JUDGMENTS OF PROPORTION WITH GRAPHS: OBJECT-BASED ADVANTAGES

Olivier St-Cyr[1] and Justin G. Hollands
Human-Computer Interaction Group
Defence Research and Development Canada – Toronto
Toronto, Ontario, Canada

This study investigated whether a stacked bar's vertical arrangement or single-object properties underlie its accuracy for proportion judgments. We hypothesized that observers would be less accurate when stacked bars were separated than when they formed a single object, according to the object-based theory of attention (Duncan, 1984). Thirty participants judged proportions with three different graph types: bars, stacked bars, and staggered stacked bars. Stacked bars produced smaller error than staggered stacked bars, while bars produced the greatest error. The results show an object-based advantage for stacked bars, but a vertical arrangement advantage for staggered stacked bars over bars was also evident.

INTRODUCTION

This study investigates how people make quantitative judgments of graphical elements. Previous research on judgments of graphical elements (e.g., Hollands and Spence, 1992, 2001; Follette 2000) have shown differences in the effectiveness of graph types (e.g. pie chart, bars, stacked bars) as a function of task type. In particular, the stacked bar graph (where bars are stacked vertically to form a single bar object) has been shown to be effective for proportion judgments. In this study, we investigated whether the causal factor is the vertical arrangement of the constituent bars or the single-object properties of the stacked bar. Beyond typical graphing applications, bar graphs of various kinds are found in a variety of visualization systems, including battlespace visualization systems (e.g., Rayson, 1999). Graphical perception research is also relevant to the display of quantitative information in such systems.

Based on the hierarchical ranking of elementary graphical-perception tasks proposed by Cleveland and McGill (1984), one would expect the bar graph to be more accurate than the stacked bar, since all graphical elements of the former are aligned on one common axis. However, Follette and Hollands (1998) presented participants with bar graphs (where values were aligned to a common axis) and stacked bars, and had them make two kinds of proportion judgments: part-to-part (A / B) and part-to-whole (A / [A + B]). They found that whereas stacked bar graphs were better than bar graphs for part-to-whole judgments, the opposite was true for part-to-part judgments. The Cleveland and McGill ranking did not apply, therefore, to the results from the part-to-whole task.

However, there were only two quantities shown in any of the graphs used by Follette and Hollands (1998) and real-world graphs typically show many data values. Follette (2000) examined participants' accuracy when making the two kinds of proportion judgments with several data values. With multiple values it is possible for bars involved in the judgment to be either adjacent or non-adjacent. The results obtained in Follette's non-adjacent condition are of greater real-world relevance than the adjacent situation, because with more values it becomes less likely that values of interest will be adjacent. Follette found that bar graphs were less accurate than stacked bars for both tasks (although more noticeably in the part-to-part task) when the bars were not adjacent. This is surprising when considering the Cleveland and McGill (1984) hierarchy, since one would expect the bar condition to be more accurate than the stacked bar, given that all graphical elements of the bar graph are aligned on one common axis.

The motivation for the current work was to account for the higher accuracy of stacked bars relative to bars obtained by Follette (2000) for proportion judgments involving non-adjacent elements. We argue that Follette's results may be explained by object-based theories of attention (Duncan, 1984). Duncan showed that judgments of object elements are more accurate when the elements belong to a single object rather than different ones. Studies conducted by Baylis and Driver (1993) and Vecera and Farah (1994) found similar results with different stimuli. Similarly, in Follette's experiment, the different segments of a stacked bar are part of a single object, which may have produced an object-based advantage relative to separate bars.

To test this hypothesis, we created a third graph type called a *staggered stacked bar* (see Figure 1). The staggered stacked bar is similar to a stacked bar except that graphical elements are separated horizontally to ensure that they are not part of a single object. The bars forming the stack are jittered right or left at random so that they are non-contiguous. The vertical alignment of each graphical element is preserved. If an object-

[1] Olivier St-Cyr is now with the Cognitive Engineering Laboratory, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Ontario, Canada.

based hypothesis is true, accuracy for staggered stacked bars should be lower than that for stacked bars. In our experiment, participants had to make judgments of proportion using three different graph types: bars, stacked bars, and staggered stacked bars. Participants made part-to-part proportion judgments on non-adjacent graphical elements.

METHOD

Participants

Thirty volunteers (male and female) served as participants in this study. The participants (17 females and 13 males) were 18 years or above (mean age = 26.4 yrs) and had normal or corrected-to-normal vision. They were recruited from the staff of a scientific institute and students of nearby universities using local advertisements. They were selected based on their willingness to participate. They were remunerated for their participation.

Apparatus and Stimuli

The experiment was programmed in an E-Prime version 1.0 (Psychology Software Tools Inc., Pittsburgh, PA) script running under the Windows 98 operating system. An 18-inch (46 cm) diagonal ViewSonic VP181 liquid crystal display running at a resolution of 800 x 600 was used to display stimuli. A personal computer was used to control the presentation of graphs and collect participants' responses.

Twenty graphs were constructed for each graph type (bar, stacked bar, and staggered stacked bar graphs). Figure 1 shows examples. Sets of 5 percentage values summing to 100 were obtained from Follette (2000) and used as the data for each graph. Bar, stacked bar, and staggered stacked bar graphs were presented on the display. The widths of each bar segment was 2 cm and each graph showed five segments whose total heights were approximately 11 cm. For the bar graph condition, bars were separated by 2 cm horizontally. Two segments, labeled A and B, were used for the proportion judgments. Segment A was always smaller than segment B. Segments A and B were always separated by three other segments.

Design and Procedure

The experiment had a one-way within-subjects design with Graph Type (bar, stacked bar, or staggered stacked bar) serving as the independent variable. The order of conditions was counterbalanced across participants using a Latin square. Dependent variables were judged proportion and completion time.

Each participant read the experimental protocol and signed the informed consent form. The experimenter answered any general questions about the design, showed the participant the experimental setup, and described the task. Each participant judged the set of graphs for the 3 graph types twice, forming 6 blocks. Since there were 20 trials per block, there was a total of 120 proportion judgments. Each block was preceded by a three-trial practice session. The participant provided his/her answers as a percentage to the nearest integer using the numeric keypad. The participant pressed the enter key to complete the trial. The correct answer was displayed after each of the three practice trials. After the practice trials, the participant judged the set of 20 graphs for that block. Correct answers were not displayed. The participant completed the study at his/her own pace. The experiment took about 30 minutes to complete.

RESULTS

Absolute Error

An absolute error score (|judged proportion – true proportion|) was computed for each trial. A mean absolute error value was computed for each participant in each condition. These data were submitted to a one-way within-subjects analysis of variance (ANOVA) with graph type serving as the independent variable. Figure 2 summarizes the results. Graph type affected absolute error, $F(2,58) = 14.00$, $p < .0001$, with greater error for staggered stacked bars than stacked bars, and the largest error for bars, Newman-Keuls, $p < .05$.

The distance between elements A and B was greater in the bar condition than the stacked bar condition. Perhaps greater distance between graphical elements reduces judgment accuracy. Since the distance between A and B varied across trials in the staggered stacked bar condition, we examined the effect of the A-B distance in that condition on the average absolute error scores. The correlation was not significant ($p > .18$), suggesting that the distance between elements A and B did not underlie the stacked bar advantage.

Completion Time

A mean response time was computed for each participant in each condition. These data were submitted to a one-way within-subjects ANOVA. Graph type had no effect on completion time, $F(2,58) = 0.01$, $p > 0.90$. Mean values are shown in Figure 2.

DISCUSSION

We predicted that judgments of proportions shown in staggered stacked bars would be less accurate than judgments of stacked bars. The results were consistent with this prediction. In addition, bar graphs were less accurate than

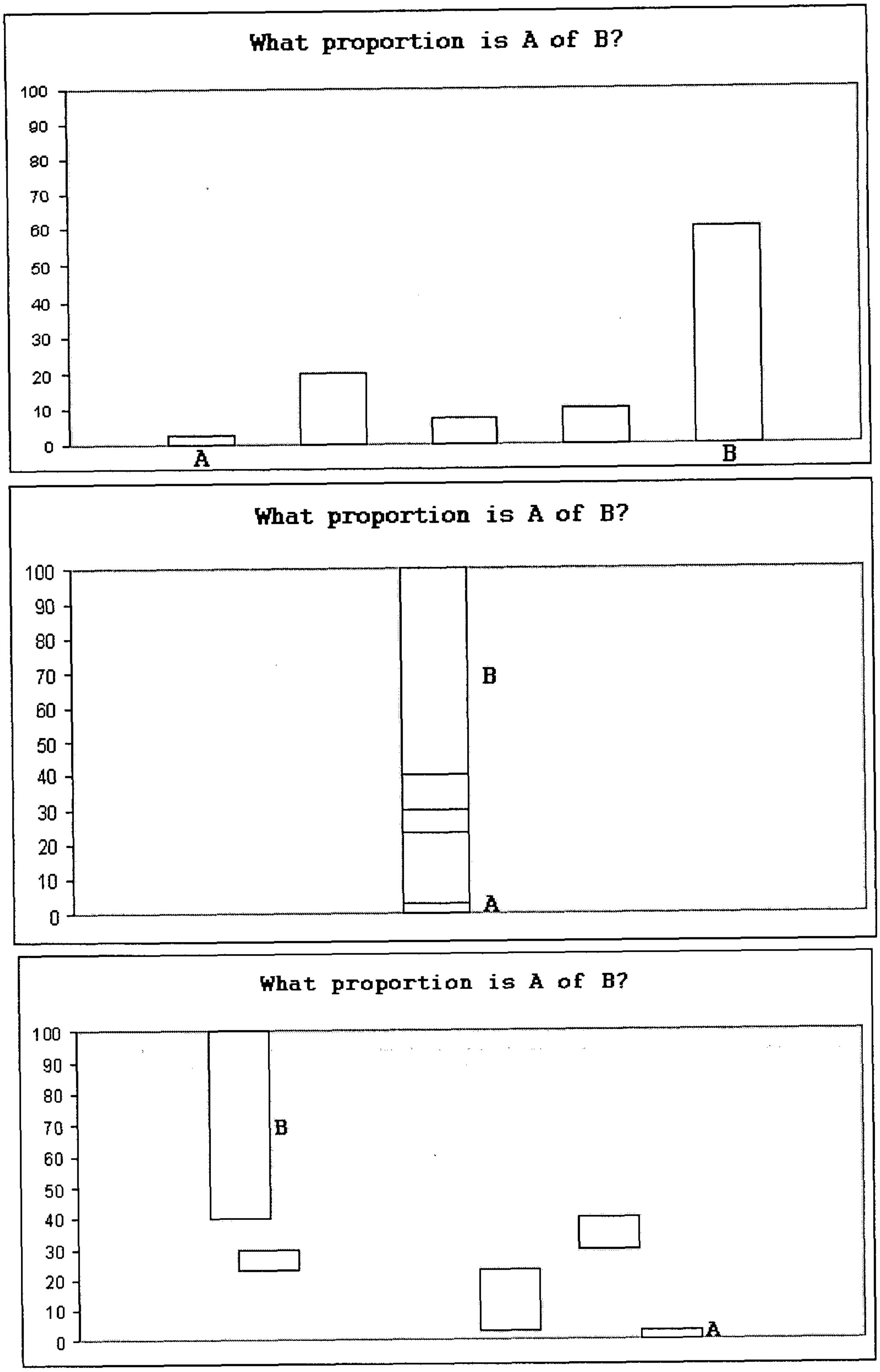

Figure 1. *Graph types used in experiment. bar graph (top), stacked bar graph (middle), and staggered stacked bar graph (bottom).*

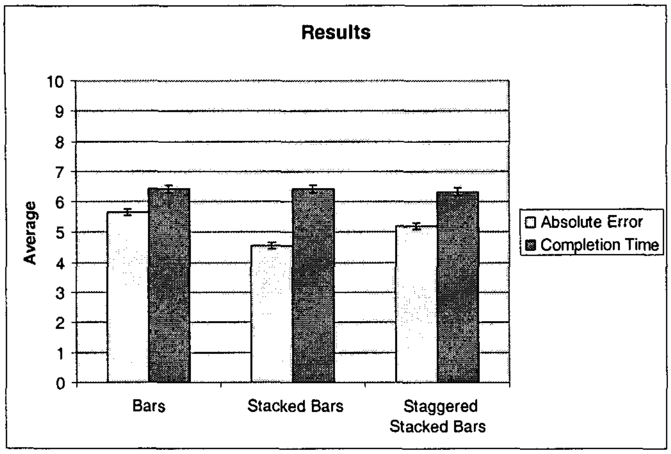

Figure 2: *Absolute error and completion time as a function of graph type. Error bars indicate the within-subjects standard error of the mean (Loftus & Masson, 1994).*

stacked bar graphs, replicating Follette (2000). Thus, a common baseline did not improve participants' judgments, despite the predictions of Cleveland and McGill (1984). An analysis of the effects of distance between elements suggests that the stacked bar advantage was not due to greater distance between bars in the bar graph condition.

The results also showed that staggered stacked bars were more accurate then bars, an unexpected result. This might be explained by the fact that staggered stacked bar graphs preserve the vertical arrangement of stacked bars, although graphical elements are not part of a single object.

Analysis for the completion times did not produced any significant results, showing no speed-accuracy tradeoff. However, this result stands in contrast to previous findings, since Follette (2000) found a difference between bars and stacked bars with respect to completion time. Our design was limited to non-adjacent graphs in a part-to-part task, whereas Follette's participants performed part-to-part and part-to-whole proportion judgments on adjacent and non-adjacent graphs. Perhaps having to switch between different arrangements and tasks led to transfer of strategies from these other conditions in Follette's experiment.

In our experiment, participants judged non-adjacent graphical elements. We believe this to be a more common situation than adjacent elements given that most graphs depict multiple elements, and the probability of adjacency is reduced with each added data point. Based on our results, this study proposes the following revision of Cleveland and McGill's (1984) ranking of perceptual tasks:

1- Vertical arrangement of graphical elements in a single object
2- Vertical arrangement of graphical elements
3- Graphical elements with common baseline

It appears that object-based advantages may indeed be an important element of graphical perception, although further investigation is necessary. In particular, the impact of task type, task switching and judged data elements on graph reading strategy needs to be better understood.

The staggered stacked bar graph introduced in this research was not intended to be a useful type of graphical representation. However, it allowed us to test our object-based stacked bars hypothesis. Other graphical methods (e.g. breaking the co-linearity of the left and right borders or connecting bars using horizontal contours) could have been used to assess our hypothesis. Testing other methods may provide converging evidence for the object-based hypothesis.

ACKNOWLEDGMENTS

This research was supported by a DRDC Technology Investment Fund grant to the 2nd author. We thank John Bowen for technical support.

REFERENCES

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 451-470.

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.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501-517.

Follette, L. (2000). *Comparing the effectiveness of bar and stacked bar graphs for judgments of proportion.* Unpublished Master's Thesis. University of Idaho, Department of Psychology.

Follette, L. and Hollands, J. G. (1998, October). *Bars vs stacked bars: Effectiveness is task dependent.* Poster presented at the 42nd Annual Meeting of the Human Factors and Ergonomics Society, Chicago, IL.

Hollands, J. G., & Spence, I. (1992). Judgments of change and proportion in graphical perception. *Human Factors, 34*, 313-334.

Hollands, J. G., & Spence, I. (2001). The discrimination of graphical elements. *Applied Cognitive Psychology, 15*, 413-431.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review, 1*, 476-490.

Rayson, J. K. (1999). Aggregate towers: Scale sensitive visualization and decluttering of geospatial data. *Proceedings IEEE Symposium on Information Visualization* (pp. 92-99). San Francisco: IEEE Computer Society.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146-160.