The Classification of Graphical Elements

J. G. HOLLANDS, *Defence R&D Canada – Toronto*

Abstract In three experiments, participants classified stimuli depicting pie charts and stacked bar graphs on two criteria: a proportion shown in the graph, and the graph's overall size (scaling). Sorting times and errors were measured. For stacked bars, performance was impaired when participants sorted on the proportion and scaling varied. No such impairment occurred for pie charts. Experiment 1 showed that varying scaling produced Garner interference in classification of proportions with stacked bars, but not pies. Experiment 2 showed that this result held when the position of the pie slice was varied; Experiment 3 results showed facilitation for particular combinations of proportion and scaling levels. In general, the results showed that proportion and scaling had an asymmetric integral relation for stacked bar graphs, but were separable dimensions for pie charts.

When we decide if a proportion shown in a graph is large or small, how is our judgment affected by changes in the scaling or overall size of the graph? Garner (1974) developed the notions of integral and separable dimensions to describe how variation in one aspect (or dimension) of a stimulus facilitates or interferes with perception of variation in another. The importance of Garner's integrality-separability distinction for applied psychology is discussed in textbook treatments (e.g., Wickens & Hollands, 2000). Some major theoretical developments in engineering psychology and cognitive ergonomics have been influenced by Garner's work, for example, the proximity compatibility principle (e.g., Carswell & Wickens, 1987; Wickens & Carswell, 1995) and the emergent feature concept (e.g., Bennett & Flach, 1992). Garner's research on the dimensional structure of stimuli has been shown to apply to the design of graphical displays (Bennett & Flach, 1992). In this paper, I employed Garner's methodological and theoretical approaches to the classification of quantities shown in two common graphical forms. The aim was to demonstrate the utility of the classification task as a metric for the effectiveness of a graphical display, and also to see if a higher-order property of a stimulus (e.g., a depicted proportion) would act like the perceptual dimensions (e.g., size, shape, colour hue) usually manipulated in classification studies.

THE CLASSIFICATION TASK

Felfoldy and Garner (1971) measured the time required to classify a deck of stimulus cards. Each card portrayed a circle with drawn diameter. Within a deck, the circle size and the orientation of the diameter line were varied in particular ways. In one deck, circle size varied while line orientation remained constant, and participants were required to sort the cards based on size. In another deck, orientation was varied and size remained constant, and participants sorted based on orientation. These sorts served as control conditions. In a third deck, dimensions were *orthogonal* so that all four possible combinations of size and orientation were in the deck. Again, participants were asked to sort on size in one condition, and to sort on orientation in another. In a fourth deck, dimensions were perfectly *correlated*, that is, the large circle was always shown with one orientation and the small circle was always shown with the other orientation. Participants sorted on size and orientation in separate conditions with this deck.

Felfoldy and Garner (1971) argued that if dimensions are separable, participants should process each dimension independently, without regard to the other. If dimensions are integral, information from each dimension is integrated, and classification is based on some mentally redefined, single dimension. Hence with integral dimensions, orthogonal dimensions produce perceptual interference (sometimes called *Garner interference*; Morein-Zamir, Henik, & Spitzer-Davidson, 2002) and correlated dimensions produce perceptual facilitation. With separable dimensions, neither effect occurs. Indeed, in the Felfoldy and Garner experiment described above, size and orientation were found to be separable dimensions (i.e., no interference with orthogonal dimensions, no facilitation with correlated dimensions). In another experiment reported by Felfoldy and Garner, saturation and hue of colour chips were found to be integral dimensions (i.e., both interference and facilitation effects occurred).

Some stimulus dimensions have an *asymmetric integral* relationship (Garner, 1974) where variation in

Canadian Journal of Experimental Psychology, 2003, 57:1, 38-47

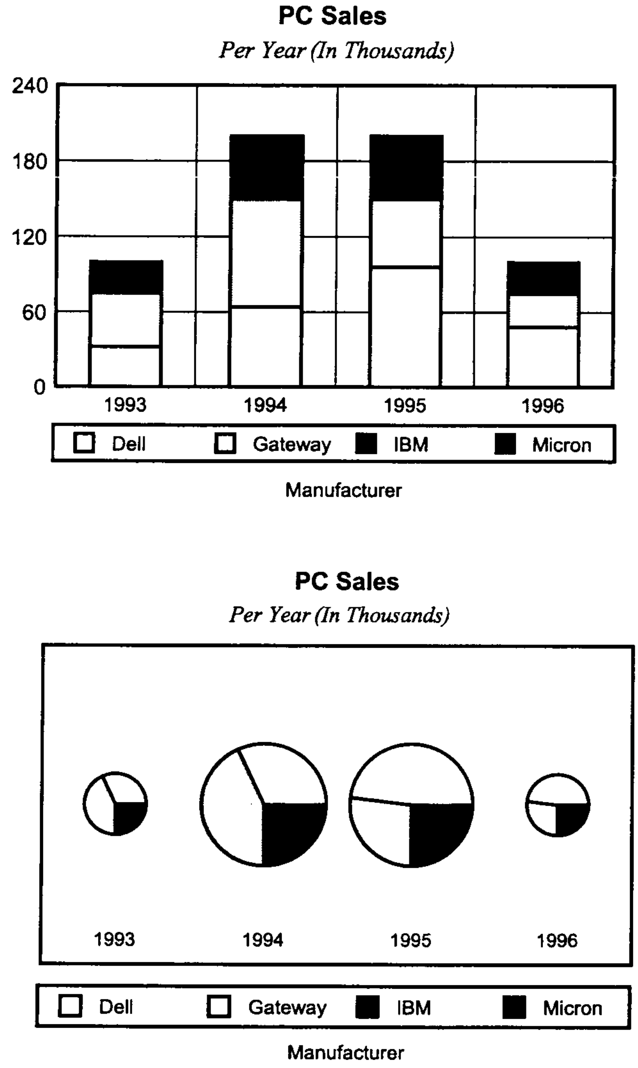

Figure 1. Examples of stacked bar graphs and pie charts showing proportional data over time. The data are fictitious.

dimension x affects classification of y, but not the reverse. This tends to occur when there is a dependent relationship between the two dimensions. More formally, if dimension x exists then dimension y exists, but if y exists, x may or may not exist. For example, a vowel phoneme cannot exist without a pitch, but pitch may exist as a dimension without any linguistic properties. Colour can exist without a form, but a form must have a colour.

RELEVANCE TO GRAPHICAL PERCEPTION

Some researchers have applied Garner's ideas to the perception of the graphical elements found in statistical graphics. Carswell and Wickens (1990) required participants to classify perceptual features commonly used in graphs. In their task, participants made absolute judgments in each classification; for example, participants classified graphical elements in terms of slopes, linear extent, areas, or colours. They found little evidence for integral dimensions with graphical stimuli. This type of perceptual discrimination is an important part of the graph-reading activity. However, there are other graph reading tasks. For instance, the graph reader often makes relative judgments, such as the estimation of proportion (e.g., what percentage is quantity *A* of the whole?) (e.g., Hollands & Spence, 1992, 1998, 2001).

Data are often plotted in absolute form but it is their relation to some larger value that is of interest. It would not be unusual graphical practice to depict market share for several manufacturers over several years using a stacked bar graph with the height of each stacked bar representing the total sales for that year, as shown in Figure 1 (the data are fictitious). Suppose the graph reader is trying to determine if Dell's market share was larger in 1996 than in 1994. The judgment is problematic because total sales vary year to year, and therefore it is possible to have greater absolute sales, but lower market share (or vice versa). No perceptual dimension (e.g., length, area) represents the proportion; it must be derived through comparison of Dell's sales with total sales for each year.

Figure 1 also shows a set of pie charts depicting the same data. The area of each pie chart represents total computer sales for one year. Despite the variation in total area, examination of Figure 1 suggests that determining Dell's market share with pies is easier than with stacked bars. The perceptual dimension of angle codes the variation in proportion regardless of the scaling changes affecting area.

GENERAL PREDICTIONS

The aim of the present experiments was to examine how variation in one dimension of a graphical form affects judgments of another dimension. With graphs, important dimensions may not be perceptual, but rather derived or computed through cognitive operations. I propose that such *cognitive dimensions* will have asymmetric integral relations with respect to a perceptual dimension. That is, a cognitive dimension (e.g., proportion) will show interference or facilitation if another perceptual dimension varies, but variation of the cognitive dimension will not affect classification along the perceptual dimension. It is probable that given our widespread exposure to various perceptual dimensions, their processing is more likely to be automatic, whereas cognitive dimensions are more likely to require the use of controlled processing subject to interference (Schneider & Shiffrin, 1977).

Experiments were designed to test these hypotheses. Participants sorted decks of stimulus cards, and the amount of time required for each sort was measured. In Experiment 1, participants classified stacked bar graphs and pie charts. The overall size of the graph (scaling) and the proportion were varied. Experiment 2 was a check to see if the results of Experiment 1 were due to participants judging slope instead of angle. In Experiment 3, the manner in which levels of scaling and proportion were assigned was examined for correlated dimensions.

Experiment 1

Consider the stacked bars shown in Figure 2 and imagine classifying the proportion as large or small. Changing the scaling of the whole graph affects the perceptual dimensions of area and height for the part being compared to the whole. It appears that participants need to estimate a part-to-whole ratio (a *ratio estimation* operation; Hollands & Spence, 1992, 1998, 2001) to obtain the cognitive dimension of proportion prior to classification. Hence, varying scaling when participants must classify the proportion should produce Garner interference relative to a control condition where the scaling remains constant and areas or heights can be judged. Conversely, variation in the proportion does not affect the perceptual dimensions of the stacked bar's area and height; hence, varying the proportion should not interfere with judgments of scaling. We would also expect facilitation when proportion and scaling are correlated and the proportion is being judged, but would not expect facilitation when the scaling was judged. If these results occur, it would suggest that proportion and scaling were asymmetric integral dimensions for stacked bars.

Now consider the pie charts shown in Figure 2. If participants were to classify large and small proportions depicted within large and small pie charts, the perceptual dimension of angle could be used to classify proportions. Thus, the cognitive dimension of proportion need not be explicitly estimated. Assuming that angle and scaling are separable dimensions (given the similarity to the stimuli used by Felfoldy & Garner, 1971), there should be no Garner interference with orthogonal dimensions, nor should there be perceptual facilitation with correlated dimensions. Further, it seems unlikely that a change in the proportion should have an effect on judgments of scaling, since variation in the proportion does not affect the area of the pie.

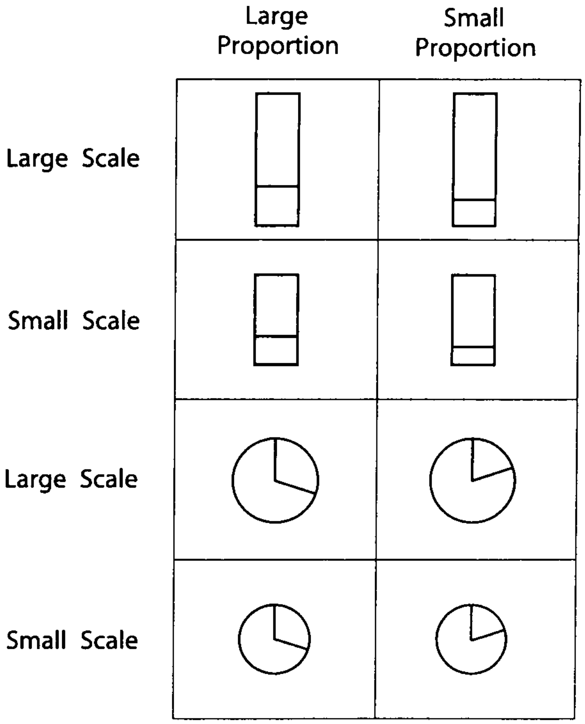

Figure 2. Stimuli used in Experiment 1.

A contrary view is based on the work of Cleveland (1985), who ordered several perceptual features used in graphs (e.g., length, angle, area) in terms of their accuracy. Cleveland's ranking is based on a series of experiments (Cleveland & McGill, 1984) in which participants estimated the proportion one graphical element was of another (e.g., what percentage is A of B?). Based on this ranking, better performance would be predicted for bars (nonaligned heights) than for pies (angles). This stands in distinction to the Garner interference predictions for the orthogonal sort, as better performance is expected with pie charts than stacked bar graphs. Hence, Experiment 1 will pit the prediction based on Garner's integrality-separability distinction against Cleveland's prediction based on the hierarchy of perceptual tasks.

METHOD

Generally, the experiment replicated Experiment 1 of Felfoldy and Garner (1971), except that different stimuli were used.

Participants. Forty-eight undergraduate students served as participants in Experiment 1. They received course credit for their participation.

Materials and apparatus. Four decks of 32 stimulus cards were constructed for each graph type. Each graph was drawn using CorelDraw software and printed on a Hewlett-Packard Laser Jet IIP before being mounted on a 10 x 15 cm index card. Large-, medium-, and small-scaled pie charts had 4-, 3.65-, and 3.3-cm diameters, respectively. Large-, medium-, and small-

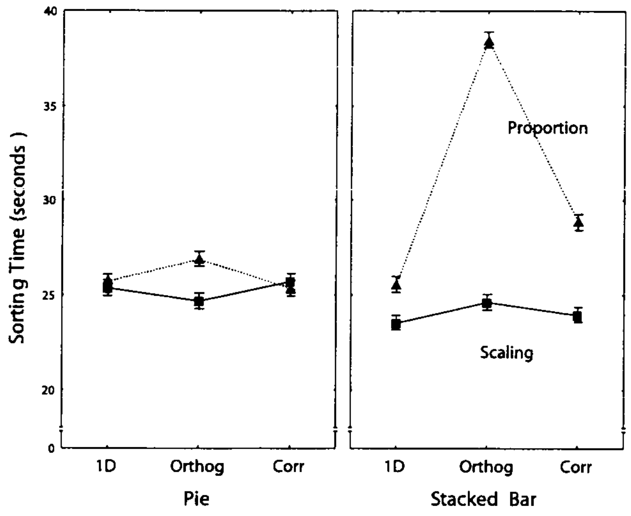

Figure 3. Experiment 1. Mean sorting time as a function of graph type, stimulus set, and sort type. Error bars in all graphs indicate the within-subjects standard error of the mean using the procedure described in Loftus and Masson (1994).

TABLE 1

Experiment 1. Mean Error Rate (Proportion of Cards Misclassified) as a Function of Graph Type, Stimulus Set, and Sort Type[a]

	Stimulus Set		
Sort Type	1D	Orthog	Corr
		Pie	
Proportion	.005	.014	.006
Scaling	.004	.003	.003
		Stacked Bar	
Proportion	.000	.137	.013
Scaling	.003	.000	.003

[a] The standard error of the mean (within-subjects, see Loftus & Masson, 1994) was 0.0052.

scaled stacked bars had 6.3-, 5.2-, and 4.3-cm heights, respectively. The width of all stacked bars was 2 cm. Hence, the areas of stacked bars and pies were approximately equal. Large, medium, and small proportions were 30%, 25%, and 20%, respectively.

In the first deck (one-dimensional size), only the scaling was varied. Large and small scalings were used, 16 of each type. The proportion was 25% for this set. In the second deck (one-dimensional proportion), the proportion was either large or small, and the scaling was medium. In the third deck (correlated), small proportions were plotted with large scaling, and large proportions were plotted with small scaling. There were 16 graphs of each type. In the fourth deck (orthogonal), small proportions were plotted with small scaling (8 graphs) and large scaling (8 graphs), and the same was true of large proportions. Figure 2 shows the various graphs used (large and small only).

A stopwatch was used to time each participant.

Design and procedure. A 2 x 3 x 2 (Graph Type by Stimulus Set by Sort Type) within-subjects design was used. The two graph types were pie and stacked bar; the three stimulus sets were one-dimensional, correlated dimensions, and orthogonal dimensions; the two sort types were classifying the overall size (scaling) of the graph, or the proportion shown. Following the procedure described in Felfoldy and Garner (1971), participants were instructed to sort each complete deck as fast as possible while avoiding any errors. Participants were allowed to examine the first few cards in the deck before each trial. Participants completed all sorts for one graph type before the other. The order of graph types was counterbalanced across participants. The order of the six conditions for each graph type was counterbalanced across participants using Latin squares.

RESULTS

Sorting times. Figure 3 shows mean sorting times for each condition. The sorting times were submitted to a 2 x 3 x 2 within-subjects analysis of variance (ANOVA). A three-way interaction between Graph Type, Stimulus Set, and Sort Type occurred, $F(2,94) = 19.4$, $MSE = 15.8$, $p < .00001$. Participants took longer with the orthogonal stimulus set when sorting on proportion with the stacked bar graph than any other condition, Newman-Keuls (NK) post-hoc test, $p < .001$. Correlated stacked bars also took longer than any other condition (except orthogonal) when sorting on proportion, NK, $p < .0005$. There were no other differences among the stacked bar conditions, NK, $p > .10$. There were no differences in sorting times among the pie chart conditions, NK, $p > .10$.

Errors. If the participant placed a card in the incorrect pile, this was classified as an error. The maximum number of errors (misclassifications) that a participant could make was 16 (piles were not defined in advance). The total number of errors in each sort was obtained for each participant and divided by 16 to obtain a proportion (error rate). In general, the pattern of error rates mirrored the sorting time data; mean values are shown in Table 1. Error rates were generally low (on average less than one misclassified card per sort) and therefore are not discussed further.

DISCUSSION

The results for stacked bar graphs showed that proportion and scaling were asymmetric integral dimensions. Participants showed interference with orthogonal dimensions when they sorted on proportion, but no interference occurred when participants sorted on scaling. When sorting on proportion with orthogonal dimensions, participants need to estimate the ratio of part to whole prior to classification. That is, the relevant dimension was not perceptual but rather derived through cognitive operations. When only proportion (or overall size) varied, the perceptual dimensions of areas or heights could be compared.

Participants took longer to classify proportion with stacked bars when scaling and proportion were correlated dimensions. If the dimensions were integral, facilitation should have occurred in this situation. In Experiment 1, large proportions were shown with small scaling, and vice versa. With correlated dimensions it may matter how levels of dimensions are combined. That is, if large proportions had been shown with large scaling, facilitation might have occurred. I examined this question further in Experiment 3.

The pie chart results showed neither perceptual facilitation with redundant dimensions nor perceptual interference with orthogonal dimensions. This result suggests that participants did not derive the cognitive dimension of proportion, but rather used the perceptual dimension of angle when asked to sort on proportion. Because area and angle are separable dimensions (Felfoldy & Garner, 1971), scaling and proportion have separable relations with pie charts.

In contrast to Cleveland's (1985) prediction based on the hierarchy of perceptual tasks, the Experiment 1 results show the pie chart's advantage when classifying proportions and scaling varies orthogonally. However, the pie slices in Experiment 1 were always drawn so that one of the radii forming the slice was vertical. It is possible that participants were not judging angle with pies, but rather judging the slope of the changing radius. Indeed, this other slope was either positive (with a 20% proportion), zero (with a 25% proportion), or negative (with a 30% proportion), making classification on the basis of slope a distinct possibility. Slope discrimination is more accurate than angle discrimination (Regan, Gray, & Hamstra, 1996; Snippe & Koenderink, 1994), which may have provided a specious advantage for pies over bars in Experiment 1. In real-world pie charts there are often many proportions plotted, and only one will be aligned with the 12 o'clock position. Therefore, graph readers cannot judge slopes and must judge angles when comparing pie slices, either in the same pie or across different pies. To ensure that participants are classifying on angle, both radii forming the angle should vary. This was investigated in Experiment 2.

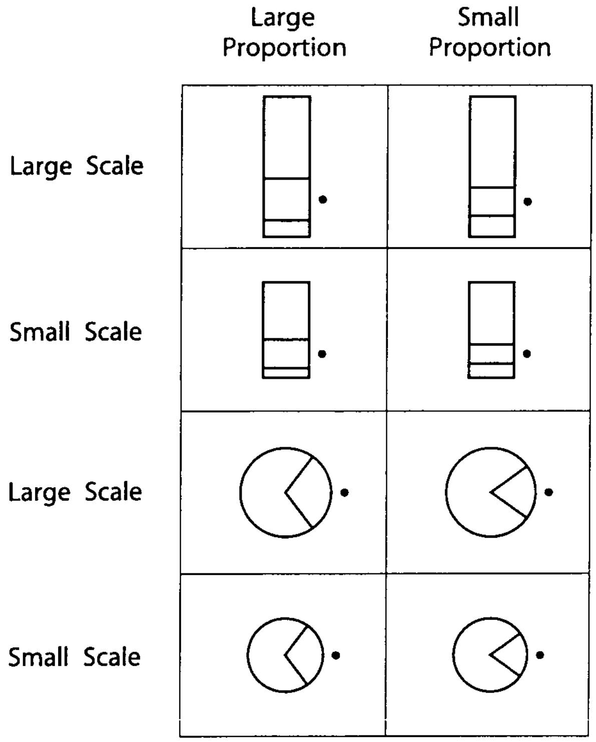

Figure 4. Shifted stimuli used in Experiment 2.

Experiment 2

In Experiment 2, proportions shown in pie charts were either drawn in a fixed-start position, as in Experiment 1, or using a horizontal bisector, as shown in Figure 4. Proportions were shown in corresponding locations for stacked bars, also shown in Figure 4. Participants always sorted on proportion in Experiment 2, with orthogonal or one-dimensional stimulus sets. If participants sorted on angle with both fixed-start and horizontal bisector positions, there would be little difference between these two conditions. If, however, participants sorted on slope in the fixed-start condition, but on angle in the horizontal bisector condition, we would expect faster sort times in the fixed-start condition. Similar results were predicted for error.

METHOD

Participants. Sixty-four undergraduate students served as participants in Experiment 2. They received course credit for their participation.

Materials and apparatus. The materials were the same as in Experiment 1, with the following exception. Only

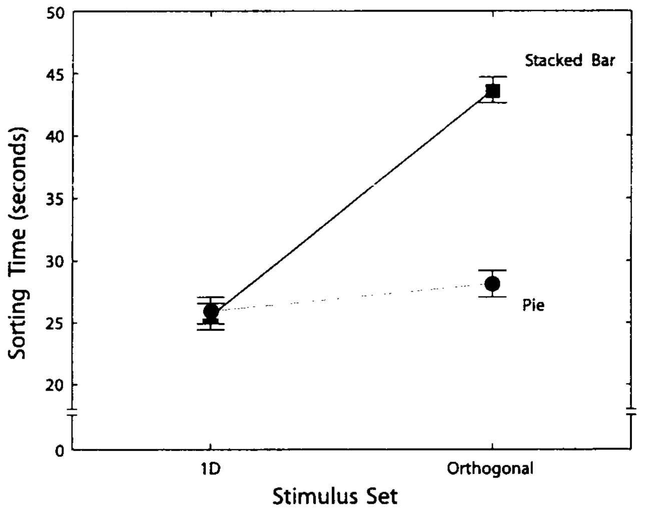

Figure 5. Experiment 2. Mean sorting time and standard error as a function of graph type and stimulus set.

TABLE 2
Experiment 2. Mean Error Rate (Proportion of Cards Misclassified) as a Function of Graph Type, Stimulus Set, and Position[b]

	Stimulus Set	
Position	1D	Orthog
	Pie	
Shifted	.005	.026
Fixed-Start	.004	.040
	Stacked Bar	
Shifted	.011	.216
Fixed-Start	.012	.158

[b] The standard error of the mean (within-subjects, see Loftus & Masson, 1994) was 0.0162.

orthogonal and one-dimensional stimulus sets were used. For half of the decks, the proportions were shifted to correspond to a horizontally bisected position for the pie chart, as shown in Figure 4; for the other half, the proportions were shown in the fixed-start position used in Experiment 1. Thus there were four decks for each graph type: orthogonal/fixed-start, orthogonal/shifted, one-dimensional/fixed-start, and one-dimensional/shifted. With stacked bars, identifying the proportion to be judged is problematic when it is shifted away from the base, as shown in Figure 4. To avoid this problem, a black dot was placed beside the to-be-judged proportion. For consistency, this was done with all stimulus sets. This technique has been used before to specify the to-be-judged proportion (e.g., Hollands & Spence, 1998).

Design and procedure. A 2 x 2 x 2 (Graph Type by Stimulus Set by Position) within-subjects design was used. The two graph types were pie and stacked bar; the two stimulus sets were one-dimensional and orthogonal; and the two positions were fixed-start or shifted. Participants sorted on the proportion dimension only. The order of the four conditions for each graph type was counterbalanced across participants using Latin squares, and the order of graph types was also counterbalanced. Other procedural details were as in Experiment 1.

RESULTS

Sorting times. The sorting times were submitted to a 2 x 2 x 2 within-subjects ANOVA. There was no main effect of Position, nor did it interact with either of the other independent variables, all $ps > .40$. A two-way interaction between Graph Type and Stimulus Set occurred, $F(1,63) = 113.1$, $MSE = 72.1$, $p < .00001$. This result is shown in Figure 5. Sorting times were greater for orthogonal stacked bars than orthogonal pies, and both one-dimensional stacked bars and pies, NK, $p < .0005$. Sorting times were greater for orthogonal pies than one-dimensional pies, NK, $p < .05$.

Errors. The total number of errors in each sort was obtained for each participant, and a proportion error value computed. Mean values are shown in Table 2. As in Experiment 1, errors were infrequent (on average less than one per sort) and the pattern of error data mirrored the sorting time data.

DISCUSSION

In Experiment 2, proportions shown in pie charts were drawn in a fixed-start position, as used in Experiment 1, or using a horizontal bisector. When proportions were drawn in the fixed-start position, participants could have judged the slope of one radius rather than judge an angle. When drawn using a horizontal bisector, both radii forming the angle varied. The Experiment 2 results showed no effect of position, suggesting that participants were judging angles in both bisected and fixed-start situations. Shifting position with stacked bars also had little effect. Regardless of position, Cleveland's (1985) ranking was incorrect even when slices were not aligned with the 12 o'clock position – bars were less effective than pies for classification of proportion in the orthogonal condition.

As in Experiment 1, participants took longer and made more errors for orthogonal sorts with stacked bars than any other condition. Participants took slightly longer for orthogonal sorts than one-dimensional sorts with pies (a similar trend occurred in Experiment 1). However, the size of the difference obtained in

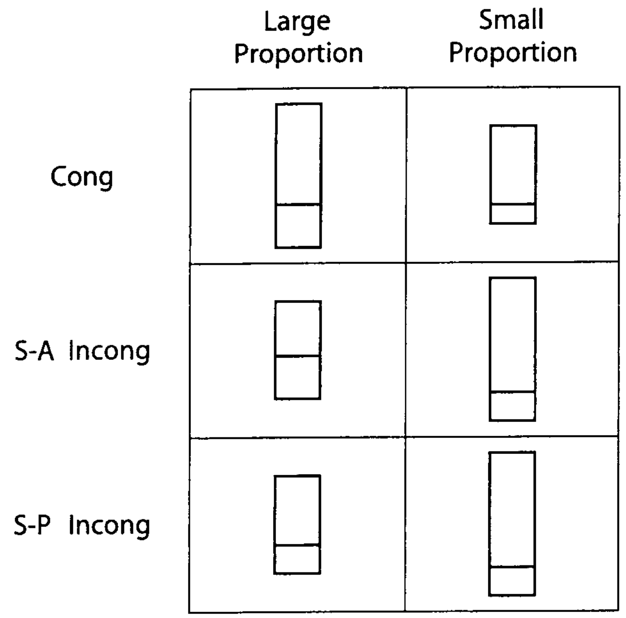

Figure 6. Correlated stimulus sets used in Experiment 3.

Experiment 2 (roughly 2 s) was much smaller than the 18-s difference between orthogonal and one-dimensional conditions for stacked bars. Many researchers (e.g., Shepard, 1991) maintain that the integrality/separability distinction should be viewed as a continuum rather than as a dichotomy. Certainly, the current evidence suggests that proportion and scaling are more separable for pies than for stacked bars.

Experiment 3

In the correlated sets used in Experiment 1, large proportions were shown with small scaling, and vice versa. There was no advantage in the correlated condition; indeed, participants took slightly longer in that condition with stacked bars. It may matter how levels of dimensions are combined. If large proportions had been paired with large scaling (and small proportions with small scaling), facilitation might have occurred for stacked bars. In Experiment 3, participants were shown correlated sets in which levels of scaling and proportion were *congruent* in this way. They were also shown correlated sets in which scaling (*s*) and proportion (*p*) were incongruent such that large proportions were shown with small scaling, and vice versa, just as in Experiment 1 (the *s-p incongruent* condition). If the congruent condition was faster then the s-p incongruent condition, it would indicate that the way in which levels of dimensions are combined is important. If the congruent condition was faster than the one-dimensional condition, it would suggest that facilitative effects do occur given the right combination of dimension levels, and would be consistent with an asymmetric integral relation between proportion and scaling dimensions with stacked bars.

In the s-p incongruent condition, the area (and height) of the smaller proportion when scaled large was the same as the area (and height) of the larger proportion when scaled small (as was the case for the correlated condition of Experiment 1). Hence, participants would need to estimate proportions and classify stimuli along that cognitive dimension. However, facilitation with incongruent stacked bars might occur if scaling was correlated with area (and height), in addition to proportion, as shown in Figure 6. The condition in which scaling is correlated with area *a* was the *s-a incongruent* condition. In this case, larger areas *and* proportions were correlated with smaller scaling, and participants could compare the areas representing the proportions and be perfectly accurate. The difference between large and small areas was larger than the difference in areas in the one-dimensional condition. Hence, if participants judged areas, there would be facilitation relative to the one-dimensional situation. Moreover there would be a difference between s-a incongruent and s-p incongruent conditions. If these results did not occur, it would suggest that participants were doing something other than simple area judgments in the s-a incongruent condition: presumably, estimating proportions.

Finally, it was predicted that participants would take longer with an orthogonal set of stacked bars than with a one-dimensional set, as found in Experiments 1 and 2.

Hence, the following 5 comparisons were planned:

1) one-dimensional and orthogonal;
2) one-dimensional and congruent;
3) one-dimensional and s-a incongruent;
4) s-p incongruent and congruent;
5) s-p incongruent and s-a incongruent.

METHOD

Participants. Fifty undergraduate students served as participants in Experiment 3. They received course credit for their participation.

Materials and apparatus. The materials were the same as in Experiment 1, with the following exceptions. Only stacked bar stimuli were used. In addition to one-dimensional and orthogonal stimulus sets, participants sorted three types of correlated sets: congruent, s-p incongruent, and s-a incongruent; stimuli from these correlated sets are shown in Figure 6.

Design and procedure. A one-way within-subjects design was used with Stimulus Set serving as the inde-

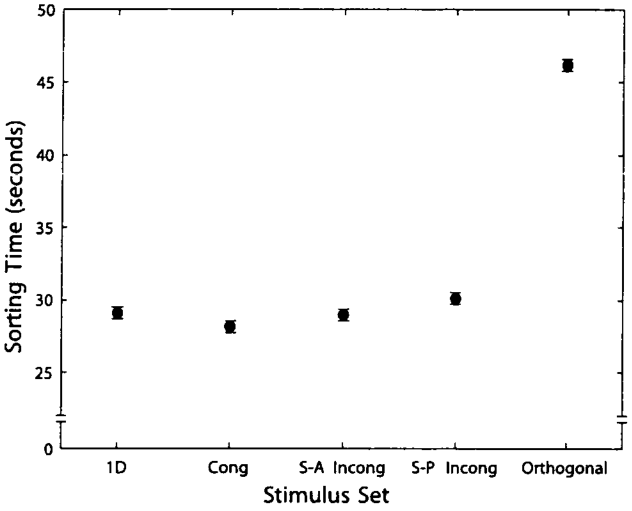

Figure 7. Experiment 3. Mean sorting time and standard error as a function of stimulus set.

pendent variable. Participants sorted stacked bars on the proportion dimension only. The five stimulus sets were: one-dimensional, orthogonal, congruent, s-p incongruent, and s-a incongruent. The order of the five conditions was counterbalanced across participants using Latin squares. Other procedural details were as in Experiment 1.

RESULTS

Sorting times. Figure 7 shows mean sorting times for each condition. The sorting times were submitted to a one-way within-subjects ANOVA. A main effect for Stimulus Set occurred, $F(4,196) = 89.4$, $MSE = 33.0$, $p < .00001$. Planned comparisons showed that orthogonal stimulus sets took longer than one-dimensional sets, $F(1,49) = 126.2$, $MSE = 57.6$, $p < .00001$, that congruent sets took less time than one-dimensional, $F(1,49) = 4.03$, $MSE = 5.8$, $p < .05$, and that there was no difference between s-a-incongruent and one-dimensional sets, $F(1,49) = 0.09$, $MSE = 9.53$, $p > .75$. Planned comparisons also indicated that participants took longer with s-p-incongruent sets than with congruent sets, $F(1,49) = 6.13$, $MSE = 15.6$, $p < .05$, and that there was no difference between s-p-incongruent and s-a-incongruent sets, $F(1,49) = 1.63$, $MSE = 21.03$, $p > .20$.

Errors. The total number of errors in each sort was obtained for each participant, and a proportion error value was computed. Mean values are shown in Table 3. As in previous experiments, the pattern of means reflected the sorting time data and errors were infrequent (on average less than one per sort).

TABLE 3
Experiment 3. Mean Error Rate (Proportion of Cards Misclassified) as a Function of Stimulus Set[c]

Stimulus Set				
1D	Cong	S-A Incong	S-P Incong	Orthog
.000	.000	.001	.021	.174

[c] The standard error of the mean (within-subjects, see Loftus & Masson, 1994) was 0.0061.

DISCUSSION

In Experiment 3, participants were shown correlated sets in which scaling and proportion were congruent, sets in which they were incongruent, and sets in which the area and height of the proportion were incongruent with scaling. The congruent condition was faster than the s-p-incongruent condition, indicating that the way in which levels of dimensions are combined was important. The congruent condition was faster than the one-dimensional condition, showing that facilitative effects can occur given the right combination of dimensions. Similar trends occurred for error. As in Experiments 1 and 2, participants took longer with orthogonal stacked bars than with the one-dimensional set. In combination with the results of the earlier experiments, the results generally suggest that proportion and scaling have an asymmetric integral relation for stacked bars.

In the s-p incongruent condition, the area (and height) of the smaller proportion when scaled large was the same as the area (and height) of the larger proportion when scaled small (as was the case for the correlated condition of Experiment 1). Hence, participants could not compare areas or heights when sorting proportions in the s-p incongruent situation. The s-a incongruent condition (in which larger areas and heights were correlated with smaller scaling) did not show facilitation with respect to the s-p incongruent or one-dimensional control conditions, suggesting that participants did not compare the areas or heights representing the proportions. Participants appeared to do something other than classify on area in the s-a incongruent condition: presumably, they estimated proportions and classified on this derived cognitive dimension.

General Discussion

ASYMMETRIC INTERFERENCE AND COGNITIVE DIMENSIONS

I have argued that a cognitive dimension should have an asymmetric integral relation with respect to a perceptual dimension. That is, a cognitive dimension should show interference (or facilitation) if another perceptual dimension varies, but variation of the cognitive dimension will not affect classification along the perceptual dimension. Experiment 1 showed that varying scaling produced interference in classification of proportions with stacked bars, but not pies. The implication is that proportion acted as a derived cognitive dimension for stacked bars, but not for pies, where the perceptual dimension of angle could be used. Experiment 2 showed that the position of the pie slice (12 o'clock vs. horizontal bisector) had no effect on classification, consistent with a judgment of angle regardless of pie slice position. Experiment 3 results showed facilitation for particular combinations of proportion and scaling levels with correlated stimulus sets. In general, the results showed that proportion and scaling had an asymmetric integral relation for stacked bar graphs. For pie charts, the perceptual dimension of angle can be used to judge proportion, so proportion and scaling act as separable dimensions for pie charts.

IMPLICATIONS

The Experiment 1 and 2 results showed that it was time-consuming to distinguish proportions drawn at different scales with stacked bars, but not with pies. This was not a trivial difference: In both experiments, participants took about one-and-a-half times as long to classify different-scale stacked bars than same-scale stacked bars or different scale pies. Participants did not use the extra time to achieve greater accuracy, since errors showed a similar pattern.

The experiments used only two proportions and two scalings, but stacked bar graphs often show many proportions with many scalings, as shown in Figure 1. When scaling varies, each comparison across years involves extra processing time for stacked bars, relative to pies. The results suggests that a set of pie charts whose area depicted total sales would better enable a graph reader to distinguish among the various proportions. In contrast to Cleveland's (1985) prediction based on the hierarchy of perceptual tasks, our experiments show the pie chart to advantage. Perhaps the angle information in the pie chart serves as a kind of emergent feature, allowing the graph reader to forego the intermediate computation of ratios.

The experiments in this paper were run when the author was at the University of Idaho. I thank Tanya Deyo, Lisa Follette, Victor Hoffer, Kelli Kludt, Clint Payton, and Candace Schmidt for running participants and preparing stimuli, and Brian Dyre for discussions about the experimental designs. I also thank Melody Carswell, Doug Gillan, Jocelyn Keillor, and Sharon McFadden for helpful comments. Results from Experiment 1 were presented at the 41st Annual Meeting of the Human Factors and Ergonomics Society, in Albuquerque, New Mexico, September 1997.

Correspondence concerning this article should be addressed to J. G. Hollands, Defence R&D Canada – Toronto, 1133 Sheppard Avenue W., P.O. Box 2000, Toronto, Ontario M3M 3B9 (E-mail: justin.hollands@drdc-rddc.gc.ca).

References

Bennett, K. B., & Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors, 34,* 513-533.

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

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

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

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphic methods. *Journal of the American Statistical Association, 70,* 531-554.

Felfoldy, G. L., & Garner, W. R. (1971). The effects of speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics, 9,* 289-292.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Hollands, J. G., & Spence, I. (1992). Judgments of change and proportion in graphical perception. *Human Factors, 35,* 313-334.

Hollands, J. G., & Spence, I. (1998). Judging proportion with graphs: The summation model. *Applied Cognitive Psychology, 12,* 173-190.

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.

Morein-Zamir, S., Henik, A., & Spitzer-Davidson, I. (2002). The importance of irrelevant-dimension variability in the Stroop flanker task. *Psychonomic Bulletin & Review, 9,* 119-125.

Regan, D., Gray, R., & Hamstra, S. J. (1996). Evidence for a neural mechanism that encodes angles. *Vision Research, 36,* 323-330.

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing I: Detection, search, and attention. *Psychological Review, 84,* 1-66.

Shepard, R. N. (1991). Integrality versus separability of stimulus dimensions: From an early convergence of evidence to a proposed theoretical basis. In *The perception of structure* (pp. 53-71). Washington, DC: American Psychological Association.

Snippe, H. P., & Koenderink, J. J. (1994). Discrimination of geometric angle in the fronto-parallel plane. *Spatial Vision, 8,* 309-328.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37,* 473-494.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Sommaire

Lorsque nous décidons si la proportion représentée dans un graphique est grande ou petite, comment notre jugement est-il perturbé par des changements apportés à l'échelle ou à la grandeur totale du graphique? Garner (1974) a élaboré les notions de dimension intégrale et de dimension séparable, pour décrire la façon dont la variation d'un aspect (ou dimension) d'un stimulus facilite ou perturbe la perception de la variation d'un autre aspect. Dans le présent article, les approches méthodologique et théorique de Garner, qui se rapportent à la classification de l'ordre de grandeur du stimulus, ont été appliquées aux quantités présentées dans deux formes graphiques courantes. L'objectif de la présente étude consistait, premièrement, à démontrer l'utilité de la tâche de classification en tant que norme de mesure de l'efficacité d'un affichage graphique et deuxièmement, à voir si le caractère d'ordre supérieur d'un stimulus (p. ex., la représentation de la proportion) jouerait le rôle des dimensions perceptuelles (p. ex., la dimension, la forme, la tonalité de la couleur) qui sont généralement manipulées lors d'études de classification. En particulier, nous affirmons qu'une dimension cognitive telle que la proportion possède une relation asymétrique intégrale par rapport à la dimension perceptuelle. C'est-à-dire qu'une dimension cognitive fera apparaître une interférence (ou une facilitation) si une autre dimension perceptuelle varie, mais la variation de la dimension cognitive n'influera par sur la classification reliée à la dimension perceptuelle.

Au cours de trois expériences, les participants devaient classer des stimuli représentant des graphiques à secteurs et des graphes à barres empilées selon deux critères : la proportion représentée par le graphique et la grandeur totale du graphique (l'échelle). Les erreurs et le temps de classification ont été mesurés. Dans le cas des graphes à barres empilées, la performance était altérée lorsque les participants devaient trier la proportion et l'échelle modifiées. Une telle altération de la performance n'a pas été observée dans le cas des graphiques à secteurs. L'expérience 1 a démontré que la modification de l'échelle entraînait une interférence dans la classification des proportions des barres empilées, mais pas dans le cas des secteurs du graphique. L'expérience 2 a démontré que ce résultat était obtenu lorsque la position du graphique à secteurs variait. Les résultats obtenus à l'expérience 3 indiquent que la facilitation de la performance s'observe lorsque des combinaisons particulières des niveaux de proportion et d'échelle sont réunies. En général, les résultats démontrent que la proportion et l'échelle possèdent une relation asymétrique intégrale dans le cas des graphes à barres empilées, mais que, conformément à nos hypothèses, elles deviennent des dimensions séparables dans le cas des graphiques à secteurs.

Les résultats des expériences 1 et 2 démontrent que la tâche consistant à distinguer les proportions présentées à différentes échelles absorbait plus de temps dans le cas des barres empilées, à l'inverse des graphiques à secteurs. La différence observée n'est pas insignifiante : dans les deux expériences, les participants ont mis une fois et demi plus de temps à classer les barres empilées présentées selon une échelle différente qu'à classer les barres empilées de même échelle ou les secteurs d'échelle différente. Les participants n'ont pas eu recours au temps supplémentaire qui leur était alloué pour répondre avec une plus grande exactitude aux items, car les erreurs démontraient un patron semblable.

Les expériences n'ont utilisé que deux proportions et deux échelles, mais les graphes à barres empilées présentent fréquemment plusieurs proportions et plusieurs échelles. Lorsque l'échelle varie, chaque comparaison, d'une année à l'autre, suppose le recours à un temps de traitement supplémentaire dans le cas des barres empilées, relativement aux secteurs. Ces résultats laissent croire qu'une série de graphiques à secteurs qui illustrerait les ventes totales permettrait davantage au lecteur du graphique de faire la distinction parmi les différentes proportions.

#519442
CA022517