APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **19**: 953–962 (2005)
Published online 24 March 2005 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1105

Designing Bar Graphs: Orientation Matters

MARTIN H. FISCHER[1]*, NELE DEWULF[1,2] and ROBIN L. HILL[1]

[1]*University of Dundee, UK*
[2]*Ghent University, Belgium*

SUMMARY

We investigated whether recent discoveries about the cognitive representation of numbers would predict performance in a graph comprehension task. Participants decided verbally whether statements of the form 'A > B' were correct descriptions of subsequently presented bar graphs. We observed longer decision times for horizontal compared to vertical bar graphs and for negative compared to positive number graphs. Comprehension was faster when the spatial layout of magnitude information matched an internal 'mental number line'. These results show that the design of graphs can benefit from research into number representations. Copyright © 2005 John Wiley & Sons, Ltd.

Graphs are a very common means of communication. They appear in newspapers, magazines, scientific articles, and schoolbooks, in various forms and for different purposes. Harris (1999) provides a detailed overview of a wide range of charts, graphs, maps, diagrams, and tables, all of which can be used to convey information. In general, two different objectives can be distinguished when using graphs (Wickens, 1992): Either the author wants to convey a specific message to the reader, or he wants to support the exploration of a data set. In both cases it is important to think carefully about the design of the graph.

Several authors have provided empirically founded principles for designing user-friendly and effective graphs (e.g. Cleveland, 1985; Kosslyn, 1994; for a recent review see Shah & Hoeffner, 2002). From this work it is clear that performance with graphs is task-dependent. For example, Meyer, Shinar, and Leiser (1997) compared the relative efficiency of line graphs, bar graphs and tables. They found that reading exact values and identifying maxima led to advantages for bar graphs, whereas trend reading led to advantages for line graphs. Similarly, when participants see bar graphs, they tend to describe concrete contrasts in the data (higher, lower, greater than, less than); when they see line graphs, they tend to describe trends (rising, falling, increasing, decreasing; Zacks & Tversky, 1999).

*Correspondence to: Dr Martin H. Fischer, Department of Psychology, University of Dundee, Dundee DD1 4HN, UK. E-mail: m.h.fischer@dundee.ac.uk

Looking at bar graphs in particular, Feeny, Hola, Liversedge, Findlay, and Metcalf (2000; see also Fischer, 2000) recently documented the importance of spatial congruency between textual and graphical information. Participants compared two bars labelled A and B, with a statement about the relation between A and B underneath (e.g. $A > B$). Verification times were faster when the spatial order of the statement matched that of the labels in the graph. In summary, bar graphs are an efficient but task-dependent means of visualizing numbers, and visual-spatial attributes of bar graphs place important constraints on performance.

Users of modern graph-generating software have many choices when they design bar graphs—for example, there are several different bar graph formats on offer in Microsoft Excel©). With respect to the selection between vertical or horizontal orientation of the bars, no empirical evaluation or recommendation has been published. Our literature search merely turned up arbitrary illustrations of the two options without empirically based selection rationale (e.g. Harris, 1999, pp. 27ff and 79ff; Kosslyn, 1994, p. 38; Kostelnick & Roberts, 1998, pp. 267f. and 274f;[1]). This is surprising in the light of recent research on the cognitive representation of numbers, which makes clear predictions for the usability of horizontal and vertical bar graphs. We provide here a brief review of some basic findings and then report an empirical assessment of the two graph formats.

Number processing has been studied extensively over the past decades. One basic performance observation is the *distance effect*: It is more difficult to decide which of two numbers is the smallest or the largest if they are numerically close. For instance, it takes longer to compare 2 and 3 than 2 and 9 (Moyer & Landauer, 1967). This effect also obtains when quantity information is task-irrelevant, as in number naming (Brysbaert, 1995) or visual matching (Dehaene & Akhavein, 1995). When numerical distance is held constant, it is easier to discriminate between two small numbers, such as 15 and 20, than between two larger numbers, such as 85 and 90. This *size effect* also obtains across various tasks, such as number comparison (Den Heyer & Briand, 1986) and number naming (Brysbaert, 1995). Together, size and distance effects suggest that we possess a mental number line that represents magnitude information on an analogue scale with logarithmic compression. Whenever we deal with numerical information this representation is automatically activated and determines our performance. Thus, the larger the numbers or the smaller their distance, the harder the task.

A third frequently observed effect is the *SNARC effect* (Spatial-Numerical Association of Response Codes). Left responses are easier for small numbers and right responses for large numbers. This result is most often reported for speeded parity judgments with lateralized response buttons (Dehaene, Bossini, & Giraux, 1993; Fischer, 2003a) but also obtains when participants merely search for a vowel in the number's name (Fias, Brysbaert, Geypens, & d'Ydewalle, 1996) or detect lights next to a task-irrelevant number (Fischer, Castel, Dodd, & Pratt, 2003). This suggests that the mental number line is spatially oriented from left to right, thus leading to the observed spatial congruency effect. It is noteworthy that the SNARC effect is task-dependent, so that the same numbers (e.g. 4 or 5) can be associated with right space if the number range is 0–5, or with left space if the number range is 4–9. Similarly, small numbers can become associated with left space if interpreted as distances on a ruler, or with right

[1]Harris used the terms bar graph and column graph for horizontally and vertically oriented bars, respectively.

space if interpreted as times on a clock face (Bächtold, Baumüller, & Brugger, 1998). Thus, the left-right association is merely a default (at least in Western cultures), possibly reflecting reading habits.

Fischer (2003b) recently used the SNARC effect to show that we have separate cognitive representations for positive and negative numbers. He displayed digit pairs from the range −9 to 9 next to each other and found faster magnitude decisions when large negative values were on the left side, which is spatially congruent with a negatively extended mental number line. Experience with negative values on the abscissae of graphs may well explain this association, although it was also clear from the results that negative numbers required more processing time than positive numbers.

To summarize, research into numerical cognition suggests that magnitude information is represented on a mental number line where smaller numbers are more discriminable than larger ones. Furthermore, small numbers are automatically associated with left space, larger numbers with right space, and these spatial associations can extend to negative numbers. These findings have clear implications for the optimal visualization of numerical information in statistical graphs. The following study adopted the speeded verification method of Feeney et al. (2000) and Fischer (2000) to assess, for the first time, the relative usability of horizontal and vertical bar graphs. From the above evidence for automatic access to a horizontal mental number line we expected that horizontal bar graphs would support a faster understanding than vertical bar graphs. Several precautions were taken to prevent contamination of this comparison by other factors.

First, it was possible that vertical bar graphs would have an indirect advantage, due to the fact that the labels printed with each bar would be easier to read from left to right, compared to vertical reading of labels in horizontal bar graphs (see Figure 1). To assess this potential confound, we measured the spatial congruency effect between statements and graphs (Feeney et al., 2000; see also Fischer, 2000). Thus, if label reading contaminates the processing speed for graphs we would expect faster decisions when the order of labels in the vertical graph is spatially congruent with their order in a previously presented statement. For the horizontal graphs we expected no such effect.

Second, horizontal bars would extend to the left or right side, thus possibly inducing unspecific spatial compatibility effects for lateralized button responses. To avoid this selective contamination of the horizontal condition we recorded voice onset latencies with a microphone.

The influence of spatial-numerical associations on performance was investigated for both horizontal and vertical bar graphs. Specifically, we expected faster decisions for graphs that present smaller or negative magnitudes on the left side and at the bottom (SNARC effect). To prevent contamination of this assessment from the distance effect we selected pairs of numbers with a constant distance of 3. An odd distance was chosen to insure that each graph displayed one odd and one even number (Hines, 1990).

METHOD

Participants

Seven students (one male, six females) with normal or corrected to normal vision were paid £5 for participating. They were between 18 and 24 years old (average: 20.4 years) and naïve with regard to the purpose of the study. They were not specializing in quantitative subjects.

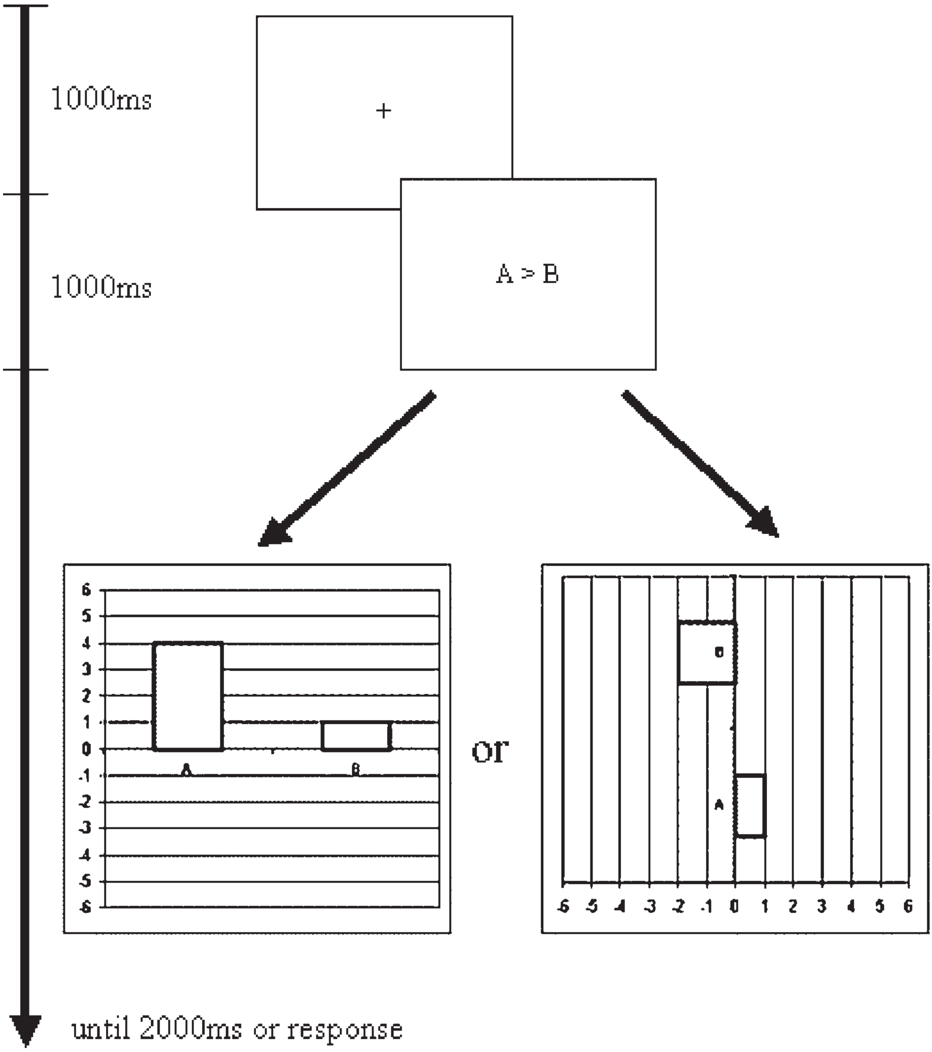

Figure 1. Flow chart of one trial. Two possible bar graphs are depicted, both requiring ' Yes' responses

Apparatus

Stimuli were delivered with DMDX experiment software (Forster & Forster, 2003) through a Pentium PC and displayed on a colour monitor (70 Hz) with 15 inch display diagonal. Voice onset times were digitally recorded with a microphone to the nearest millisecond.

Stimuli

All bar graphs (82 mm height, 87 mm width) were generated in black on a white background, using Microsoft Excel© (see Figure 1 for examples). Each contained two bars (14 mm wide, between 6 and 30 mm long) that were labelled 'A' and 'B' (Arial, size 8,

Table 1. The 12 pairs of digits used to design the bar graphs, and their average decision time for vertical and horizontal bar graphs in milliseconds. Condition labels refer to match (+) or mismatch (−) with numerical (N) or absolute (A) versions of the mental number line. See text for details

Condition		Displayed values		Vertical bars	Horizontal bars
N+A−		−5	−2	1,056	1,098
		−4	−1	1,058	1,079
		−2	1	1,006	1,032
	Mean			**1,040**	**1,070**
N+A+		−1	2	996	996
		1	4	929	1,035
		2	5	952	980
	Mean			**959**	**1,004**
N−A+		−2	−5	1,063	1,133
		−1	−4	1,042	1,125
		1	−2	995	1,013
	Mean			**1,033**	**1,090**
N−A−		2	−1	942	1,022
		4	1	934	1,057
		5	2	949	1,015
	Mean			**945**	**1,031**

bold) on the category axis. The value axis ranged from −6 to 6, with every integer including zero (Arial, size 10, bold) having its own gridline (< 0.5 mm wide). For vertical bar graphs the value axis was on the left side, for horizontal bar graphs it ran below the chart. The data pairs used to generate the graphs ranged from 'A = −5, B = −2' to 'A = 2, B = 5', and conversely from 'A = −2, B = − 5' to 'A = 5, B = 2'. Pairs containing 0 were excluded. This resulted in 12 pairs, of which four contained negative digits, four positive digits, and four both a positive and a negative digit (see Table 1). The fixation cross and all statements were printed in Arial, size 14, bold.

Task and procedure

Participants were tested individually. After obtaining informed consent, they were given a sheet with the instructions, and encouraged to ask questions if there was anything unclear. After tuning the sensitivity of the microphone, there were four practice trials with feedback.

Each experimental trial showed a fixation cross for 1,000 ms, followed by a statement about A and B (1,000 ms) and a graph for 2,000 ms or until response (see Figure 1). Participants responded 'Yes' if the graph corresponded with the statement and 'No' otherwise. Response latencies were recorded to the nearest millisecond. The experimenter typed responses on a different computer. There were four blocks of 96 trials, with breaks in between. The experiment lasted about 45 min.

Design

A 12 (digit pair) × 2 (graph orientation) × 2 (congruency of labels in statement with labels in graph) × 2 (order of labels: AB or BA) × 2 (Correct response: yes or no) within subjects design was used. This resulted in 192 different trials, which were presented twice in random order to each participant.

RESULTS AND DISCUSSION

Accuracy was high (average $= 97\%$, range 92–99%) and there was no significant correlation between accuracy and decision times ($r = -0.33$, $p = 0.472$). There was thus no speed accuracy trade-off, and errors were not analysed. Reaction times outside the range from 400 ms to 2,000 ms were excluded, as well as wrong answers and time-outs, which left 89% of the raw data for analysis.

Average decision time was 1,020 ms. An overview of results appears in Table 1. Contrary to the prediction of a horizontal number line, we found significantly *faster* decision times for vertical compared to horizontal bar graphs, $t(6) = 3.82$, $p = 0.009$. Average decision times were 992 and 1,048 ms, respectively. This result was not contaminated by differential reading times for the labels. A 2 (labels: congruent, incongruent) $\times$ 2 (orientation: vertical, horizontal) repeated measures analysis of variance (ANOVA) showed only the main effect of orientation, $F(1, 6) = 14.04$, $p = 0.01$, $\text{eta}^2 = 0.70$, with all other p values > 0.24.

To assess the impact of the SNARC effect on the processing of positive and negative magnitudes from horizontal and vertical bar graphs, the 12 displays were classified according to the schema proposed by Fischer (2003b; see Table 1). This schema codes whether the spatial layout of the display matches an absolute version of the mental number line (where only positive numbers are cognitively represented), and also whether the spatial layout of the display matches a numerical version of the mental number line (where both positive and negative numbers are cognitively represented, with negative numbers to the left of positive numbers). The assessment of spatial congruity with the *absolute* version of the mental number line is coded as A+ (or matching) when numbers with small absolute values are on the left of numbers with larger absolute values (e.g. 2/5 or $-2/-5$). The assessment of spatial congruity with the absolute version of the mental number line is coded as A-(or mismatching) when numbers with small absolute values are on the right of numbers with larger absolute values (e.g. 5/2 or $-5/-2$). Similarly, the assessment of spatial congruity with the numerical version of the mental number line is coded as N+ (or matching) when numbers with small numerical values are on the left of numbers with larger numerical values (e.g. 2/5 or $-5/-2$). The assessment of spatial congruity with the numerical version of the mental number line is coded as N-(or mismatching) when numbers with small numerical values are on the right of numbers with larger numerical values (e.g. 5/2 or $-2/-5$).

A 2 (orientation: vertical, horizontal) $\times$ 2 (numerical number line: match, mismatch) $\times$ 2 (absolute number line: match, mismatch) ANOVA replicated the main effect of orientation, $F(1, 6) = 14.00$, $p = 0.01$, $\text{eta}^2 = 0.70$, but also found a marginal interaction between orientation and numerical match, $F(1, 6) = 4.86$, $p = 0.07$, $\text{eta}^2 = 0.45$. This was due to a 23 ms benefit of numerical matches over mismatches in the *horizontal* bar graphs, $t(6) = 4.17$, $p = 0.01$, and an insignificant 11 ms cost of numerical matches compared to mismatches in the vertical orientation, $t(6) < 1$. The highly significant interaction between numerical and absolute match, $F(1, 6) = 138.58$, $p = 0.001$, $\text{eta}^2 = 0.96$, reflected the fact that displays where both factors matched or mismatched were easier to classify than mixed matches. As can be verified in Table 1, this advantage was associated with the presence of only a single negative number in these displays, whereas there were five negative numbers in the mixed match conditions. There were no other reliable effects, all p values >0.14.

The processing cost associated with negative number displays was further investigated by regrouping the data according to whether the displays contained only positive, only

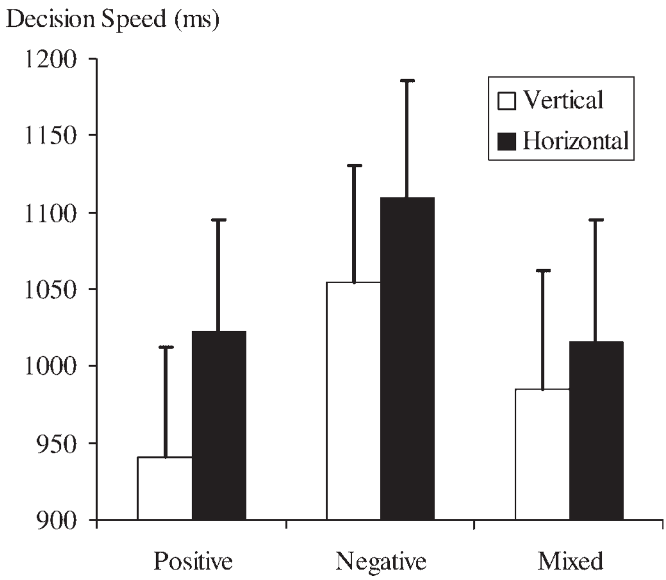

Figure 2. Interaction effect of display by orientation. Error bars are one standard error of the mean

negative, or both positive and negative numbers. A 3 (display: positive, negative, mixed) $\times$ 2 (orientation: vertical, horizontal) ANOVA showed significant main effects of display, $F(2, 12) = 55.79$, $p = 0.001$, eta$^2 = 0.90$, and orientation, $F(1, 6) = 14.12$, $p = 0.009$, eta$^2 = 0.70$. The average decision times for positive, negative, and mixed displays were 982, 1,082, and 1,000 ms, respectively. All pairwise differences were reliable, $p \leq 0.05$. The marginal interaction of both factors, $F(2, 12) = 3.58$, $p = 0.06$, eta$^2 = 0.37$, is shown in Figure 2. It was due to a significantly larger orientation effect for positive compared to mixed displays, $t(6) = 7.99$, $p = 0.001$, whereas the orientation effects did not differ between the other conditions, all $p > 0.30$. Moreover, the orientation effect was reliable for positive (81 ms, $t(6) = 5.12$, $p = 0.002$) and negative displays (54 ms, $t(6) = 2.57$, $p = 0.043$), but not for mixed displays (31 ms, $t(6) = 1.76$, $p = 0.129$).

CONCLUSIONS AND RECOMMENDATIONS

Research into the cognitive representation of numbers suggests that we represent magnitude information on a mental number line that runs from left to right for increasing magnitudes. Thus, we expected horizontally displayed magnitudes to be more congruent with this cognitive representation and to yield faster decision times than vertically displayed magnitudes. The results revealed the opposite, showing faster decisions for vertically oriented bar graphs. This finding was not due to a speed-accuracy tradeoff, or faster label reading in the vertical graphs. We believe that our finding indicates the prevalence of vertical over horizontal bar graphs, thus making vertical bar graphs more familiar to our participants (cf. Kosslyn, 1994, p. 38). This hypothesis was supported by informal comments made by participants after the experiment. In addition, our result can be taken to indicate that the *vertical* dimension can also become associated with number magnitudes (see Dehaene, 1997, p. 82f. for a similar suggestion). This possibility receives support from a recent eye movement study: Schwarz and Keus (2004) found faster downward eye movements in response to smaller numbers and faster upward eye movements in response to larger numbers, thus reflecting a vertical SNARC effect. In addition, cross-cultural developmental studies by Tversky, Kugelmass, and Winter (1991) also suggest a universal preference to associate more with upward.[2]

Despite the falsification of a key prediction, two further results of the present study illustrate the mutual benefit from research into numerical cognition for the design of user-friendly graphs, and vice versa. First, we found that decisions were overall fastest for displays containing only positive numbers, followed by displays with mixed numbers, and slowest decisions for negative number displays. This result was not due to differences in the visual discriminability of the bar lengths because we kept the distances between bar endings constant. The penalty for negative numbers was predicted by Fischer's (2003b) report that also showed that negative number comprehension is less automatized and might require an additional processing step when compared with positive or mixed numbers (Fischer & Rottmann, in press).

And second, we found a processing advantage when magnitude information was spatially arranged so that it matched a negatively extended mental number line. This observation supports Fischer's (2003b) finding of a cognitive representation of negative numbers. Interestingly, this advantage was only present for the *horizontal* bar graphs, indicating that a processing benefit for horizontal over vertical bar graphs can emerge when negative numbers are displayed in a way that is congruent with the (numerical) mental number line.

What are the practical implications of our finding of small decision speed advantages when comparing the comprehension of horizontal and vertical bar graphs? It is unlikely that repeated processing of suboptimally designed bar graphs leads to an accumulation of more substantial processing costs in real-life tasks. But it must be noted that processing speed and accuracy are two related aspects of performance. Our method required that response accuracy was held constant at a fairly high level. The observed effects of poor design on decision speed indicate that suboptimally designed graphs are likely to induce accuracy problems in tasks with time pressure. These could result in erroneous encoding or poorer memory for these graphs. Future research should look at this prediction because it is also conceivable that more processing time leads to better memory for the poorly designed graphs, due to the need for deeper processing (Craik & Lockhart, 1972). In general, more complex graph comprehension tasks should be studied to insure that the present results generalize to real world situations.

In conclusion, bar graphs are a suitable visualization medium to support the reading of exact values and identification of maxima (Meyer et al., 1997), as well as to describe concrete contrasts in the data (Zacks & Tversky, 1999). The use of irrelevant depth cues should be avoided (Fischer, 2000; Siegrist, 1996). The results of the present study further suggest that vertical bar graphs are user-friendlier than horizontal bar graphs. If possible, the values to be displayed should be linearly transformed to be all positive. If this is not possible, negative values should be displayed to the left of or below positive values, and smaller positive values to the left of or below larger positive values.

[2]We thank an anonymous reviewer for pointing this observation out to us.

ACKNOWLEDGEMENTS

ND conducted this work while she was sponsored through an Erasmus student exchange from the European Union. We thank Dan Wright and two anonymous reviewers for their helpful comments.

REFERENCES

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus–response compatibility in representational space. *Neuropsychologia, 36*(8), 731–735.

Brysbaert, M. (1995). Arabic number reading: on the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General, 124*, 434–452.

Cleveland, W. S. (1985). *The elements of graphing data.* Monterey, CA: Wadsworth.

Craik, F. I. M., & Lockhart, R. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behaviour, 11*, 671–684.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(2), 314–326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.

Den Heyer, K., & Briand, K. (1986). Priming single digit numbers: automatic spreading activation dissipates as a function of semantic distance. *American Journal of Psychology, 99*, 315–340.

Feeny, A., Hola, A. K. W., Liversedge, S. P., Findlay, J. M., & Metcalf, R. (2000). How people extract information from graphs: evidence form a sentence-graph verification paradigm. *Theory and Application of Diagrams, Proceedings*, 149–161.

Fias, W., Brysbaert, M., Geypens F., & d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: evidence from the SNARC effect. *Mathematical Cognition, 2*, 95–110.

Fischer, M. H. (2000). Do irrelevant depth cues affect the comprehension of bar graphs? *Applied Cognitive Psychology, 14*, 151–162.

Fischer, M. H. (2003a). Spatial representations in number processing—Evidence from a pointing task. *Visual Cognition, 10*(4), 493–508.

Fischer, M. H. (2003b). Cognitive representation of negative numbers. *Psychological Science, 14/3*, 278–282.

Fischer, M. H., & Rottmann, J. (in press). Do negative numbers have a place on the mental number line? *Psychology Science*.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience, 6*(6), 555–556.

Forster, K. I., & Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*(1), 116–124.

Harris, R. L. (1999). *Information graphics: A comprehensive illustrated reference*. New York, Oxford: Oxford University Press.

Hines, T. M. (1990). An odd effect: lengthened reaction times for judgments about odd digits. *Memory and Cognition, 18*, 40–46.

Kosslyn, S. M. (1994). *Elements of graph design*. New York: Freeman.

Kostelnick, C., & Roberts, D. D. (1998). *Designing visual language*. Boston: Allyn & Bacon.

Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors, 39/2*, 268–286.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*, 1519–1520.

Schwarz, W., & Keus, I. (2004). Moving the eyes along the mental number line: comparing SNARC effects with manual and saccadic responses. *Perception and Psychophysics, 66*(4), 651–664.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: implications for instruction. *Educational Psychology Review, 14/1*, 47–69.

Siegrist, M. (1996). The use or misuse of three-dimensional graphs to represent lower-dimensional data. *Behaviour and Information Technology, 15/2*, 96–100.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic production. *Cognitive Psychology, 23*, 515–557.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd edn). New York: HarperCollins Publishers.

Zacks, J., & Tversky, B. (1999). Bars and lines: a study of graphic communication. *Memory & Cognition, 27/6*, 1073–1097.

