# Toward a Comprehensive Model of Graph Comprehension: Making the Case for Spatial Cognition

Susan Bell Trickett and J. Gregory Trafton

Naval Research Laboratory, 4555 Overlook Drive,
Washington DC 20375
trickett@itd.nrl.navy.mil, trafton@itd.nrl.navy.mil

**Abstract.** We argue that a comprehensive model of graph comprehension must include spatial cognition. We propose that current models of graph comprehension have not needed to incorporate spatial processes, because most of the task/graph combinations used in the psychology laboratory are very simple and can be addressed using perceptual processes. However, data from our own research in complex domains that use complex graphs shows extensive use of spatial processing. We propose an extension to current models of graph comprehension in which spatial processing occurs a) when information is not explicitly represented in the graph and b) when simple perceptual processes are inadequate to extract that implicit information. We apply this model extension to some previously published research on graph comprehension from different labs, and find that it is able to account for the results.

## 1 Introduction

Until recently, models of graph comprehension have mostly focused on simple graphs and tasks, for which information is explicitly represented in the graph. Examples of this type of graph/task combination include extracting trends from bar graphs, reading off values from bar and line graphs, comparing values in bar graphs, and the like (e.g, (Cleveland, 1985). Recently, however, researchers have begun to question the extent to which these simple, context-lean graphs and tasks represent the true nature of graph use beyond the psychology laboratory. In reality, graphs may be used to make predictions, and thus require information that is not explicitly represented (DeSanctis & Jarvenpaa, 1989) and graphs may be used as problem-solving tools (Cheng *et al.*, 2001; Scaife & Rogers, 1996; Tabachneck-Schijf *et al.*, 1997; Trafton *et al.*, 2000; Trafton & Trickett, 2001). Researchers have thus been attempting to understand more complex graph/task interactions, as well as the use of more complex graphs themselves. These newer models have begun to account for such factors as the iterative nature of graph interpretation (Carpenter & Shah, 1998; Peebles & Cheng, 2003), the familiarity of the graph (Peebles & Cheng, 2003), the role of graph and domain knowledge (Freedman & Shah, 2002), and the importance of expertise (Roth & Bowen, 2003; Tabachneck-Schijf et al., 1997).

In their recent review of the graph comprehension literature, Shah, Freedman, and Vekiri draw a distinction between perceptual and conceptual processes (Shah *et al.*, 2005). In their interpretation, perceptual processes are “bottom-up encoding

D. Barker-Plummer et al. (Eds.): Diagrams 2006, LNAI 4045, pp. 286 – 300, 2006.

mechanisms," which focus on the visual features of the display, whereas conceptual processes equate to "top-down encoding processes," which influence interpretation. They propose that perceptual processes account for performance on "simple, fact-retrieval tasks," but they further argue that "If visual features do not automatically evoke a relationship, either because the relationships are not visually integrated in a graph or because the graph viewer does not have the prior knowledge required to make an interpretation, information must be retrieved by complex inferential processes."

Although several models agree that these "complex inferential processes" are an essential part of the graph comprehension process under some circumstances, they remain largely unspecified. Indeed, in taking into account both perceptual and conceptual processes, Shah et al. identify five factors that play a role in predicting performance on graph comprehension: display characteristics, data complexity, the viewer's task, the viewer's prior domain knowledge, and the viewer's knowledge about graphs. What current models lack is a means to specify precisely how these factors will influence the type of complex inferential processes that will be engaged.

We investigated what happens in complex, problem-solving domains when scientists are unable to extract the information they need from the visualization they are using (Trafton & Trickett, 2001). Based on an in-depth analysis of several hours of verbal protocols, we found not only that it was extremely common for the scientists to confront situations where they were unable to directly extract the information they needed, but also that in these cases, they used spatial transformations more frequently than any other strategy to generate this information. We concluded that models of graph comprehension should be expanded to include spatial processing, particularly in complex domains for which complex visualizations are required.

Our most recent research has focused on complex graphs in another complex domain (meteorology). Common tasks in this domain certainly include fact retrieval (e.g., temperature, wind-speed, etc., at a specific location), in which information is explicitly represented and little specialized domain knowledge is required to extract it. However, equally commonly, forecasters use graphs to draw inferences (e.g., finding the pressure at location C when it is given for locations A and B), for which information is not explicitly represented and some domain knowledge is required. They must also make complex weather predictions, for which information is not explicitly represented and a great deal of domain knowledge is needed. Our approach has been to observe and record experts and journeymen using weather graphs as part of their regular work, and to interview them about their strategies. Our results have been consistent: the verbal protocols show that in this domain, at least, people use a great deal of spatial processing to extract and use information from data visualizations (Trafton et al., 2000; Trafton & Trickett, 2001). Further evidence of spatial processing is found in meteorologists' gestures when they talk about how they performed the task (Trafton *et al.*, in press). Additionally, in keeping with the important role of domain knowledge in graph comprehension (Freedman & Shah, 2002; Roth & Bowen, 2003; Tabachneck-Schijf et al., 1997), experts use far more spatial processing than journeymen (Trafton et al., in press). This general result—that spatial processing is prevalent in complex graph comprehension—has been replicated in studies of fMRI research, in addition to the original work in astronomy and computational fluid dynamics. All these domains share some important characteristics with the meteorology domain (in terms of the complexity of the visualizations, the task, and the domain) (Trafton & Trickett, 2001).

Based on these data, we have become convinced that spatial processing is an important component of a comprehensive model of graph comprehension, and specifically, that it plays a crucial role in guiding the "complex inferential processes" discussed above that are involved when information cannot be directly extracted from the graph. Yet, curiously, spatial processing is not explicitly included in any of the current models of graph comprehension. This is something of a puzzle, because these models are highly successful in accounting for graph comprehension behavior in many graphs and tasks, including some in which information must be inferred rather than extracted directly. Indeed, some models (Lohse, 1993; Peebles & Cheng, 2003; Pinker, 1990) are explicitly and exclusively propositional; others are simply noncommittal (Freedman & Shah, 2002; Roth & Bowen, 2003). The goal of this paper is to investigate why this should be so—that is, why current theories of graph comprehension, so successful in analyzing many of the behaviors associated with graph interpretation, do not account for our data. We propose a refinement of current theories that enables them to predict when spatial processing will occur, and how it will guide graph comprehension in these complex situations.

First, we provide a brief definition of spatial processing. Though this definition is a simplification, it is nonetheless useful because it provides operational means by which we can identify spatial processing in our verbal protocol data and in our analyses of the requirements of graph tasks in the graph comprehension literature. Second, we briefly describe a generic model of graph comprehension, based on recent analyses by Shah, Freedman, and Vekiri (2005). We analyze two situations, one in which needed information can be directly extracted from the graph, and a second in which it must be inferred using perceptual processes. The purpose of this description is to establish a terminology that applies to graph comprehension tasks specifically addressed by this model, which we can then use to describe the tasks and actions involved in the more complex meteorological task. Third, we summarize the results of several studies in complex domains that show spatial processing is used in these tasks. We also analyze several specific instances of forecasting tasks, to show *when* and *how* spatial processing is used. We show that when the desired information was not explicitly represented in the graph *and* when perceptual processing could not generate the type of information needed, spatial processing was used. Finally, we apply our model to an analysis of a graph comprehension tasks from the graph comprehension literature, and show how it provides a viable interpretation of performance on these tasks.

## 2 Spatial Processes

Baddeley was instrumental in establishing the distinction between verbal and spatial processing (Baddeley, 1999) and in further distinguishing between spatial and visual processing (Baddeley & Liebeman, 1980). Spatial processing involves "the internalized reflection and reconstruction of space in thought" (Hart & Moore, 1973).

Operationally, we define spatial processing in two ways. Spatial processing involves maintaining spatial information (e.g., the relative locations of objects) in working memory (so-called spatial working memory). Instances of spatial processing can therefore be identified by means of task analysis (Gray *et al.*, 1993). Spatial processing can also be identified via the use of mental spatial transformations, which occur when a spatial object is transformed from one mental state or location into another mental state or location. Mental spatial transformations—which we refer to simply as spatial transformations—occur in a mental representation that is an analog of physical space and are frequently part of a problem-solving process. There are many types of spatial transformations: creating a mental image, modifying that mental image by adding or deleting features, mental rotation (Shepard & Metzler, 1971), mentally moving an object, animating a static image (Bogacz & Trafton, 2005; Hegarty, 1992), making comparisons between different views (Kosslyn *et al.*, 1999; Trafton *et al.*, 2005), and any other mental operation which transforms a spatial object from one state or location into another.

Note that we distinguish between spatial processing (i.e., the use of spatial transformations) and purely perceptual processing, in which graph users are able to make *direct* or *explicit* comparisons from the graph itself, without the need to hold spatial information in working memory. Thus, comparing two adjacent bar heights on a graph requires only perceptual processing, whereas comparing a bar height on a displayed graph with one on a remembered graph requires spatial processing, because the remembered bar height would have to be projected onto the displayed graph for the comparison to occur (assuming that specific values had not been remembered).

Most graph comprehension research is not designed to specifically identify the type of processing—verbal or spatial—used. However, while people are doing graph-related tasks, it is possible to elicit verbal protocols, that "dump" the contents of working memory during problem-solving (Ericsson & Simon, 1993). These verbal protocols can then be coded, and instances of spatial processing can be identified. We have conducted several studies in which meteorologists (and scientists) give verbal protocols while making their forecasts, and we have coded the spatial transformations in those protocols. Our IRR for this coding has been consistently good.

## 3 A Generic Model of Graph Comprehension

Several general models of graph comprehension are based on one proposed by Pinker (Pinker, 1990), in which the visual features of the display, gestalt processes, and the graph schema all interact to allow the user to extract the conceptual message of the graph. To summarize this model: 1) the user has a goal (which is provided) to extract a specific piece of information 2) the user looks at the graph, and the graph schema and gestalt processes are activated 3) the salient features of the graph are encoded, based on gestalt principles 4) the user now knows which cognitive/interpretive strategies to use, because the graph is familiar (graph knowledge)—that is, the user knows where to look at what steps to take 5) the user extracts necessary goal-directed visual chunks 6) the user may compare two or more visual chunks and 7) the user extracts the relevant information to satisfy the goal (answer the question).

Figure 1 shows a simple bar graph, depicting information from the US Census Bureau (www.census.gov). Suppose the goal is to extract the amount of lifetime earnings for woman with a high school diploma. This information is explicitly represented in the graph, and can be directly extracted when the user executes steps 1 through 5, and step 7. Suppose, however, the goal is to extract how much more women earn if they complete a bachelor's degree. In this case, the information is not explicitly represented; however, it can be easily extracted by repeating steps 1 through 5 for the bachelor's degree bar, and calculating the difference. Other information that is not explicitly represented can also be extracted in a straightforward manner—for example, the trend of earnings as education increases can be extracted by using the perceptual process of noticing that each successive bar is a bit taller than the bar to its left (i.e., by comparing visual chunks). Similarly, the answer to the question of who earns the most can be extracted by locating the tallest bar (again, comparing the bar heights, or visual chunks). None of these questions, which are typical of questions posed for this type of graph, requires the use of spatial processes; all can be answered by using the perceptual processes built in to the generic model.

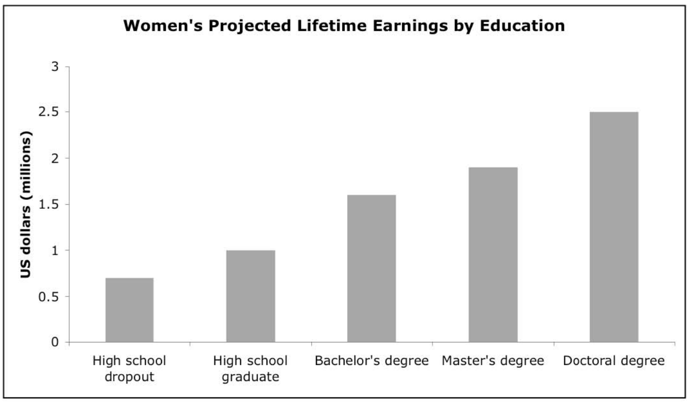


**Fig. 1.** Simple bar graph

Building on this model, Shah and Carpenter have shown that the processes involved are iterative, rather than serial (Carpenter & Shah, 1998), and Freedman and Shah have further adapted the model to account for the role of domain knowledge, which may influence the last stage, in accordance with the user's expectations (Freedman & Shah, 2002). Nonetheless, these basic perceptual processes have been sufficient to account for many tasks in the graph comprehension literature, such as fact-retrieval from line or bar graphs, trends for line and bar graphs, making proportion judgments from pie charts, making comparison judgments, determining the slope of a regression line, and so on. Variability in performance on these tasks is likely due to other factors, such as graph or domain knowledge (Freedman & Shah, 2002).

In this generic model, although graphs depend on spatial arrays, the processes by which information is extracted are largely perceptual (whether the information is represented explicitly or implicitly). Graphs depict relationships by means of the strategic arrangement of spatial elements, and those relationships can be easily extracted because spatial attributes are automatically encoded relationally, e.g., a higher line is encoded as meaning a greater value (Pinker, 1990). What makes some graphs better than others for particular tasks is precisely this characteristic of graphs—e.g., trend information is more easily extracted from line graphs than bar graphs (Tversky, 1995) because a line with an increasing slope is encoded as going from less to more. The trend can thus be interpreted by means of virtually effortless perceptual processing.

## 4 Tasks in Complex Domains

In this section, we examine how well the generic model outlined above accounts for performance in the meteorological and scientific domains. The simplest type of task in meteorology is straightforward fact retrieval. Figure 2 shows a typical meteorological graph. Note that this graph is significantly more complex than the bar graph in Figure 1: it shows four variables (wind speed, wind direction, temperature, and sea level pressure) in addition to latitude and longitude lines overlaid onto a map of the Southeastern U.S. Despite the graph complexity, however, in such cases, when the given goal is to extract a specific piece of information, the generic model is quite adequate. For example, if a forecaster wanted to know the current temperature at Pittsburgh, he would take the same steps as those outlined above: first, he would look at the weather graph, activating the graph schema, then he would find Pittsburgh and encode the color, thereby extracting the required visual chunk; the graph schema would guide him to translate the color into a temperature value, by looking at the legend, and he would "read off" the appropriate value from the legend. The forecaster might make several iterations between looking at Pittsburgh and the legend (Trafton *et al.*, 2002),, but the basic mechanisms from the generic model can easily account for performance on this task. The model also supports some tasks in which information is not directly represented, such as the comparison "Which is hotter, Pittsburgh or Washington?" As in the comparison task in the bar graph, the forecaster could answer this question by using perceptual processes, locating both Pittsburgh and Washington and their associated colors, and either looking up the values on the legend or (a more likely expert strategy) comparing the colors to see which represents the hotter temperature.

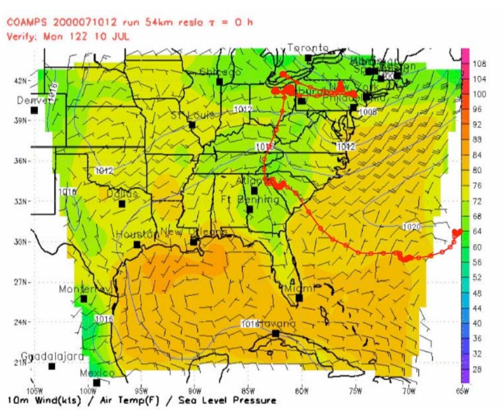


**Fig. 2.** Meteorological graph (note that the original graph is in color)

In many forecasting tasks, however, information cannot be directly extracted from the graph, and in these cases, it may not be possible to extract the information by perceptual processes, such as directly comparing visual chunks. Trafton, Marshall, Mintz, and Trickett (Trafton et al., 2002) conducted an eyetracking study in which forecasters were asked to perform a number of routine forecasting tasks. The tasks were designed to have certain characteristics: asking for quantitative information, where the answers were explicitly represented in the graph; asking for quantitative information that was imprecisely represented (i.e., the values were represented by a symbology that the user must know in order to extract the needed information); and asking for quantitative information that was entirely *implicit* in the graph (e.g., what is the pressure at location C, when it was given for locations A and B, so that values must be inferred).

In another study involving scientists rather than forecasters, Trafton, Trickett, and Mintz (2005) investigated the use of mental imagery in scientific visualization. Two astronomers and a physicist were observed using complex visualizations to analyze data. Trafton et al. compared the number of spatial transformations in the verbal protocols with the number of physical transformations they performed on the visualizations (i.e., creating a new visualization or adjusting a current one). There were significantly more spatial transformations than physical transformations, suggesting that the scientists frequently used spatial processing in preference to the computer's visualization capabilities. More interestingly, Trafton et al. found that comparisons were extensively used to tie the internal and external visualizations together. They divided these comparisons into two types: comparing two external visualizations and comparing an external and an internal visualization. They also coded the type of comparison made: comparing features (such as color or size) or aligning (i.e., making an estimation of "fit" between the two images). They found that the type of comparison was related to the type of visualization being compared. When two external visualizations were compared, the scientists most frequently compared features of the visualization; however, when an internal and an external visualization were compared, the scientists most frequently made alignments. In terms of our model, when the information was explicitly represented (in the external visualizations), the scientists used a perceptual strategy of comparing visible features; when the information was not explicitly represented, they used a spatial strategy of aligning one visualization with another, in order to estimate the overall "fit" between the two.

In addition to these studies, further data from our study of forecasters shows a large number of situations in which data must be inferred, e.g. resolving discrepancies between models or between a representation of certain phenomena and the forecaster's conflicting belief about the phenomena, integrating large amounts of disparate information into a comprehensive mental model, comparing visual chunks that are no longer on the visible display, but must be recalled in memory and mentally juxtaposed, as well as projecting the changes in the visualization that will likely occur over time, given current conditions. We analyze detailed examples of some of these situations below, in order to show where the information came from and how it was generated. We consistently find that when forecasters cannot directly extract the information from the display, they do one of two things: either they recall it, and given the nature of the domain, much of this recalled information is spatial, or they generate it by means of spatial transformations, a form of spatial processing.

**Table 1.** Resolving a discrepancy between two weather models

| Utterance | Information Explicitly Available? | Action |
|---|---|---|
| You also have a 12 max 14 | Yes | Extract information |
| winds are not supporting that | Yes | Note discrepancy |
| The next chart has it moving down further to the south | No | Recall from spatial memory (previous visualization, no longer on screen) |
| **there is a low coming off the coast that is propagating around** | No | Spatial transformation (accompanied by hand gesture tracing location of imagined low) |
| **so I would move it further to the south** | No | Spatial transformation |
| and that just supports what I said about ours, OK | N/A | Resolve discrepancy |

Table 1 shows an instance of a forecaster trying to resolve a discrepancy between two visualizations. Her goal was to determine whether or not to maintain a high-seas warning, and the chart on display showed the projected sea heights in different locations of a particular model for the period of interest. She begins by reading off the projected sea height in the location she is interested in, "12 max 14" (i.e., high enough to be of concern to her). However, information about wind speed conflicts with this information. She then recalls another visualization that showed the high seas area in transition further to the south (at this stage, she is using her memory of relative location, i.e., spatial processing). Her next utterance is a spatial transformation, in which she mentally creates and moves a low pressure system (it is not represented on the current visualization, but is recalled from memory and projected onto the current visualization). This is followed by another spatial transformation, in which she mentally moves the area of high seas further to the south. Implicitly, she performs a mental comparison of these transformed mental representations, and finally indicates that in this new location, the high seas makes sense and the discrepancy is resolved. It is important to note that, just as in the use of more simple graphs, there is an interaction between the demands of the task and the type of visualization. In the current case, however, this interaction is complex and requires several different types of processing: information must be extracted directly from the current visualization, both spatial and non-spatial information must be recalled from previously accessed visualizations, and spatial information must be both superimposed and manipulated on the current visualization.

**Table 2.** Discrepancy between weather model and forecaster's mental model

| Utterance | Information Explicitly Available? | Action |
|---|---|---|
| I can't buy an 82 out of the weather bureau at all… | N/A | Note discrepancy |
| and having a hard time understanding why they're coming up with what they got. | N/A | Reiterate discrepancy (no easy resolution) |
| They have um Brunswick a max temperature of 78 for Friday, | Yes | Extract information |
| **we push a front through** | No | Spatial transformation |
| and we go to 82 degrees | No | Non-spatial projection |
| **That's just no way you would think anything like that would hap-pen….** | No | Spatial transformation (mental comparison) |
| I'm not buying that | N/A | Reiterates disbelief |
| So again what I'm gonna do | N/A | Conclusion |
| I'm gonna more or less stay with what I had yes-terday, | N/A | Conclusion |
| I'm going 77 | N/A | Conclusion |

Table 2 illustrates a similar case of discrepancy resolution, only in this instance the discrepancy is between the National Weather Service prediction and the forecaster's own mental model. The forecaster begins by doubting the NWS temperature prediction and puzzles over how it was constructed. He then attempts to reconstruct the process, recalling the prediction for the day prior to the disputed forecast, and performs a spatial transformation on that model, mentally moving a front through the relevant area (the front is not represented on the current visualization). He updates the mental representation after the front has hypothetically moved through, and projects the disputed 82 degrees on that update. He performs another spatial transformation by mentally comparing the two relevant chunks—the updated mental representation and the representation containing the 82 degrees, and finds them still discrepant. As a result, he cannot believe the NWS forecast is valid, and resolves the discrepancy by sticking with his own mental model. As in the previous example, spatial transformations are a crucial part of how he uses the visualization to resolve the problem.

Similarly, Table 3 shows an instance of a forecaster needing to compare visual chunks that are not visible on any current display. In this situation, the forecaster is attempting to update a paper chart, by integrating information from all the previous visualizations (from disparate models) she has viewed. She first notes a discrepancy between one of these models and the others, by performing a mental comparison of the two representations (a spatial transformation). Her second utterance indicates that one of the remembered models did display the lows that are absent from the Canadian model. In each of the next three utterances, she performs some form of mental comparison between her memory of the Canadian model and her memory of the ENSAP model. In the end, she determines a placement for the low on the third representation, a paper chart that she is attempting to update.

**Table 3.** Compare visual chunks not on the visible display

| Utterance | Information Explicitly Available? | Action |
|---|---|---|
| **Also, one thing I'm noting is that the Canadian model is having a problem picking up the two lows** | No | Spatial transformation: comparison (2 mental representations) |
| that are circulating around this cut-off low off of the coast of Greenland | No | Recall from memory |
| **They do have something there** | No | Spatial transformation: Comparison (alignment) |
| **But they're not putting a central pressure on it, as ENSAP is** | No | Spatial transformation: Comparison (alignment) |
| **And it…they're definitely there** | No | Spatial transformation (projection) |
| So I'll put an X where I think that low should be | N/A | Resolution |

From these examples, it appears that spatial transformations serve a particular purpose in this domain, namely, they provide a means whereby the forecasters can *generate* needed information that is not explicitly represented in the visualization. This information is constructed by performing spatial transformations on the information that is explicitly represented, and developing new mental representations based on those (mentally) transformed visualizations. These new mental representations can be further manipulated or used as the basis for comparisons.

We thus propose that current models of graph comprehension do not include spatial processing because for the tasks and graphs used in most studies of graph comprehension, either the information can be directly extracted from the display, or if not, it can be inferred using direct perceptual processing of the available visual chunks. We propose that models of graph comprehension should account for these conditions, in that when the information can neither be extracted directly nor inferred from perceptual processes, spatial processing will be used. We now test this aspect of this model by applying it to a graph task in the graph comprehension literature.

## 5 Tasks in the Graph Comprehension Literature

In this section, we turn to a re-analysis of a previously published graph comprehension study, in order to test our model of spatial processing. We have chosen to focus our analyses on one of the tasks investigated by Simkin and Hastie (Simkin & Hastie, 1987) because their stated aim was to establish elementary codes that can account for the processes that operate "when people decode the information represented in a graph," that is, that can apply across different graph/task combinations and account for differences in performance across different graphical representations.

Simkin and Hastie used three different graph types—simple bar, divided bar, and pie charts—and three tasks—discrimination, comparison judgment, and proportion judgment. The graphs were similar to those in Figure 3. We focus only on the comparison judgment task, because according to Simkin and Hastie's task analysis, all their different elementary codes are involved in this task across the different graph types. For the comparison task, participants were asked to assess the percentage the smaller visual chunk was of the larger. This information is not explicitly represented in the graph; the question of interest to us is whether it can be extracted by purely perceptual means, or whether spatial processing must be used.

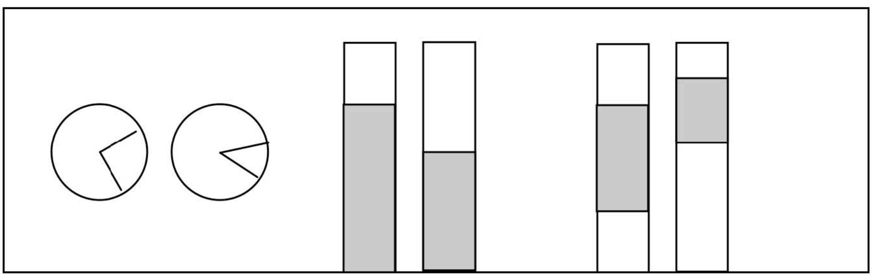

**Fig. 3.** Left to right: pie, bar, and divided bar graphs, of the type used by Simkin and Hastie

Simkin and Hastie developed four elementary processes by which people extract information from graphs: anchoring, scanning, projection, and superimposition. Anchoring involves selecting a portion of the graph as a baseline, or anchor, against which other judgments can be made (e.g., 50% of a bar). Scanning involves moving the eye from the anchor to the edge of the distance to be judged (e.g., from the midpoint of a shaded area to its edge). Projection involves mentally drawing a line from a point in one image to a point in anther. Superimposition involves moving elements of the graph to a new position, to create overlap with other elements in the graph.

In terms of our model, anchoring, scanning, and projection can be considered perceptual processes, at least in these very simple tasks. Although projection involves mentally drawing a line, this extension most likely does not make much, if any, demand on spatial working memory, because of the direct juxtaposition of the start and end points. Superimposition, however, does involve spatial processing, because a spatial element, in which size and/or angle are crucial, must be mentally moved from one spatial location to another. Superimposition is thus a spatial transformation.

Simkin and Hastie's task analysis showed that for the simple bar graph, projection, anchoring, and scanning are required, whereas for the divided bar and pie graphs, superimposition, anchoring, and scanning are needed. In this case, our model predicts that the task would take longer and be less accurate for divided bar and pie graphs than for simple bar graphs, because it involves at least one spatial transformation, and spatial processing is more effortful than purely perceptual processing. This is indeed what Simkin and Hastie found: reaction time was significantly shorter for simple bar than divided bar and pie graphs, though these did not differ significantly from each other, and accuracy for simple bar graphs was greater than for divided bar or pie graphs, though again, these two were not significantly different from each other.

This analysis suggests some validation of our model, in that we are able to match Simkin and Hastie's independently established elementary processes to our distinction between perceptual and spatial processing, and show that graph/task combinations that involve spatial processing were found to be more difficult than one that did not. The analysis also shows that there are simple graph tasks that can be—and have been—performed in the psychology laboratory for which perceptual processing alone cannot account for performance on the task. Although Simkin and Hastie did not analyze elementary graphing processes according to their perceptual or spatial nature, it is clear that at least one of their processes—superimposition—involves spatial cognition. Since Simkin and Hastie's goal in establishing these four elementary codes was to "develop a vocabulary of elementary mental processes that can be combined to build information-processing models of performance in common graph-perception tasks," it would be a simple matter to perform task analyses on other tasks to determine which of the elementary codes (perceptual or spatial) is involved.

## 6 General Discussion

We have proposed that a comprehensive model of graph comprehension needs to include spatial processing as an important component of the graph comprehension process, and we have proposed a model that predicts when spatial processing is required. In accordance with prior models of graph comprehension, we have argued that when information can be extracted directly from the graph, the task can be accomplished by perceptual processing alone. However, we have proposed an extension to those models, such that when information is not explicitly represented and the information cannot be extracted by perceptual processes, spatial processing will likely be used. Whether perceptual processing is sufficient in such cases depends on the interaction of the task and graph. We have suggested that in more complex domains, in which the graphs and tasks are more complex, it is more likely that spatial processing will be required. However, we have also shown that spatial processing can be needed for simple tasks performed on simple graphs, depending on the graph type.

We believe that the reason spatial processing has not been part of graph comprehension models is that the focus on simple tasks and graphs has made it unnecessary, since in general, simple perceptual processes are sufficient to account for performance. Differences in performance on different graph types have been attributed to a better or worse match between the task and the graph. In addition, the strength of graphs as a form of representation is that they can make implicit things explicit (at least, good graphs do), so graphs are designed and selected so that it *is* possible to make direct comparisons between visual chunks. However, as tasks, domains, and visualizations become more complex, this transparency may not always be possible or even desirable (forecasters, for example, don't want simpler graphs, they want many variables represented). As graph comprehension research moves out of the psychology laboratory into the "real world" of practice, it will be more important for graph comprehension models to include spatial processing.

Although it may be rare for simple graph/task combinations to require spatial processing, such cases do exist. In addition to Simkin and Hastie's comparison judgment task, consider the case of a simple two-by-two psychology experiment, represented by a simple bar graph as in Figure 4. In this instance, the interaction is explicit, but in order to determine whether there is a main effect, the spatial strategy of mentally averaging the relevant bar heights is most likely used. Other problem-solving uses of graphs (e.g., (Cheng et al., 2001; Scaife & Rogers, 1996; Tabachneck-Schijf et al., 1997) are likely to involve spatial processing as well.

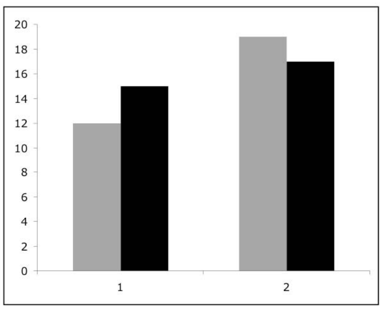


**Fig. 4.** Simple bar graph: Depending on the task, spatial processing may be used

In general, many spatial cognition tasks can be solved by a non-spatial strategy, for example, mathematically. If this strategy is chosen, there is obviously no spatial cognition at work. However, we believe that people will prefer spatial strategies in graph comprehension tasks, in part because the graph is a spatial array, and using spatial cognition does not require "translation" of the spatial code. Our model makes testable predictions as to when spatial processing will be used, and the type of strategy used in these instances where multiple strategies are possible is, in fact, an empirical question.

Our model extension embraces more recent refinements of the generic model discussed earlier. For example, eyetracking data from our studies (Trafton et al., 2002) confirms the iterative nature of perceptual processing (Carpenter & Shah, 1998). Furthermore, expertise appears to be an important factor in spatial processing of graphical information (c.f.,(Freedman & Shah, 2002)—both domain and graph knowledge. Without these two important types of knowledge, it is unlikely that a graph user would be able to generate the necessary spatial transformations to extract the needed information.

Without incorporating spatial processing, we believe current models of graph comprehension will be incomplete. Including spatial processing in our models will help us to understand why some representations might be better than others at a cognitive level, by shedding light on processes that underlie different graph/task interactions. It can help identify situations in which spatial processing is unavoidable and can help us make predictions about performance using these graphs. In some situations, it can help us design better graphs, by developing creative ways to reduce the number of spatial transformations required (for example, by facilitating direct comparisons that can be performed perceptually). In sum, we propose that including spatial processing is an important step in building a comprehensive model of graph comprehension.

## References

Baddeley, A. D. (1999). *Essentials of human memory*. Hove, UK: Psychology Press, Ltd.

Baddeley, A. D., & Liebeman, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 521-539). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bogacz, S., & Trafton, J. G. (2005). Understanding dynamic and static displays: Using images to reason dynamically. *Cognitive Systems Research 6*(4), 312-319.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology:Applied, 4*(2), 75-100.

Cheng, P. C.-H., Cupit, J., & Shadbolt, N. R. (2001). Supporting diagrammatic knowledge acquisition: An ontological analysis of cartesian graphs. *International Journal of Human-Computer Studies, 54*(457-494).

Cleveland, W. S. (1985). *The elements of graphing data*. Monterey, CA: Wadsworth.

DeSanctis, G., & Jarvenpaa, S. L. (1989). Graphical presentation of accounting data for financial forecasting: An experimental investigation. *Accounting, Organizations adn Society, 14*, 509-525.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* (2nd ed.). Cambridge, MA: MIT Press.

Freedman, E. G., & Shah, P. (2002). *Toward a model of knowledge-based graph comprehension.* Paper presented at the Diagrammatic Representation and Inference, Second International Conference, Diagrams 2002, Callaway Gardens, GA, USA.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project ernestine: Validating a goms analysis for predicting and explaining real-world performance. *Human-Computer Interaction, 8*(3), 237-309.

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition: A review. In R. M. Downs & D. Stea (Eds.), *Image and environment: Cognitive mapping and spatial behavior* (pp. 246-295). Chicago: Aldine Publishing Company.

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

Kosslyn, S. M., Sukel, K. E., & Bly, B. M. (1999). Squinting with the mind's eye: Effects of stimulus resolution on imaginal and perceptual comparisons. *Memory and Cognition, 27*(2), 276-287.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction, 8*, 353-388.

Peebles, D., & Cheng, P. C.-H. (2003). Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors, 45*(1), 28-46.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Roth, W.-M., & Bowen, G. M. (2003). When are graphs worth ten thousand words? An expert-expert study. *Journal of the Learning Sciences, 21*(4), 429-473.
Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*(185-213).
Shah, P., Freedman, E. G., & Vekiri, I. (2005). The comprehension of quantiative information in graphical displays. In P. Shah & A. Miyake (Eds.), *Cambridge handbook of visuospatial thinking* (pp. 426-476). New York, NY: Cambridge University Press.
Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.
Simkin, D., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association, 82*(398), 454-465.
Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). Camera: A computational model of multiple representations. *Cognitive Science, 21*(3), 305-350.
Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: Extracting and generating information from complex visualizations. *International Journal of Human Computer Studies, 53*(5), 827-850.
Trafton, J. G., Marshall, S., Mintz, F. E., & Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualizations. In M. Hegarty, B. Meyer & H. Narayanan (Eds.), *Diagramatic representation and inference* (pp. 206-220). Berlin Heidelberg: Springer-Verlag.
Trafton, J. G., Trickett, S., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S. (in press). The relationship between spatial transformations and iconic gestures. *Spatial cogntion and computation*.
Trafton, J. G., & Trickett, S. B. (2001). A new model of graph and visualization usage. In J. D. Moore & K. Stenning (Eds.), *The proceedings of the twenty third annual conference of the cognitive science society* (pp. 1048-1053). Mahwah, NJ: Erlbaum.
Trafton, J. G., Trickett, S. B., & Mintz, F. E. (2005). Connecting internal and external representations: Spatial transformations of scientific visualizations. *Foundations of Science, 10*, 89-106.
Tversky, B. (1995). Cognitive origins of graphic productions. In F. T. Marchese (Ed.), *Understanding images: Finding meaning in digital imagery.* (pp. 29-53). New York: Springer-Verlag.