# The effect of graphical format and instruction on the interpretation of three-variable bar and line graphs

David Peebles
Department of Behavioural and Social Sciences,
University of Huddersfield,
Queensgate, Huddersfield,
HD1 3DH, UK.
D.Peebles@hud.ac.uk

July 3, 2011

**Abstract**

This study investigates how graph format (bar or line graph) can affect undergraduate psychology students' ability to interpret three-variable 'interaction' data typically found in undergraduate psychology courses. The study revealed that significant differences in interpretation are produced by graph format in four key subtasks; identifying average values of variable levels, finding maximum and minimum values, recognising main effects and interactions, and identifying dependent and independent variables. These results are interpreted as arising from Gestalt principles of perceptual organisation that make different aspects of the data more salient and draw users' attention to different variables. Based on the results of this experiment, it is argued that, in general, undergraduate psychology students' interpretations of such three-variable interaction data are more accurate when using the bar form.

## Introduction

Analysing and interpreting data is a key skill taught in all scientific undergraduate degree courses—including psychology—because communicating and interpreting analyses of data appropriately is a fundamental activity in the sciences (Friel, Curcio, & Bright, 2001). Studies have shown however that statistics is considered by many students to be the most challenging, anxiety producing and therefore least popular of their subjects (e.g., Gal & Ginsburg, 1994; Onwuegbuzie, Daros, & Ryan, 1997).

For statistics teaching therefore, it is important that continual efforts are made to study students' learning practices and how these relate to understanding and performance. By bringing the theories and methods of experimental cognitive psychology to bear on the processes involved in learning and practicing statistics, we can identify specific features that are not optimal, and suggest possible improvements to make the activities easier to perform or the concepts more readily understandable.

One of the core skills learned by statistics students is the ability to produce and interpret graphical displays of data accurately and appropriately. Although many experienced graph users take their abilities for granted, the knowledge and skills required are far from trivial and require considerable training and practice to be mastered.

A substantial body of research has been conducted over the last thirty years to investigate the perceptual and cognitive processes involved in graph interpretation and the representational and computational properties of various graphical formats (e.g., Carpenter & Shah, 1998; Kosslyn, 1989; Lohse, 1993). This research has significantly advanced our understanding of graph-based reasoning and has had practical consequences in the form of guidelines and examples of best practice for appropriate graph design in various contexts (e.g., Kosslyn, 2006; Tufte, 2001).

## Factorial research designs

Factorial research designs are widely used in all branches of the natural and social sciences as well as in engineering, business and medical research. The efficiency and power of such designs to reveal the effects and interactions of multiple independent variables (IVs) or factors on a dependent variable (DV) has made them an invaluable research tool and, as a consequence, the teaching of such designs, their statistical analysis and interpretation lies at the core of all natural and social science curricula.

The simplest form of factorial design is the *two-way factorial design*, containing two factors, each with two levels, and one DV—for example the differences in word recall (DV) between amnesics and a control group ($IV_1$) in an implicit versus explicit memory task ($IV_2$). Statistical analysis of these designs most often results in a $2 \times 2$ 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 such bar and line graphs (taken from the experiments reported here) are shown in Figure 1. Bar and line graphs such as those in Figure 1 can display the same data set in the same coordinate system and are *informationally* equivalent (Larkin & Simon, 1987).

In terms of their visual and conceptual structure, bar and line graphs have a great deal in common, the key difference being the way in which the data points are represented in the coordinate system. However this relatively minor difference has been shown to have a remarkable effect on which features are made salient, which in turn influences the type of information extracted from the display.

In line graphs, lines integrate individual plotted points into single objects, features of which (e.g., slope, height relative to other lines, etc.) can indicate relevant information about the entire data set (Carswell & Wickens, 1990, 1996). This feature has been found to lead people to encode the lines in terms of their slope (e.g., Simcox, 1983, reported by Pinker, 1990) and interpret them as representing continuous changes on an ordinal or interval scale (Kosslyn, 2006; Zacks & Tversky, 1999). For this reason line graphs are typically regarded as a form of configural or object display.

In contrast, bar graphs are an example of a separable display as each data point is represented by a single, separate bar. Because of this, people are more likely to encode bars in terms of their height and interpret them as representing the separate values of nominal scale data (Culbertson & Powers, 1959; Zacks & Tversky, 1999).

These differences in encoding and interpretation can result in significant performance variation for different tasks; people are typically better at comparing and evaluating specific quantities using bar graphs (Culbertson & Powers, 1959; Zacks & Tversky, 1999) whereas people are generally better at identifying trends and integrating data using line graphs (Schutz, 1961). This situation is therefore a prime, real-world example in which two informationally equivalent and relatively similar representations are widely used, but which are known to be *computationally* inequivalent (Larkin & Simon, 1987) in certain circumstances. It seems appropriate to ask therefore, whether these computational differences significantly affect the ease and efficiency with which people interpret them and the depth and accuracy of the interpretations produced.

According to the proximity compatibility principle (Carswell & Wickens, 1987), graph format should correspond to task requirements, so that configural displays should be used if information needs to be integrated, whereas separable displays are more appropriate if specific information needs to be located. In the case of interaction data however, there are reasonable arguments for using either format.

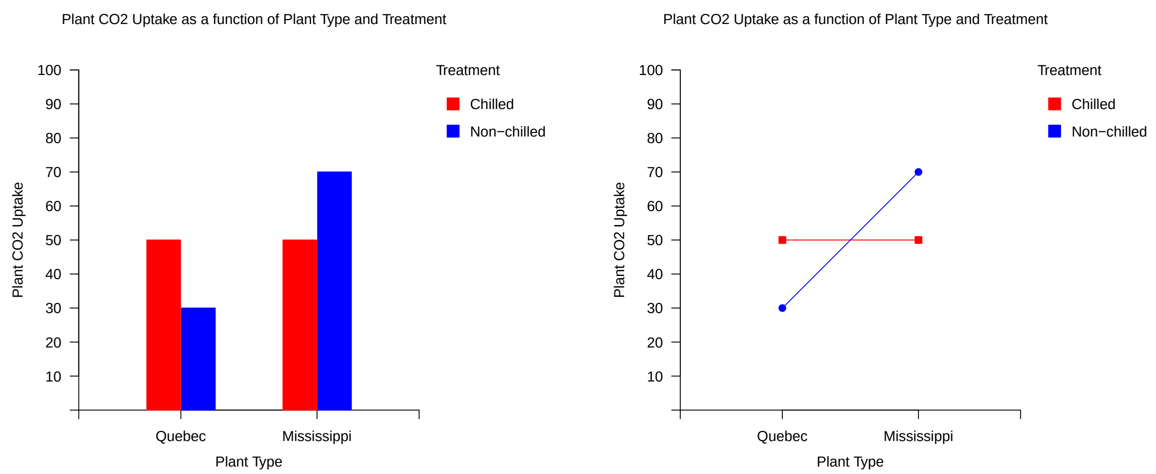
Interaction graphs differ from more conventional line graphs in that the variables plotted on the x axis are categorical, regardless of whether the underlying scale could be considered as continuous (e.g., hot/cold) or categorical (e.g., male/female). The argument for using bars for interaction graphs is that, because people encode bars as separate entities, they are less likely to misinterpret the levels of the x axis variable as representing two ends of a continuous scale (Aron, Aron, & Coups, 2006; Zacks & Tversky, 1999). By contrast, line graphs are more likely to be interpreted as representing continuous data with points on the lines representing intermediate values on the scale. Proponents of the line graph (e.g., Kosslyn, 2006) have argued however that the risk and costs of misinterpreting line graphs are outweighed by the benefit of lines for producing easily recognisable patterns that can be associated with particular effects or interactions.

## Understanding undergraduate students' graphical reasoning

Our understanding of how the mental representations of various graphs are constructed during learning and how they relate to each other and affect our interactions with diagrams is still far from complete however. Reasoning with graphs is a complex skill that involves the appropriate representation of abstract concepts relating spatial and numeric entities and a set of procedures for interpreting and producing different diagrams accurately. It also relies on fundamental properties of the visual system that rapidly identify patterns and search for visual features.

This complex process can go wrong for a number of reasons. Interpretive schemas and procedures may be incorrectly or inadequately learned, leading to misinterpretation or unwillingness to produce or interpret graphs. In a number of studies Peebles has also demonstrated how the perception and interpretation of different commonly used graphs can be significantly affected by their visual and representational properties (e.g., Peebles & Cheng, 2003; Peebles, 2008).

The effects of these cognitive and perceptual phenomena may have important implications for the teaching of graphical literacy and how we represent data most effectively. For example, in a recent study Peebles and Ali (2009) investigated how people interpreted three-variable interaction data in bar or line graph form. The graphs present exactly the same information but use different graphical features to represent the data points. Despite these relatively minor differences, the experiment revealed that line graph users were significantly more likely to misinterpret or be unable to interpret the data represented than bar graph users. More specifically, 39% of the line graph users were not only unable to integrate the information appropriately, but were unable to understand them at even an elementary level (e.g., read individual points accurately) and made a large number of common errors. By contrast, no bar graph users exhibited such low levels of performance.




(a) Bar graph

(b) Line graph

Figure 1: Bar and line graphs used in session 1.

Peebles and Ali (2009) hypothesised that these observed differences result from a combination of two factors: (a) the inadequate knowledge structures and procedures of some novice users and (b) Gestalt principles of perceptual organisation (Wertheimer, 1938) that made data points and their relationships more visually salient in the bar graphs, thereby making their interpretation, particularly by novices, much more detailed and accurate. Specifically, the visual salience of the lines in the line graphs drew attention to them and made the identification and interpretation of the specific data points more difficult. This was not found in the bar graphs because each data point is represented by a unique, readily identifiable bar. This finding is consistent with those from an earlier study which also revealed significant differences in the interpretation of bar and line graphs (Peebles, 2008).

One potential implication of these findings is that three-variable data of this type may be more effectively taught to undergraduate students in the form of bar graphs than with the more traditional line graphs. A possible longer term implication may be that this recommendation is more generally applicable to other forms of data.

# Experiment

There still remains however an important question regarding the robustness of this effect and its relationship to training. It may be the case that it only manifests in the absence of appropriate instruction. If so, the conclusion to be drawn would be that explicit and rigorous teaching of line graph interpretation is essential to prevent it being hampered by the potentially confusing features of the format. Alternatively, it is possible that the visual salience of the lines in line graphs is so high that its effect on interpretation is still found after explicit training has occurred. If this is the case, then it may be wise to conclude that such data would be best taught and communicated in bar graph form. The primary aim of this study is to conduct an experiment designed to answer this question.

## Method

### Participants

Participants were 86 foundation level undergraduate psychology students at the University of Huddersfield. There were 77 females and 9 males.

### Design

The study consisted of three elements; two 16-question questionnaires and a 25 minute lecture on graph interpretation that was given immediately after the first questionnaire. A period of 14 days separated the two questionnaires.

### Materials

Two questionnaires were constructed using two data sets that produced two pairs of bar and line graphs. The graphs for the session 1 questionnaire are shown in Figure 1 while those for the session 2 questionnaire are shown in Figure 2.

The variables in the graphs were chosen so that no prior knowledge of the domain or relationships would influence interpretation and the values of the conditions were devised to present and test the fundamental items of knowledge required to produce an appropriate knowledge structure or schema for each graph and to produce patterns that would test the various hypotheses under investigation.

Both session 1 and session 2 graphs represented a number of similar relationships. Both showed a possible main effect of the x variable (with the session 2 graph showing a more pronounced effect) and an interaction between the x and legend variables (with a more prominent interaction effect in the session 1 graphs).

The questions in both questionnaires were essentially identical, with only minor changes in wording to account for the different content. The questions were devised to examine students' knowledge of relevant concepts (i.e., dependent (Q2) and independent (Q3) variables, main effects (Q13) and interactions (Q14), their ability to extract basic information from the graph (i.e., condition markers (Q4a, Q4b, Q6, Q12), causal relationships (Q1)), maximum (Q5) and minimum (Q9) and mean (Q4c, Q4d) values, and to identify the various relationships between the variables (Q7, Q8, Q10, Q11, Q15). The final question asked students to summarise the meaning of the graph.

The questions were devised to elicit students' knowledge and skills, ranging from knowledge of the concepts of *dependent* and *independent* variables and the causal relationships between them, the ability to associate variable labels with their graphical referents and to extract basic information, to


(a) Bar graph

(b) Line graph

Figure 2: Bar and line graphs used in session 2.

higher-level tasks like recognising complex relationships and using the concepts of *main effect* and *interaction*.

A 25 minute presentation introducing the basic concepts and skills required to interpret 2 × 2 factorial graphs was also produced. Two versions were constructed which were identical apart from the graph format used.

### Procedure

In the first session, conducted in the students' ninth week of teaching, 123 students of a cognitive psychology class completed the first questionnaire individually using either a bar or line graph (which were distributed randomly throughout the class). After the questionnaires were all completed and collected, students were divided into two groups according to which graph they had used, taken to separate rooms where they were taken through the slide-based lecture by the experimenters. To ensure that the lectures were identical, the experimenters used prepared lecture notes that were identical apart from where bars or lines were referenced. Each lecturer stayed to the script of the lecture and read the same text at appropriate stages in the lecture.

In the second session, conducted two weeks after the first, 116 students from the same class completed the second questionnaire individually using the same graph format they had used in session 1.

## Results

86 students in total completed both questionnaires, 45 bar graph users and 41 line graph users. This required the removal of 37 session 1 questionnaires and 30 session 2 questionnaires.

Figures 3 and 4 present the proportion of correct responses for the two graph conditions for each of the questions 1–15 in session 1 and 2 respectively while the mean proportions of correct responses are displayed in Table 1. The data show that in a large number of cases graph format had little effect on the ability of students to answer the question. The data also reveal however that students' understanding of this data representation (whichever format it is in) is patchy and far from complete.

The data in Table 1 show that, taken as a whole, the training session had little overall effect on the proportions of correct responses from either condition. This mean figure hides a number of interesting and significant differences for individual questions that reveal the effects of graphical format on novices' interpretation of the data.

Table 1: Mean proportion correct responses for the two graph conditions for questions 1–15 in sessions 1 and 2. Standard deviations are in brackets.

| | Graph condition | |
|---|---|---|
| Session | Bar | Line |
| 1 | 64.45 (22.80) | 59.49 (25.18) |
| 2 | 60.24 (24.24) | 59.76 (29.00) |

Analysis of the pattern of responses across the range of questions should be able to reveal the typical contents of undergraduate psychology students' graph schemas and limitations in their reasoning with these graphs. This wider analysis is currently underway for future publication. The primary focus of this study however is to investigate differences between graph formats and so the following analysis will focus on specific questions where significant differences between graph conditions were found. This occurred for a number of the same questions in both the pre and post tests.

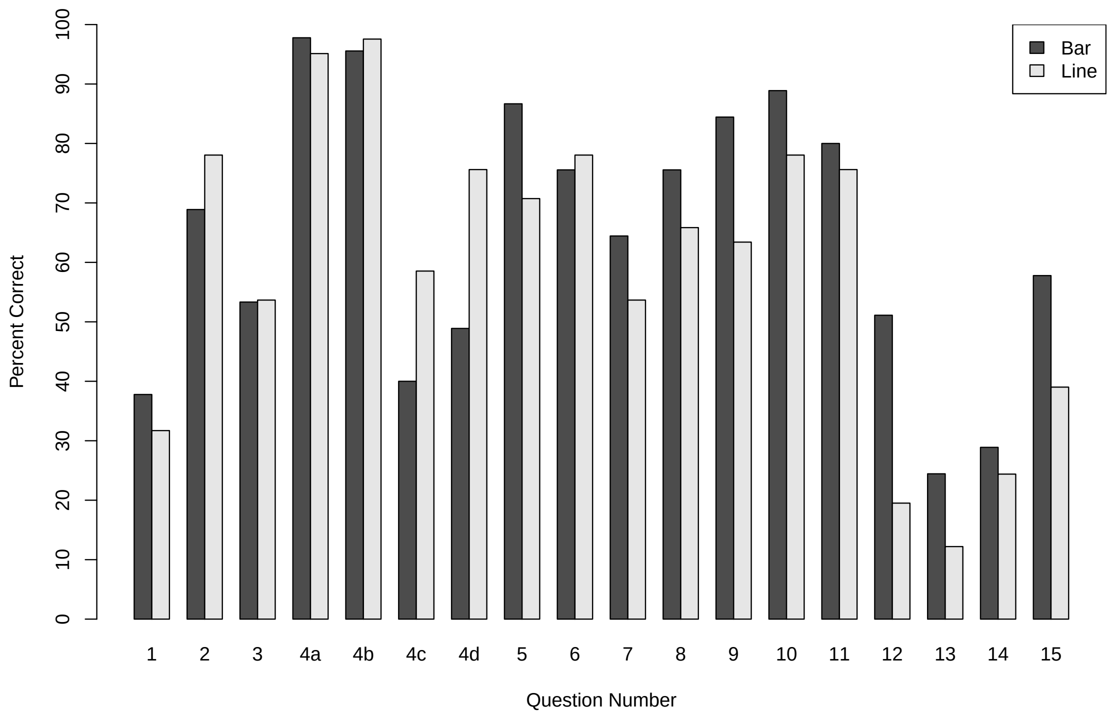



Figure 3: Percent correct responses for the two graph conditions, questions 1–15, session 1

## Identifying average values of variable levels

A clear distinction to emerge between the two graph conditions was in how the two formats facilitated the identification of mean values of different variable levels. The task involves identifying two condition values and combining them. This can be achieved in a number of ways involving varying degrees of calculation by the user. A calculation-intensive method is to read off the two condition values from the y axis and then mentally calculate the mean. Although requiring explicit calculation, this method has the benefit of only requiring knowledge of how to compute the mean of two numbers, together with the ability to read off values from the y axis.

A more perceptual and less calculation-intensive method is to locate the mid-point between the two conditions and then read off the value of this mid-point from the y axis. In contrast to the calculation method, the knowledge used in this strategy primarily concerns the fact that the mean of two points on a linear axis is to be found at the location equally between them (i.e., it is knowledge concerning the properties of Cartesian coordinate systems rather than purely mathematical knowledge). This particular feature of Cartesian graphs may be emphasised by the graphical features of a graph—for example, if the mid-point between two conditions is made visually salient (e.g., by the intersection of a line).

Two questions required students to identify the location of average values of levels of the x axis (Q4c) and legend (Q4d) variables in the pattern in the centre of the graph by writing a label next to it. The expectation was that while there should be no difference in the ability of bar and line graph users in identifying the mean of the x axis level, the by the crossed line pattern should facilitate line graph users to identify the average value of the legend level.

For both the pre and post tests, this was found to be the case. While no significant differences were found between bar and line graph conditions for Q4c, significantly more line graph users identified the legend variable mean than bar graph users in the pre-test (76% compared to 49%; $\chi^2 = 6.477$, $df = 1$, $p < .05$) and in the post-test (73% compared to 42%; $\chi^2 = 7.168$, $df = 1$, $p < .05$).

This facilitation effect was reversed however when users are asked to compare the levels of the legend variable, as they were in Q12 in both pre and post tests. In both first and second session graphs the average values of the legend variable levels is the same.

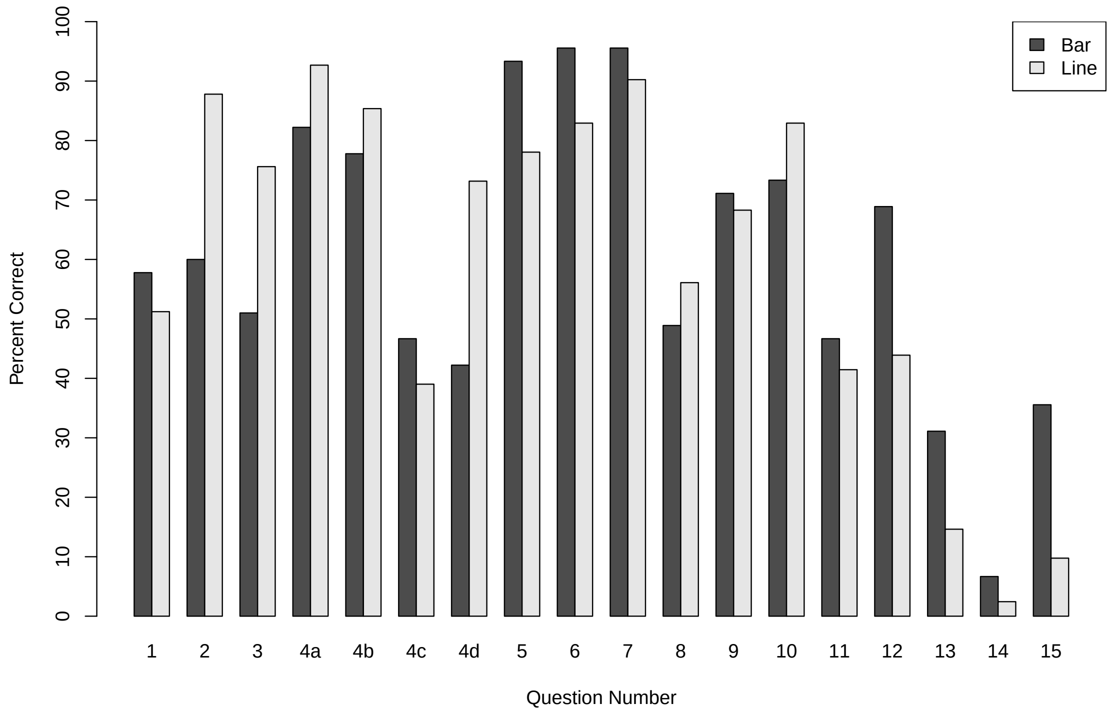



Figure 4: Percent correct responses for the two graph conditions, questions 1–15, session 2

Based on previous studies, the prediction is that line graph users will focus attention on the lines (rather than the individual condition points) and be drawn to comparing them, either in terms of their lengths or (more likely) their highest points. In contrast, bar graph users' attention will more likely be drawn to the individual bar tops and their associated values. This will facilitate a more directly comparative or computational procedure that produces the average values of the two levels (and the conclusion that they are equal).

The results of both sessions supported these assumptions. In session 1, 51% of bar graph users replied that the overall levels were the same, whereas only 20% of line graph users did so ($\chi^2 = 7.97$, $df = 1$, $p < .005$). This pattern was replicated in session 2, where 69% of bar graph users answered the question appropriately negatively but only 44% of line graph users did so ($\chi^2 = 4.49$, $df = 1$, $p < .05$).

The erroneous responses of the line graph users support the suggestion that their attention was drawn to the lines, which they then compared. In the pre-test significantly more line graph users (22.0%) responded with the level with the highest point (non-chilled Mississippi plants) compared to only 2.2% of bar graph users; $\chi^2 = 8.13$, $df = 1$, $p < .005$) while 24.4% responded with the highest condition (non-chilled Mississippi plants) compared to only 13.3% of bar graph users. In the post-test, 17.1% of line graph users answered with the level with the highest point (fasting and relaxation training) compared to only 4.4% of bar graph users.

## Finding maximum and minimum values

The most basic information extraction operations that people can perform with interaction graphs are (a) to read off individual condition values from the y axis and (b) to generate qualitative value comparisons between conditions (e.g., propositions involving predicates such as “greater than”, “smaller than”, “largest” and “lowest”). Conducting these tasks demonstrates an understanding of the basic representational principles of the graph but requires little else in terms of mathematical inference or calculation.

Students’ grasp of these basic operations was examined in Q5 and Q9. These questions were very similar in terms of the information requested and the processes required to obtain it. They required users to identify the maximum (Q5) or minimum (Q9) value of the dependent variable and then obtain the condition associated with that value. This represents a slight increase in complexity compared to Q4 (which supplied a condition to be located) in that the information given was a location which had to be identified by the user and the information required was the associated condition.

For both questions in the pre-test more bar graph users produced the correct answer than line graphs users. For Q9 this difference (84.4% compared to 63.4%) was statistically significant ($\chi^2 = 3.940$, $df = 1$, $p < .05$) but the difference for Q5 (86.7% compared to 70.7%), although substantial, did not reach statistical significance ($p = .07$). This superior performance in the bar graph condition is most likely due to the fact that bar graph users are more able to locate individual conditions and identify maxima and minima because they are represented by individual bars.

A range of erroneous responses was produced by line graph users, with many incorrect answers being produced by only one or two people. This suggests that because the line graph representation is not allowing rapid identification of the target information, many participants are searching for candidate conditions or variable levels. Taking the two questions together however, the most common single response was a level of the legend variable (e.g., “Chilled treatment” or “Non-chilled treatment”) with approximately 12% of respondents producing this error for each question. The prevalence of these responses is most likely due to the visual salience of the coloured lines.

A similar pattern was found for Q5 in the post-test where 93.3% of bar graph users provided the correct answer (fasting and relaxation training) compared to only 78.1% of line graph users (although this difference failed to reach statistical significance; $\chi^2 = 3.00$, $df = 1$, $p = 0.08$).

## Recognising main effects and interactions

In both the pre- and post-test graphs the patterns indicate that there could be a main effect of the x axis variable. Identifying main effects and interactions is an advanced skill that requires explicit training in what the concepts mean and how to identify them in graphs. As a consequence therefore, correct responses (i.e., those that gave only the correct response) to Q13 and Q14 in the pre-test which asked about these concepts were predictably low (Q13 bar = 24.4%, line = 12.2%; Q14 bar = 28.9%, line = 24.4%). For Q13, 60% of bar graph users and 66% of line graph users answered ‘I don’t know what a “main effect” is’.

After the training (which included instruction on the identification of main effects and interactions), the number of correct responses to the equivalent questions in session 2 were still low (Q13 bar = 31.1%, line = 14.6%; Q14 bar = 6.7%, line = 2.4%). For Q13, far fewer students in both conditions (22.2% bar, 24.4% line) responded ‘I don’t know what a “main effect” is’.

However in both pre- and post tests, bar graph users were roughly twice as likely to answer that there was a main effect of the x variable. In the post test, this difference was significant $\chi^2 = 5.69$,

$df = 1$, $p < 0.05$.

For Q14 however, which asked participants if there was an interaction in the graph, line graph users were nearly twice as likely to answer that there was an interaction between the legend variable and the dependent variable (53.7% compared to 28.9%; $\chi^2 = 4.48$, $df = 1$, $p < 0.05$.

Taken together, these results indicate that students' responses can be skewed by the representation they are using—bar graph users to the x axis variable and line graph users to the legend variable.

### Identifying dependent and independent variables

One unexpected result found in session 2 was a significant difference between bar and line graph users in identifying the dependent and independent variables. Specifically, significantly more line graph users than bar graph users responded correctly to questions Q2 ($\chi^2 = 7.11$, $df = 1$, $p < 0.05$) and Q3 ($\chi^2 = 4.74$, $df = 1$, $p < 0.05$).

In Q2, which asked for the dependent variable, although 60% of bar graph users correctly identified the y axis variable, 33.3% identified the x axis variable (compared to only 12.2% of line graph users; $\chi^2 = 4.25$, $df = 1$, $p < 0.05$).

This error then led to a similar proportion of bar graph users identifying the y axis variable as an independent variable in the subsequent question (35.56% compared to only 9.76% of line graph users; $\chi^2 = 6.62$, $df = 1$, $p < 0.05$). Further analysis of the data revealed that 22.2% of bar graph users switched the labels, calling the x axis and legend variables the dependent variables and the y axis variable the independent variable (compared to only 7.3% of line graph users).

This result is unexpected in that knowledge of the concepts of dependent and independent variables are not tied to any particular graph format and should not therefore be expected to be influenced by particular graphical elements. It seems however, that for those students who did not know the answer, the strong salience of the bars drew their attention to the x axis variable which was then given as the answer. It is not clear why a similar bias towards the legend variable in line graphs was not found however.

## Discussion

Gestalt principles of perceptual organisation (Wertheimer, 1938) are regarded as playing a crucial role in the visual processing of graphical representations. Pinker (1990) for example, argues that Gestalt laws are one of the four key principles that determine the nature of the mental representations that users generate when reading a graph. According to Pinker, the Gestalt laws of *proximity*, *similarity*, *connectedness*, *good continuation* and *common fate* all determine how individual graphical features are grouped together to form coherent wholes and so relate patterns to variables and their values together.

The results of this experiment support the suggestion that such Gestalt principles can have strong effects on the interpretation of data presented in graphical form and provide new insights into the details of some of these effects. The study revealed that significant differences in interpretation are produced by graph format in four key subtasks; identifying average values of variable levels, finding maximum and minimum values, recognising main effects and interactions, and identifying dependent and independent variables.

These effects were found in novice users, and more significantly, were still present when students had received an introductory lecture on the appropriate interpretation of these graphs. This suggests that the effect can be quite strong, and may require significant training to eradicate.

The data also suggest that, in general, bar graphs are more likely to be interpreted more accurately and thoroughly by novices than line graphs. Because the bars do not incorporate potentially confusing and distracting graphical elements users are encouraged to identify the values of individual conditions and are then better able to compare and combine them to produce more accurate identifications (e.g., of maximum and minimum values) and comparative judgements (e.g., about the of the relative sizes of variable levels).

The lack of connecting lines between conditions allows bar graphs also to direct users' attention to the x axis variable and allow users to more readily identify the two separate levels of that variable.

In this experiment this resulted in bar graph users' improved ability to identify a main effect of the x variable.

It could be argued that this latter facilitation may also be provided by line graphs in situations where there is a main effect of the legend variable (e.g., in patterns involving to widely separated parallel horizontal lines). Although this requires empirical testing, based on the results of this experiment (specifically the improved bar graph performance in comparing the levels of the legend variable) it is unlikely that the lines will provide the same degree of facilitation and that bar and line performance will be similar.

This does suggest however that where there is a main effect of one variable in data then placing that variable on the x axis of a bar graph will allow it to be identified more readily.

As part of their training, psychology students are expected to develop sophisticated graphical literacy skills as much of their work will involve the production and interpretation of graphical displays of data. Interaction graphs form a significant proportion of this experience and it is vital therefore that the processes involved in their use are understood so that skills may be taught appropriately and the best graphical formats used.

It has been assumed that students can interpret both bar and line interaction graphs equally well and that the benefits of line graphs enjoyed by experts can readily be acquired by novices. This study demonstrates the limitations of this assumption and shows that students' interpretations are sometimes accurate but quite often limited and subject to influence by the type of graph used.

These limitations implies that students' understanding of these three-variable interaction data should not be assumed as not all of their properties cannot be inferred by reference to previously learned graph schemas. It also suggests that interpretive biases may only be eradicated by substantial training and practice.

# References


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.

Carswell, C. M., & Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics*, *30*(3), 511–527.

Carswell, C. M., & Wickens, C. D. (1990). The perceptual interaction of graphical attributes: Configurality, stimulus homogeneity, and object interaction. *Perception & Psychophysics*, *47*, 157–168.

Carswell, C. M., & Wickens, C. D. (1996). Mixing and matching lower-level codes for object displays: Evidence for two sources of proximity compatibility. *Human Factors*, *38*(1), 1–22.

Culbertson, H. M., & Powers, R. D. (1959). A study of graph comprehension difficulties. *Audio-Visual Communication Review*, *7*, 97–110.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*, 124–158.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, *2*(2).

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.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–100.

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

Onwuegbuzie, A. J., Daros, D., & Ryan, J. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in Mathematics*, *19*(4), 11–35.

Peebles, D. (2008). The effect of emergent features on judgments of quantity in configural and seperable displays. *Journal of Experimental Psychology: Applied*, *14*(2), 85–100.

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., & 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.

Schutz, H. G. (1961). An evaluation of formats for graphic trend displays—experiment II. *Human Factors*, *3*, 99–107.

Simcox, W. A. (1983). *A perceptual analysis of graphic information processing.* Unpublished doctoral dissertation, Tufts University, Medford, MA.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

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*(6), 1073–1079.