Strategy and pattern recognition in expert comprehension of 2 × 2 interaction graphs

David Peebles

Department of Behavioural and Social Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK

Abstract

I present a model of expert comprehension performance for 2 × 2 "interaction" graphs typically used to present data from two-way factorial research designs. Developed using the ACT-R cognitive architecture, the model simulates the cognitive and perceptual operations involved in interpreting interaction graphs and provides a detailed characterisation of the information extracted from the diagram, the prior knowledge required to interpret interaction graphs, and the knowledge generated during the comprehension process. The model produces a scan path of attention fixations and a symbolic description of the interpretation which can be compared to human eye movement and verbal protocol data respectively, provides an account of the strategic processes that control comprehension, and makes explicit what underlies the differences between expert and novice performance.

Keywords: Graph comprehension, ACT-R, Computational modelling

1. Introduction

Working with graphs is a complex skill that requires specific knowledge of the representational system being used together with a set of procedures to map spatially represented information in the graph with a set of propositions that specify quantitative and qualitative relationships between the entities represented. Providing a detailed account of this skill therefore requires one to specify a number of core assumptions including: what and how information is encoded in the diagram, what and when information is obtained from the

Email address: `d.peebles@hud.ac.uk` (David Peebles)

diagram by the user during a task, what and how prior graph knowledge is stored and utilised, and what new knowledge is created during the process. In addition, one must also specify the strategies people employ to carry out different tasks and how much these strategies use information in the diagram and in stored internal representations.

There have been several attempts to provide detailed process models of different aspects of graph use. Models are constructed from sets of perceptual and cognitive operators (e.g., encode the value of an indicator, make a spatial comparison between indicators (Gillan, 1994), compare two digits in working memory, or make a saccade (Lohse, 1993)), obtained either from task or verbal protocol analyses. Lohse (1993) and Gillan (1994) have produced models of question answering with several different graph types (including line graphs, bar charts and scatter plots) by constructing sequences of operators (each of which has an associated execution time) to generate predicted scan paths across the graph and total task completion times which can be compared to human data.

Other researchers have procedurally analysed graph use for different purposes. For example, Casner (1991) identified a set of perceptual and cognitive operators to construct models of several graph-based tasks which informed an automated system that generated graphical representations most suited to the tasks commonly undertaken with them. A similar method was adopted by Tabachneck-Schijf et al. (1997) in their analysis of an economics expert's construction of a graph while explaining the principle of supply and demand which they then used to develop a computational model incorporating both diagrammatic and propositional representations.

More recently, the cognitive modelling of reasoning with information displays has been advanced by the development of *cognitive architectures*; computational theories of the large-scale structure of the mind providing accounts of how cognition is controlled and how knowledge is encoded, stored, retrieved and utilised (e.g., ACT-R (Anderson, 2007), EPIC (Meyer & Kieras, 1997), and Soar (Laird et al., 1987)).

The first two of these architectures incorporate theories of visual processing and motor control which allows modellers to produce more detailed accounts of the information obtained from the display during the task. For example Peebles & Cheng (2003) used ACT-R to produce a computational model of question answering using two different types of line graph. Their model generated saccades and fixations as it answered each question which, together with task completion times, were compared to human data. In addition, the model was able to account for human scan paths in terms of the varying demands on memory imposed by different questions.

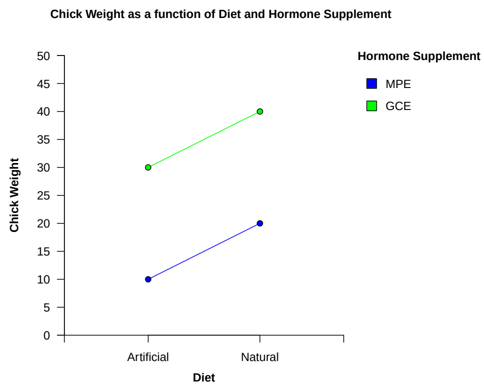
The Peebles and Cheng study, as did those by Lohse (1993) and Gillan (1994), investigated question answering in which participants were given items of information and were required to produce associated information using different processes, including identification (e.g., “In 1997, what was the value of gas?” (Peebles & Cheng, 2003)), comparison (e.g., “In 1977 did tin cost less than sulphur?” (Lohse, 1993)), and arithmetic computation (e.g., “What is the sum of A, B, and C?” (Gillan, 1994)).

While these are important tasks, particularly for investigating sequences of elementary processes, it could be argued that they do not necessarily reflect how many people normally work with graphs and that they do not address the important prior comprehension stage where labels and graphical features are encoded, associated, and interpreted (Carpenter & Shah, 1998).

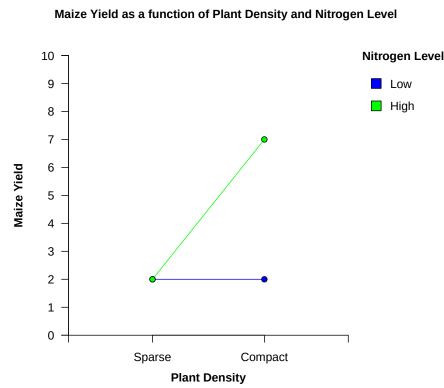
Comprehension requires knowledge of the conventions used in the graph to represent data and other facts such as how labels are to be interpreted based on their location. The output of the process is assumed to be a set of knowledge structures that represent the variables and graphical features together with structures that encode knowledge about the quantitative or qualitative relationships between the variables depicted.

A prime example of a scenario where people encounter a graph with the sole aim of comprehending the relationships between variables (as opposed to identifying trends or individual values for example) is the analysis of data from factorial experiments. The simplest form of factorial design is the *two-way factorial design*, containing two factors, each with two levels, and one DV. Statistical analysis of these designs typically results in a 2×2 matrix of mean values of the DV corresponding to the pairwise combination of the two levels of each IV. Interpreting the results of even these simplest of designs accurately and thoroughly is often not straightforward however, but requires a significant amount of conceptual understanding—for example the concepts of simple, main, and interaction effects. As with most other statistical analyses however, interpretation can be eased considerably by representing the data in diagrammatic form.

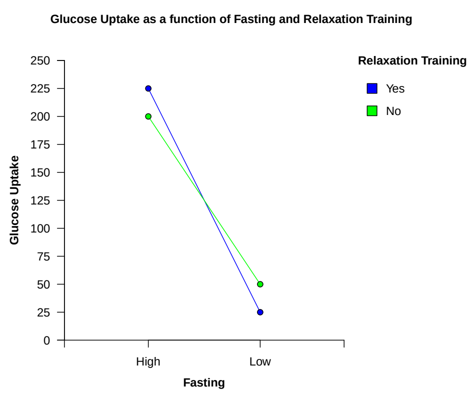
Data from two-way factorial designs are most often presented as either line or bar graphs—variously called *interaction* or *ANOVA* graphs. Examples of line graphs used in Peebles & Ali (in preparation) and this study are shown in Figure 1. Interaction line graphs differ from more conventional line graphs because the data represent pair-wise combinations of the IV levels so that



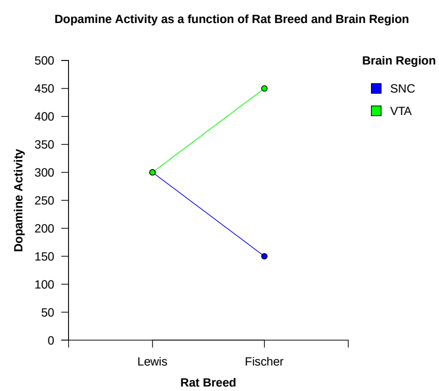
(a) Line graph 1



(b) Line graph 2



(c) Line graph 7



(d) Line graph 8

Figure 1: Four of the eight line graphs used in the experiment.

the variables plotted on the x axis are categorical, regardless of whether the underlying scale could be considered as continuous (e.g., hot/cold) or categorical (e.g., male/female).

The rules for interpreting interaction graphs are quite specific therefore and sufficiently different from other more frequently encountered line graphs that simply applying general interpretive rules will not prove particularly helpful (other than for obtaining the DV values of specific conditions etc.) and may lead to the misinterpretation of the x axis variable levels as representing two ends of a continuous scale (Aron et al., 2006; Zacks & Tversky, 1999).

It has been argued however (Kosslyn, 2006, e.g.,) that the risk and costs of misinterpreting line graphs are outweighed by the benefit of lines for producing distinct and easily recognisable patterns that indicate key features of the data such as main effects or interactions. These patterns will be discussed in detail below.

In a series of studies, Peebles and Ali have observed and recorded novices (undergraduate psychology students) and experts (cognitive science professors and postgraduate researchers) interpreting interaction graphs like the ones in Figure 1 (Peebles & Ali, 2009; Ali & Peebles, in press; Peebles & Ali, in preparation). These studies have shown that without knowledge of the appropriate interpretive rules, novices' interpretations are often limited to qualitative descriptions of differences between conditions and can be skewed by the different Gestalt principles of perceptual organisation (Wertheimer, 1938) operating in the graph. In contrast, expert users are able to employ their knowledge of which graphical features represent which effects to identify relationships between variables much more rapidly and accurately with no prior knowledge of the domain variables being represented in the graph.

An example of this is shown in the verbal protocol below which contains a verbatim transcription of a (not atypical) expert participant interpreting the graph in Figure 1c (taken from Peebles & Ali (in preparation)).

1 (Reads) "Glucose uptake as a function of fasting and relaxation training"
2 Alright, so we have... you're either fasting or you're not...
3 You have relaxation training or you don't...
4 And so... not fasting... er...
5 So there's a big effect of fasting...
6 Very little glucose uptake when you're not fasting...
7 And lots of glucose uptake when you are fasting...
8 And a comparatively small effect of relaxation training...
9 That actually interacts with fasting.

The protocol (which lasted 43s) shows the initial identification of the IVs and their levels followed by a rapid identification of the key features of the data; the main effect of the x axis variable and the interaction between the two IVs.

The purpose of the research reported here is to develop a computational model of graph comprehension that specifies the processes underlying both expert and novice behaviour with sufficient detail and comprehensiveness to satisfy all of the criteria outlined at the beginning of this paper. Specifically, the model aims to provide a precise account of the minimum information required to interpret interaction graphs appropriately together with a hypothesis as to the nature of the processes involved in representing and interpreting that information. The model is developed within the ACT-R cognitive architecture and therefore embodies assumptions about the nature of the mental representations and the computations that form the strategies used to generate new representations. Finally, the model provides an explanation for the differences between expert and novice interpretations.

2. A model of expert graph comprehension

Space limitations preclude a detailed description of ACT-R here. However a comprehensive account of the cognitive architecture can be found in Anderson (2007). In summary, ACT-R consists of a set of modules that acquire information from the environment, process information, and execute motor actions to achieve goals. ACT-R has memory stores for declarative and procedural knowledge. The former consists of a network of knowledge chunks while the latter is a set of production rules. Cognition proceeds via a pattern matching process that attempts to find production rules with conditions that match the current state of the system and tasks are performed through the successive actions of production rules.

ACT-R also incorporates a subsymbolic level of computations that govern memory retrieval and production rule selection and which allow models to account for widely observed recency and frequency effects on retrieval and forgetting. Subsymbolic computations also underlie ACT-R's different learning mechanisms.

For tasks involving displays and other devices, task environments can be defined to be acted upon by the model. The graphs used in this study are defined as sets of visual objects (lines, circles, rectangles, and text) with certain features (size, colour) at specific x-y coordinates on a 2D window.

The graph comprehension model is based on verbal protocol data from novice and expert users (Peebles & Ali, 2009; Ali & Peebles, in press; Peebles & Ali, in preparation). In these studies, verbal statements recorded during the comprehension task were coded and categorised in terms of their functional role and content (e.g., "an association between a level and its identifier"; "a comparison between the two legend variable levels for one of the levels of the x axis variable") to produce a set of common interpretive operations.

The verbal protocols indicate that comprehension is typically carried out in two main phases: (a) a variable identification stage followed by (b) a pattern recognition and description stage. The protocols also reveal that experts and a large proportion of novices rarely report specific DV values, but typically produce qualitative descriptions of the differences between conditions (Trafton et al., 2000, cf.).

In the first stage, the three variables are identified, categorised as dependent or independent according to location, and the latter associated with their levels, which in turn are associated with identifiers (left or right position for the x axis variable and colour for the legend variable).

In the second stage, the plot region is scanned and the pattern produced by the plot points is interpreted. This interpretation is typically done by comparing distances between plot points and using the comparison to probe long-term declarative memory for interpretive knowledge. If successful, the retrieved knowledge is used to construct an interpretation. If no interpretation is available however, the model will simply describe the identification or comparison process being carried out. Interpretive operations are carried out until either a full interpretation is produced or until no other operations are available or identified.

2.1. Representing and encoding information in the graph

The key information that the model encodes from the display is the set of four x-y coordinate locations and the distances between them. The perceptual processes by which this spatial information is obtained and initially represented are not specified in detail, although it is assumed to be acquired using a subset of the *elementary perceptual tasks* (e.g., judgement of length, direction, area, position on a common scale etc.) identified by Cleveland & McGill (1984).

Two elementary perceptual tasks are used extensively to encode information from the display. The first—*judgement of position on a common scale*—is used to encode the distance between two plot points, initially as a numerical value (the proportion, p, of the distance to the overall length of the y axis). This numerical code is not used directly in reasoning however but is converted to a symbolic qualitative size description: "no" ($p = 0$), "very small" ($0 < p < 0.2$), "small" ($0.2 =< p < 0.4$), "moderate" ($0.4 =< p < 0.6$), "large" ($0.6 =< p < 0.8$), and "very large" ($0.8 =< p =< 1.0$).

The second elementary perceptual task—*judgement of length*—is used in the comparison of distances required to evaluate differences between variable levels (e.g., comparing the distance between the two High Fasting values with that between the two Low Fasting values in Figure 1c). The elements formed for these comparisons are assumed to be the result of Gestalt processes of perceptual organisation (Ali & Peebles, in press; Kosslyn, 1989; Pinker, 1990) by which users group objects by colour or proximity.

It is also assumed that in such comparisons the "direction" (i.e., the relative ordering of levels) of a length is encoded. For example, when comparing the High Fasting and Low Fasting distances in Figure 1c, the fact that they have different levels of the Relaxation Training variable as their higher value will be noted. This additional information is essential for the identification of various global patterns such as crossed, parallel and diverging lines.

2.2. Prior graph knowledge

Two forms of declarative knowledge are involved in the task: prior knowledge relating to how the graph represents information and the knowledge of the variables and their relationships generated during the comprehension process itself.

There are three core items of knowledge required to interpret interaction graphs. Two are common to many Cartesian graphs and concern (a) the typical allocation of the dependent and independent variables to the graph axes and legend and (b) the principle that the distance between two graphical elements encodes the magnitude of a relationship between the concepts represented by those elements.

The third set of facts required are specific to the graph type and concern the spatial indicators of the three key relationships; simple effects, main effects, and interactions. These indicators are: (a) the distance between two plot points indicating the size of the *simple effect* of the level jointly represented by those points, (b) differences in the y-axis location of the midpoints between two pairs of plot points indicating the size of the *main effect* of the variable, and (c) differences in the inter-point distances between levels, combined with information about their point ordering, indicating the size, and type of any interactions that may exist.

Some relationships form distinctive and relatively common patterns however which experts learn to identify rapidly, either through explicit instruction (Aron et al., 2006, e.g.,) or simply through repeated exposure. Four patterns indicating the existence (or otherwise) of interaction effects are particularly common and readily identified: the "crossover interaction" shown in Figure 1c, the "sideways V" pattern shown in Figure 1d, and a related

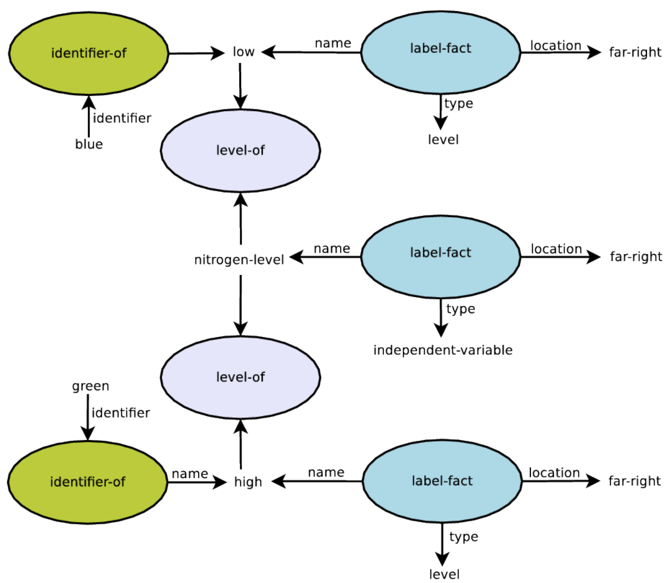


Figure 2: Graphical representation of knowledge generated after processing the legend of Figure 1b.

pattern formed by a horizontal and a sloped line (Figure 1b). In contrast, parallel lines (e.g., Figure 1a) signal that there is no interaction between the IVs.

In addition to these interaction patterns, two patterns indicating substantial main effects can also be recognised by experts (and are often rapidly identified by novices due to their visual salience). These patterns are shown in Figures 1a and 1c. The large gap between the mid-points of the two lines in Figure 1a shows a large main effect of the legend variable while the large difference between the mid-points of the two values representing each x axis level in Figure 1c reveals a large main effect of the x axis variable.

2.3. Generated knowledge

Several declarative knowledge structures are also generated during comprehension. The first is a set of related chunks that represent each variable, the levels associated with it, and the identifiers of each level. A graphical representation of such a structure that combines seven knowledge chunks to represent the legend variable of Figure 1b is shown in Figure 2.

Three other knowledge structures are generated as graph and interpretive information is accumulated and associated during comprehension. The current expert model contains all the prior interpretive knowledge described in the previous section so that each knowledge structure generated combines a qualitative description of the elements being analysed or compared together with a complete and accurate interpretation. The information in these structures is then output in a form that may then be compared to verbal protocols produced by human experts.

For example, one knowledge structure represents the interpretation of an individual variable level (e.g., the Compact level of the Plant Density variable in Figure 1b). The information combined in this structure can then produce symbolic output equivalent to the statement: "The difference between the two values for compact plant density is very large so there is a very large simple effect of high plant density".

A second knowledge structure records the comparison of two levels of one variable (e.g., the levels of the Hormone Supplement variable in Figure 1a) to produce output comparable to the statement: "There is a large difference between the hormone supplement levels; GCE generally resulted in greater chick weight than MPE, which indicates a large main effect of hormone supplement".

The third knowledge structure represents a comparison of the lengths and point ordering of two levels (e.g., the levels of the Fasting variable in Figure 1c) which can be used to produce output that can be translated as: "Although the effect size of the fasting levels is the same, the direction of their effects is different, indicating a crossover interaction between the two independent variables".

Finally, a knowledge structure is produced when two plot points are compared without any further interpretation (e.g., high nitrogen level in Figure 1b). Information in this structure can be used to produce output that can be translated as: "When nitrogen level is high, maize yield is much greater for compact plants than for sparse plants".

2.4. The comprehension process

Appendix B contains an output trace produced by the model as it carries out the comprehension task using the graph in Figure 1c, with each line in the trace representing one step in the process[1]. Text in square brackets is

[1]A video of the model interpreting all eight graphs from the expert study (Peebles & Ali, in preparation) can be viewed at `http://youtu.be/qYY_No0i1Hc`

information currently being processed that has either been obtained from the graph or retrieved from declarative memory.

Numbers in square brackets (e.g., in lines 28 and 33) represent the perceptual difference between two objects in the display which are subsequently translated into qualitative size judgements (e.g., lines 29 and 33) according to the categories described above. Other text in the output is simply to indicate other events (e.g., goal setting or memory retrieval failures) or to clarify to human readers what a particular knowledge element represents.

As previously intimated, the model assumes that comprehension proceeds after an initial phase of variable identification; a process usually initiated by reading the title (lines 1–4). Currently when the model reads the title the three words that name variables are identified by retrieving previously defined word category information from declarative memory. This mechanism is undoubtedly simplistic and currently substitutes for a more complex knowledge retrieval process that is assumed to take place.

The model then seeks items of text at the left (lines 5–6), lower (lines 7–13), and right (lines 14–18) regions of the display. When each variable label is located, the model identifies it as a particular type according to its location and then associates the independent variables with their level labels by identifying nearby text. The model also associates each of the four levels with its physical attribute; left, right, blue and green, and uses these labels when processing the graph. This is consistent with verbal protocol and eye movement data from our studies showing that graph readers often produce an interpretation and then must re-read the appropriate label in order to identify which particular level is being processed.

When the three variables have been processed, the model attends the central region of the display and processes the pattern produced by the four coordinate points in the plot region (line 26).

The model represents the interpretation process by a set of production rules for the various patterns and features in the graph. When the appropriate condition occurs (i.e., the model is directing attention to the plot region), individual production rules fire to draw attention to specific indicators. The indicator (a spatial distance, difference or order comparison), is extracted from the pattern and (together with information about what the indicator is) used to probe declarative memory for an interpretation consisting of the name and size of the effect. For example on line 39 of the trace the model identifies that there is 0.8 difference between the plot points at either end of the blue line (i.e., the gap between them is 0.8 of the y axis) and then retrieves the knowledge that this indicates a very large simple effect of the Yes level of the Relaxation Training variable.

Once a recognition production rule fires to initiate the process, a chain of subsequent productions is triggered which obtains further information from the graph and declarative memory until an interpretation is produced. If a memory retrieval attempt fails, the model simply describes the feature being attended to (behaviour observed quite often in novices) but in the current expert model, such retrieval failures do not occur.

To capture the rapid pattern recognition behaviour of experts observed and described above the model contains six productions, one for each pattern type, which fire and initiate an interpretive sequence when a pattern is identified. Two such patterns are present in the example protocol. The first one recognised is the crossed lines (lines 27–31) which leads to the identification of a crossover interaction (line 29). The second pattern is the substantial difference between the x axis levels (lines 32–37) which leads to the identification of a main effect of the x axis variable (line 37).

Once the patterns have been processed (or if there are no such patterns in the graph), the model samples the display region for further features that may indicate other important relationships. This is shown in lines 38–43 of the trace where the model identifies the simple effects of the two legend variable levels but that there is no main effect of the legend variable.

Comparing the expert verbal protocol and the listing in Appendix B, one can identify the equivalence between the global structure and information content of the expert and model outputs. The expert's reading of the title (line 1) and identification of the x axis (line 2) and legend (line 3) variables are captured in lines 1–4, 7–13, and 14–18 respectively of the model output. The key elements of the interpretation are also very similar. The large and salient main effect of the x axis variable is identified in lines 4–7 of the expert protocol and lines 32–37 of the model protocol. The interaction is also rapidly identified by both expert (line 9) and model (lines 27–31) and both also note the relatively minor effect of the legend variable (line 8 in the expert protocol and lines 38–43 in the model output).

3. Discussion

Comprehending and reasoning with graphs requires a wide range of perceptual and cognitive operations sequenced together in various combinations to perform specific tasks. The type and sequence of operators involved in a task may differ depending on a number of factors, including the graph or domain knowledge of the user, the type of graph being used, or individual cognitive factors such as working memory capacity (which may determine the relative frequency of memory retrieval requests and saccades to graph labels etc.).

Graph comprehension is an important area to study therefore because it provides an opportunity to investigate how environmental and internal factors interact to produce behaviour. In addition, graph-based tasks can be analysed using behavioural measures such as eye movements and concurrent verbal protocols to provide insights into what and when information is being processed during the course of the activity.

Computational modelling is a valuable tool for developing and testing hypotheses about the representations and mechanisms necessary for cognitive tasks as it provides a formalism for characterising them, requires one to be explicit about the boundaries of one's model in terms of which processes are being defined precisely and which are not, and allows one to explore the consequences of particular assumptions (McClelland, 2009).

Developing models within a cognitive architecture such as ACT-R provides the additional benefit of allowing the model to incorporate a large number of assumptions regarding issues such as knowledge representation, cognitive control, visual attention, learning and forgetting etc., all of which are supported by previous empirical research. In addition, ACT-R's vision module includes mechanisms that allow models to simulate certain Gestalt principles of perceptual organisation, which are regarded as playing a crucial role in the visual processing of graphical representations (Kosslyn, 1989; Pinker, 1990). Specifically, the comprehension model associates variables and their levels, and levels with their colour identifiers using mechanisms that are functionally equivalent to the Gestalt laws of *proximity* and *similarity* respectively.

The model described above represents an initial attempt to specify at a detailed algorithmic level the representations, cognitive processes, and strategies involved in comprehending interaction graphs. It provides a precise account of the graph knowledge required and the spatial information necessary to interpret the graph accurately and specifies a control structure that determines the flow of information during the task to generate a set of knowledge representations, saccades and fixations over the graph, and a sequence of output statements which are largely consistent in terms of order, function and content with verbal protocols produced by expert users.

The assumptions of the model imply that to interpret interaction graphs accurately, novices must acquire three forms of graph-specific knowledge: an understanding of what effects the different distances and spatial differences in the graph indicate, the relationship between distance and effect size, and how the various combinations of distance differences and point orders can be interpreted in terms of the interactions between the IVs. The model provides a precise specification of the relatively small amount of knowledge required and a clear demonstration of its sufficiency to interpret the graphs.

Previous studies have shown that comprehension performance varies widely, even between experienced users (Peebles & Ali, 2009; Ali & Peebles, in press; Peebles & Ali, in preparation). For example, the order in which effects are identified varies, either due to the effects of particular Gestalt principles of perceptual organisation (Ali & Peebles, in press), as a result of experts' familiarity with common patterns, or the relative visual salience of the graphical features being displayed (e.g., very large main effects). In addition to the core interpretative knowledge therefore, the current model also incorporates explicit pattern recognition rules to account for the speed and sequential order of expert interpretations.

Previous studies have also compared expert and novice performance on both bar and line graph formats and showed that the interpretations of all users (but novices in particular) were affected by the format used. Specifically, line graphs users are influenced to attend to the legend variable while bar graph users attend to the two IVs more equally (Peebles & Ali, 2009; Ali & Peebles, in press; Peebles & Ali, in preparation). Broadening the scope of the model further, other factors such as domain knowledge and the number of variable levels (Shah & Freedman, 2011) should also be addressed.

The current production set is sufficient to process any 2×2 data set of three variables to produce an appropriate interpretation similar to the trace in Appendix B. The model therefore provides a solid basis from which to explore hypotheses concerning the mechanisms underlying a broader range of behaviour. These hypotheses will take the form of enhanced or reduced declarative graph or domain knowledge, additional recognition productions, and mechanisms to represent visual salience. A more comprehensive model must also bring ACT-R's subsymbolic mechanisms that govern memory retention, retrieval, and learning processes into play as these no doubt have a significant effect on strategy choice and eye movement patterns (Peebles & Cheng, 2003).

Finally, the current model does not attempt to provide a detailed account of the perceptual processes by which spatial information is encoded or represented during the execution of elementary perceptual tasks. There are currently several attempts to develop mechanisms for spatial representation and processing within cognitive architectures—including ACT-R—however (a number of which are presented in (Gunzelmann, 2011)) and it may be possible for the current functions to be replaced in a future model by ones more conforming with theory and empirical evidence.

Beyond the goal of developing the model to account for the full range of observed behaviour with an increasing number of interaction graph formats, lies the larger aim of constructing a model of comprehension for a broader class of graphs. As discussed earlier, interaction graphs embody a specific set of interpretive rules that are not shared by other graphs. The current model clearly identifies and characterises these rules and distinguishes them from the knowledge and procedures that can be applied to other graphs. It is hoped that in so doing, the model will simplify the task of identifying graph-specific operators and form a basis upon which to develop and explore a range of graph comprehension models for other graphical formats. As it stands however, the model provides a valuable demonstration that the assumptions it currently embodies are sufficient to produce an expert interpretation of the relationships depicted in 2×2 interaction graphs.

References

Ali, N., & Peebles, D. (in press). The effect of Gestalt laws of perceptual organisation on the comprehension of three-variable bar and line graphs. *Human Factors*, *doi:10.1177/0018720812452592*.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?*. New York, NY: Oxford University Press.

Aron, A., Aron, E. N., & Coups, E. J. (2006). *Statistics for Psychology*. (4th ed.). London: Pearson.

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

Casner, S. M. (1991). A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, *10*, 111–151.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*, 531–554.

Gillan, D. J. (1994). A componential model of human interaction with graphs: 1. linear regression modelling. *Human Factors*, *36*, 419–440.

Gunzelmann, G. (Ed.) (2011). *Modeling spatial cognition [Special issue]* volume 3. Topics in Cognitive Science.

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

Kosslyn, S. M. (2006). *Graph design for the eye and mind*. New York: Oxford University Press.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence*, *33*, 1–64.

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

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*, 11–38.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance 1. basic mechanisms. *Psychological Review*, *104*, 3–65.

Peebles, D., & Ali, N. (2009). Differences in comprehensibility between three-variable bar and line graphs. In *Proceedings of the Thirty-first Annual Conference of the Cognitive Science Society* (pp. 2938–2943). Mahwah, NJ: Lawrence Erlbaum Associates.

Peebles, D., & Ali, N. (in preparation). Comparing novice and expert comprehension of three-variable bar and line graphs.

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*, 28–45.

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.

Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction between top-down and bottom-up processes. *Topics in Cognitive Science*, *3*, 560–578.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science*, *21*, 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*, 827–850.

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology*. London: Routledge & Kegan Paul.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory and Cognition*, *27*, 1073–1079.

3.1. Appendix A: Model output for the graph in Figure 1c

1 seek text at top of display...
2 [glucose-uptake] [= variable]
3 [as] [a] [function] [of] [fasting] [= variable]
4 [and] [relaxation-training] [= variable]
5 seek text at far left of display...
6 [glucose-uptake] at [far-left] is the [dependent] variable
7 seek text at bottom of display...
8 [fasting] at [bottom] is the [independent] variable
9 look to nearest text...
10 [low] is a level of [fasting]
11 [low] is [right]
12 [high] is a level of [fasting]
13 [high] is [left]
14 seek text at far right of display...
15 [relaxation-training] at [far-right] is the [independent] variable
16 look to nearest text...
17 [yes] is a level of [relaxation-training]
18 [no] is a level of [relaxation-training]
19 seek objects in plot region...
20 a [blue] [line]
21 no associate for [blue] so look to legend...
22 found [blue] [rectangle]. looking for nearest text...
23 [blue] represents [yes]
24 found [green] [rectangle]. looking for nearest text...
25 [green] represents [no]
26 variables identified. look to plot region...
27 pattern: values reversed for x axis levels...
28 [0.0] difference in distance between points. [neither] bigger
29 [no] difference and [different] point order = [crossover-interaction]
30 for [low] [fasting] [no] [relaxation-training] greater than [yes] [relaxation-training]
31 for [high] [fasting] [yes] [relaxation-training] greater than [no] [relaxation-training]
32 pattern: substantial difference between x axis levels...
33 [0.1] difference [left] = [very-small] [simple] effect [high] [fasting]
34 [0.1] difference [right] = [very-small] [simple] effect [low] [fasting]
35 compare [left] and [right] levels...
36 [large] difference. [high] [fasting] greater than [low] [fasting]

37 [large] [main] effect of [fasting]
38 identify legend levels...
39 [0.8] difference [blue] = [very-large] [simple] effect [yes] [relaxation-training]
40 [0.6] difference [green] = [large] [simple] effect [no] [relaxation-training]
41 compare [blue] and [green] levels...
42 [no] difference. both levels of [relaxation-training] are the same
43 [no] [main] effect of [relaxation-training]