning: deduction rules or mental models? *J Exp Psychol: General* **115**, 16–25.

Gurr, C. A. (1998) On the isomorphism, or lack of it, of representations. In: Marriot, K. & Meyer, B. (eds.) *Theory of Visual Languages* (pp. 289–301) Berlin: Springer.

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

Johnson-Laird, P. N. & Bara, B. (1984) Syllogistic inference. *Cognition* **16**, 1–61.

Johnson-Laird, P. N. & Byrne, R. M. J. (1991) *Deduction*. Hove and London: Lawrence Erlbaum Associates.

Knauff, M., Rauh, R. & Schlieder, C. (1995) Preferred mental models in qualitative spatial reasoning: a cognitive assessment of Allen's calculus. In *Proceedings of the 17th Annual Meeting of the Cognitive Science Society* (pp. 200–205) Pittsburg, PA, July 1995.

Oberlander, J., Cox, R., Monaghan, P., Stenning, K., & Tobin, R. (1996) Individual differences in proof structures following multimodal logic teaching. In *Proceedings of the 18th Annual Meeting of the Cognitive Science Society* (pp. 201–206) La Jolla, CA, July 1996.

Rauh, R. & Schlieder, C. (1997) Symmetries of model construction in spatial relational inference. In *Proceedings of the 19th Annual Meeting of the Cognitive Science Society* (pp. 638–643) Stanford, CA, August 1997.

Roberts, M. (1993) Human reasoning: deduction rules or mental models, or both? *Q J Exp Psychol* **46**, 569–589.

Stenning, K., Cox, R. & Oberlander, J. (1995) Contrasting the cognitive effects of graphical and sentential logic teaching: reasoning, representation and individual differences. *Language and Cognitive Processes* **10**, 333–354.

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

Sternberg, R. J. & Weil, E. M. (1980) An aptitude x strategy interaction in linear syllogistic reasoning. *J Edu Psychol* **72**, 226–239.