How Graphs Mediate Analog and Symbolic Representation

Merideth Gattis
Department of Psychology
University of California, Los Angeles
Los Angeles, California 90024-1563
merideth@psych.ucla.edu

Keith J. Holyoak
Department of Psychology
University of California, Los Angeles
Los Angeles, California 90024-1563
holyoak@psych.ucla.edu

Abstract

Three experiments are reported that examine the impact of people's goals and conceptual understanding on graph interpretation, in order to determine how people use graphical representations to evaluate functional dependencies between continuous variables. Subjects made inferences about the relative rate of two continuous linear variables (altitude and temperature). We varied the assignments of variables to axes, the perceived cause-effect relation between the variables, and the causal status of the variable being queried. The most striking finding was that accuracy was greater when the Slope-Mapping Constraint was honored, which requires that the variable being queried -- usually the effect or dependent variable, but potentially the cause instead -- is assigned to the vertical axis, so that steeper lines map to faster changes in the queried variable. This constraint dominates when it conflicts with others, such as preserving the low-level mapping of altitude onto the vertical axis. Our findings emphasize the basic conclusion that graphs are not pictures, but rather symbolic systems for representing higher-order relations. We propose that graphs provide external instantiations of intermediate mental representations, which enable people to move from pictorial representations to abstractions through the use of natural mappings between perceptual properties and conceptual relations.

Introduction

Graphs provide an external representational medium that connects analog and symbolic internal representations. Although graphs present information spatially, they are more abstract than simple pictorial representations. Graphs are able to omit non-essential details and to highlight central higher-order relations, such as rate of change in a variable, based on a mapping from visuospatial properties of the graph to symbolically-interpreted concepts. In addition to being more complex than pictures, graphs are simpler than other symbolic systems. By engaging built-in perceptual processing mechanisms, graphs function at an intermediate level of abstraction. Graphs mediate between analog and symbolic representations by exploiting natural properties of the visual system in at least two ways. First, graphs permit computational efficiency for abstract problems by reducing the number of solution steps that sentential or mathematical procedures would require (Larkin & Simon, 1987) or by reducing the number of dimensions or variables that must be explicitly compared. Utilizing the visual dimension to express information allows graphs to coordinate information on two or more dimensions. Integrating information across dimensions normally imposes a heavy cognitive load, but graphs reduce this load by integrating values on dimensions via points or lines which simultaneously represent a value or set of values on more than one dimension. This visual integration, or chunking, allows us to reason about relations between two or more sets of data on those two or more dimensions.

Second, graphs also exploit vision by utilizing perceptual properties to convey conceptual relations. The "natural" mapping between percepts and concepts is so automatic for some pairings, such as "greater area (more dots, thicker bars, or higher lines) equals greater quantity," that we often fail to recognize the relationship as a mapping rather than as an inherent equivalence of meaning. Other natural mappings involve more abstract pairings of perceptual and conceptual relations, such as "steeper equals faster," which is the focus of the present experiments. Mappings between perceptual and conceptual relations allow graphs to support reasoning about abstract concepts that are difficult to grasp directly, such as slope and function.

Psychological research on graphs has been sparse, and has primarily focused on encoding processes, but all reported work seems to reflect a consistent theme of the tight linkage found in graphs between perceptual properties and conceptual relations. Pinker (1990) reported several experiments by himself and Simcox indicating that global trends are best represented by global features readily encoded by the visual system, while local values are best represented by local features. Tversky and Schiano (1989) and Schiano and Tversky (1992) found evidence that the rules used for encoding graphs are not simply properties of the visual system, but rather result from interpreting a figure as a graph. They demonstrated that the label given to a diagrammatic representation invokes a particular set of rules or biases for encoding. Calling the representation a "graph" created a particular perceptual bias (the interpretation of a line orientation was biased toward 45 degrees), whereas calling it a "map" or a "figure" either did not create any bias or biased the interpretation away from all orienting angles (including 45 degrees). These results indicate that graphs constitute a symbol system that is associated with a specialized set of rules connecting the perceptual to the conceptual.

In the present study we examine the impact of people's goals and conceptual understanding on graph interpretation in order to determine how graphs are used to connect analog and symbolic representations. The central question we address concerns how people use graphical representations to evaluate functional dependencies between continuous variables. A key problem is that different constraints on graphical representations may be in conflict. In particular, there may be tension between the goal of preserving low-level pictorial correspondences between the domain being represented and the visuospatial display, versus the need to capture systematic correspondences between higher-order relations in the represented domain and visuospatial relations in the display. Figure 1 illustrates this problem with a graph typical of those used in teaching atmospheric science, which represents functional dependencies between altitude and temperature. Textbooks and teachers invariably plot such altitude/temperature graphs with altitude on the y axis, because this preserves the low-level, pictorial correspondence between *up* in the world and *up* in the graph.

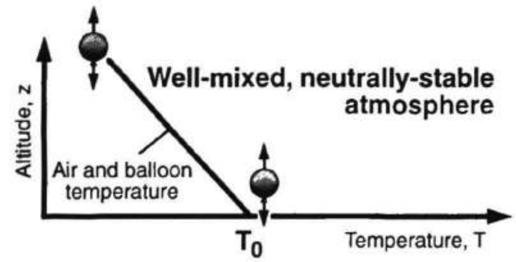

Figure 1. A typical altitude/temperature graph used in atmospheric science.

However, the cost of this decision is that the functional dependency between altitude (the causal variable) and temperature (the effect) violates cognitive constraints derived from the analogical relationship between the rate of change in the effect (dependent variable) and the slope of a line. Specifically, if the independent variable (IV) is mapped to the x axis, and the dependent variable (DV) to the y axis, then rate of change in DV with respect to IV maps to rate of change in y with respect to x (i.e., the slope of the line). This assignment, a graphing convention that we term the Slope-Mapping Constraint, ensures that judgments about rates can be based on the visually transparent mapping *steeper = faster*. However, the standard pictorial-based assignment of altitude (IV) to the vertical axis violates the Slope Mapping Constraint (since rate of change in the DV will equal the reciprocal of the slope, instead of the slope). Our experiments utilize conflicts between pictorial and symbolic representations to reveal component processes of graph interpretation.

Experiment 1: Conflicting Constraints

Method

Subjects: Sixty-six undergraduate psychology students at the University of California, Los Angeles (UCLA) participated in exchange for course credit. Approximately half of the subjects were science majors and half were non-science majors, distributed randomly across conditions.

Materials, Procedure, and Design: Subjects were given graphs representing the functional dependency between atmospheric altitude and air temperature, along with a brief explanation of the relationship. This dependency is a loose causal relation in that altitude does not directly cause temperature; however, an interaction of several factors, including altitude, determines air temperature. In the case of a moving air parcel, the change in altitude causes a change in temperature. The graphs either represented altitude on the y axis in accordance with the atmospheric-science tradition of preserving verticality, but in violation of the Slope-Mapping Constraint (Figure 2A); or on the x axis, in accordance with the graphing convention of mapping the IV to the x axis that follows from the Slope-Mapping Constraint (Figure 2B). Because of the relation displayed in these graphs, altitude was the IV and temperature the DV. On the first page, a single data line was drawn across the graph representing a single value of y for each value of x. The line displayed an inverse linear relationship between altitude and temperature, with higher temperatures at lower altitudes and lower temperatures at higher altitudes. All subjects were first asked to identify a particular point on the graph, and then to make a judgment about the direction of change for one of the variables represented in the graph. These questions were used to screen out subjects who were not capable of basic graph interpretation.

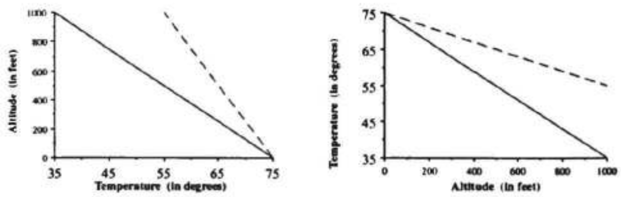

Figure 2. Graphs that violate (Panel A) or conform with (Panel B) the Slope-Mapping Constraint.

On the second page subjects were shown another graph and asked to perform one of two rate-judgment tasks. In the drawing task, subjects were given the original graph accompanied by a description of a second data set in which temperature changed more slowly relative to altitude, and were asked to draw a line that could represent such a data set. In the multiple-choice task, subjects were given the original graph with the addition of a dotted second line, and were asked to judge whether the second set of data represented a slower or faster rate of change of temperature than the first.

Subjects were assigned in approximately equal numbers to four conditions, defined by the factorial combination of axis assignment (altitude on x or else y axis) and task (drawing or multiple choice). Assigning altitude (IV) to the y axis preserved the pictorial correspondence *up = up*, at the cost of violating the Slope-Mapping Constraint of *steeper = faster*, whereas assigning altitude to the y axis reversed these properties. We predicted that judging rate of change would be more dependent on the higher-order mapping than on the pictorial correspondence, so that accuracy would be higher when altitude was on the x axis.

Results and Discussion

As predicted, subjects were more accurate in both rate tasks when the IV (altitude) was assigned to the x axis. All analyses of solution frequencies were based on the G^2 statistic (maximum likelihood chi square). For the drawing task, the percentage of subjects who drew a line showing the correct rate relation was 78% when altitude was on x and 33% when it was on y, G^2 (1) = 5.60, $p < .025$. Similarly, the percentages of correct choices in the multiple-choice task by subjects in the two conditions were 100% and 64%, respectively, G^2 (1) = 6.86, $p < .01$. These results reveal that when the two pressures conflict in a graphical representation of rate of change, preserving the pictorial relation of verticality is less important than preserving the more abstract correspondence between steepness of lines and speed of rate change.

Experiment 2: Testing Alternative Interpretations

Experiment 2 was designed to distinguish the Slope-Mapping Constraint from other factors with which it was correlated in Experiment 1. We examined whether the results depended on existence of a conceptual relation of cause and effect between the IV and DV, and whether there is a preference for a particular causal relation.

Method

Subjects: One hundred thirty-six students in undergraduate psychology courses at UCLA served as subjects in exchange for course credit.

Materials, Procedure and Design: The two graphs used in Experiment 1 were also used in Experiment 2. In the causal condition the x and y axes were labelled in the same way as in Experiment 1. In the non-causal condition the x and y axes were labelled simply as "A" or "B". In order to manipulate the causal direction between altitude and temperature, two cover stories were created for the causal condition. One was the version used in Experiment 1, in which altitude loosely causes temperature. A second version described the movement of a hot air balloon, a scenario in which atmospheric air temperature (relative to a constant balloon air temperature) causes changes in altitude (of the balloon). The functional relations described in the latter story were nearly identical to those described in the former, except that the causal direction was reversed. In the case of a hot air balloon, change in the air temperature loosely causes change in the altitude of the balloon. No cover story was used for the non-causal conditions.

Subjects were assigned to one of eight conditions. Four conditions received one of the two causal cover stories (making either altitude or temperature the cause), in each case varying which variable was assigned to the x axis. The other four conditions did not receive any causal cover story, so that the graph simply represented a functional dependency. Subjects in these non-causal conditions were given graphs identical to those in the four causal conditions but without the cover story or the altitude and temperature labels. The x and y axes were labelled "A" and "B", with the label assignments counterbalanced across each condition.

The two graph interpretation questions used in Experiment 1 were presented on the first page, along with a graph of a single set of data. On the second page all subjects were presented with that same graph together with an additional set of data represented as a dotted line. As in the multiple-choice task used in Experiment 1, subjects were instructed to choose the correct rate comparison description for the second data line from two possibilities (faster or slower than the first line). This question always probed the rate of change in the DV (the effect in the causal conditions) with respect to the IV. In the noncausal conditions, the DV was defined as the variable (A or B) being queried. The more difficult drawing task was not used.

The design was such that among the four causal conditions, two assigned the cause (IV) to the x axis, in accord with the Slope-Mapping Constraint, and two made the reverse assignment. When altitude was the cause, the Slope-Mapping Constraint conflicted with the pictorial correspondence of *up = up*, as was the case in Experiment 1; whereas when altitude was the effect the two pressures converged. Among the four matched non-causal conditions, two satisfied the Slope-Mapping Constraint (IV on the y axis) and two violated it.

Results and Discussion

As in Experiment 1, subjects made rate judgments more accurately when the Slope-Mapping Constraint was satisfied (DV on y axis) than when it was not, 93% versus 46%, G^2 (1) = 39.6, $p <. 0001$. This effect held regardless of whether a casual cover story was provided, and whether altitude or temperature was the cause. These results thus rule out all of these alternative interpretations of the results obtained in Experiment 1.

Experiment 3: Querying Cause Versus Effect

Although the results of Experiment 2 did not reveal any influence of a causal cover story on rate judgments, it would be premature to rule out any role of causal relations in graph interpretation. Prior experience with graphs in which the "cause" variable was always placed on the x axis could unconsciously carry over to graphs without any explicit causal interpretation, so that subjects may have assumed that the IV in the non-causal condition was in fact a cause. In addition, Experiments 1 and 2 only tested subjects' ability to make inferences about the rate of change in an effect as its cause varied. This type of question involves a *predictive* inference from a cause to its effect. However, people are also capable of making *diagnostic* inferences, judging variations in a cause by variations in its effect (Waldmann & Holyoak, 1992). Assuming that people tend to code causal knowledge in the cause-to-effect direction, diagnostic inferences that depend on "backwards" reasoning may be more difficult, especially when cast as judgments about higher-order relations such as rate of change. It is therefore possible that people will be especially sensitive to violations of graphing conventions in the context of judging the rate at which a cause changes with its effect. Accordingly, Experiment 3 asked subjects to reason about a change in cause as a function of a change in effect, as well as the opposite case. Experiment 3 examined the combined influence of three factors on graph interpretation: the Slope-Mapping Constraint, preserving pictorial correspondences with the real world, and a possible assymmetry in causal inference.

Method

Subjects: One hundred thirty-nine students in undergraduate psychology courses at UCLA served as subjects in this experiment in exchange for course credit.

Materials, Procedure, and Design: The four causal conditions from Experiments 2 were repeated in this experiment, with the same graphs and axes labelling. The two cover stories used in Experiment 2 to manipulate the causal direction between altitude and temperature were again used in Experiment 3, with slight changes to create a context in which subjects would find it plausible to be asked about the change in the cause as a function of the change in the effect. The multiple-choice task used in the previous experiments was used once again in Experiment 3.

Subjects were assigned to one of eight conditions, defined by the factorial combination of axis queried (x versus y), causal status of the queried variable (cause or effect), and the variable interpreted as the effect (altitude or temperature). Experiment 3 was identical to Experiment 2 in all other respects.

Results and Discussion

Table 1 presents the percentages of correct rate judgments for each of the eight conditions. As in the previous experiments, accuracy was higher when the queried variable was assigned to the y axis rather than the x axis, 92% versus 54%, G^2 (1) = 26.2, $p < .0001$. Overall, this effect did not vary with the identity or causal status of the queried variable. This result thus yields a further generalization of the Slope-Mapping Constraint. If we follow the conventional interpretation of the cause as the IV and the effect as the DV, the present findings demonstrate that it is not essential for the IV to be assigned to the x axis and the DV to the y axis. Rather, the crucial requirement for satisfying the Slope-Mapping Constraint is that the *variable being queried* (whether DV or IV) be assigned to the y axis.

Overall, querying the cause did not significantly reduce accuracy relative to querying the effect. The interpretation of relative rate in a graphical representation did not appear to depend on whether the question required predictive reasoning (from cause to effect) or diagnostic reasoning (from effect to cause).

Table 1. Percentage correct rate judgments in each condition of Experiment 3.

Rate comparison about	Causal relation	Axis Assignment	% correct
Effect	Altitude is effect	Altitude on x	64
		Altitude on y	89
	Temp. is effect	Temp. on y	90
		Temp. on x	61
Cause	Temp. is cause	Temp. on y	93
		Temp. on x	72
	Altitude is cause	Altitude on x	22
		Altitude on y	95

General Discussion

The results of the present study help to illuminate the component processes in graph interpretation, and the manner in which these external representations serve to mediate between analog and symbolic internal representations. We found that the ease of drawing inferences about rate of change depended on the nature of the mapping between an analog conceptual dimension (rate of change in a continuous variable) and the analog visuospatial dimension of slope of a line. Furthermore, the optimal mapping was dependent on the reasoner's goal. Our most consistent and powerful finding was that accuracy was greater when the Slope-Mapping Constraint was honored, which requires that the variable being queried -- usually the effect or dependent variable, but potentially the cause instead -- is assigned to the vertical axis, so that steeper lines map to faster changes in the queried variable. This constraint dominates when it conflicts with others, such as preserving the low-level mapping of altitude onto the vertical axis.

These results emphasize the basic conclusion that graphs are not pictures. Rather, they are symbolic systems for representing higher-order relations. Graphs mediate between these two representational levels by utilizing visuospatial relations to convey conceptual information. Sometimes this dependency of conceptual expression on perceptual properties can lead to confusion between graphs and pictures (Clement, 1990), but for the most part it is an efficient representational mechanism. Our findings have pedagogical implications, in that they demonstrate that the near-universal textbook convention of plotting altitude on the vertical axis, thereby preserving a pictorial mapping, is extremely detrimental to the solution of problems that require inferences about relative rates.

An important question for future research is to determine whether the Slope-Mapping Constraint observed here is based on a deep property of the cognitive system, or whether it develops as a product of experience with graphs and the conventions of axis assignment. Lakoff and Johnson (1980) have argued that the metaphorical mapping between the vertical dimension of space and the "upward" direction of abstract dimensions is likely to be a cognitive universal. From this perspective, our results show that when we must choose between maintaining this spatial mapping either at a pictorial level (*up = up*) or at the level of more abstract relations (*steeper = faster*), the latter mapping is more important for inferring rates. If this view is correct, we would expect that the nature of the cognitive system has led to this graphing convention, which is non-arbitrary.

This issue is related to the larger question of whether other perceptual-conceptual linkages found in graphs arise from mental representations that mediate the pictorial and the symbolic. Graph theorists have fostered the belief that rules for graphing are the products of explicit design decisions by those schooled in the art and science of representation, imagery, or symbols (e.g., statisticians, graphical designers, and semioticians). But perhaps in constructing graphs people have basically been enhancing natural representations that already exist in the human mind. Evaluating the latter possibility will require searching for other rules of graphing that may depend on analogical mappings between perceptual properties and conceptual relations. It will also be necessary to investigate the differences between relatively direct visual displays, such as pictures, drawings, figures, maps, and some diagrams, and more abstract visual displays, such as graphs and mental models. In addition, developmental studies of the relative ease of learning to interpret graphs under alternative conventions, such as those for assignment of axes, may cast light on the origins of "natural" rules for graph construction.

Acknowledgments

This research was supported by NSF Grants USE-9150811 and SBR-9310614.

References

Clement, J. (1989). The concept of variation and misconceptions in cartesian graphing. *Focus on Learning Problems in Mathematics, 11*, 77-87.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: Chicago University Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp.73-126). Hillsdale, NJ: Erlbaum.

Schiano, D. J., & Tversky, B. (1992). Structure and strategy in encoding simplified graphs. *Memory & Cognition, 20*, 12-20.

Tversky, B., & Schiano, D. J. (1989). Perceptual and cognitive factors in distortions in memory for graphs and maps. *Journal of Experimental Psychology: General, 118*, 387-398.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222-236.