

Running Head: Shedding Light on the Graph Schema

Shedding Light on the Graph Schema: Perceptual Features vs. Invariant Structure

Raj M. Ratwani

George Mason University

J. Gregory Trafton

Naval Research Laboratory

Document Statistics:

Abstract (p.2): 146 words

Body, references, and author note (pp. 3-19): 4,085 words

Tables (pp. 20-22): 2

Figures (pp. 23-24): 1

Correspondance:

Raj M. Ratwani
Department of Psychology
George Mason University
Fairfax, VA 22030

310-483-9494 (voice)

703-993-1359 (fax)

rratwani@gmu.edu

Abstract

Most theories of graph comprehension posit the existence of a graph schema to account for people's prior knowledge of how to understand different graph types. The graph schema is, however, a purely theoretical construct: there are no empirical studies that have explicitly examined the nature of the graph schema. We sought to determine whether graph schemas are based on perceptual features or on a common invariant structure shared between certain graphs. The process of activating the graph schema was isolated as participants responded to graphs presented in pure and mixed blocks. Any differences in reaction time between the blocks could be attributed to loading the appropriate schema. Results from a series of experiments using five types of graphs suggest graph schemas are based on the graphical framework, a common invariant structure among certain types of graphs. These results provide insight into the comprehension of novel graphs.

Shedding Light on the Graph Schema: Perceptual Features vs. Invariant Structure

Graphs provide useful and efficient ways in which to display information.

When asked to extract information from a graph, people generally have some stored knowledge which is used to comprehend the graph, despite the fact that various graph types represent information differently. How do theories of graph comprehension account for our prior knowledge of how to interpret different graphs?

By focusing on how specific information is extracted from a graph, several theories of graph comprehension have been developed (Lewandowsky & Behrens, 1999; Lohse, 1993; Peebles & Cheng, 2001, 2003; Pinker, 1990; Shah & Carpenter, 1995; Simkin & Hastie, 1987); collectively these theories suggest the following processes: (1) Construct relationships among the graph elements using early visual processes, (2) Build a propositional representation, (3) *Activate the graph schema*, (4) Devise the conceptual question (determine desired information), and (5) Devise the conceptual message (the answer). Activation of the graph schema (step 3) is the critical process accounting for the prior knowledge and specific operations used to comprehend the graph. The schema is a mental representation in long term memory activated by matching the early visual input to a stored mental representation used to interpret the graph.

Although researchers agree upon the graph schemas as a construct, there is little to no agreement on how it is organized and there is little to no empirical evidence investigating the defining features of the schema. This lack of empirical research is partially because there has not been a strong methodology to examine

such a construct. The goal of this study is to begin to understand the features that underlie the graph schema, and to illustrate a methodology that can be used to understand schemas in other domains.

Differing Views of the Graph Schema

Although there is no consensus on the nature of the graph schema, the graph comprehension theories coupled with support from other research domains leads to two distinct hypotheses: a *perceptual feature view* and an *invariant structure view*. The perceptual feature view suggests that surface level features and distinct perceptual properties of a graph determine the schema (Lohse, 1993). In Lohse's model of graph interpretation, each unique type of graph instantiates a different schema and consequently a different set of procedures to interpret the graph. For example, bar graphs, line graphs, and pie charts would all activate different schemas from one another. The scene perception literature, which shares many characteristics with graph comprehension in regards to time scale and the process of matching a visual array to a stored mental representation (Potter, 1993), also provides support for this view. Theories of scene perception suggest perceptual characteristics, such as the identity of select key objects, may drive activation of the appropriate schema (Henderson & Hollingworth, 1999). The most visible and distinguishing perceptual characteristic of a graph is the graphical pattern (Kosslyn, 1989), thus, this may be the underlying feature that defines the graph schema. The graphical pattern is the pictorial object of the graph, for example, the actual lines in a line graph or the actual bars in a bar graph.

The invariant structure view suggests certain broad categories or types of graphs may rely on the same schema; however, the defining features of these categories that determine which graphs rely on the same schema are not clear. Pinker (1990) suggested graph schemas are organized hierarchically; there is a general graph schema which contains common properties of all graphs, and graph specific schemas for commonly encountered graph types. Peebles and Cheng (2001; 2003) suggest that certain graphs that share similar features may rely on a similar graph schema; however, Peebles and Cheng did not specify those features. The notion of certain graphs sharing similar schemas based on a common set of properties is similar to the problem-type schemata in the problem solving literature (Bernardo, 1994). Thus, a common invariant structure among certain graphs may be the underlying principle that determines the graph schema.

We suggest that the graphical framework may be the common invariant structure. The graphical framework is the structural component of a graph that often represents the basic concepts and operations for extracting information from the graph (Kosslyn, 1989). For example, looking at Figure 1, in the line graphs, bar graphs and horizontal bar graphs the framework is represented by the x and y axes; these axes represent the Cartesian coordinate system upon which these graphs are based. Similarly, in the pie charts and doughnut graphs the circle encompassing each graph is the framework; this structure represents the Polar coordinate system upon which these graphs are based. This framework may be the critical structural feature that determines the schema. Line and bar graphs may have different graphical patterns, yet they share the same Cartesian coordinate

framework and may share some underlying mental representations. Likewise, Polar coordinate based graphs may rely on a common mental representation.

The Mixing Costs Paradigm

In order to distinguish between these two views, different graphs were examined in a mixing costs paradigm (Los, 1996). This paradigm consists of presenting stimuli (graphs in this case) in both pure and mixed blocks and then comparing reaction times across the different blocks. A reaction time difference between the pure and mixed blocks can be attributed to a mechanistic difference in processing the stimuli. In this study, we attempted to isolate the activation of the graph schema as the only difference between stimuli.

To examine the graph schema, we held all of the stages of processing for each graph constant, except for the activation of the graph schema. Early visual processing and the construction of the propositional representation of the graphs (i.e. steps 1-2 from the task analytic theories) were equated by examining the same graph type across the pure and mixed blocks. Although these processes may vary slightly based on the orientation of the graphical pattern, these differences should not impact the mean reaction time. The conceptual question (step 4) was equated for all graphs by asking the same question for each graph, and although the conceptual message (step 5) may slightly differ from graph to graph, reaction time differences should be minimal. Thus, the activation of the graph schema is the only step that may vary depending on the graph being examined and the cost of activating the appropriate schema should be apparent using the mixing costs paradigm.

Five graph types were examined: line graphs, bar graphs, pie charts, doughnut graphs, and horizontal bar graphs. Two graphs were examined at a time to determine what characteristics underlie the graph schema. If a given set of graphs rely on different schemas, it should take more time to load the appropriate schema in the mixed block as compared to the pure block, resulting in a time cost. In the mixed block, the appropriate schema needs to be activated resulting in a longer reaction time as compared to the pure block where a new schema does not have to be activated (because it was activated on previous trials). If, however, the graphs being examined rely on a similar schema, the appropriate schema is already activated in the mixed block just as it is in the pure block, resulting in no time cost. According to the perceptual feature view, pairs of graphs with unique graphical patterns should rely on different schemas and result in time costs. The invariant structure view suggests that graphs with a common graphical framework will share similar schemas resulting in no time costs.

General Method

Participants

Each experiment included 18-31 undergraduate students (see Table 1 for number of participants by experiment). The participants were primarily first or second year psychology students participating for course credit.

Materials

Two graph types were tested in each experiment; each experiment contained 80 samples of each graph type resulting in a total of 160 graphs per experiment. Each randomly generated graph depicted the number of “Widgets”

located in three different trays (A, B, and C), ranging from 1-9. In the line and bar graphs, the x-axis contained labels for the three different trays, and the y-axis was a scale ranging from 0-10; in the horizontal bar graphs the axes were reversed. The order of tray labels was randomized for every graph. The doughnut graphs and pie charts contained a legend that assigned each tray (A, B, or C) to a specific colored segment (black, white or gray), and a number between 1-9 was assigned to each segment to indicate the number of widgets it represented. Examples of each of the graphs can be found in Figure 1. The color and tray association was randomized in every graph. To hold the conceptual question constant, participants were asked, “How many Widgets are there in Tray B?” for each graph. This type of extraction question was used because most of the theories of graph comprehension were based on this type of question (Lohse, 1993; Pinker, 1990). In addition to the graphs, 80 text sentences (e.g. There were five ships in the bay.) which required participants to enter the number appearing in the sentence served as a filler task.

Design

In each experiment, two different graphs and text were presented in six different blocks; each block contained forty stimuli. There were three pure blocks with all the same stimuli type (2 pure graph, 1 pure text) and three mixed blocks. One mixed block contained twenty of each graph type; the other two mixed blocks contained twenty of each graph type with twenty sentences. The stimuli in each block were randomly ordered. The three blocks with text were filler tasks and not analyzed resulting in three blocks of interest: two pure graph and one

mixed graph. For example, in Experiment 1a where line and bar graphs were examined, the blocks of interest were a pure block of line graphs, a pure block of pie charts and a mixed block of line graphs and pie charts.

Procedure

Blocks were presented using a Latin squares design. Stimuli were presented on a computer screen and reaction time (RT) data were collected. Participants were asked to respond to each stimulus as quickly and accurately as possible using the keypad. Upon responding, the next stimulus appeared immediately. Three practice trials were completed prior to each block.

Data Analysis

The two pure blocks of graphs were used to calculate two means by averaging the RT of all forty stimuli in each block, respectively. The mixed block of graphs was segmented by trial type and graph yielding switch trials and non-switch trials (Los, 1999). The switch trials consisted of the RT to a particular kind of graph when it was preceded by a different kind of graph (e.g. line preceded by pie), whereas the non-switch trials consisted of the RT to a particular graph that was preceded by the same kind of graph (e.g. line preceded by line). For example, in Experiment 1a where line and pie charts were examined, analysis of the pure blocks resulted in a pure line and pure pie mean. Analysis of the mixed block resulted in four means: switch line, switch pie, non-switch line, and non-switch pie.

The pure block RT means reflect a highly activated schema for the particular graph being viewed since the same kind of graph is seen throughout the

block. The mixed block switch RT means reflect the process of having to activate the appropriate mental representation required to process the particular graph being viewed since the previous graph viewed was different. Thus, the switch means capture the time cost of activating the appropriate schema. The non-switch RT means reflect a more activated schema as compared to the switch RT means since the previous graph viewed is consistent with the current graph. For the purposes of this study we focused on the pure means and switch means since we wanted to compare highly activated schemas to situations where the schema needed to be activated. This way of calculating time costs is different from the task switching literature (Monsell, 2003)¹. Table 1 shows the pure, switch and non-switch means for all experiments.

In each experiment, the graph and block type were compared using a two-way ANOVA. A main effect of graph reflects differences in the time required to process a particular graph. An effect of block type reflects the time costs between the pure and mixed blocks. There were no significant interactions in any of the experiments. Incorrect responses and RTs greater than three standard deviations from the mean (less than 5% of the data) were excluded. Table 2 shows the actual time costs for pairs of graphs across all experiments.

Experiments 1 a,b,c

¹ In the task switching literature (Monsell, 2003) mixing costs are calculated by taking the difference between pure and non-switch means; we have used pure and switch means. We have run our analyses both ways and the results are consistent in all but one case.

Line graphs, bar graphs and pie charts were examined in the first set of experiments. Experiment 1a compared line graphs to pie charts and Experiment 1b compared bar graphs to pie charts. Both the perceptual feature and invariant structure view suggest different mental representations for each experiment since graphs with different patterns and frameworks (i.e. Cartesian and Polar coordinate) were compared. Thus, there should be time costs, reflected as a significant block type effect, in each experiment as well as a significant graph effect.

In Experiment 1c, line and bar graphs were compared. The perceptual feature view suggests time costs, there should be a significant block type effect, whereas the invariant structure view suggests no time costs since line and bar graphs share the same Cartesian framework. Because bar graphs are best for extracting discrete data (Zacks & Tversky, 1999), they should be faster than line graphs, reflected in a significant graph effect.

Results

1a. Line and Pie. Responses to line graphs were significantly faster than to pie charts, $F(1,19) = 19.1$, $MSE = 47562$, $p < .001$. The main effect of block type was significant, $F(1,19) = 12.8$, $MSE = 23459$, $p < .01$; the mixed blocks were slower than the pure blocks.

1b. Bar and Pie. Responses to bar graphs were significantly faster than to pie charts, $F(1,21) = 115.2.1$, $MSE = 20142$, $p < .001$. The main effect of block type was significant, $F(1,21) = 14.3$, $MSE = 33256$, $p < .01$; the mixed blocks were slower than the pure blocks.

1c. Line and. Bar. Responses to bar graphs were significantly faster than to line graphs, $F(1,20) = 38.02$, $MSE = 12532$, $p < .001$. The main effect of block type, however, was non-significant, $F(1,20) = .15$, $MSE = 30391$, $p = .7$; there was no difference between the pure and mixed blocks. Based on a power analysis the probability of detecting an effect was greater than 85%, suggesting that the null effect is not likely to be an issue of power.

Discussion

Each experiment demonstrated a significant effect of graph, suggesting different overall processing times for each graph. Further, given the task of extracting specific values, the rank ordering of processing times (i.e. bar graphs being the fastest, followed by line and then pie) is in agreement with the graph comprehension research that suggests bar graphs are best for extracting discrete values (Zacks & Tversky, 1999).

The block type effects in Experiments 1a and 1b suggests that both line and bar graphs may rely on different mental representations than pie charts. In Experiment 1c when two Cartesian coordinate graphs were examined there were no time costs. Together, these results provide support for the invariant structure view and the role of the graphical framework in determining the schema. In Experiment 2 we sought to replicate these findings with different types of graphs.

Experiment 2 a,b,c

In order to further test the invariant structure view, doughnut graphs were examined, a Polar coordinate based graph type that relies on the same framework as pie charts. If the graphical framework determines the schema, doughnut graphs

paired with line graphs (Experiment 2a) and bar graphs (2b) should incur time costs since these combinations pair graphs with different frameworks. However, doughnut graphs paired with pie charts (2c) should not incur time costs since both graphs rely on the same framework.

Results

2a. Line and Doughnut. Responses to line graphs were significantly faster than to doughnut graphs, $F(1,30) = 37.1$, $MSE = 31065$, $p < .01$. The main effect of block type was significant, $F(1,30) = 9.3$, $MSE = 36178$, $p < .01$; the mixed blocks were slower than the pure blocks.

2b. Bar and Doughnut. Responses to bar graphs were significantly faster than to doughnut graphs, $F(1,22) = 61.8$, $MSE = 53795$, $p < .001$. The main effect of block type was significant, $F(1, 22) = 15.3$, $MSE = 30852$, $p < .01$; the mixed blocks were slower than the pure blocks.

2c. Pie and Doughnut. Responses to pie charts were significantly faster than to doughnut graphs, $F(1,20) = 9.1$, $MSE = 26090$, $p < .01$. The main effect of block type was non-significant, $F(1,20) = .75$, $MSE = 19610$, $p = .4$; there was no difference between the pure blocks and mixed blocks. Based on a power analysis, the probability of detecting an effect was greater than 85%, suggesting that the null effect is not likely to be an issue of power.

Discussion

The results of Experiment 2 further support the invariant structure view. In Experiments 2a-b when Polar and Cartesian coordinate graphs were examined in the same block they resulted in time costs. In Experiment 2c when two Polar

coordinate graphs were examined there were no time costs. Similar to Experiment 1, the graph effect was significant in all cases suggesting a difference in processing times for each graph.

Although Experiments 1 and 2 provide support for the invariant structure view, one possible explanation for the null block type effect in Experiments 1c and 2c is that these graphs required the same visual procedures to extract information. For example, in Experiment 1c, participants may have realized they can look to the x-axis to find the desired tray, then scan up and over to the y-axis regardless of the graph being presented. Perhaps the absence of a block type effect was because the same visual procedures were being used and not because of the shared mental representation. While the consistent significant graph effects in each experiment provides some evidence against this argument, Experiment 3 specifically examined this issue.

Experiment 3

To examine whether the null block type effect in Experiments 1c and 2c was due to the same visual procedures, graphs that share the same framework but required different visual procedures were tested. Line graphs and horizontal bar graphs both share the same graphical framework; however, a different set of visual procedures is required to extract information. Horizontal bar graphs require one to first look to the y-axis as opposed to the x-axis in line graphs. If it is the graphical framework that is the underlying structural component that determines the schema, there should be no time costs.

Results and Discussion

Responses to horizontal bar graphs were faster than to line graphs, $F(1,17) = 15.9$, $MSE = 10784$, $p < .001$. The main effect of block type was non-significant, $F(1,17) = .08$, $MSE = 25634$, $p = .8$; there was no difference between the pure blocks and the mixed blocks. Based on a power analysis the probability of detecting an effect was greater than 85%, suggesting that the null effect is not likely to be an issue of power.

The main effect of graph suggests different processing times for each graph. No time costs were incurred in this experiment, despite the fact that different visual procedures were required to extract information from the graphs. This suggests that using the same visual procedures does not account for the null effects observed in previous experiments; we believe it is the activation of the graph schema (or lack thereof) that accounts for the block type effects found in these experiments. The results of this experiment provide further evidence for the invariant structure view; the line and horizontal bar graphs share a common framework and consequently rely on a similar mental representation.

General Discussion

Using the mixing costs paradigm, we have systematically examined the reaction times to different graphs and have shown that the graph schema seems to be based on the invariant structure shared by certain categories of graphs. Specifically, it is the graphical framework that is the invariant feature of graphs that determines the schema; graphs that share the same framework rely on a similar mental representation. The framework represents the conceptual knowledge necessary to extract information from the graph. We argue that once

the graphical framework for a particular graph is perceptually identified, it is matched to the mental representation for the particular graph type (e.g. bar graph is matched to Cartesian coordinate), thus activating the concepts and operations required to extract information.

Our view of the graph schema is not far from that of Pinker (1990) and Peebles and Cheng (2001). Pinker suggested a hierarchical organization to the graph schema and our data align with that view. Our results suggest that a critical component of this hierarchy is the graphical framework. Further, we have more clearly specified the Peebles and Cheng description of the graph schema by illustrating that the framework is the property that determines the schema.

Our results provide a foundation for understanding graph schemas; however, one of the limitations of this study was that the same specific extraction question was asked in each experiment. Although this was purposefully done to isolate the activation of the graph schema, the role of schemas in more complex graphs with more complex questions requires further investigation. The mixing costs methodology proved to be a powerful technique to isolate the activation of the graph schema in these experiments; a similar methodology can be used to extend research on graph schemas and to investigate mental representations in other domains.

An obvious question that arises is how people are able to comprehend novel graphs for which they do not have a stored mental representation. We believe two processes are at work to comprehend novel graphs. Upon examining the graph, partial matching between the early visual array and a stored mental

representation may occur. This partial matching will activate a set of operations that may be applicable to the graph. In addition to a partial matching mechanism, similarities in relational structure between concepts and the novel graph are likely to play a role in interpreting novel graphs (Gattis, 2002).

References

- Bernardo, A.B.I. (1994). Problem specific information and the development of problem type schemata. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 379-395.
- Gattis, M. (2002). Structure mapping in spatial reasoning. *Cognitive Development*, 17, 1157-1183.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243-271.
- Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185-225.
- Lewandowsky, S., & Behrens, J. T. (1999). Statistical graphs and maps. In F. T. Durso (Ed.), *Handbook of applied cognition* (pp. 513-549).
- Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction*, 8, 353-388.
- Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica*, 94(2), 145-188.
- Los, S. A. (1999). Identifying stimuli of different perceptual categories in mixed blocks of trials: evidence for cost in switching between computational processes. *Journal of Experimental Psychology: Human Perception and Performance*, 25 (1), 3-23.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Science*, 7, 13-140.

- Peebles, D., & Cheng, P. C.-H. (2001). Graph based reasoning: From task analysis to cognitive explanation. In *Proceedings of the twenty-third annual conference of the cognitive science society*: Lawrence Erlbaum and Associates.
- 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.
- Potter, M.C. (1993). Very short-term conceptual memory. *Memory and Cognition 21*(2)156-161.
- Shah, P., & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124*(1), 43-61.
- Simkin, D. K. & Hastie, R. (1987). An information processing analysis of graph perception. *Journal of the American Statistical Association, 82*, 454-465.
- Zacks, J. & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory and Cognition, 27*, 1073-1079.

Author Note

This work was supported by ONR grant N0001403WX30001. The authors thank Sander Los and Matthew Peterson for their comments and suggestions.

Table 1. Average pure, switch, and non-switch reaction times (in milliseconds) for

Experiment	<i>N</i>	Pure Block RT	Mixed Block Switch RT	Mixed Block Non Switch RT
1a- Line	20	2010 (291)	2136 (275)	2043 (359)
Pie		2227 (322)	2346 (296)	2233 (290)
1b- Bar	22	1798 (222)	1914 (305)	1876 (235)
Pie		2092 (235)	2270 (325)	2177(309)
1c- Line	21	1916 (324)	1941 (302)	1890 (255)
Bar		1775 (278)	1780 (282)	1770 (274)
2a- Line	31	1971 (302)	2087(307)	1997 (357)
Doughnut		2176 (246)	2268 (301)	2183(296)
2b- Bar	23	1871 (218)	2010 (265)	1893 (240)
Doughnut		2247 (289)	2395 (306)	2269 (300)
2c- Pie	21	2293 (343)	2308 (337)	2237 (357)
Doughnut		2440 (343)	2373 (323)	2293 (301)
3- Line	18	1959 (247)	1981 (278)	1985 (267)
Horizontal Bar		1876 (231)	1874 (254)	1818 (247)

Table 2. Time costs for the pairs of graphs examined in Experiments 1-3, shaded cells represent a significant time cost. Each column in the table represents the graph being judged. For example, the first column of numbers represents the time cost to line graphs when mixed with each of the other types of graphs. The cost to line graphs when mixed with bar graphs was 25ms (N.S) and the cost to line graphs when mixed with pie charts was 126ms.

	Line	Bar	Pie	Doughnut	Horizontal Bar
Line	--	5	119	92	-2
Bar	25	--	178	148	--
Pie	126	116	--	-67	--
Doughnut	116	139	15	--	--
Horizontal Bar	22	--	--	--	--

Figure Caption

Figure 1. Examples of graphs used in Experiments 1, 2 and 3; *upper left:* line graph, *upper right:* bar graph, *lower left:* pie chart, *lower right:* doughnut graph, *center:* horizontal bar graph.

