Int. J. Human–Computer Studies (1996) **45**, 185–213

External cognition: how do graphical representations work?

MIKE SCAIFE AND YVONNE ROGERS

School of Cognitive and Computing Sciences, University of Sussex, Brighton BN1 9QH, UK. email:mikesc/yvonner@cogs.susx.ac.uk

(Received 9 March 1995, and accepted in revised form 11 January 1996)

Advances in graphical technology have now made it possible for us to interact with information in innovative ways, most notably by exploring multimedia environments and by manipulating three-dimensional virtual worlds. Many benefits have been claimed for this new kind of interactivity, a general assumption being that learning and cognitive processing are facilitated. We point out, however, that little is known about the cognitive value of *any* graphical representations, be they good old-fashioned (e.g. diagrams) or more advanced (e.g. animations, multimedia, virtual reality). In our paper, we critique the disparate literature on graphical representations, focusing on four representative studies. Our analysis reveals a fragmented and poorly understood account of how graphical representations work, exposing a number of assumptions and fallacies. As an alternative we propose a new agenda for graphical representation research. This builds on the nascent theoretical approach within cognitive science that analyses the role played by external representations in relation to internal mental ones. We outline some of the central properties of this relationship that are necessary for the processing of graphical representations. Finally, we consider how this analysis can inform the selection and design of both traditional and advanced forms of graphical technology.

The Speaker, Betty Boothroyd, rebuked an M.P. for using a cardboard diagram in the Commons to explain overseas aid figures. She said "I have always believed that all Members of this House should be sufficiently articulate to express what they want to say without diagrams." (Guardian, 7 December 1994)

1. Introduction

Virtual reality and visualization,† as means of representing and interacting with information, are very much at the forefront of technological development. An overriding intuition is that much can be gained computationally from interacting with virtual reality simulations or visualizing from three-dimensional dynamic images (Fairchild, Serra, Hern, Hai & Leong, 1993). Many benefits for industrial and educational applications have been claimed, such as powerful visualization tools for designers, architects and chemists (e.g. Rheingold, 1991; Earnshaw & Watson, 1993). However this is merely the latest in a long line of assumptions about graphical technological advancements, each claiming better ways of facilitating cognitive tasks. These include the ideas that:

† The term visualization is defined as "mechanisms by which humans perceive, interpret, use and communicate visual information" (McCormick, DeFanti & Brown 1987).

1071-5819/96/080185 + 29$18.00/0

- static pictures and diagrams are better than sentential representations†
- three-dimensional representations are better than two-dimensional ones
- solid modelling is better than wire-frame modelling
- colour is better than black and white images
- animated diagrams are more effective than static images
- interactive graphics are better than non-interactive graphics
- virtual reality is better than animation.

Such generalizations about the benefits of advanced graphical technologies over good old fashioned representations, however, beg the question of what is actually gained cognitively from having more explicit, dynamic and interactive representations of information. Why, for example, should an animated diagram that changes in response to user interaction be more effective at facilitating problem-solving than static diagrams? Why not the other way round, where static diagrams are more effective than animations or non-interactive graphics are better than interactive graphics and so on? Given this uncertainty, how can researchers and designers decide whether to take on board the immense cost and effort to develop a virutal reality appliction, for example, when a static diagram might be more effective for the task in hand?

The value of different graphical representations,‡ be they good old-fashioned or technologically-advanced, cannot be assessed adequately from our intuitions. To be effective a number of interdependent factors need to be considered, such as the level of experience with the graphical representation, the knowledge domain and the type of task. Whilst there have been numerous empirical studies investigating different aspects of graphical representations there has been little attempt to integrate the findings into an analytic framework. What is needed, therefore, is a more systematic approach for evaluating the merits of different kinds of graphical representations, one that is theoretically-driven and which accounts for the cognitive processing when people interact with them. Without such an approach we have no principled way of either making sense of the vast empirical literature on the benefits of graphical representations or of making predictions about the value of new forms, such as animation and virtual reality.

The current state of understanding is not encouraging. Most of the theoretically-based research on the role of external representations in cognitive science has been concerned with how we learn to read, write and understand written text. For graphical representations there is an obvious imbalance in terms of the work that has been done. On the one hand their value in helping to understand tasks/concepts presented verbally is well-documented but on the other there is no evidence of a detailed theory that explains this, e.g. "The literature is overflowing with work investigating the facilitative effects of pictures on text comprehension. And yet, no one has a clear idea of the cognitive processes underlying these effects" (Glenberg & Langston, 1992: p. 129). In fact the quote applies equally to any learning or problem-solving situation utilizing graphical representations whether text-comprehension is significantly involved or not.

† Epitomized by the widely-used proverb, "a picture is worth a thousand words".

‡ Graphical representations include diagrams, maps, plans, animations and virtual reality and are distinct from propositional/sentential representations and formal notation (cf. Larkin & Simon, 1987).

Part of the problem may stem from the large variation in graphical representational forms, associated with a correspondingly wide range of functions. Past research spans a wide area from map design to technical illustration to the value of pictures for children learning science, with a mélange of methodologies, explanatory frameworks and mechanisms. Recent reviews are consistent in pointing out the lack of integration in the field. The problems here are severe for any attempt to provide an overall picture. For example, as Molitor, Ballstaedt and Mandl (1989) point out, a large number of studies have been concerned with the manipulation of task variables within highly-specific situations, reporting mainly on the success or failure of graphical representations to affect performance. As Winn (1993) notes, it is even difficult to make (practical) generalizations within this "kind" of study, precisely because of their idiosyncrasies. In addition different authors have frequently ploughed their own furrow and have been highly selective in assimilating what other researchers have done. What has largely been absent, therefore, has been any attempt to explain how these experimental effects are produced psychologically, frequently ignoring recent work in cognitive science. Molitor *et al.* (1989: p. 27) comment that much of the empirical work on graphical representation has been "... usually formulated into *ad hoc* questions, and (is) not grounded in a cognitive processing theory".

Why is there a lack of a suitable, explicit processing model? One reason may be that the form of graphical representation does not lend itself to systematic computational analyses. The theoretical frameworks and formal notations that have been developed for analysing verbal language are not applicable to the syntactically- and semantically-dense properties of graphical representations (Goodman, 1968). Another reason, as we argue later, may be that there seems to be pervasive (and possibly unwarranted) assumption that graphical representations must work in a certain way because of their figural nature. Thus many studies are almost "black-box" in their approach to psychological mechanisms. Some, however, have attempted to look systematically at the effective perceptual features of graphical representations. For example Winn (1993) analysed diagrams in terms of a model of visual search, focusing on strategies for extracting information. His analysis identifies the importance of external features such as the spatial distribution and discriminability of elements of the diagram. He also points to important cognitive processes such as knowledge of content and symbol conventions in the reading process. Winn points to a lack of graphical representation-specific research on search strategies but we would emphasize equally the paucity of work on determining how graphical representations are themselves represented and how this interacts with the kinds of high-level cognitive processes, such as applying knowledge of content, that Winn rightly emphasizes.

We argue that an alternative approach is needed to understanding graphical representations: we need to ask what is the nature of the relationship between graphical representations and internal representations and to consider how graphical representations are used when learning, solving problems and making inferences. Such an enterprise means working towards a detailed description of cognitive mechanisms. In this respect we would point to the kind of account offered by theorists such as Larkin and Simon (1987) and Koedinger and Anderson (1990). These offer a model of diagram use which explicitly presents the main elements of a theoretical account, viz. a statement of (i) the properties of the graphical representation; (ii) the way that these properties may be represented internally and (iii) computational processes that mediate between the two. These models thus demonstrate the kind of systematic approach to graphical representation we are advocating. However, we would also argue that focusing primarily on internal representations, as these models do, is not enough—it misses out much of the cognitive processing that goes on when interacting with graphical representations and, hence, a useful account of their value. These models, thus do not match a second desideratum: an account which analyses more fully the interplay between internal and external when carrying out a cognitive task.

1.1. INTERNAL AND EXTERNAL REPRESENTATIONS

Within cognitive science, in general, there has been a move towards promoting the need to analyse the interaction between internal and external representations. In a special issue on situated action in the journal of Cognitive Science, Vera and Simon (1993) stress that, "A fundamental problem for cognitive modellers is to interleave internal and external states in order to achieve naturalistic behaviour" (p. 12). Norman (1988, 1993) has for several years been describing cognition in terms of "knowledge in the head" and "knowledge in the world". Larkin (1989) has also shifted her thinking from Larkin and Simon's (1987) earlier computational model of diagram use—that focused primarily on internal representations—to considering the role played by external displays in cognitive problem-solving. Others, like Cox and Brna (in press) have been examining specifically the cognitive effects of external representations in reasoning tasks. External representations, here, may refer to both linguistic and graphical forms.

What we see emerging from this trend—to broaden and situate the base from which to explain cognitive behaviour—is external representations being given a more central functional role in relation to internal cognitive mechanisms. This is a substantial step away from traditional cognitive modelling and, significantly, an important theoretical advancement, that potentially allows us to account more adequately for how graphical representations work. Thus instead of trying to adapt internally-based processing models of cognition we can begin to specify characteristics of the internal/external relationship in the cognitive processing of graphical representations. The value of this is to focus our attention more on the *cognitive processing* involved when interacting with graphical representations, the *properties of the internal and external structures* and the *cognitive benefits* of different graphical representations. In addition to enabling us to develop more appropriate cognitive models, we believe that his new perspective—which we have coined *external cognition*—allows us to begin to assess more effectively how technological innovation in graphical representations should be approached.

In our examination of the emerging literature on internal/external representations we have abstracted three central characteristics which we consider as a useful analytic framework from which to explain aspects of external cognition. These are *computational offloading, representation* and *graphical constraining.*

(i) *Computational offloading.* This refers to the extent to which differential external representations reduce the amount of cognitive effort required to solve informationally equivalent problems. For example, Larkin and Simon (1987) point to the greater efficiency in geometry problem-solving for diagrams over sentential forms through their ability to provide direct perceptual recognition of geometric relations. Explicitly representing the problem state in diagrams in this way enables solutions to be more readily "read-off". In contrast, solutions for the same problems represented as sentential descriptions typically are implicit and so have to be mentally formulated. This requires a greater computational effort.

(ii) *Re-representation.* This refers to how different external representations, that have the same abstract structure, make problem-solving easier or more difficult. For example, Zhang and Norman (1994) describe carrying out the same multiplication task using roman or arabic numerals. Both represent the same formal structure, but the former is much harder for people, used to working with the decimal system, to manipulate to reach the solution (e.g. LXVII $\times$ X is much more difficult to solve than 68×10).

(iii) *Graphical constraining.* This refers to the way graphical elements in a graphical representation are able to constrain the kinds of inferences that can be made about the underlying represented world. This characterization is a term developed in recent work on the value of diagrams for solving formal logic problems by Stenning and colleagues (e.g. Stenning & Oberlander, 1995; Stenning & Tobin, in press). A central idea is that the relations between graphical elements in a graphical representation are able to map onto the relations between the features of a problem space in such a way that they restrict (or enforce) the kinds of interpretations that can be made. The closer the coupling between the elements in the visual display and the represented world, the more tractable the inferencing. Although the characterizations might appear to overlap, we see them more as complementary; computational offloading highlights the cognitive benefits of graphical representations, re-representation relates to their structural properties and graphical constraining to possible processing mechanisms. In the paper we show how these characterizations can be employed to begin to assess how graphical representations work.

1.2. AIMS OF THE PAPER

A main aim of this paper is to take stock and examine why our theoretical understanding of graphical representations is so impoverished, especially given the vast body of empirical literature. Following this, we consider how we might begin to redress the situation through analysing interactivity, building on the above internal/external characterizations for different cognitive activities and graphical representations.

In attempting to do so, we wish to clarify three theoretical issues: (i) the erroneous equivalence of external and internal structure, particularly where it justifies an implicit processing model; (ii) to identify existing, explicit processing models which are of the right type although they do not go far enough in the direction we want and (iii) to clarify the concept of external cognition as referring to the totality of the relationship between external representation, internal representation and their interaction (processing).

We focus our analysis on two main classes of graphical representations which span traditional and developing technologies: (i) static diagrams and other illustrations that have a role as adjuncts to text (or oral language) and (ii) animations. We then extend the lessons of the analysis to (iii) the area of virtual reality. Thus the purposes of our analysis are both to outline the relevant theoretical questions that need to be considered in understanding graphical representation applications and to suggest how they might apply to the design of innovative graphical technologies.

Before we move on to our analysis, however, we need to address potential sources of misunderstanding by considering the referential scope of key terms in our discussion.

2. A note on representation, static diagrams, animation and virtual reality

The term "representation" has a variety of different meanings, depending on the context. A common distinction is between representation as process, and representation as product, as the outcome of this process. Process concerns the transformations and preservations that occur in deriving the representation from what is being represented. Description of product is typically concerned with structural characterizations of the representation, for example as image-like, mental model or propositional. Confusion might arise since the two senses, process and product, may be used interchangeably. In fact the two cannot always be easily separated, since characterization of structural properties is usually related to a particular processing model. Here we shall discuss representation in both senses.

The classes of static diagrams and animations are considered distinctive, in so far as they have been identified as having different characteristics in the literature. It is acknowledged, however, that there is likely to be some overlap between what constitutes a diagram and an animation, expecially for displays that are comprised of both static and animated components. There are many different exemplars of diagrams and no single accepted taxonomy that can be conveniently employed to describe them, although there is good evidence emerging for stable classification strategies (Lohse, Walker, Biolsi & Rueter, 1991; Cox & Brna, 1993). There may, in fact, not be a single, critical feature for the term "diagram". Some authors seek to draw a distinction between representations like graphs, describing quantatitive data in two dimensions, and other, less-constrained types. We would probably subscribe to that view but it is not crucial for us here. We would rather adopt a position similar to that of Winn (1987: p. 153) who treats diagrams as representations with the function of being "... to describe whole processes and structures, often at levels of great complexity".

Animations are equally difficult to define and, again, there is, as yet, no single theoretically- or even empirically-grounded classification scheme available. Animations—be they computer, film, video or other media-based—differ from static diagrams in presenting a series of rapidly changing static displays, giving the illusion of temporal and spatial movement. This can be achieved through a range of techniques. For example, in "multi-dimensional" animation, interdependent objects appear to move in relation to each other; in "partial" animation certain parts of a display move whilst the rest of the display remains static; in "artificial" animation, implicit movement is made explicit or processes normally invisible to the eye are made visible. While not an exhaustive classification, we can see the diversity of animation "types".

The third class of graphical representations that we examine is virtual reality or virtual environments. These are computer-generated graphical simulations, intended to create "the illusion of participation in a synthetic environment rather than external observation of such an environment" (Gigante, 1993: p. 3). Images are displayed stereoscopically to the user, via a head-mounted display. Objects within this field of vision can be interacted with via a dataglove or other input device, use of a virtual reality headset can change the field of vision in the virtual world and users can "fly" around the virtual world through gesturing. A major motivation for virtual reality systems is to enable people to become "immersed in the experience" of interacting with external representations (Kalawsky, 1993). However this is difficult to operationalize (Sheridan, 1992) and there is no taxonomy of types of virtual reality immersion. Most virtual reality classifications are based on the types of graphical techniques used for rendering three-dimensional objects and in terms of applications that may benefit from being represented in virtual reality (Kalawsky, 1993).

3. Empirical work on graphical representations involving diagrams and animations

Having set out some desiderata for a study of graphical representations we will now try to make our ideas more concrete. Rather than attempt a global review, we shall concentrate on a small number of influential studies to draw out some general issues pertinent to our aim of assessing the pros and cons of different kinds of display. We shall examine two studies that concern the use of static diagrams and two that have investigated animated displays. These have been chosen as examples which have clear aims to show how graphical representations might work and what processes are involved. We shall adopt the format of first describing the findings and then offering a critique before making some general comments on the theoretical issues that surround graphical representation research.

3.1. RESEARCH ON STATIC DIAGRAMS

Work on static diagrams represents a considerable corpus of research from which it is hard to make generalisations. Winn (1987), reviewing the field, notes that there is an interaction between (at least) ability level, diagram format and task type to be considered in drawing conclusions across studies. We shall consider here two studies that have looked at the value of diagrams for problem-solving: Larkin and Simon's (1987) study of physics and geometry problems and Bauer and Johnson-Laird's (1993) study of logic problems.

Larkin and Simon (1987) analysed examples taken from classic physics (pulleys and weights) and geometry (theorem proving) textbooks. Their aim was to develop computational models that allowed a contrast between processing of "sentential" and "diagrammatic" representations which contained the same information about the problem. In the first case elements appear in a single sequence, while in the second they are indexed by their location in two-space. Their theoretical anslysis suggests that a diagram "preserves explicitly the information about the topological and geometric relations among the components of the problem, while the sentential representation does not" (p. 66).

The approach taken by Larkin and Simon provides an explicit formalism. The elements of their system are (i) data structures that represent the problem to be solved (ii) productions that contain knowledge of the laws of the domain (the "program") and (iii) an attention manager. They propose that a diagrammatic data structure may differ markedly from an informationally-equivalent sentential one through affording the possibilities of easier search. For example, similar attributes may be clustered at the same spatial location making it easier to recognize them.

Diagrams, therefore, provide simultaneous information about the location of components in a form that enables objects and their relations to be easily tracked and maintained. This greatly reduces the need to search and recognize. In contrast, sentential representations of the same problem can not provide the same external memory cues. Solving certain kinds of problems using sentential representations thus incurs a much greater computational load, particularly to keep track of how the solution is progressing. Moreover, much of the necessary information to derive a solution is not available from the sentential descriptions and so has to be formulated explicitly. Hence, far more has to be computed mentally to determine possible states and their consequences, than when carrying out the same tasks using problem diagrams such as the one shown in Figure 1.

In describing their model, Larkin and Simon (1987) noted that there was one respect in which diagrams could not be supposed to have an intrinsic advantage—inference-making. They observed that when data structures are informationally equivalent "Inference is largely independent of representation . . " (p. 71). Bauer and Johnson-Laird (1993), however, postulated that for certain kinds of problems, diagrams should help reasoning, a claim based on Johnson-Laird's (1983) model theory of deductive reasoning. They investigated the role of external representations, in the form of schematic diagrams, on the solving of deductive reasoning tasks. The problems were double-disjunctive reasoning, which require reasoners to keep

1. Two transversals intersect two parallel lines and intersect with each other at a point X between the two parallel lines.

2. One of the transversals bisects the segment of the other that is between the two parallel lines.

3. Prove that the two triangles formed by the transversals are congruent.

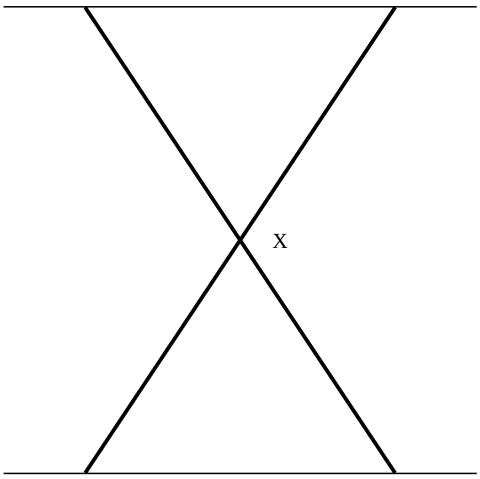


FIGURE 1. Verbal and diagramatic representation of a geometry problem (after Larkin & Simon, 1987).

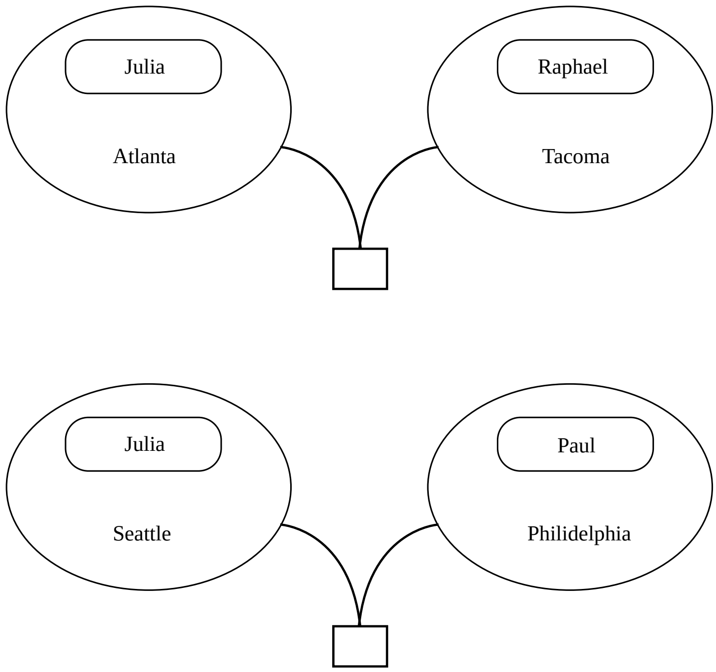
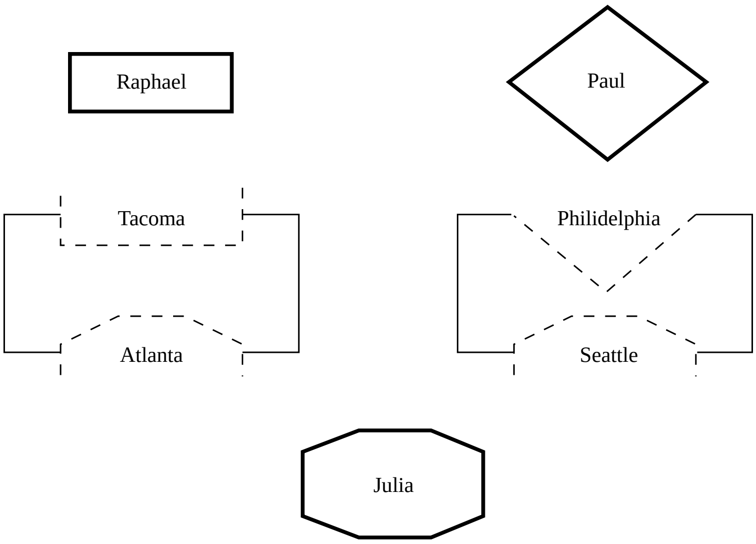


FIGURE 2. Diagram showing arbitrary and abstract icons used in Bauer and Johnson-Laird (1993) study on graphical reasoning with double subjunctives.

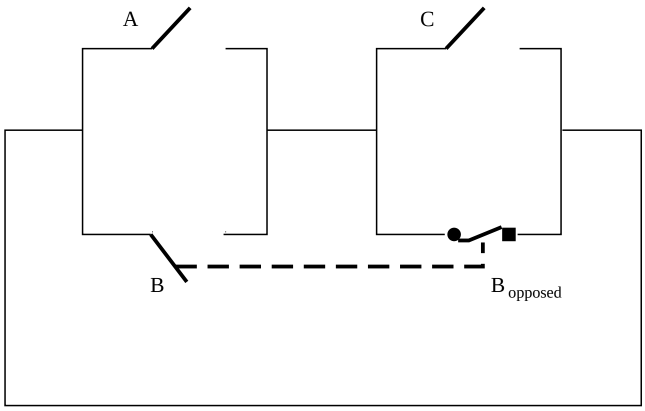
track of various alternative states in order to solve them. Because of the difficulty of taking into account many models of the premises, subjects are known to perform poorly on these types of problems. Bauer and Johnson-Laird hypothesized that providing diagrams should enable reasoners to keep track of alternative models.

Bauer and Johnson-Laird's (1993) first reported attempt at developing a schematic diagram to make explicit the alternative possibilities was largely unsuccessful. This they attributed to their using arbitrary and abstract icons for representing explicitly the alternatives, which were found to be of no help to the reasoner (see Figure 2). Their second attempt, however, was more successful. Two types of more concrete diagrams were constructed: one based on an electrical circuit and the other a jigsaw. In both examples, a particular problem-solving context was provided from which to make the deductions. The instructions for the circuit representation of the problem was couched in terms of switches and lights being on or off in the circuit whilst the instructions for the jigsaw representation were expressed in terms of completing a path from one side of the figure to the other. This involved inserting shapes, corresponding to specific people specified in the reasoning problem, into slots in the path, corresponding to particular places [see Figures 3(a) & (b)]. In both examples, therefore, the subjects were required to solve the reasoning task by mentally transforming parts of the diagram. In doing so, the solvers no longer need to solve the problems entirely in their head but can work them out by interacting with the diagrams.

Indeed, the results showed that performance was significantly better and faster when using the diagrammatic representations than when solving the same problems using sentential representations. The findings seem to provide further support for the important role of diagrams as external memories, enabling a picture of the whole



The event is occuring.
What follows?



The light is on.
What follows?

FIGURE 3. (a) Diagram involving insertion of shapes into "jigsaw" slots used in Bauer and Johnson-Laird (1993) study on graphical reasoning with double disjunctives. (b) Diagram involving circuits used in Bauer and Johnson-Laird (1993) study.

problem to be maintained simultaneously, whilst allowing the solver to work through the interconnected parts (cf. Larkin & Simon, 1987; Larkin, 1989; Zhang & Norman, 1994). Although it could be argued that a sentential representation can also act as an external memory aid, the extent of the "computational offloading" is considerably less. This suggests that having the problem states and its solution more explicitly represented in the diagram than in the sentential representation means that less inferencing is required.

Both of the above studies show the potential value of graphical representation for aiding problem solution in terms of search, recognition and inference. However we need to ask the related questions: how much light do they shed on the role of the external representation and how does this mesh with details of internal representations (cognitive mechanisms)? Consider first the Larkin and Simon (1987) account. Their principal concern is with diagrammatic internal representations, which as Parkes (1993: p. 37) points out provide "... access to the properties of the pictures which are posited to facilitate more efficient computations ...". Larkin and Simon (1987: p. 66) describe such representations as having the property of corresponding "... on a one-to-one basis, to the components of a diagram describing the problem". However, their account leaves open the question of (i) how this is produced in human beings, and (ii) what work is being done by the external and the internal representations respectively. Consider the following quote (p. 92): "We have seen that formally producing perceptual elements does most of the work of solving the geometry problem. But we have a mechanism—the eye and the diagram—that produces exactly these perceptual results with little effort. We believe that the right assumption is that diagrams and the human visual system provide, at essentially zero cost, all of the inferences we have called 'perceptual'". It is hard to get a precise understanding of how "perceptual inferences" might work but Larkin and Simon do offer a suggestive analysis of the metaphorical sense of (mentally) "seeing" as referring to inferences that are akin to perceptual experiences in being based on "productions with great computational efficiency" (p. 71). This conceptual issue notwithstanding there is a further problem. As Koedinger and Anderson (1990: p. 518) note, the Larkin and Simon argument "... is based on an assumption that perceptual inferences are generally easier than symbolic inferences" but "... it is possible that perceptual inferences appear easier because, in general, they have been much more highly practised than symbolic inferences". In other words the value of diagrams in such situations is strongly related to experience and expertise of the individual having "operators" that match the display (cf. Larkin, 1989). Novice physicists, for example, will not make the same inferences as experts from the same diagram (cf. Anzai, 1991).

The Bauer and Johnson-Laird (1993) explanation of what subjects are doing when using diagrams to solve the logic problems is also debatable. It seems unlikely that they are using the diagrams to compare alternative models, as postulated by Johnson-Laird's model theory. Alternatively, it appears that the external representations (the combination of the diagram and analogical instructions) have re-represented the original deduction problems into simpler and different tasks. In their own words, Bauer and Johnson-Laird (1993) explain: "in the case of the diagrammatic problems, the subjects form a visual representation of the diagram, and in their mind's eye they can imagine moving the pieces or switches (i.e. they carry out visual transformations of images). Bypassing the construction of the meanings of verbal premises and manipulating visual images appears to reduce the load on working memory and to speed up the process of inference." (p. 373).

Therefore, rather than facilitating the reasoner to envisage all the alternative possibilities inherent in the premises it seems more likely that the combination of instruction and diagram is constraining a particular way of conceptualizing and solving the logic problem. This account fits in better with Stenning and Oberlander's

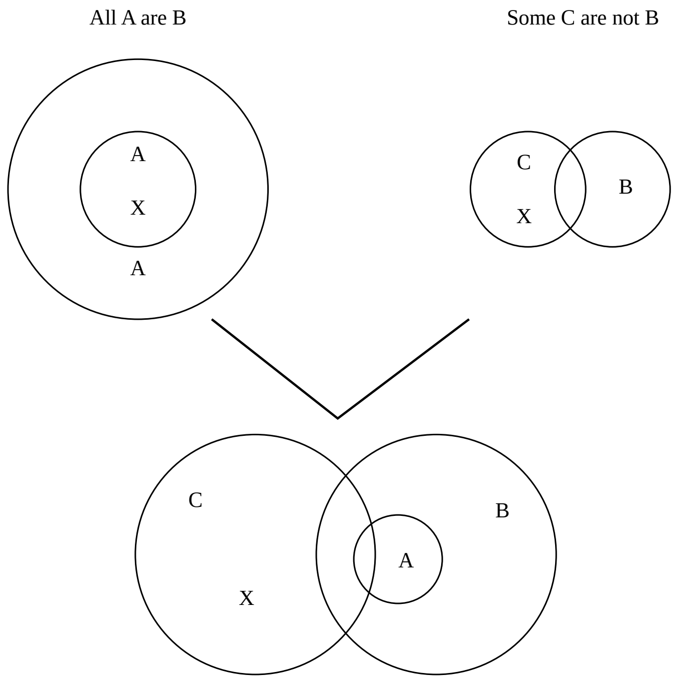


FIGURE 4. An example of the use of Euler's circles to constrain the solution for the syllogism: *All A are B. Some C are not B* (after Stenning & Tobin, in press).

(1995) theory of specificity which differs from the model theory and by postulating that "graphical representations such as diagrams limit abstraction and thereby aid processibility" (p. 2). This is achieved through the information available in the diagram restricting the possible interpretations of the problem and in so doing guiding the reasoner to make the correct solution. Thus, certain diagrams are more effective than others because they exploit better the constraining properties of varying graphical forms. For example, Stenning and Tobin (in press) claim that Euler's Circles (see Figure 4) are more effective than three-dimensional cube diagrams in helping subjects solve logic problems because the geometrical constraints of the intersecting circles represent the logical constraints much better. In other words, a diagram is more likely to afford a particular reading of the problem and way of solving it than a sentential representation because it is less expressive (i.e. decreases indeterminacy). Having built a mental model of the combined external representations (the instructions and the diagram) that satisfies the premise in their own minds, it is unlikely that the solvers will then build alternative, but equally plausible, models of the problem (Cox & Brna, 1995).

3.2. RESEARCH ON ANIMATIONS

One of the strengths of studies, such as Larkin and Simon (1987), is the postulation of an explicit model of cognitive processing. However, in reviewing research into the role of animations in learning and problem-solving contexts we failed to find any similarly detailed models. Thus, in our attempt to consider in more detail how animation is processed we decided to critique two empirical studies that sought to investigate the mechanisms by which animations are effective in making inferences from graphical representations of physical systems. These are Hegarty (1992), which focuses on mental animation and Kaiser *et al.* (1992), which focuses on external animation.

In Hegarty's study, the graphical representations used were static canonical diagrams of pulley systems whilst in Kaiser, Proffitt, Whelan and Hecht's (1992) study, both static and animated canonical graphical representations were used to depict objects falling, being severed, or being displaced from various dynamical systems (e.g. pendulums and moving planes). the primary aim of Hegarty's study was to ascertain the extent and form of mental animation that occurs when making judgements about the motion of pulley systems. In contrast, the main objective of Kaiser *et al.*'s (1992) study was to determine how external animations enable more effective judgements to be made about the trajectories of moving objects compared with static diagrams. In both studies, subjects were required to reason through initially comprehending a verbal problem together with a static or dynamic graphical representation used to convey the problem state, and then predict correctly future states or trajectories of part of the system depicted in the graphical representation.

Hegarty's central idea is mental animation, which involves simulating mentally, in a serial manner, components in the graphical representation of the pulley system. An obvious reason for this is that we are unable to animate all parts of the diagram at once, due to the constraints of working memory. It also seems plausible, given that we can only perceive the working of certain aspects of a real world pulley system at any one time, depending what is in our field of view at that time. With real-workd pulleys, however, the motion of each part is always available; we need only to follow the way the components move to make judgments about them. Moreover, we can do this in a haphazard way. With diagrams, however, Hegarty argues we make inferences about the motion of the static parts by following the temporal order of the causal chain of events from input to output.

This level of theorizing seems intuitively plausible for explaining how people reason with relatively simple pulley system problems and is to some extent supported by her empirical findings. For more complex systems, Hegarty suggests that other mental strategies are likely to be used. However, the form that these alternative forms might take, how they develop and whether they are used in combination with mental animation or separately is beyond the scope of her paper. Likewise, the actual functional role of the graphical representation is not discussed in her theory of incremental animation, although she does acknowledge that it needs to be researched further.

In contrast to Hegarty's approach, Kaiser *et al.* (1992) explain reasoning about mechanical systems in terms of what the external representation does for the learner. Like Hegarty, they stress the importance of information being processed sequentially, but in terms of the external representation being able to "temporally parse a multi-dimensional problem into unidimensional components" (p. 671). In doing so, they propose that the distinct state changes that have to be recognized to make correct judgements about the system are made more obvious through an animation than with a static display. The idea that the external representation does the "temporal parsing", rather than the problem-solver having to do it, is illustrated with an example of common-sense reasoning about the C-shaped tube physics problem (based on McCloskey, Caramazza & Green, 1980). The main finding is that when the problem is represented as a static two-dimensional representation (see

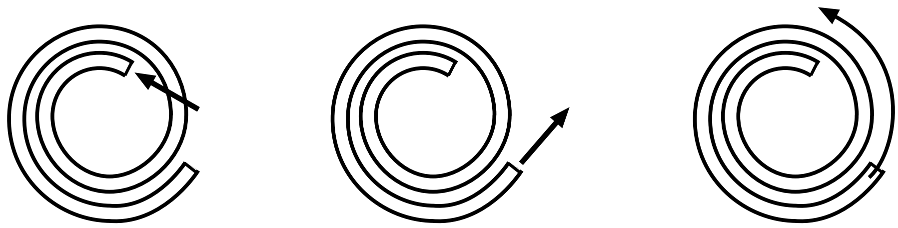


FIGURE 5. Diagrams for the spiral tube problem with correct response A and incorrect response B (after McCloskey *et al.,* 1980).

Figure 5), students often incorrectly infer that the projected motion of the ball on exiting the curved tube continues in a curvilinear trajectory. Kaiser *et al.* (1992) found the same effect for both free choice and forced choice conditions. However, when shown various incorrect and the correct animation sequences in a forced choice situation, students invariably selected the correct "straight" trajectory. Kaiser *et al.* explain this performance shift in terms of the animation temporally segregating the ball's behaviour in the extended body system (whilst in the tube) from when it behaved as a point particle (after exiting the tube). It is this visible temporal parsing that is claimed to make the change in states more obvious to the subjects and allow them to make the correct inference.

The animation studies present a different form of the problems encountered in the static diagram research. In particular they assume that running mental models and parsing external animations use the same structures and functional processes as when perceiving real-world dynamic systems. For example, in an earlier study Hegarty, Just and Morrison (1988) proposed that people decided which attributes of a system were relevant to judging mechanical advantage on the basis of "causal models" of mechanical systems arising from relevant (physical) experience with such systems. The appeal to an equivalence in processing, however, does not help us in understanding the merits of different forms of graphical representation in terms of how they are processed and interacted with for various tasks. A more pertinent question to address, then, is how do we understand and make connections between the static and animated forms that represent the dynamic processes of real systems such that we can make inferences about them both?

The same lack of specificity is seen in Kaiser *et al.*'s (1992) explanation of the superiority of animated over static forms in terms of the visible temporal parsing of the ball in the container system, making change in states more obvious. An alternative explanation for the difference in performance could be in terms of experience with the two representational formats, reflecting more a difficulty with interpreting the canonical forms in the diagram in relation to the problem that had to be solved rather than one of not being able to recognize the significance of the temporal parsing of the objects in different states. This objection reiterates our concern that there is a crucial role for expertise and practice that is not being recognized. Most importantly, whilst providing further support for the value of the explicitness inherent in animations, Kaiser *et al.*'s (1992) study offers no explanation of the cognitive mechanisms involved in learning and reasoning with animations. Pedagogically, too, it is unclear how animation can facilitate learning or problem-solving. In particular, we would emphasize the absence of an analysis of how a better level of understanding can result from seeing objects moving explicitly as opposed to having to imagine how they move. Indeed Kaiser *et al.* (1992) comment on how subjects who had been shown the animation first and then a static diagram of the same problem performed no better than those who had just been shown the static diagram. Here again, we have further evidence that the benefit of viewing an animation is transitory and not readily mapped onto the static representation with its more arbitrary conventions.

3.3. CONCLUSIONS FROM THE STUDIES

In all of these four studies, then, we are left with questions about the mechanisms by which diagrams and animations are effective. How do viewers identify the "key" features and constraints of a graphical representation and then map them onto the relevant aspects of the problem to be solved? Further, there seems to be a real issue about what problem the subjects were "really" solving and how far expertise and experience are central factors. It is often hard to separate general claims about graphical representations *per se* from factors that have to do with individual differences in ability in the subject or understanding of the domain-specific genre of the diagrams involved. In domains with highly evolved notations such as geometry or physics, diagrams are not "merely" aids for solution but play an essential part in the process of knowledge acquisition and depiction. A circuit diagram, an architectural plan or a mathematical notation comprise a set of meaningless symbols to the uninitiated; they only take on their intended meaning through learning the conventions associated with them. In a real sense it is impossible to develop expertise in these subjects without the ability to both read and produce diagrams of a particular sort (cf. Anzai, 1991). This strongly suggests that—in such domain-specific cases at least—diagrams can only trade on established domain knowledge to be effective. The point is well-put by Larkin and Simon (1987: p. 71): "If students lack productions for making physics inferences from diagrams, they . . . will find them largely useless".

4. Processing mechanisms

The studies we have reviewed are, of course, a tiny selection from a vast range. However they serve to illustrate two of our major themes, that there is a lack of an adequate cognitive processing model and that focusing on the externality of graphical representations to see how they work is of crucial importance to a better understanding. Below we examine further why the former is so and contrast this with a more detailed analysis of the few studies that have begun to analyse external/internal representations in cognitive processing.

4.1. PROCESSING AND THE RESEMBLANCE FALLACY?

One problem with the Kaiser *et al.* (1992) study noted above was the lack of explanation of how subjects recognize the temporal segregation of the objects as being central to an understanding of mechanical systems. This seems to be because the external and internal representations are assumed simply to have the same characteristics. This is an example of what we shall call the "resemblance fallacy", which has a much wider appearance in the graphical representation literature and may help to explain something of the apparent unwillingness to specify processing models. It is prevalent, we believe, because the structure of graphical representations, their spatial/iconic/figural qualities, promotes an *intuition* as to their value as an imput for perception/cognition whereas the reality is that we have no well-articulated theory as to how such an advantage might work. Evidence for our over-reliance on such intuitions can be seen by examining the kinds of arguments that have been made for the links between graphical representations, perception and internal representations. The possibility of different representational formats—image/proposition/mental model—and their properties interacts with the issue of how graphical representation might work.

Consider first, work on conventional, static graphical representations. By design these are typically well-suited to the information pick-up capacities of the visual system—object perception, search, pattern-matching, etc. Analysis of their syntactic properties in these terms, e.g. how elements are grouped, what affects discriminability etc., is therefore productive. However there is a problem in going beyond these data to identifying the form of internal representations that result. A particularly common argument has been that the quasi-pictorial qualities of images suggests a privileged link with pictorial input. For example Winn (1987: p. 159) summarizes the relevance of work on imagery to graphical representation thus: "These studies exemplify a body of research that leads to the following conclusion: Graphic forms encourage students to create mental images that, in turn, make it easier for them to learn certain types of material". And Reed (1993: p. 299) claims a "substantial similarity between the functional equivalence of pictures and images", stating that: "We would have a better understanding of how images aid problem solving if we had a better understanding of how pictures aid problem solving".

The problem with this line of argument is that it does seem to rest on intuition. What can "encourage" and "easier" mean in terms of mechanism? Further, while pictures can undoubtedly serve to stimulate imagery under certain circumstances (e.g. Finke, 1990) it is by no means clear that they are *necessarily* represented in this way. Halford (1993) points out that we do not have to accept any more than a mapping between relations for an external representation-internal representation pair. In addition there is some doubt about the extent to which imagery is computationally important and processing may be better explained in terms of other representational forms (e.g. Pylyshyn, 1973; Molitor *et al.,* 1989; Anderson, 1990). In short the case for an intimate relationship between graphical representation and images may not be logically compelling and is currently heavily under-specified.

The prevalence of image-based explanations for graphical representation is, presumably, based on the premise that images facilitate cognitive operations on analogic representations. This has benefits for basic processes like optimizing search and reducing the load on working memory (e.g. Larkin & Simon, 1987). However these same benefits have also been invoked for mental models and these, too, have been proposed as representations for graphical representation. Since the claims for images and mental models are close it is unsurprising to see some link proposed between the two, e.g. Hegarty (1992) for people working on pulley system problems: ". . . running a mental model involves transforming mental images . . . people infer motion of mechanical systems by transforming mental images . . .". Similarly Bauer and Johnson-Laird (1993) talk about the "manipulation of visual images". Thus, while the mental model resulting from interaction with graphical representations need not be thought of as image-like (cf. Glenberg & Kruley, 1992), there is often an apparently close relationship posited between the two and, hence, the possibility of another intuitive and unsubstantiated link between picture input and representation type. Molitor *et al.* (1989: p. 10) describe the situation thus: "In mental models reality is represented in an analogous, predominantly imaginative form". However the same authors also note that the mental model construct is "very fuzzy and used in a different sense by each author" (p. 10), a conclusion echoed by Hong and O'Neill (1992). As with images such a range of possibilities argues for caution in positing a necessary relationship between the pictorial nature of graphical representation and a particular representational format. At the very least, different graphical representations and/or tasks may engender different kinds of representations and this must remain an issue for future research.

The discussion so far of the resemblance fallacy has been exemplified by work done with static diagrams. However there is evidence that some of the same kinds of problem are occurring with work on animations. Here, as we noted previously, there is an assumption that "adding" animation to an equivalent static display will be advantageous. But why should this be so? In articles about animation we commonly find an intuition-led chain of assumptions, echoing the same causal chain of reasoning used to account for the efficacy of diagrams (external representation producing a mental image or mental model which in turn results in better learning or reasoning) that was criticised earlier. Here, an illustrative line of reasoning goes something like this: animations can show motion explicitly and "directly" and hence provide more accurate information (Kaiser *et al.,* 1992); this reduces processing demands on working memory allowing other tasks to be performed (Rieber & Kini, 1991) and enables more "useful" mental models to be formed for solving problems (Park & Gittleman, 1992); these in turn facilitate learning or reasoning. Whilst the first part may be factually accurate, the rest does not logically follow.

As argued above in relation to diagrams, we cannot simply assume a privileged relationship between a graphical representation of a system—in this case an animation—and someone's understanding or ability to reason about it, by virtue of its resemblance, albeit highly simplified and schematized, to the dynamic properties of a real-world system. As with diagrams used in specialized domains, e.g. physics or geometry, a person has to learn to "read" and comprehend the significance of the content of the animations in relation to other information that is being presented verbally or as text and to assimilate this to their current understanding of the domain. This requires making multiple connections between what the animations are intending to convey and the abstract concepts that are being learned about. How students integrate information arising from different representations of knowledge is crucial (Laurillard, 1993).

4.2. PROCESSING AND THE INTERNAL/EXTERNAL

The force of our comments, however, is not solely to do with being less intuitive in our accounts. Consider the claim by Larkin and Simon (1987: p. 97) that: "mental imagery—the uses of diagrams and other pictorial representations . . . held in human memory . . . play a role in problem solving quite analogous (to) external diagrams . . . and that *this role is also played by the internal and the external in* *concert.*" (our emphasis). How do the internal and the external act "in concert"? Clearly if we set up our processing model so that both representations have the same structure/characteristics the problem seems less apparent—and the solution becomes obvious too. But it seems clear that what is really needed is an integrated approach that analyses the dynamic cognitive processing of graphical representations; how this understanding is integrated with existing knowledge, what information gets lost, misinterpreted or correctly interpreted, and how is it subsequently *re-represented,* externally.

The importance of considering internal and external representations in concert is beginning, as mentioned previously, to be more of a central concern in cognitive science. Larkin (1989) has tried to tackle the problem by outlining a computational model called DiBS, that represents information available in external displays as data structures that enable internal operators to be cued as to know what to do next. The model's central searching mechanism is based on the observation that "each step requires only looking at the display, and doing what it suggests, without more effortful mental calculation or storage" (p. 319). DiBS therfore, works largely by manipulating attributes of the external display. The examples that Larkin has chosen to represent in her model are well suited to the transformation of external data structures. They include simple everyday problems (e.g. brewing coffee) and textbook problems (e.g. linear equations) that once learned become highly routinized and error-free. Hence, for these kinds of tasks there is no need to activate any internal representations other than a very general mechanism that is characterized as knowing "where an object wants to go" in each step of the task. As such, DiBs is more a model of how external representations cue a set of automatic procedural-based actions for solving clearly defined problems. In its current state, therefore, DiBs does not account for other kinds of learning and less well-defined problem-solving tasks, involving more complex cognitive processing, where combinations of interacting internal/external mechanisms are likely to come into play.

Inspired by Larkin's (1989) "move to the external", a number of empirical studies have been carried out since, in an attempt to analyse in more detail the properties of display-based problem-solving. For example, O'Malley and Draper (1992) differentiate between the knowledge users need to internalize when learning to use display-based word processors (e.g. MacWrite) with that which they can always depend upon being available in the external display. The tendency, therefore, appears to be for users to learn only what is necessary to enable them to find the information they require in the interface display. Information represented in such displays is viewed as an external memory aid "which 'fill the gaps' in users' internalised representations when they interact with the system" (p. 86).

How the gaps are filled and what cognitive mechanisms are involved, however, is not clear. This is also true of Zhang and Norman's (1994) recent analysis of distributed cognitive tasks (between internal and external representations), where they argue that an analysis of the relation between different external forms of representing the same abstract problem (in this case the Tower of Hanoi) is necessary before considering the processes that are activated when solving the problem. The study was designed so that in certain conditions subjects had to internalize several rules to carry out the task whilst in others the same rules were embedded in the external display. Their findings indicated that the fewer rules subjects had to internalize the easier it was for them to perform the problem-solving task. The implication is that external representations can significantly change the nature of a task through constraining the permissible moves allowed in solving the task. Furthermore, this form of "computational offloading", i.e. implicitly embedding rules in the external representation as opposed to making subjects internalize them, is thought to reduce the load on internal working memory providing more "space" for planning subsequent moves. Although Zhang and Norman (1994) did not investigate the processes involved in interacting with graphical representations in their experiment, they do suggest that the nature of the relationship is likely to be uni-directional: "perceptual processes are activated by external representations while cognitive processes are usually activated by internal representations" and that moreover "different processes are activated by different representations (p. 118). We would argue, however, that the interplay between internal and external representations in problem-solving is likely to be more complex, involving cyclical, interacting processes, especially when considering how graphical representations are both perceived and acted upon.

Before moving onto our final discussion of how we can consider the design of graphical representations based on an analysis of the relationship between internal/external representations we briefly introduce our third category of external representations, virtual reality. As with our analysis of the previous two categories of graphical representations, diagrams and animations, we consider the cognitive benefits, mechanisms and structures that are involved when interacting with virtual reality representations. We shall identify some of the same problems emerging here as for the diagrams and animation work.

5. Empirical work and research issues on graphical representations involving virtual reality

In our earlier introduction to virtual reality we noted that the central concept of "immersion" was an elusive one. One way of characterizing it is in terms of realism. It is often suggested that a main advantage of virtual reality is that simulations can be constructed to have a higher level of fidelity with the objects they represent, compared with other kinds of external representations, e.g. static diagrams, in terms of having more functional, physical and spatial resemblance. The illusion afforded by the virtual reality technology can make virtual objects appear to be more realistic and to behave according to physical laws. For example, terrains developed for flight simulators can seem very life-like, giving a spatial awareness that can closely approximate that which could be obtained in the real world. Virtual reality researchers are inclined to believe that learning and training applications, e.g. aviation and defence can be improved through having a greater fidelity wilth the represented world (e.g. Cover, Ezquerra, O'Brien, Rowe, Gadaez & Palm, 1993). One suggestion is that the transfer of training to the real world could be easier, with less errors.

The push towards developing life-like simulations, however, raises the question of what are the cognitive benefits of representing the world at higher levels of realism than at higher levels of abstraction. The intuition that perceiving and acting in an artificial environment that is designed to simulate a physical world as closely as possible, will provide a better learning or training experience through the development of better spatial cognition is reminiscent of the "resemblance fallacy" discussed in the previous sections. Similarly, it may prove to be as incorrect. For example, results from a recent study investigating transfer of training in virtual reality systems found that subjects learnt performance characteristics specific only to the virtual reality context, which were of no use when carrying out the same task in the real world (Kozak, Hancock, Arthur & Chrysler, 1993). These preliminary findings suggest, therefore, that the actual experience of being immersed in a virtual reality world is quite distinct from interacting with real world artifacts. The value of virtual reality, therefore, should not be assumed to come about through a structural and spatial equivalence between the virtual reality simulation and the real world.

One of the problems of the move "towards the virutal", therefore, is that learners may gain an inappropriate or artificial understanding of the world that is being modelled. Another problem is the degree of interactivity within the virtual environment. Currently there exists largely indirect means of interacting with the virtual graphical representations—typically via selecting objects through moving customized joysticks and trackballs, gesturing through data gloves or more crudely through using a keyboard. This contrasts sharply with the high level of interactivity that exists with six degree full motion flight simulators. Not only do they have a real set of flight controls but they also have highly realistic feedback—all of which is well integrated with the animated graphical representations of the physical terrains that are flown over.

Instead of considering virtual reality immersion in terms of the value gained from attaining higher levels of perceptual fidelity with the real world it would be more useful to reconceptualize it as a question of how best to constrain virtual reality simulations to provide external representations that are effective for training (cf. to the earlier idea about diagrams and graphical representations). In particular, the virtual environments need to be designed to guide the learner to the crucial aspects that are necessary for performing the appropriate activity for a given task at a given time. The issue then becomes one of determining what aspects of the represented world need to be included and how they should be represented, what aspects should be omitted and what additional information needs to be represented that is not visible in the real world but would facilitate learning. From a cognitive perspective, it enables us to assess the benefits of virtual reality in terms of the processing mechanisms that operate at differing levels of abstraction of information. For example, we can analyse differences in task demands and performance characteristics for specific tasks, e.g. taking off or landing for different virtual reality simulations, ranging from presenting simple canonical structures (e.g. schematic outlines) to more fully rendered depictions of scenes. Hopefully, this way the pitfalls of the resemblance fallacy can be avoided.

Another way in which the notion of virtual reality immersion has been characterized is in terms of "steering" the interaction. Here, the intuition is that virtual reality simulations provide more opportunities to visualize and manipulate the behaviour of abstract data structures or processes which are not normally visible to the naked eye. For example, NASA have developed a Virtual Wind Tunnel, whereby a scientist (who is a computational fluid dynamicist) controls the computation of virtual smoke streams by using the finger tips (Gigante, 1993). Abstract equations for the computed airflow around a digital model of an aircraft are translated into visible smoke streams. By moving around the virtual aircraft in the virtual reality environment and visualizing the smoke streams, the fluid specialist is thought to be able to "discover areas of instability, separation of the flow from the aircraft's surface, and other interesting phenomena" (Gigante, 1993: p. 9). An assumption is that better mental models of the abstract processes will develop through making these kinds of processes more concrete (Gigante, 1993). It is also assumed that on-line problem-solving will be facilitated. Whilst not making any explicit links between the resulting mental structures and cognitive processes there is a sense that the two are connected. What we do not get a sense of, however, is how experts, who are highly familiar with abstract representations and have to interact with them in their work, are able to transfer between these forms of representation and the concretized visual representation of the same problem space in the virtual reality simulation. The value of being able to "steer" a physical simulation should be analysed, therefore, in relation to how it integrates with ways of interacting with other existing forms of external representations in professional practice.

6. Informing the design and selection of graphical representations

At the beginning of this paper we asked the important practical question of how designers could determine which kind of external representation to choose from—be it text, diagram, multi-media or virtual reality—for the domain or task they are designing for. As has become clear through the paper, however, to answer this question depends on having a better understanding of internal representation/external representation interactivity. In addition, it requires addressing specific issues, such as what form the display should take, what information should be made explicit, how this should be represented, how this maps onto the object/concept being represented and which graphical style to use. Our analysis of the graphical representation literature across a variety of disciplines, however, led us to the conclusion that despite a plethora of empirical studies on how different graphical representations affect performance and a few theoretical analyses, the findings are difficult to generalize beyond the specific features investigated in each study. Moreover, the majority of studies were largely silent about the criteria used in the design (sic) of the graphical materials for the experiments. It appears, therefore, that we need to be more explicit about the selection and design of graphical representations for both applications and experiments investigating cognitive aspects of interacting with graphical representations.

6.1. WHAT CHARACTERIZES GOOD DESIGN?

As Bauer and Johnson-Laird (1993) discovered, facilitation of problem-solving depends on the kind of graphical representation being used. So what are the attributes of a "good" graphical representation? The issue of "good design" has been the subject of a number of different studies which have attempted to give guidelines for producing graphical displays (e.g. Kosslyn, 1989; Goettl, Wickens & Kramer 1991) but these have been concerned primarily with the syntax of the display, consistent with information-processing principles. Typically they are applicable to stylized displays, such as graphs and charts, and to domains where diagrams have highly evolved notations (e.g. electronic circuits, program flow charts, etc.). However the process of prescribing for a diagram structure in domains where strong constraints on form do not apply usually means that we rely on a few common conventions and heuristics. Domains such as physics have knowledge structures which are well-understood, capable of precise modelling. By contrast formalizing an understanding of, say, a diagram of a food chain is more problematic with a number of different representations possible.

In such circumstances two factors are clearly crucial: the kind of notation (e.g. the symbols of visual programming languages, universal icons) and the visual organization that is used to structure them. The notation that is employed should be appropriate (easily understandable). This is true both for the reading and for the production of diagrams. Merrill and Reiser (1993: p. 12), describing students learning LISP, observe that "the requirement to translate into an external notation system that does not match well with the structure of the students' plan imposes an additional working memory load of continually mapping between the two representations". Secondly we should recognize the importance of the canonical forms of diagrams, e.g. recognizing a diagram comprising of a set of images connected by arrows as a "cycle" may be critical. The existence of such forms (capable of supporting many different kinds of content) has long been recognized by researchers (e.g. Amigues & Caillot, 1990; Anzai, 1991) and is obvious in the culture at large (cf. Lohse *et al.,* 1991). One advantage of conventional two-dimensional diagrams is that they can trade on such recognition by activating appropriate "readability rules" (Sugiyama & Misue, 1991) and by cueing appropriate kinds of inferences in the reader. Cox and Brna (1991) noted that there was a strong correlation between a student's ability to identify types of graphical representation (e.g. diagram, map, table) and their ability to use them to solve problems. Pedagogically, encouraging as wide an acquaintance with different forms of diagrams is clearly important for their effective use but it may also be that experience with static forms itself may be a useful precursor to the ability to read more dynamic ones.

Under such circumstances we need insight into how people read and interact with diagrams. The stress here is important for our ideas about developing good diagram skills. In the vast majority of studies and analyses of static diagrams the assumption has been that the subject does nothing to change the external form. This may well be true for cases such as library books or slides in a presentation but it may not be the optimal case. Koedinger and Anderson (1990) observe that high school students frequently use annotations to problem diagrams to hold together information needed for inferences. Such a strategy may also be performed mentally but, given previous arguments, is probably more efficient for learners when external. In fact the logical conclusion of these arguments is to maximize the load on the external representation. As has been observed in several cases making a "cognitive trace" available for problem-solving is of great benefit (e.g. Merrill & Reiser, 1993). One lesson for diagram use might, therefore, be to promote opportunities for external manipulation (i.e. cognitive tracing) as well as encouraging production skills, as we noted above.

Good diagram design also has the crucial requirement that the degree of abstraction of material should be appropriate to the varying demands of the task and learner's ability. Levonen and Lesgold (1993) describe SHERLOCK, a computer-based electronics coaching system, which has the facility for representing both realistic (picture-like) diagrams of the system and schematized expert representations of the same domain. Switching between the two enables a kind of apprenticeship learning. Consistent with this approach Cheng (1993) advocates the availability of multiple representations, from specific examples to overviews, which learners could choose to look at as they wished. While this raises issues about integration between different views it also emphasizes the importance of learner control. There is no reason to doubt, for many situations, that multiple representations could be made available within a two-dimensional diagram. However it may also be that this is something that may well be better achieved in other forms.

6.2. CONCEPTUAL DESIGN ISSUES AND FUTURE DEVELOPMENTS

The above discussion points to a number of factors which designers should be aware of. We suggest that it is useful to begin to formalize them as a set of general conceptual design issues, akin to the set of cognitive dimensions of notations that Green (1989, 1990) has advocated, for describing important features of the design of programming languages and software tools to support users' tasks. Firstly, they can help bridge the gap between our conceptual understanding of how graphical representations work and the practial concerns of designing graphical representations. Secondly, requirements for future technological developments also can be assessed in relation to cognitive processing. Thirdly they can help us reframe design questions. We could ask what is required to design advanced graphical representations that can be of "added" cognitive value for particular users, domains and tasks? Below we present an initial attempt to identify some of the key conceptual design issues.

6.2.1. Explicitness and visibility

Diagrams, animations and virtual reality can in their respective ways all make salient certain aspects of a display. A design objective, therefore, should be to facilitate perceptual parsing and inferencing, through directing attention to key components that are useful or essential for different stages of a problem-solving or a learning task. In addition, the various graphical representations can represent "hidden" processes which underlie complex phenomena. The aim, here should be to facilitate higher level understanding, i.e. cognitive inferencing but also in relation to how this interacts with perceptual processing. As with "perceptual inferences", the users may need much prior knowledge in knowing how to interpret what is shown.

6.2.2. Cognitive tracing and interactivity

Diagrams that have been already constructed allow the user to leave cognitive traces, i.e. mark, update and highlight information. However this is a limited function. There is no possibility of interaction or feedback—the user cannot test new configurations. In contrast, when interacting with animations and virtual reality objects there is more scope for providing feedback but less for leaving cognitive traces. For example, various parameters of a computer-based model can be set in a virtual reality or three-dimensional simulation (cf. microworlds) and the outcome circtly observed. Graphical representations should be designed with a view towards how they support different kinds of cognitive tracing and levels of interactivity.

6.2.3. Ease of production

Related to the above issue is ease of production of a graphical representation. It appears that diagram production and comprehension are intimately related. A history of being taught to draw diagrams makes for fewer problems with understanding new ones. This is particularly important for domains where evolved notations are crucial. However, where the possibility of acquiring expertise is limited, the demands of reading the diagram efficiently may be too great. Recent software developments now make it possible for users to select alternative or partially animated views of the same process, and to play (and replay) them at different speeds, thus enabling multiple abstractions to be interpreted. Furthermore, software is being developed that will allow novice users easily to construct their own animations through compiling components from a toolbox of animations or modify pre-designed animations. The hypothesis about diagram production and comprehension could be tested for these more interactive forms of animations: having a better understanding of how to create animations will enable people to have a better understanding of how they work and what they are trying to convey.

6.2.4. Combining external representations

The conventions of constructing two-dimensional diagrams have largely evolved to be complementary to textual expositions. In some cases it may be that text is indispensable for understanding the function of a particular diagram. However fairly mundane factors such as spatial separation of text and diagram may significantly increase the computational load involved in comprehension (e.g. Sweller, Chandler, Tierney & Cooper, 1990). In contrast, animations and virtual environments have been designed to be largely graphical, although they may be accompanied by spoken narration or verbal text. Studies have shown that it can be more difficult to integrate written text with these kinds of graphical representation than with static diagrams. For example, response times from Rieber's (1989) study of combining text with animations to represent Newton's Laws of motion indicated that the subjects simply viewed the animations and then moved immediately onto the next screen of information without reading any of the accompanying text. Other studies which have combined spoken narration with animations, however, have fared better, showing that this combination is more effective. For example, Mayer and Anderson's (1991) study of subjects' understanding of the operation of a bicycle tyre pump, showed comprehension to be better when the information was depicted as an animation with concurrent narration, than when presented just as an animation. Having parallel auditory narration could also be effective for virtual reality to guide users in exploring and interacting with the environment. Hybrid graphical representations could also be developed that allow users to interact with static diagrams on a computer display by adding animations or conversely, allowing users immersed in a dynamic virtual reality environment to interact with static objects (e.g. jotting notes into a virtual notebook). The objective here, would be to provide support for different kinds of interactivity.

6.2.5. Distributed graphical representations

While we have not commented on this aspect in the paper, diagrams offer the possibility of joint evolution of representations, e.g. in idea-sketching, where planning can be facilitated using notations in any framework that suits the task. Here temporary conventions can be set up, which also has the drawback of being potentially unintelligible at some later date or to others. Various shared drawing tools also have been developed to support collaborative sketching and designing (e.g. Ishii & Kobayashi, 1992; Scrivener, Harris, Clark, Rockoff & Smyth, 1992). Collaborative design sessions can be recorded which, when played back, reconstruct the collaborative drawings as animations. Virtual reality environments can also provide opportunities for virtual construction of graphical representations for users in geographically dispersed locations. However, the value of enabling collaborative construction and editing of graphical representations in terms of enhancing task performance is only beginning to be researched.

7. Overview and discussion

A major aim of this paper was to examine the strength of claims for the value of advances in graphical technology for facilitating cognitive tasks. We have seen that these claims are often underpinned by assumptions which have little empirical support and/or insufficient theoretical grounding. In addition there has been little progress towards a framework, either methodological or theoretical, that might allow the designer to produce and evaluate new forms of graphical representation or even improve on existing ones.

Part of our argument has been to pick over the bones of previous studies to see *why* there has been so little progress and/or intergration despite an enormous volume of research. The answer seems to be in several parts. Firstly, the studies have been highly detailed and do not generalize. Secondly, they have failed to produce adequate rationales for the material tested, making it difficult to determine what is actually being assessed. Thirdly they make assumptions about the kinds of linkages between external representation and internal representation which are rarely articulated (cf. the resemblance fallacy) or, if they are, may not give sufficient weight to the role of the external. Fourthly, articulating the links require theoretical analyses, of which those that might seem appropriate, are only beginning to emerge as theoretical developments in cognitive science (e.g. Hutchins, 1995).

Most existing accounts of how graphical representations are effective, therefore, have been black box in nature—there exists a gap in terms of explaining adequately any cognitive processes involved. For example the account of internal process may be couched in terms of "applying knowledge of content" but give us little about what kind of internal representation is mediating task performance. Likewise, as we argued in describing the resemblance fallacy, making assumptions that the internal representation is a mental model or image-like may simply give the illusion of solving the processing-internal representation-external representation riddle. But instead, the problem of explaining the value of graphical representations is shifted simply from an external to an internal account. In contrast we promote an alternative approach that analyses how different graphical representations work in terms of core "external cognition" processes and properties of the graphical representation, e.g. computational offloading, re-representation and graphical constraining. We believe that such an enterprise is central to evolving a more adequate account of the cognitive benefits and mechanisms involved.

Related to this is a further, critical and under-acknowledged theme, that of interactivity. Specifying how people interact with graphical representations, when learning, solving problems and making inferences, is complex since it will involve not only a specification of the cognitive mechanisms alluded to above but also some sense of the behavioural aspects. For example the fact that students prefer to mark diagrams as they work, the established value of cognitive traces and the dialectic between graphical representation production and use all point to a need to conceptualize graphical representations as more than passively observed, with obvious implications for design and innovation. In turn the potential significance of such activity will be a function of variables such as the level of experience with the graphical representation and knowledge domain, type of task and abstractness of information being represented. Many of the presumed benefits of good-old fashioned graphical representations (i.e. static diagrams) were considered to be due to years of practice of perceptual processing of visual stimuli and the learning of graphical conventions. This may help us to understand why advanced graphical technologies (e.g. animations and virtual reality) have not, as yet been able to demonstrate comparable performance or learning benefits. Similarly we have even less understanding of how (and if) computational offloading works in such dynamic environments. There may also be contingent problems such as the fact that the temporal constraints of existing kinds of "passive" animations and "immersed" virtual reality may only allow for shallow processing of information (e.g. Philips, 1986; Palmiter, Elkerton & Baggett 1991), thereby preventing them from having the equivalent computational benefits that static diagrams offer.

In sum, we propose a new agenda for research into graphical representations that is based on an analysis of interactivity and, thus, considers the relationship between different external and internal representations. Such an approach should help us to better understand, design and select graphical representations—be they "old fashioned" or technologically advanced—which are appropriate for the learning environment, problem-solving task or entertainment activity in question.

The authors would like to gratefully acknowledge support for this work from the ESRC (Cognitive Engineering Programme award L127251033). We should also like to thank the anonymous referees of this paper as well as Ben du Boulay, Ron Chrisley, Graham Hole and other members of the Socio-Cognitive Research Group at COGS, Sussex University for their comments on early drafts.

References

AMIGUES, R. & CAILLOT, M. (1990). Les representations graphiques dans l'enseignement et l'apprentissage de l'electricite. *European Journal of Psychology of Education,* **5,** 477–488.

ANDERSON, J. R. (1990). *Cognitive Psychology and its Implications.* (3rd edn) San Francisco, CA: W. H. Freeman.

Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith, Eds. *Towards a General Theory of Expertise.* pp. 64–92. Cambridge: Cambridge University Press.

Bauer, M. I. & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science,* **4,** 372–378.

Cheng, P. C.-H. (1993). Multiple interactive structural representation systems for education. In R. Cox, M. Petre, P. Brna & J. Lee, Eds. *Proceedings of the AI-ED93 Workshop on Graphical Representations, Reasoning and Communication.* pp. 25–27. August, Edinburgh.

Cover, A. S., Ezquerra, N. F., O'Brien, J. F., Rowe, R., Gadacz, T. & Palm, E. (1993). Interactively deformable models for surgery simulation. *IEEE Computer Graphics & Applications,* **13,** 68–85.

Cox, R. & Brna, P. (1993). Analytical reasoning with external representations. In R. Cox, M. Petre, P. Brna & J. Lee, Eds. *Proceedings of the AI-ED93 Workshop on Graphical Representations, Reasoning and Communication.* pp. 33–36. August, Edinburgh.

Cox, R. & Brna, P. (in press). Supporting the use of external representations in problem solving: the need for flexible learning environments. *Journal of Artificial Intelligence in Education,* **6.**

Earnshaw, R. A. & Watson, D., Eds. (1993). *Animation and Scientific Visualization: Tools and Applications.* London: Academic Press.

Fairchild, K. M., Serra, L., Hern, N. H., Hai, L. B. & Leong, A. T. (1993). Dynamic fisheye information visualizations. In R. A. Earnshaw, M. A. Gigante & H. Jones, Eds. *Virtual Reality Systems.* pp. 161–177. London: Academic Press.

Finke, R. A. (1990). *Creative Imagery: Discoveries and Inventions in Visualization.* Hillsdale, NJ: Lawrence Erlbaum.

Gigante, M. A. (1993). Virtual reality: definitions, history and applications. In R. A. Earnshaw, M. A. Gigante & H. Jones, Eds. *Virtual Reality Systems.* pp. 3–15. London: Academic Press.

Glenberg, A. M. & Kruley, P. (1992). Pictures and anaphora: evidence for independent processes. *Memory and Cognition,* **20,** 461–471.

Glenberg, A. M. & Langston, W. E. (1992). Comprehension of illustrated text: pictures help to build mental models. *Journal of Memory and Language,* **31,** 129–151.

Goettl, B. P., Wickens, C. D. & Kramer, A. F. (1991). Integrated displays and the perception of graphical data. *Ergonomics,* **34,** 1047–1063.

Goodman, N. (1968). *Languages of Art: An Approach to the Theory of Symbols.* Indianapolis, IN: Bobs-Merrill.

Green, T. R. G. (1989). Cognitive dimensions of notations. In A. Sutcliffe & L. Macaulay, Eds. *People and Computers V.* Cambridge: Cambridge University Press.

Green, T. R. G. (1990). The cognitive dimension of viscosity: a sticky problem for HCI. In D. Diaper, D. Gilmore, G. Cockton & B. Shakel, Eds. *Human–Computer Interaction—INTERACT'90.* Amsterdam: North Holland.

Halford, G. S. (1993). *Children's Understanding: the Development of Mental Models.* Hillsdale, NJ: Lawrence Erlbaum.

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

Hegarty, M., Just, M. A. & Morrison, I. M. (1988). Mental models of mechanical systems: individual differences in quantitative and qualitative reasoning. *Cognitive Psychology,* **20,** 191–236.

Hong, E. & O'Neill, F. (1992). Instructional strategies to help learners build relevant mental models in inferential statistics. *Journal of Educational Psychology,* **84,** 150–159.

Hutchins, E. (1995). *Cognition in the Wild.* Cambridge, MA: MIT Press.

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

Ishii, H. & Kobayashi, M. (1992). Integration of inter-personal space and shared workspace: clearboard design and experiments. *Proceedings of the Conference on Computer Supported Cooperative Work,* pp. 33–42. New York, NY: ACM.

KAISER, M. K., PROFFITT, D. R., WHELAN, S. M. & HECHT, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 669–690.

KALAWSKY, R. S. (1993). *The Science of Virtual Reality and Virtual Environments.* Workingham: Addison-Wesley.

KOEDINGER, K. R. & ANDERSON, J. R. (1990). Abstract planning and perceptual chunks: elements of expertise in geometry. *Cognitive Science,* **14,** 511–550.

KOSSLYN, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology,* **3,** 185–226.

KOZAK, J. J., HANCOCK, P. A., ARTHUR, E. J. & CHRYSLER, S. T. (1993). Transfer training from virtual reality. *Ergonomics,* **36,** 777–784.

LARKIN, J. H. (1989). Display-based problem solving. In D. KLAHR & K. KOTOVSKY, Eds. *Complex Information Processing: The Impact of Herbert A. Simon.* pp. 319–341. Hillsdale, NJ: Lawrence Erlbaum Associates.

LARKIN, J. H., & SIMON, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science,* **11,** 65–100.

LAURILLARD, D. (1993). *Rethinking University Teaching.* London: Routledge.

LEVONEN, J. J. & LESGOLD, A. M. (1993). Visual reasoning in coached practice environments: visual meaning making in SHERLOCK In R. COX, M. PETRE, P. BRNA & J. LEE, Eds. *Proceedings of the AI-ED93 Workshop on Graphical Representations, Reasoning and Communcation.* pp. 21–24. August, Edinburgh.

LOHSE, G., WALKER, N., BIOLSI, K. & RUETER, H. (1991). Classifying graphical information. *Behaviour and Information Technology,* **10,** 419–436.

MAYER, R. E. & ANDERSON, R. B. (1991). Animations need narrations: an experimental test of a dual-coding hypothesis. *Journal of Educational Psychology,* **83,** 484–490.

MCCLOSKEY, M., CARAMAZZA, A. & GREEN, B. (1980). Curvilinear motion in the absence of external forces: naive beliefs about the motion of objects. *Science,* **210,** 139–1141.

MCCORMICK, B. H., DEFANTI, T. A. & BROWN, M. D. (1987). Visualization in scientific computing—a synopsis. *IEEE Computer Application Graphics,* **7,** 61–70.

MERRILL, D. & REISER, B. J. (1993). Scaffolding the acquisition of complex skills with reasoning-congruent learning environments. In R. COX, M. PETRE, P. BRNA & J. LEE, Eds. *Proceedings of the AI-ED93 Workshop on Graphical Representations, Reasoning and Communication.* pp. 9–16. August, Edinburgh.

MOLITOR, S., BALLSTAEDT, S. & MANDL, H. (1989). Problems in knowledge acquisition from text and pictures. In H. MANDL & J. R. LEVIN, Eds. *Knowledge Acquisition from Text and Pictures.* pp. 3–30. Amsterdam: North-Holland.

NORMAN, D. (1988). *The Psychology of Everyday Things.* New York, NY: Basic Books.

NORMAN, D. (1993). Cognition in the head and in the world. *Cognitive Science,* **17,** 1–6.

O'MALLEY, C. & DRAPER, S. (1992). Representation and interaction: are mental models all in the mind? In Y. ROGERS, A. RUTHERFORD & P. BIBBY, Eds. *Models in the Mind.* London: Academic Press.

PALMITER, S., ELKERTON, J. & BAGGETT, P. (1991). Animated demonstrations vs written instructions for learning procedural tasks: a preliminary investigation. *International Journal of Man–Machine Studies,* **34,** 687–701.

PARK, O. & GITTLEMAN, S. S. (1992). Selective use of animation and feedback in computer-based instruction. *Education Technology Research & Development,* **40,** 27–38.

PARKES, A. (1993). Intelligent learning environments and reified AI. In R. COX, M. PETRE, P. BRNA & J. LEE, Eds. *Proceedings of the AI-ED93 Workshop on Graphical Representations, Reasoning and Communication.* pp. 37–42. August, Edinburgh.

PHILIPS, R. J. (1986). Computer Graphics as memory aid and a thinking aid. *Journal of Computer Assisted Learning,* **2,** 37–44.

PYLYSHYN, Z. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin,* **80,** 1–24.

REED, S. K. (1993). Imagery and discovery. In B. ROSKOS-EWOLDSEN, J. J. INTONS-PETERSON & R. E. ANDERSON, Eds. *Imagery, Creativity, and Discovery: A Cognitive Perspective. Advances in Psychology, 98.* pp. 287–312. Amsterdam: North-Holland.

RHEINGOLD, H. R. (1991). *Virtual Reality.* London: Secker and Warburg.

REIBER, L. P. (1989). The effects of computer-animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Computing Research,* **5,** 431–444.

RIEBER, L. P. & KINI, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals. *Journal of Computer-Based Instruction,* **18,** 83–88.

SCRIVENER, S. A. R., HARRIS, D., CLARK, S. M., ROCKOFF, T. & SMYTH, M. (1992). Designing at a distance: experiments in remote-synchronous design. *Proceedings of OZCHI'92, CHISIG Annual Conference.* pp. 44–53. Gold Coast Australia, November.

SHERIDAN, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence,* **1,** 120–125.

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

STENNING, K. & TOBIN, R. (in press). Assigning information to modalities: comparing graphical treatments of the syllogism. In E. EJERHED & S. LINDSTROM, Eds. *Action, Language and Cognition. Proceedings of International Conference on Dynamic Semantics,* Umea, Sweden. September.

SUGIYAMA, K. & MISUE, K. (1991). Visualization of structural information: automatic drawing of compound digraphs. *IEEE Transactions on Systems, Man, and Cybernetics,* **21,** 876–892.

SWELLER, J., CHANDLER, P., TIERNEY, P. & COOPER, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology General,* **119,** 176–192.

VERA, A. H. & SIMON, H. A. (1993). Situated action: a symbolic interpretation. *Cognitive Science,* **17,** 7–48.

WINN, W. (1987). Charts, graphs, and diagrams, in educational materials. In D. M. WILLOWS & H. A. HOUGHTON, Eds. *The Psychology of Illustration. Volume 1 Basic Research.* pp. 152–198. New York, NY: Springer-Verlag.

WINN, W. (1993). An account of how readers search for information in diagrams. *Contemporary Educational Psychology,* **18,** 162–185.

ZHANG, J. & NORMAN, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science,* **18,** 87–122.

Paper accepted for publication by Associate Editor, Dr. D. Boehm-Davis