Title
When Graph Comprehension Is An Insight Problem

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Fox, Amy Rae
Hollan, James D.
Walker, Caren M.

Publication Date
2019

Peer reviewed

When Graph Comprehension Is An Insight Problem

Amy Rae Fox (amyraefox@ucsd.edu)
Department of Cognitive Science, University of California San Diego
The Design Lab—San Diego, California, USA

James D. Hollan
Department of Cognitive Science, University of California San Diego
The Design Lab—San Diego, California, USA

Caren M. Walker
Department of Psychology, University of California San Diego
Early Learning & Cognition Lab—San Diego, California, USA

Abstract

How do you make sense of an unconventional graph? Building on research demonstrating that prior knowledge of graphical conventions is difficult to overcome, we reconstrue graph reading as an insight problem. We hypothesize that imposing a *mental impasse* during a particular type of graph reading task will improve comprehension by inducing a sense of puzzlement, prompting learners to reconsider their interpretation. We find support for this proposal in a between-subjects experiment in which participants presented with an impasse-formulated version of graph reading questions are significantly more likely to correctly interpret a graph featuring an unconventional coordinate system. We characterize the differential patterns of mouse movements for learners between conditions and discuss implications for the use of novel graphical forms in science communication.

Keywords: graph comprehension; diagrammatic reasoning; insight; problem solving; representation; external representation; information visualization; mouse tracking

Introduction

The adage, "a picture is worth ten thousand words," surely applies to graphs. But what about a graph you don't know how to read? As Larkin and Simon note, "a representation is useful only if one has the productions that can use it," (1987, pg. 71). If we lack the ability to draw inferences from a graph, it is rendered useless. How is it then, that we develop such productions for new graphical forms?

Techniques for supporting graph comprehension have been a focus of research in the learning, cognitive and computer sciences for the past two decades. The most minimal interventions involve "graphical cues": visual elements that guide attention, akin to gesture and pointing. Acartürk (2014) investigated the influence of lines, arrows and point markers, finding that—used appropriately—such cues can help readers interpret the emphasis and temporal scope of a graph in alignment with a designer's intention. Kong and Agrawala (2012) surveyed the use of "graphical overlays" finding that reference structures (e.g. added gridlines), redundant encodings (e.g. data value labels), highlights, summary statistics, and annotations, are all commonly used to reduce cognitive load for particular graph reading tasks. Drawing inspiration from the literature in reading comprehension, Mautone & Mayer (2007) successfully demonstrated that animations, diagrams and drawings could help geology students connect the features of graphs to their geological referents. Each of these techniques serves to reinforce the semiotic connection between a graph, the world, and the reader's understanding, or to guide attention to information designers wish to emphasize. Importantly however, the techniques explored in this literature do not support learners in discerning *how* to read the graphs: the "rules" for their representational systems. Rather, it is assumed that the reader already has some familiarity with the type of graph being read (e.g. scatterplot, line graph, bar chart). In this way, the literature fails to differentiate between two types of prior knowledge brought to bear on a graph reading problem: knowledge of the domain, and knowledge of the graphical formalism.

In recent work (Fox & Hollan, 2018) we have taken up this challenge by investigating self-directed comprehension of an unconventional graph. In our paradigm, learners answer simple graph reading problems about a familiar domain—events in time—using an obscure graphical formalism. In an observational study, we found that readers struggled to make sense of the graph, misinterpreting the coordinate system as Cartesian. In a subsequent experiment, we evaluated four sets of instructional scaffolds aimed at overcoming the Cartesian misconception. We found that only an interactive version of the graph was effective for most learners. It seems that learners' expectations for the graphical formalism were so strong, even explicit text or image instructions failed to alert them to erroneous interpretations.

We argue this can be viewed as a sort of "graphical fixedness." Akin to Duncker's classic candle problem (1945), the learners in our previous studies were fixated on the conventional functions of the tools at their disposal: the marks on the page, and their assumptions about how axes and gridlines are meant to function. In the present work, we reconceptualize our graph reading task as an insight problem. We test the hypothesis that intentionally inducing a state of puzzlement in learners—posing a *mental impasse*—will improve their ability to extract information from a simple unconventional graph.

Background

Graph Comprehension

Process models of graph comprehension describe an integration of top-down and bottom-up processing (Shah, Freedman, & Vekiri, 2005). Following the information processing tradition, these models invoke the concept of a *schema*: a structured representation of knowledge in long term memory that guides processing of new information in a "top- down" fashion (*see* Alba & Hasher, 1983; Anderson & Pearson, 1984). A number of theories describing graph comprehension have posited the existence of a graph schema that guides an individual's interpretation on the basis of their prior knowledge of similar external representations (Freedman & Shah, 2002; Pinker, 1990; Tabachneck-Schijf, Leonardo, & Simon, 1997).

Unsurprisingly, there is no consensus on the format or content of graph schemata. One important question that has been addressed is what features of a stimulus trigger activation of a particular graph schema. According to the "invariant structure view" certain general characteristics are shared across a number of graph types that then rely on a shared schema (Peebles & Cheng, 2003). Ratwani & Trafton (2008) proposed that the structural components of a graph that represent basic concepts and operations for extraction—the *graphical framework*—may be that invariant structure. In a scatterplot, for example, the graphical framework includes the x and y axes. From this formulation, one can predict that bar, line and scatterplot graphs (all relying on a Cartesian coordinate system) might invoke a single graph schema, while pie charts (relying on a polar coordinate system) might invoke a different schema. It is unclear what (if any) schema might be activated in order to comprehend a novel representation. Pinker (1990, p. 105) theorizes that upon encountering a novel graph, a reader will instantiate a "general graph schema", likely based on a combination of the graph's coordinate system and most predominate graphical forms (e.g. points, lines, bars, etc.) The exact mechanism of construction for this general schema is unknown, but Pinker suggests it may be related to the cognitive processes that represent abstract concepts like space and the movement of objects within it.

Prior Knowledge and Graphical Sensemaking

While the marks on a page invoke our prior knowledge of graphical formalisms, the context of the marks activate our knowledge of the domain (Shah & Hoeffner, 2002). We argue that scarcity of each type of prior knowledge impedes comprehension in different ways.

Limited prior knowledge. If presented with an unfamiliar graph depicting information in an unfamiliar domain, you will be unable to use knowledge of one to bootstrap inferences for the other. Imagine you are a novice physics student reading a Feynman diagram: without some understanding of particle physics, you cannot reverse-engineer the formalisms of the diagram. Without these formalisms, you cannot draw inferences about particle physics.

Limited prior *domain* knowledge. Alternatively, if presented with a familiar graph depicting data in an unfamiliar domain, you might draw on your knowledge of that graph type to learn something new about the content. If you know that a straight line represents a linear relationship, you can infer this relationship between unfamiliar variables connected by a straight line. It is this situation that we aim to optimize in STEM education. To this end, Mautone & Mayer (2007) demonstrated that animations, arrows, diagrams and photographs can all help students connect their prior knowledge of graphs to represented variables, improving their ability to draw inferences about related scientific concepts.

Limited prior *graphical* knowledge. Here, we are interested in the reciprocal case: an unfamiliar representation depicting information about a familiar domain; perhaps that strange-looking graph you saw in your favorite academic journal. Importantly, by "graphical knowledge", we are not referring to knowledge of graphs in general (graphicacy), but rather knowledge of the rules governing a *particular* graphic form. Can you figure out how to read the graph, if you know enough about the domain? (Test yourself! See Figure 1)

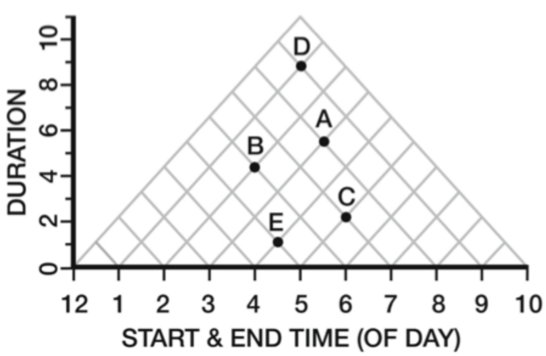
Reverse Engineering Formalisms. If the typical function of graphs is to use their formalisms as vehicles to learn something about the data (i.e. the domain) they represent, is the reverse also true? With sufficient domain knowledge, can readers reverse-engineer the formalisms governing a graph? Our data suggest this reciprocity of does not exist (Fox & Hollan, 2018). Despite extensive domain knowledge and personal experience with time, learners failed to correctly interpret the formalism of our graph with an unconventional coordinate system. Explicit instructions (text and images) were ineffective in supporting this reverse engineering, suggesting the need for a different scaffolding approach.

Problem Solving & Insight

In our earliest observational study with the Triangular Model graph (Figure 1), the vast majority of participants made the "Cartesian mistake": misinterpreting the graph as a Cartesian scatterplot (Fox & Hollan, 2018). However, for the few successful outliers, their production of the correct interpretation was accompanied by a protracted struggle, a sudden clap of their hands and ecstatic exclamation, "Oh! *That's* how it works!"

What we observed were moments of *insight.* This insight came during the study debrief when we gave the learner feedback that their answers were incorrect. In some cases, this feedback alone was sufficient to produce a moment of insight. According to Ohlsson (1992), insight results when one breaks free from an impasse*:* "a mental state in which problem-solving has come to a halt; all possibilities have been exhausted and the problem-solver cannot think of any way to proceed" (pg. 4). But unlike traditional problems in the insight literature, the state of impasse in graph comprehension is not readily apparent. We must therefore craft the state of impasse to intentionally draw a learner's attention to their own misconception. The function of our feedback in the verbal debrief was to alert the learner to the fact they had made a mistake. While we cannot provide verbal feedback as a passive scaffold, we *can* indicate to learners that they've made a mistake by anticipating their mistaken response, and designing the graph reading question to exploit this error: relying on the convention that a multiple-choice question should have at least one response.

An Unconventional Graph: The Triangular Model of Interval of Relations



This graph depicts a schedule of events (A through E)
At what time does event B begin? [non-impasse]
What event(s) begin at 3? [impasse]
**see answers in acknowledgements*

Figure 1. A Triangular Model Graph (TM)

This line of research requires a very special stimulus: one that represents information about a familiar domain but is sufficiently obscure to be unrecognizable by most learners. We selected the Triangular Model graph (Figure 1) to depict information about schedules of events using a novel coordinate system. It has an informationally equivalent analogue, the Linear Model which, as the conventional external representation for intervals of time, is the basis for many scheduling artifacts including Gantt Charts. Both models indicate the start and end time, duration, and relations between intervals, which we present to participants as "events in time."

Based on work by Kulpa (2006) extended by (Qiang, Delafontaine, Versichele, De Maeyer, & Van de Weghe, 2012) the Triangular Model (*hereafter* TM) represents intervals as points in 2D metric space (Figure 1). Each point represents an interval of time. In the vertical dimension, the height of the point indicates its duration. The intersection of the point's triangular projections (using diagonally oriented grid lines) onto the *x*-axis indicate the start (leftmost) and end (rightmost) times. In this way, every interval is represented as a unique point in the 2D graph space, and each of its elementary properties are explicitly encoded by the location of the point. Although the graph's computational efficiency is best realized with a large number of data points, and tasks that require judgement about the relation between intervals (e.g. "starts-with", or "during" relations), first order readings (i.e. reading the start, end or duration) are readily available and directly reveal the reader's interpretation of the coordinate system. (See Qiang et. al (2012) for a thorough review of the computational efficiency of the Triangular Model, and elaboration of use cases for which it is preferable to more conventional interval graphics.)

A brief inspection of the TM by even the most experienced graph reader demonstrates its relative obscurity. However, while the coordinate system is unconventional, the graph depicts information about a domain in which we all share substantial prior knowledge: events in time.

The Present Study

Results of two prior studies (Fox & Hollan, 2018) give us reason to suspect that conventional graph knowledge may hinder comprehension of unconventional representations. In the case of the TM graph, Cartesian expectations for the structure of the coordinate system interfere with our ability to follow perceptual cues provided by the graph's diagonal gridlines. In the present study, we test the hypothesis that constructing a mental impasse will improve comprehension of this unconventional graph.

Methods

Participants and Design. Sixty (55% female) undergraduate STEM majors at a public American University participated in exchange for course credit (age: 18 - 33 years). We utilized a between-subjects design with two groups and one independent variable (scaffold: none [control] vs. impasse). Participants were randomly assigned to an experimental group, yielding 30 students per condition. Prior to analysis, data from six participants were excluded based on their failure to correctly answer an attention check question.

Procedure. Participants completed the study in person, seated at a desktop computer. After a brief introduction, they were randomly assigned to an experimental condition and completed the Graph Reading Task, after which they received a short debrief. The session lasted approximately 30 minutes.

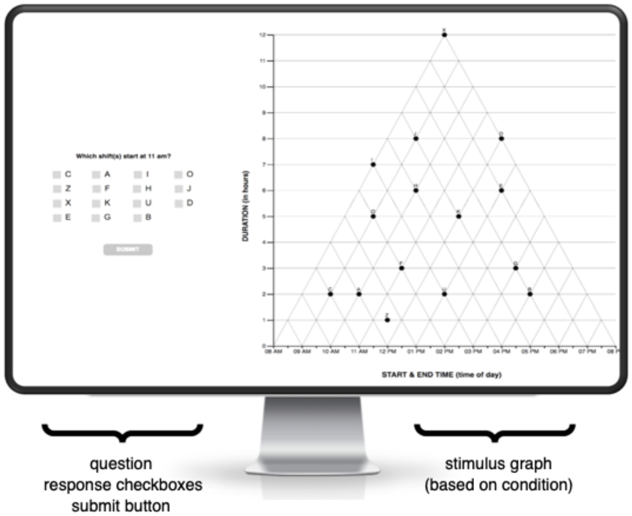
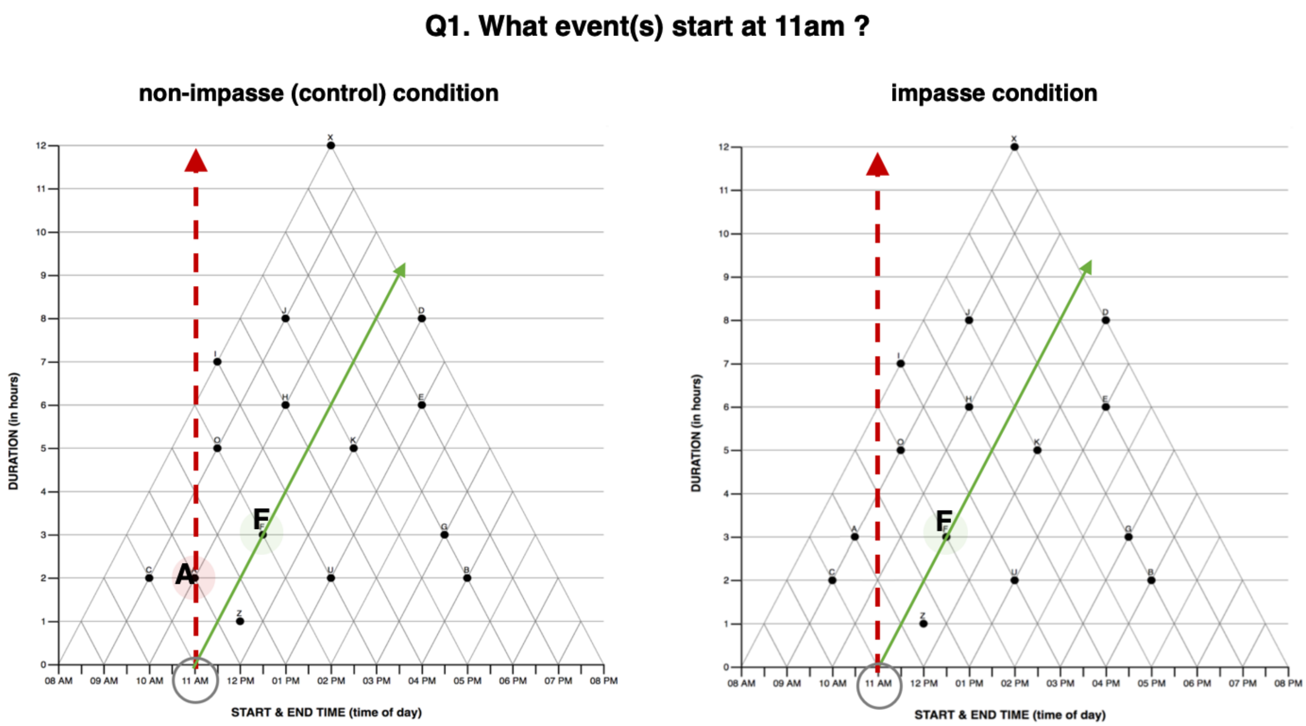


Figure 2. Sample stimulus

Materials. The Graph Reading Task consisted of a sequence of fifteen trials, each featuring a TM graph and multiple-choice question (Figure 2) about the temporal relationship between data points in the graph (i.e. "*Which event(s) start at 11am? At what time does event B end?*) Learners responded by clicking a checkbox corresponding to the data point(s) they wished to select. Trials were presented one at a time without feedback, in the same order for both conditions. Learners could not skip ahead nor return to previous questions. To assess the stability of student strategies over time, the first five trials included the assigned scaffold condition (none-control or impasse), while the following ten trials were identical (none-control). Questions were identical for both experimental conditions; however, the data sets rendered in the graph were slightly different for the first five trials. This allowed us to construct impasse problems with minimal differences between conditions. For each question in the non-impasse (control) condition, there was always a data point in the position where the participant would search if they interpreted the graph as Cartesian (Figure 3—*left*). Alternatively, in the impasse condition, the learner would find no data point in the expected position (Figure 3—*right*). For the final ten trials learners saw the same graph and questions. *See Figure 3 inset for a detailed description of the impasse structure.*



In both conditions, event F (solid green) is the single correct answer. In the non-impasse (control) condition at left, data point A crosses the Cartesian projection (dashed red line) from 11am, the most likely strategy taken by a learner misinterpreting the coordinate system. In this case, we expect the learner to select answer A. In the impasse condition (right) there is no data point intersecting the Cartesian projection. We expect this to pose a mental impasse. *note: extra lines and labels added for clarity, actual stimuli contain no such cues*

Figure 3. Experimental Conditions for Graph Reading Task Question #1

Data and Analysis. For each participant, we calculated a cumulative comprehension score [0-15], which served as the dependent variable. For further exploration of learner strategies, we integrated a JavaScript-based service (Mouseflow) to record all mouse-movements made by participants during the experiment session. Comprehension data were analyzed via inferential statistics, while mouse data were subject to exploratory qualitative analysis.

Results

Performance Accuracy. The mean comprehension score across the sample (n = 54) was approximately 6 points with a standard deviation of 0.68, and values ranging from 1 to 15 (max) points. On average, participants in the *impasse* group had higher scores (M = 7.6, SD = 5.2) than those in the non-impasse control group (M = 3.9, SD = 4.2), yielding a statistically significant difference $t(49.7) = -2.8$, $p = 0.006$; a moderate-sized effect $r = 0.37$.

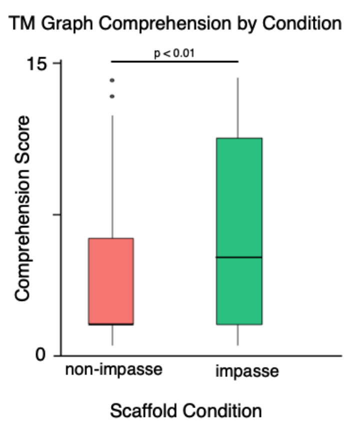


Figure 4. Results for graph reading task, by condition

Mouse Tracing Behavior. While raw comprehension scores can indicate whether learners correctly interpret the graph, they cannot reveal the strategies employed to answer the questions. To explore the mechanisms behind our results, we captured mouse tracing data. Similar to eye tracking data, mouse tracing provides an imperfect proxy for visual attention of the learner during the problem-solving session. This is a particularly rich source of insight for our graph reading problems as learners frequently used the mouse to navigate across the graph, the mouse acting like fingers tracing down or across gridlines. Of course, not all learners utilize the mouse to the same extent, and so we limit the present analysis to qualitative observation of gestalt patterns of graph traversal.

Figure 5 contains a set of heatmaps generated from raw path and dwell time data depicting the mouse movements of all participants on the *first* question of the Graph Reading Task. In the left column, we see data for learners in the control condition, and on the right, the impasse condition. The top row of heatmaps were generated from only those participants who correctly answered the question, while the bottom row from participants with a variety of incorrect answers. Visual inspection of these heatmaps reveal that across both conditions (top row), learners who correctly interpreted the coordinate system traversed the graph in a similar fashion, with the most prominent patterns following the relevant diagonal gridlines. Inspecting those with incorrect answers (bottom row), we see dramatically different patterns of tracing across conditions. While those in the control condition (bottom left) follow the expected Cartesian projection, learners in the impasse condition (bottom right) exhibit no single discernible pattern. While these learners did not arrive at the correct answer, their tracing behavior may be an indication of puzzlement.

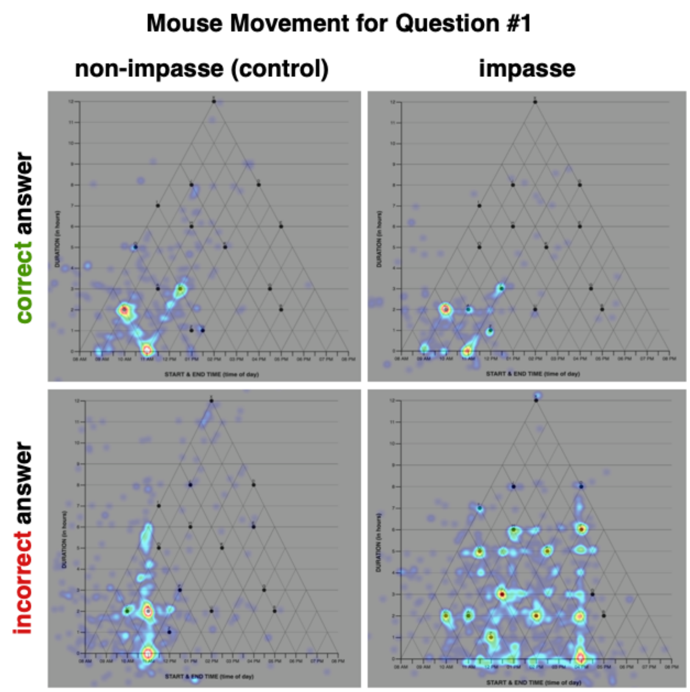


Figure 5. Mouse movement as heatmap for all participants, question #1

Discussion

The essence of functional fixedness, according to Ohlsson (1992), is that the experience of using an object in a particular way lowers the probability of finding a solution in which one uses the object in a different way. The strength of our association of the function to the object sets the strength of fixedness. In this light, we can see how substantial experience with common graphical forms serve to fix our expectations of axes, and coordinate systems in general, toward a Cartesian interpretation. The results of this study support our hypothesis that constructing a problem to present a learner with a mental impasse yields significantly better performance on the unconventional graph reading task. Of course, not all graph reading tasks need be construed as insight problems. Most often the challenge we face concerns second-order readings—the inferences to be drawn from available information— and there is a close relationship between the nature of a graph-reading task and the suitability of the graph design (Shah & Hoeffner, 2002). However, we argue that these readings of trends and relationships between data points are unlikely to be made if the reader does not understand the nature of the graphical formalism itself, and this is where insight comes into play.

Lockhart, Lamon & Gick (1988) characterize difficulties in problem solving as a failure to *access* available information. This certainly seems applicable to the difficulties we observe with the Triangular Model graph where the reader need only perceive and recognize the importance of the diagonal gridlines to extract information from the graph (first-order readings). Lockhart et. al. propose that learners must often reconceptualize a problem in order to solve it, and simply giving students information may not be enough to achieve this effect. Presenting information in a form that induces puzzlement is significantly more effective in facilitating conceptual transfer and subsequent problem solving. We argue that the puzzlement induced by finding 'no available answer' in our impasse condition worked by leaving learners with no recourse but to reconsider their strategy (or give up). This conclusion is further supported by learners' failure to interpret this graph when provided with explicit information. While the text and image scaffolds in (Fox & Hollan, 2018) did not improve performance with the TM graph, a simple manipulation of the availability of answers to the first problems in a scenario for this study did.

We expect this technique should generalize to other representations with unconventional coordinate systems, though it is unclear whether the same attention-directing mechanisms would be appropriate for forms utilizing alternative markings. This is one of several open questions we are presently pursing. In ongoing analysis of mouse tracing data, we are exploring the strategies employed by learners in the impasse state and how they may reflect learner's graphical intuitions. In particular, we're interested in the strategies employed by learners in the impasse condition that provide non-Cartesian, but nonetheless incorrect responses. How are these learners reasoning about the graph elements, and does their behavior remain consistent after the scaffold phase (first 5 questions) when the remaining 10 questions have possible Cartesian answers? Based on ongoing analysis of the time course of response accuracy, we suspect that for impasse to be effective, the learner must confront the impasse in the initial phase of graph interpretation—when the graph schema is instantiated. In ongoing work, we address this question by varying the timing of impasse vs. non-impasse questions with analysis of the time course of correct and incorrect responses. We are also investigating which components of the design and layout of the graph are most influential in triggering a Cartesian interpretation, by manipulating the layout and saliency of axes, gridlines, and rotation of the figure in graph space.

While we hope this line of research will shed light on the elusive graph schema and how we develop graphical knowledge, the most immediate implications of our findings address the presentation of graphics in publications like this one. As communicators of science, we face an inevitable tension between communicating in what we believe to be the most revealing or expository fashion, and the way a community has come to expect. This makes innovation difficult. Nonetheless, the popularity of information visualization as a research area means that novel graphical forms are ever more present in our discourse. If you choose to utilize an unconventional representation in a traditional publication format, posing a carefully designed question (in perhaps, the figure caption) may aid the motivated reader to persevere in correctly reading the new graphic, and discovering your insights.

Acknowledgments

**answers to Figure 1: event B begins at 2; event A begins at 3.* Sincerest thanks are offered to Research Assistant Evan Barosay and The UCSD Design Lab. This work was supported in part by the United States Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

References

Acartürk, C. (2014). Towards a systematic understanding of graphical cues in communication through statistical graphs. *Journal of Visual Languages and Computing*, *25*(2), 76–88. https://doi.org/10.1016/j.jvlc.2013.11.006

Alba, J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231. https://doi.org/10.1037//0033-2909.93.2.203

Anderson, R., & Pearson, P. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. Pearson (Ed.), *Handbook of Reading Research* (pp. 255–291). New York: Longman.

Duncker, K. (1945). "On problem solving". *Psychological Monographs, 58:5 (Whole No. 270).*

Fox, A. R., & Hollan, J. (2018). Read It This Way: Scaffolding Comprehension for Unconventional Statistical Graphs. In P. Chapman, G. Stapleton, A. Moktefi, S. Perez-Kriz, & F. Bellucci (Eds.), *Diagrammatic Representation and Inference* (pp. 441–457). Springer International Publishing.

Freedman, E. G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. *Diagrammatic Representation and Inference*, 18–30. https://doi.org/10.1007/3-540-46037-3_3

Kong, N., & Agrawala, M. (2012). Graphical overlays: Using layered elements to aid chart reading. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2631–2638.

Kulpa, Z. (2006). A diagrammatic approach to investigate interval relations. *Journal of Visual Languages and Computing*, *17*(5), 466–502. https://doi.org/10.1016/j.jvlc.2005.10.004
Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *99*, 65–99.
Lockhart, R. S., Lamon, M., & Gick, M. L. (1988). Conceptual transfer in simple insight problems. *Memory & Cognition*, *16*(1), 36–44.
Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, *99*(3), 640–652. https://doi.org/10.1037/0022-0663.99.3.640
Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In *Advances in the psychology of thinking* (pp. 1–44).
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. https://doi.org/10.1518/hfes.45.1.28.27225
Pinker, S. (1990). Theory of Graph Comprehension. In R. Freedle (Ed.), *Artificial Intelligence and the Future of Testing* (pp. 73–126). Hillsdale, NJ: Erlbaum.
Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P., & Van de Weghe, N. (2012). Interactive Analysis of Time Intervals in a Two-Dimensional Space. *Information Visualization*, *11*(4), 255–272. https://doi.org/10.1177/1473871612436775
Ratwani, R. M., & Trafton, J. G. (2008). Shedding light on the graph schema: perceptual features versus invariant structure. *Psychonomic Bulletin & Review*, *15*(4), 757–762. https://doi.org/10.3758/pbr.15.4.757
Shah, Freedman, E. G., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In A. Miyake & P. Shah (Eds.), *Cambridge Handbook of Visuospatial Thinking*.
Shah, & Hoeffner. (2002). Review of Graph Comprehension Research: Implications for Instruction. *Educational Psychology Review*, *14*(1), 47–69.
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. https://doi.org/10.1207/s15516709cog2103_3