

Numerical comparison processes

PETER DIXON

Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

The relative lengths of two rows of objects systematically affected the time to make judgments of the relative number of objects. The time to decide which row contained more objects generally increased with the amount of incongruity between length and number (e.g., reaction time was long when the numerically larger row was the shorter of the two). In addition, the time to decide that the rows contained an equal number of objects increased with the disparity in length between the rows. These results suggest that length information was used in the numerical comparison process. A model that uses an internal transformation of row length is proposed to account for the results.

A frequent mental process in everyday life consists of comparing two quantities and deciding whether they are equal in number, or if not, which of the quantities is greater. Such comparisons are usually made in the presence of a variety of irrelevant stimulus attributes. For instance, the physical extent of a group of objects (e.g., how much space it occupies) does not affect the number of objects. Nevertheless, several studies (cited below) have shown that irrelevant dimensions such as physical extent often affect performance in numerical comparison tasks. The experiments that follow will investigate the specific problem of how the time to compare two rows of objects is affected by the relative lengths of the rows. The results suggest that subjects can use length information to make the numerical comparisons quickly and do not have to actually count the objects in each row. A model of this comparison process will be proposed to account for the results.

An explanation of the effects of irrelevant dimensions on numerical comparisons must take into account the internal representation of number. Multidimensional scaling of similarity judgments suggests that there are at least two levels of number representation (Shepard, Kilpatrick, & Cunningham, 1975). For instance, when digits are compared in terms of their physical appearance, the scaling solutions indicate that features such as the presence or absence of enclosed spaces and the amount of curvature in the numeral are important in the internal representation. Thus one level of representation consists of physical features of the stimuli. However, subjects were also able to compare "abstract number concepts." In that case, numbers were represented in terms of their numerical magnitude and odd or even parity, and not in terms of physical properties. In fact, the scaling solutions for abstract

number concepts were almost identical no matter what form the stimuli were presented in. Thus, a numerical stimulus can be represented both in terms of its physical features and in terms of an abstract number code.

Both levels of representation might be used in numerical comparisons. For instance, the abstract level of representation would have to be used if one were deciding which of two digits is greater; the physical representation of digits in terms of curves and enclosed spaces offers no clue to relative numerical magnitude. On the other hand, if the stimuli were two rows of objects, the physical feature of row length might be used to make a numerical comparison. This would be an effective strategy whenever the longer row could be assumed to contain more objects. For instance, checkout lines at the supermarket can often be judged on length alone without actually counting the number of people in each line. Similarly, young children often attend solely to length when deciding which of several rows of objects is numerically larger (Gelman, 1972). Thus, both the abstract number representation and the physical feature representation can be used to make numerical comparisons.

The comparison of abstract number codes can be affected by irrelevant stimulus dimensions, such as physical size. The physical size of a digit obviously does not affect the numerical quantity it stands for. Thus physical size is logically an irrelevant dimension when comparing the magnitudes of two digits. Nevertheless, physical size affects the time to make digit comparisons (Hinrichs, Note 1). Reaction times are slower when the numerically larger digit in a pair is physically smaller. This may be referred to as an incongruity effect, since reaction time increases when the relative numerical magnitude and the relative physical size are incongruent. The incongruity effect may be due to interference between the dimensions of numerical magnitude and physical size. For instance, the two dimensions may be subject to semantic confusion during encoding (cf. Banks, Clark, & Lucy,

This research was supported in part by National Institute of Mental Health Grant MH-29617. The author thanks Marcel Just for advice and assistance, and Patricia Carpenter and David Klahr for comments on earlier drafts. Requests for reprints should be sent to Peter Dixon, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

1975). Alternatively, the interference may be due to response competition similar to that found in the Stroop task (Keele, 1972). Both of these interference explanations assume that abstract number codes are used in the comparison process.

Similar kinds of incongruity effects have been found with other stimuli. For instance, when deciding which of two pictured animals is larger, reaction time is slower if the larger animal is depicted by a smaller picture (Paivio, 1975). In Paivio's study, the referential size of the animal was the criterial dimension used to make the response, while the size of the picture was an irrelevant dimension. An incongruity effect may also occur in a speeded classification task, in which stimuli are classified with respect to a given criterial dimension. Clark and Brownell (1976) presented up arrows or down arrows either high or low in the visual field. Subjects took longer to classify an arrow as up or down when its position was incongruent (e.g., an up arrow low in the field) than when it was congruent. Thus, incongruity effects of one sort or another occur in a variety of situations.

An incongruity effect also occurs when making numerical comparisons of two rows of objects. For instance, young children make more errors with incongruent stimuli (in which the numerically larger row is shorter) than with congruent stimuli (Pufall & Shaw, 1972). Presumably, this effect occurs because length cues tend to be used to make the numerical comparison. This kind of explanation of the incongruity effect is different from the dimensional interference hypothesis, in that abstract number codes are not used in the comparison process.

Adults may also show an incongruity effect when comparing two rows of objects. Although adults would probably not make as many errors as children, they may still have difficulty with the same kinds of stimuli. In that case, their reaction time performance in this task would parallel the accuracy effects found with children. Thus, an incongruity effect might be expected when adults compare two rows to determine which one contains more objects.

EXPERIMENT 1

Experiment 1 investigated the effect of relative row length on numerical comparisons. The subject's task was to choose the numerically larger of two rows of symbols (the relative judgment task). The rows varied in terms of incongruity, where incongruity is defined as the length of the numerically smaller row minus the length of the larger.¹ Thus, in a very incongruent stimulus pair, the numerically larger row would actually be shorter than the numerically smaller row. In a congruent stimulus pair, the numerically larger row would also be the longer row. It was hypothesized that the time to make the relative judgment would be longer

for incongruent stimuli, even though length was logically irrelevant to which row contained more symbols.

Method

A stimulus display consisted of two rows of symbols shown on a single horizontal line, one on the left and one on the right. One row was made up of pluses and the other was made up of stars. Length was a function of the spacing of the symbols in each row. Each row contained two, three, four, or five symbols, separated by .5, 1.0, or 1.5 deg of visual angle. Each symbol subtended about .5 deg, so that symbols were adjacent with the smallest spacing. Six numerical combinations were used: 2-3, 2-4, 2-5, 3-4, 3-5, and 4-5. Six spacing combinations were used: .5-.5, .5-1.0, .5-1.5, 1.5-1.5, 1.5-1.0, and 1.5-.5 deg. Thus the length of a row ranged from 1 deg (two symbols with .5-deg spacing) to 6.5 deg (five symbols with 1.5-deg spacing). Each spacing combination could be paired with each numerical combination in two ways. The larger of the two spacings was used with the larger number to produce a relatively congruent display (e.g., * * * * +++) and the reverse number-spacing combination was used to produce a relatively incongruent display (e.g., +++++ * * * *). Two identical displays were constructed if the spacings were equal. The amount of incongruity is defined as the length of the smaller row minus the length of the larger. The range of incongruity possible varied with the numerical combination. The most incongruent display had an incongruity of 2.5 deg (the combination 4-5 with spacings of 1.5 and .5 deg), while the least incongruent display had an incongruity of -5.5 deg (the combination 2-5 with spacings of 5 and 1.5 deg).

A total of 72 displays were determined by the six numerical combinations, six spacing combinations, and congruent or incongruent assignments of number and spacing. The left-to-right ordering of the rows and the symbols used in each (pluses or stars) were determined randomly for each subject.

The stimuli were presented on a standard video monitor at a distance of approximately 50 cm. The centers of the rows were separated by about 10 deg. The display onset and offset was controlled by a Honeywell DDP-116 computer. When the computer was ready to start a trial, the word "ready" appeared. The stimulus presentation began 500 msec after the subject pressed a ready button. The display offset occurred when the subject responded by pressing a button on the left or right to indicate the larger row. The computer recorded the response and the reaction time in milliseconds and stored them for subsequent analysis. Subjects were told to respond as quickly as possible without making errors. An experimental session consisted of 10 practice trials followed by four blocks of 72 test trials in a random order and took less than 30 min. Error trials were rerun at a random point during the block. Ten undergraduates at Carnegie-Mellon University served as subjects in partial fulfillment of course requirements. One subject was not used because of chance performance with some stimuli.

Results

The results (shown in Figure 1) demonstrated a clear effect of incongruity [$F(66,528) = 5.49, p < .001$]. The analysis was performed on the means for correct trials averaged over blocks. Outliers greater than 3 standard deviations above their means (less than 1%) were excluded and the means recalculated. The error rates for the combinations 2-3, 3-4, and 4-5 were 1.4%, 2.3%, and 3.9%, respectively. Other numerical combinations had fewer than 1% errors. The overall error rate was 1.4%.

In general, increasing the length of the numerically

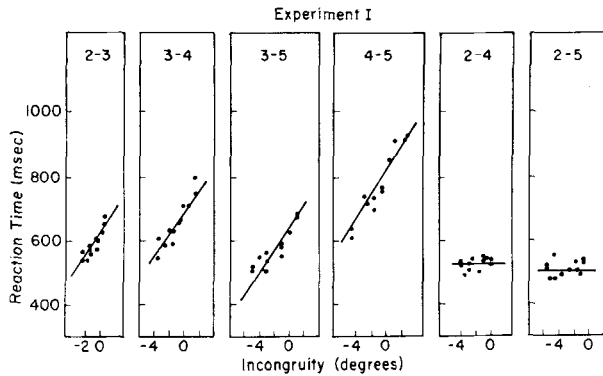


Figure 1. Mean reaction times for correct responses as a function of incongruity between length and number in Experiment 1. The best-fitting parallel lines for the numerical combinations 2-3, 3-4, 3-5, and 4-5 have slopes of 36 msec/deg of visual angle. Horizontal lines were fit to the combinations 2-4 and 2-5.

smaller row of symbols or decreasing the length of the larger increased reaction time. This result was clear for the numerical combinations of 2-3, 3-4, 3-5, and 4-5, as shown in Figure 1. However, the numerical combinations 2-4 and 2-5 did not show an incongruity effect. The other numerical combinations were parallel linear functions with a slope of 36 msec/deg of length incongruity [$F(1,528) = 289.58, p < .001$]. The residual was not significant [$F(65,528) = 1.12$].

Figure 1 also shows differences among the various numerical combinations [$F(5,40) = 65.31, p < .001$]. Generally, reaction time increased as a function of numerical magnitude. It is plausible to attribute this result to encoding processes that take longer with numerically larger rows (cf. Buckley & Gillman, 1974). Note, however, that the combination 3-5 was faster than would be expected on the basis of numerical magnitude. This discrepancy will be discussed in the context of a process model of this task. No other effects or interactions were significant.

Discussion

The results clearly indicate that an incongruity effect occurs when adults compare two rows of objects. Qualitatively similar results have been obtained with children (Pufall & Shaw, 1972); performance is more accurate with congruent stimulus pairs than with incongruent stimulus pairs. Pufall and Shaw (1972) suggest that children often use length cues to make their numerical judgments. Thus, they will tend to make more errors when the numerically larger row is shorter. However, adults do not make many errors in this task and consequently cannot be relying exclusively on length cues. Adults may be able to supplement length information with other information (such as the density of the symbols in the rows) in order to avoid errors. An explicit model of how such a strategy might work will be described in a subsequent section.

An alternative to using length and density information would be to simply count the objects in each row and compare the results. The result of counting would be an abstract number code. Since these codes would be independent of the relative lengths of the rows, the comparison of the number codes could not directly be the cause of the incongruity effect. One possible explanation for the incongruity effect hypothesizes a second, implicit comparison of length cues that interferes with the comparison of the abstract codes. The model that embodies this type of explanation may be referred to as the *interference* model.

The interference model hypothesizes that both rows are represented in terms of an abstract number code independent of length. These two codes are compared to determine which is larger. However, at the same time, the lengths of the rows are also compared. If the two comparisons have different results (as would be the case if the dimensions were incongruent), an added amount of time would be needed to resolve the conflict. Conversely, the reaction time would be reduced if the comparisons agreed. The size of these effects is assumed to vary with the relative lengths of the rows in order to obtain a continuous incongruity effect. This kind of model is sufficient to account for the incongruity effect in Experiment 1.

There are a number of difficulties with the interference model. For example, there does not seem to be any obvious reason why length should have to be compared at all. Since abstract number codes must always be compared anyway, it would seem to be a more efficient strategy for the subject to base the response on that comparison alone and not consider relative length. Thus, the interference model must also include an explanation of why length is not simply ignored in this task. In addition, the model would seem to require two separate mechanisms, one that would increase reaction time when length and number disagreed, and another that would decrease reaction time when they agreed. It does not seem very parsimonious to use two different processes to explain results that can be described with a single slope parameter. Finally, it is not clear why interference (or facilitation) should be graded; that is, why a larger difference in length should lead to more interference. In particular, it is difficult to imagine how the interference model could predict a linear relationship between reaction time and length incongruity. Thus, there are a number of details that must be considered before the interference model can provide an adequate account of these results. Although none of these problems are insurmountable, they do suggest that an alternative approach might be more parsimonious. Consequently, this paper will entertain the possibility that relative judgments can be made on the basis of length and density information, without having to use abstract number codes at all.

The numerical combinations 2-4 and 2-5 were compared quickly relative to the other combinations and were not subject to an incongruity effect. This suggests that these judgments were made by considering "low-level" features prior to the comparison process used for other combinations. For instance, they may have been compared by considering the total luminance, or energy, in each row. Since energy is approximately proportional to the number of symbols, the row with substantially more energy could be identified as numerically larger. It is perhaps surprising that such a mechanism was not used with all stimulus pairs. Conceivably, it only works if the difference between rows is sufficiently large (e.g., if one stimulus in a pair has twice as much energy as the second). A comparison of total energy may have been too inexact to be useful with other numerical combinations.

EXPERIMENT 2

There are three possible outcomes when two numerical quantities are compared: The first quantity could be greater, the second quantity could be greater, or they could be equal. Experiment 1 investigated the effect of length variation on the first two outcomes. Experiment 2 examined the effect of length variation on the judgment of equality. Subjects were asked to decide whether two rows of symbols were the same or different in number. The number of symbols in a row was the criterial dimension; the length of the row was logically irrelevant. Same-different tasks with other kinds of stimuli have shown that the presence of irrelevant differences between stimuli slows reaction time (e.g., Krueger, 1973). This effect may be continuous. That is, reaction time may increase monotonically with the amount of disparity in the irrelevant dimension (Dixon & Just, 1978). Thus, a continuous length-disparity effect might be expected in this task.

This kind of disparity effect might be explained by the same mechanism as the incongruity effect. Both can be loosely described as an increase in reaction time when the irrelevant dimension does not agree with the criterial dimension. In Experiment 1 this meant that the numerically larger row was shorter than the numerically smaller row. However, in a same-different task, this might mean that for "same" trials (i.e., those with two rows of equal number), reaction times would be longer if the lengths of the rows were dissimilar. A continuous effect would occur if reaction time increased with the amount of length disparity and, hence, how much "disagreement" there was between length and number. Thus, both the incongruity and disparity effects can be described as a continuous increase in reaction time as a function of how much the two dimensions disagree.

Method

A stimulus display consisted of two rows of symbols, as in Experiment 1. The subject's task was to decide if the display

contained an equal number of pluses and stars ("same" pairs) or a different number of pluses and stars ("different" pairs). Stimuli were analogous to those in Experiment 1. A row had either three or four symbols, each of which subtended .5 deg of visual angle. Sixty-four stimulus displays were constructed, consisting of 32 pairs of rows and their mirror images. In half of the displays, one row had a spacing of .5 deg, while in the other half, one row had a spacing of 2.0 deg. In either case, the second row had a spacing of .5, 1.0, 1.5, or 2.0 deg, for a total of eight combinations. Thus, the amount of length disparity between the two rows of a pair ranged from 0 to 4.5 deg for "same" pairs and from .5 to 5 deg for "different" pairs. For each spacing combination, there were four displays, consisting of two displays with the same number in each row (three paired with three and four paired with four), and two displays with a different number in each row (three paired with four and four paired with three).

The apparatus and procedure were generally the same as in Experiment 1. Subjects responded "same" by pressing a button on their right and "different" by pressing a button on their left. They were told to respond as fast as possible without making errors.

An experimental session consisted of 10 practice trials followed by four blocks of 64 test trials and took less than 30 min. Error trials were rerun as in Experiment 1. Subjects were 12 volunteers who were paid \$2 for their services. One subject was not used because of chance performance with some stimuli.

Results

Figure 2 shows that reaction time for "same" trials increased with the disparity between the lengths of the two rows [$F(3,30) = 15.84, p < .001$]. The slope of the best-fitting parallel lines was 25 msec/deg. This analysis was performed on the means for correct "same" responses to each number and spacing combination. A total of 1.1% of the data points were excluded as

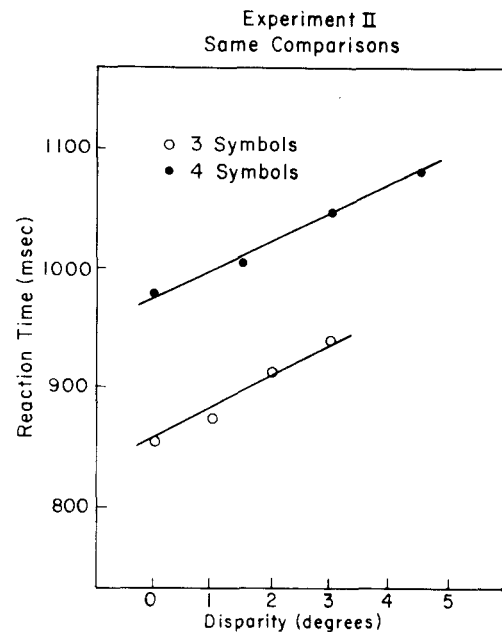


Figure 2. Mean reaction times for correct responses on "same" trials as a function of disparity in length in Experiment 2. The best-fitting parallel lines have slopes of 25 msec/deg of visual angle.

Table 1
Reaction Times (in Milliseconds) and Error Rates in Experiment 2

Symbols	Same Comparisons			Different Comparisons		
	Average Length*	Reaction Time	Percent Errors	Incongruity*	Reaction Time	Percent Errors
3	2.3	879	2.3	2.5	1,064	4.0
	3.8	910	1.4	1.5	1,055	10.2
4	3.1	1,067	7.9	1.0	1,044	5.7
	5.4	987	2.0	.5	1,064	7.9
				-.5	1,047	4.9
				-2.0	1,021	2.6
				-3.0	988	4.5
				-3.5	1,037	2.3
				-4.0	1,024	1.1
			-5.0	1,042	3.4	

*In degrees.

outliers greater than 3 standard deviations above the mean. Error rates for "same" trials are shown in Table 1.

"Same" responses for three symbols were 132 msec faster than for four symbols [$F(1,10) = 17.64, p < .005$], as shown in Figure 2. This is similar to the numerical magnitude effect found in Experiment 1 and may be attributed to encoding processes. A somewhat different effect is shown in Table 1. A given amount of disparity (e.g., 1 deg) could have been created by using two relatively short rows (e.g., 1.5 and 2.5 deg) or two relatively long rows (e.g., 3.5 and 4.5 deg). Thus, stimulus pairs can be classified as either long or short on the basis of the average lengths of the rows. There was a significant interaction between average length and the number of symbols in a "same" pair [$F(1,10) = 35.07, p < .001$]. No other effects or interactions were significant.

The analysis of "different" trials was analogous to the analysis of "same" trials; 1.3% of the reaction times were identified as outliers. Error rates for "different" trials are shown in Table 1. No effect of length disparity was found for "different" trials [$F(3,30) = 1.00$]. However, "different" trials tended to be somewhat slower if the numerically smaller row was longer than the numerically larger row. In other words, reaction times were slower if length and number were incongruent. A linear trend due to incongruity was significant [$F(1,150) = 5.58, p < .05$; see Table 1]. This effect was small and accounted for only 29% of the variance. However, neither the residual nor any other effect was significant. The overall error rate for both "same" and "different" trials was 3.9%.

Discussion

Experiment 2 demonstrated that "same" reaction times increased linearly with the amount of disparity in length. The similar slopes obtained in Experiments 1 and 2 (36 and 25 msec/deg, respectively) for an equivalent manipulation of row length suggest that a common mechanism may underlie both effects.

The results of Experiment 2 are similar to the results

found with children on tests of number conservation. The classic conservation task as developed by Piaget (e.g., 1952) consists of two judgments of quantity. First, the subject is presented with two identical rows of objects and asked to decide whether or not the two quantities are the same in number. If he answers correctly, the experimenter transforms one row, perhaps by increasing the spacing of the objects. The subject is then asked to decide again whether or not the number is the same. When the lengths of two numerically equal rows differ, children at a particular stage of development tend to respond "different." Thus, length disparity increases the probability of an error, analogous to the reaction time increase found with adults. Often children make errors in this task because they are basing their responses on a comparison of length; the longer row is chosen as numerically larger (Gelman, 1969). Adults also may be using length, but the fact that they perform accurately indicates that their responses are not based only on a simple comparison of length. The adult strategy must be more sophisticated.

One way adults could use length information without making errors would be to use an internal transformation of row length. The subject could mentally expand the shorter row and compress the longer until the lengths are equal. Such a mental transformation would be analogous to a physical transformation of length, such as that used in the conservation task. If the densities of the rows were equal after this transformation, the rows must have also been equal in number. This method may be referred to as *normalization*, since it normalizes the two rows with respect to length. Normalization overcomes the limitations of a simple length comparison by using density information as well as length information. The essential feature of this process is that the amount of transformation necessary increases with the amount of length disparity. Thus the normalization process can account for the disparity effect found for "same" trials in Experiment 2.

Similar normalization transformations have been found with other kinds of stimuli. For instance, when

comparing the shapes of two figures, subjects may normalize differences in orientation (Shepard & Metzler, 1971) or size (Bundesen & Larsen, 1975). Dixon and Just (1978) argue that these kinds of normalization operations are conceptually equivalent and have common properties. For instance, they all are incremental or continuous, they do not change the nature of the internal representation, and they all operate by reducing disparity on a dimension that is logically irrelevant to the correct response. In a same-different task, the normalization model predicts that reaction time for "same" trials should increase linearly with the amount of disparity on the irrelevant dimension. This is because the amount of transformation necessary is proportional to the amount of disparity. Since the transformation is incremental and time consuming, reaction time should increase with disparity.

The hypothesized normalization of length in Experiment 2 is entirely consistent with this more general conception of normalization processes. Since it is assumed to be a mental counterpart of a physical transformation, length normalization will progressively reduce disparity on the irrelevant length dimension without affecting quantity. It would not produce any qualitative changes in the internal representation. In addition, this strategy has the advantage that the subject does not have to compute the number of objects in a row by counting or subitizing, which may be relatively inefficient. The subject may save time by using a normalization process and then comparing the rows on the basis of length and density. A comparison process involving a similar transformation of length can also account for the results of Experiment 1.

Unlike "same" trials, "different" trials were not affected by length disparity. However, they did show a weak incongruity effect. Thus it might be proposed that "different" stimuli were processed like the stimuli in Experiment 1. In that case, a "different" response would be initiated after the subject had decided which row was larger. Although the results are suggestive, the variability of the "different" reaction times makes such hypotheses tentative.

A PROCESS MODEL

Experiments 1 and 2 have demonstrated that relative length systematically affects the numerical comparison of two rows of symbols. Different effects occur in different tasks. An incongruity effect was found in the relative judgment task, while a disparity effect was found in the same-different task. However, a single mechanism may underlie both effects. In particular, length normalization may have been used in the same-different task and a similar transformation of length in the relative judgment task. This would be consistent with the developmental literature, which indicates that children also use length information to make numerical

comparisons. Thus there is some support for the view that the relative length effects found here were caused by a process that used length information to make numerical comparisons. An explicit model of this process will be described in this section.

A flow chart of the proposed model is shown in Figure 3. Each stimulus row is encoded in terms of three quantities: the length of the row, the dispersion of the symbols in the row, and the total luminance or energy for that row. Dispersion is defined as the reciprocal of the density, that is, the length of the row divided by the number of symbols. As will be seen below, one aspect of the model depends on a property of dispersion that is not a property of density. Energy is assumed to be proportional to the number of symbols in a row. One other internal code is necessary: the average dispersion of the two rows. Average dispersion is assumed to be the mean of the dispersions for the two rows. These codes may be any monotonic function of

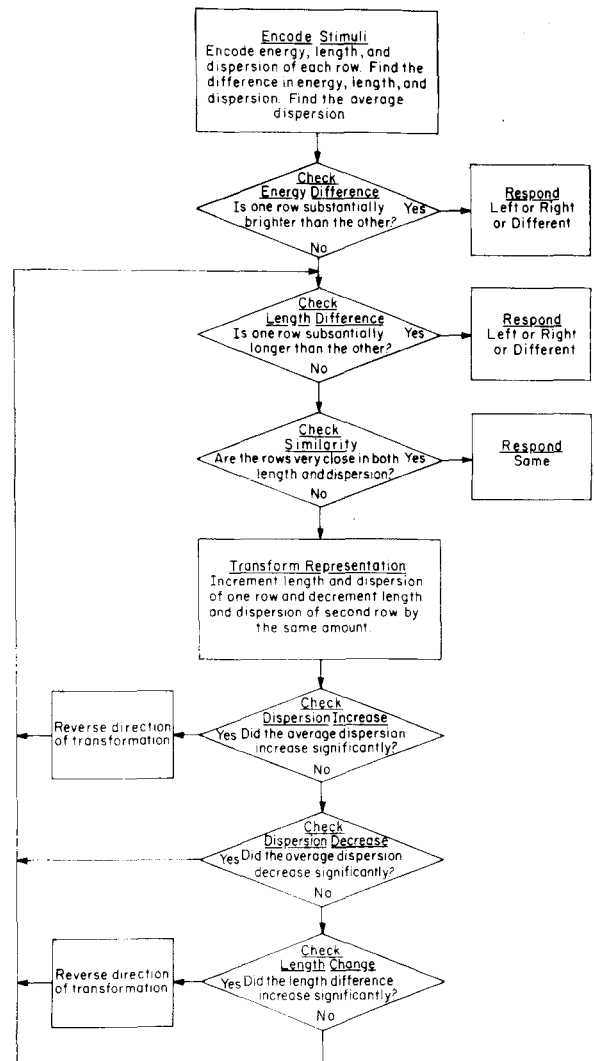


Figure 3. A process model for the numerical comparison of two rows of objects.

the actual physical values and are subject to some degree of random error. Presumably, features like energy, length, and dispersion are extracted by the visual system quickly. Since they are relatively primitive perceptual codes, they can be manipulated and compared efficiently. The proposed model can generate correct responses on the basis of these physical features alone and does not need more abstract, symbolic number codes.

As shown in Figure 3, a response can be initiated if there is a large difference in total energy. This response can indicate either which side has greater energy (in a relative judgment task), or that the two rows are different in number (in a same-different task). Energy codes are assumed to be highly variable, so the criterion for what is a sufficient difference is fairly high. The results from Experiment 1 indicate that only the numerical combinations 2-4 and 2-5 have sufficiently large differences in total energy to be detected consistently. However, the combination 3-5 may have also been compared quickly on the basis of energy some of the time. This would explain why 3-5 had faster reaction times than might be predicted on the basis of numerical magnitude. However, in that case 3-5 should have also had a smaller slope. Although there was no significant deviation from the parallel curves fit in Experiment 1, a visual inspection of Figure 1 indicates that the slope for 3-5 does tend to be somewhat smaller than the other numerical combinations.

After the initial comparison of energy, the lengths of the rows are compared. If one row is substantially longer than the other, the subject can respond either "left" or "right" (to indicate the larger row) or "different" (in a same-different task). The difference in length would be misleading if the numerically larger row was shorter, or if a "same" pair had a large length disparity. To avoid errors in these cases, the criterion for a large difference is again fairly high. Consequently, only stimulus pairs with a very small incongruity (e.g., -5 deg) can be detected at this point.

If the difference in length is not large enough to initiate a response, a test is made to detect "same" stimuli, that is, stimuli that have an equal number of symbols in each row. The test would consist of checking whether both the length difference and the dispersion difference are simultaneously small. In other words, a "same" response is initiated whenever the rows are physically similar in both length and dispersion.

Clearly, these three tests do not account for all possible situations. A method is needed for dealing with physically dissimilar "same" pairs in the same-different task and incongruent pairs in the relative judgment task. The method proposed is a transformation process, in which the internally represented row lengths and dispersions are systematically changed. These changes are such that the previously described tests of length and dispersion can lead to an appropriate response based on the altered representation.

What kind of changes would these have to be? For "same" pairs, this change would have to reduce the disparity between the two rows in length and dispersion. When they are both approximately equal, a "same" response could be made. For "different" pairs and relative judgments, the changes would have to reduce incongruity, that is, increase the length of the numerically larger row and decrease the length of the smaller. In that way, a test of the difference in length would give the correct response. It will be shown below that decreasing incongruity in this way is equivalent to decreasing average dispersion. Consequently, length and dispersion are altered in order to reduce average dispersion.

The transformation consists of incrementing the length of one row by some small quantity, q , and decrementing the length of the second row by the same amount. Corresponding changes are made in the dispersions of the rows. If q is positive, the transformation corresponds to expanding the first row and contracting the second; if q is negative, it corresponds to contracting the first row and expanding the second. Thus the sign of the transformation increment, q , indicates the direction of the transformation. Only one transformation direction will lead to the correct response. For example, expanding the longer row in a "same" pair will only make the situation worse and will not lead to a "same" response. Thus it is important that the transformation be monitored to insure that the direction is correct. If it is not, q should be changed to $-q$.

Changes in the average dispersion and the length difference are used to monitor the transformation. The first test is to verify that average dispersion has decreased. It can be shown that this is equivalent to verifying that the numerically larger row has increased in length and the smaller has decreased. Dispersion can be thought of as the length of the row divided by the number of symbols, n . Thus a length change $q > 0$ will change dispersion by q/n . The change in average dispersion will be the average of the change for the two rows, or $[(q/n) + (-q/m)]/2$, where m is the number of symbols in the second row. Since this is the same as $q(m-n)/2mn$, the change in average dispersion will be negative whenever $n > m$. Thus, increasing the length of the numerically larger row and decreasing the length of the smaller is equivalent to minimizing average dispersion.

It is possible that there will be no reliable change in the average dispersion. This could occur either because of random error, or because the stimuli are the same in number [i.e., $n = m$ and $q(m-n)/2nm = 0$]. In either case, the transformation will be assumed to be correct if the length difference decreases. Thus, if there is no reliable change in average dispersion, length disparity will be minimized. This insures that "same" pairs will be transformed correctly.

After the lengths and dispersions have been transformed and the transformation has been verified as

correct, tests are again made of the relative lengths of the rows. If still no response can be made, another incremental transformation is performed. This process continues until a test of differences in length and dispersion is sufficient to initiate a response. Thus, the central part of the model is a transformation cycle consisting of three parts: a check of the length and dispersion differences, an incremental transformation, and a transformation check. The time to complete this cycle determines the slope of the incongruity and disparity effects. This slope is the major prediction of the model.

The model as stated qualitatively predicts the results of Experiments 1 and 2. In a relative judgment task, incongruity is reduced by increments. The larger the initial incongruity, the more cycles would be needed to find the correct response. Thus, the model accounts for the main result of Experiment 1: Reaction time increases as a function of incongruity. If the stimuli are equal in number, as in Experiment 2, length disparity will be reduced, since average dispersion will not change systematically. More cycles would be needed to find a response if the stimuli have a large length disparity. According to the model, then, reaction times for "same" trials should increase as a function of length disparity. In addition, "different" trials would be treated in the same way as the stimuli in the relative judgment task. Thus, an incongruity effect is also predicted for "different" trials in a same-different task. The criteria used in the model may vary as a function of the stimulus set and task. These variations may explain some of the differences between Experiments 1 and 2 (for instance, the slightly smaller slope in Experiment 2).

The model demonstrates how length information can be used to make numerical comparisons, even though length is logically irrelevant to the task. An incremental length transformation is used in conjunction with other physical features of the stimuli (such as dispersion) in order to find the correct response. Thus there is no need to compute an abstract number code to insure accuracy. The incongruity and disparity effects are an integral part of this comparison process, not a side effect produced by interference. In its present form the model is limited to the numerical comparison of rows of objects. However, other transformations of a similar nature have been found elsewhere (Bundesen & Larsen, 1975; Dixon & Just, 1978; Shepard & Metzler, 1971). Thus, the model may be a particular instantiation of a more general transformation strategy used to compare many different kinds of stimuli.

REFERENCE NOTE

1. Hinrichs, J. V. *Physical and numerical size in number comparisons*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1976.

REFERENCES

- BANKS, W. P., CLARK, H. H., & LUCY, P. The locus of the semantic congruity effect in comparative judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **104**, 35-47.
- BUCKLEY, P. B., & GILLMAN, C. B. Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, 1974, **103**, 1131-1136.
- BUNDESEN, C., & LARSEN, A. Visual transformations of size. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **1**, 214-220.
- CLARK, H. H., & BROWNELL, H. H. Position, direction, and their perceptual integrality. *Perception & Psychophysics*, 1976, **19**, 328-334.
- DIXON, P., & JUST, M. A. Normalization of irrelevant dimensions in stimulus comparisons. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, **4**, 36-46.
- GELMAN, R. Conservation acquisition: A problem of learning to attend to relevant attributes. *Journal of Experimental Child Psychology*, 1969, **7**, 167-187.
- GELMAN, R. The nature and development of early number concepts. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 7). New York: Academic Press, 1972.
- KEELE, S. W. Attention demands of memory retrieval. *Journal of Experimental Psychology*, 1972, **93**, 245-248.
- KRUEGER, L. E. Effects of irrelevant surrounding material on speed of same-different judgment of two adjacent letters. *Journal of Experimental Psychology*, 1973, **98**, 252-259.
- PAIVIO, A. Perceptual comparisons through the mind's eye. *Memory & Cognition*, 1975, **3**, 635-647.
- PIAGET, J. *The child's conception of number*. London: Routledge & Kegan Paul, 1952.
- PUFALL, P. B., & SHAW, R. E. Precocious thought on number: The long and short of it. *Developmental Psychology*, 1972, **7**, 62-69.
- SHEPARD, R. N., KILPATRICK, D. W., & CUNNINGHAM, J. P. The internal representation of numbers. *Cognitive Psychology*, 1975, **7**, 82-138.
- SHEPARD, R. N., & METZLER, J. Mental rotation of three-dimensional objects. *Science*, 1971, **171**, 701-703.

NOTE

1. One might also define congruity as the length of the numerically larger row minus the length of the smaller. In that case, incongruity would be the same as negative congruity. For consistency, the effects in this paper will be described in terms of incongruity, even though incongruity may sometimes be less than zero.

(Received for publication January 24, 1978;
accepted May 2, 1978.)