

Thurstone's Theory of Comparative Judgment

Thurstone's theory of comparative judgment is the first to which the theory of conjoint measurement will be applied. This is fitting, if only because Thurstone's theory is now something of a classic within psychology. Few theories endure within psychology, but interest in Thurstone's theory has been fairly high since it was first proposed in a set of papers published in 1927 (see Thurstone, 1927 a, b, c, and d). Its aim was to explain pair comparison judgments in psychophysics and Luce (1977) has given a recent review of its success in that area. However, Thurstone was keen to apply the theory to the measurement of attitudes and this quickly became the major area of its application. A considerable number of attitude "measurement scales" were consequently constructed (see Shaw & Wright, 1967, for a collection). However, very little foundational research was done into its suitability in this area. Fortunately, the theory of conjoint measurement provides a means of remedying that deficiency.

As already mentioned, the aim of the theory is to explain pair comparison judgments. It may be as well then to begin this chapter with some brief comments about the method of pair comparisons and its significance for psychological measurement.

THE METHOD OF PAIR COMPARISONS

According to the method of pair comparisons stimuli are ordered pair by pair. Within each pair of stimuli the subjects are instructed to judge which is the greater in some prescribed sense. For example, the stimuli may be lights of different brightnesses and the subjects may be instructed to judge which is the brighter

within each pair. Or, the stimuli may be political candidates and the subjects may be asked to judge for each pair which candidate they prefer the most. If there are n stimuli then the complete set of pairs numbers $n \times n = n^2$. This, of course, includes those pairs made up of each stimulus paired with itself. An order on the stimuli is obtained if the judgments prove to be transitive, asymmetric, and connected (or, if the relation being judged includes identity as well as greater than, transitive, antisymmetric, and strongly connected).

There are a number of procedural issues relevant to this method of data collection that must be mentioned. In the first place it is customary not to present the full set of n^2 pairs. Obviously, if n is around ten or more the number of judgments to be made is large and extraneous motivational factors will interfere with performance. The reduction in the number of pairs presented generally proceeds as follows. First, all the pairs made up of each stimulus with itself are excluded. This reduces the number of pairs to $n(n-1)$. Second, asymmetry is assumed and so for each couple of stimuli, x and y , only one of the pairs (x, y) and (y, x) is presented. This further reduces the number of pairs to be presented to $n(n-1)/2$.

Because asymmetry is assumed true, only transitivity and connexity are open to test. However, a further procedural convention of the method of pair comparisons is to require the subject to make a judgment in relation to each pair. That is, for each stimulus pair the subject is instructed to choose one as the greater in the prescribed sense. So connexity is also assumed, leaving only transitivity to be tested. These assumptions are generally made without empirical support. If there is any doubt about their truth the procedures must be altered in order that these assumptions can be tested.

The method of pair comparisons does not often produce an order on stimuli. Even when the attribute being judged is a physical quantity, such as length or brightness, it often fails. In this case it is because humans find it difficult to perceive small differences between stimuli. Forced choices between stimuli that cannot be clearly distinguished from one another, produce violations of transitivity. This difficulty is compounded with others when the attribute being judged is not of a physical kind (e.g., preference for a political candidate). Then it often is not obvious that the attribute is even ordinal. Nevertheless, psychologists have tended to take the former case as a model for the latter.

In the former case, the subject's confusion increases as the difference between the stimuli decreases. Consider the situation where the same subject repeatedly judges between the same pair of stimuli. If the stimuli cannot be perfectly discriminated from each other then the subject will sometimes judge x greater than y and sometimes y greater than x . Assume for the moment that such repeated judgments are independent of each other. Then the proportion of times that x is judged greater than y ($p(x > y)$) should be related to the difference between x and y on the quantitative attribute being judged. As this difference decreases and x and y become more and more similar on the relevant attribute $p(x > y)$ should approach .5. As the difference between x and y increases and x and y become

more and more distinct $p(x > y)$ should move from .5 to either 1 or 0, depending on which of x or y is the greater. Hence, in the quantitative case, $p(x > y)$ may be an index of the difference between x and y on the relevant attribute.

In this chapter attention is restricted to a specific version of Thurstone's theory, called "case 5 of the law of comparative judgment." It is this version that has been most widely applied in attempts at psychological measurement. The theory is stated below in seven propositions, which are then discussed in more detail.

AN OUTLINE OF THURSTONE'S THEORY

- (1) The attribute relative to which subjects make their pair comparison judgments (henceforth called the "relevant attribute") is represented psychologically as a quantitative attribute (henceforth called the "psychological quantitative attribute").
- (2) On any occasion when a stimulus, i , is presented to a subject it gives rise to a representation, ψ_i , on the psychological quantitative attribute. (Thurstone calls the ψ_i values "discriminal processes").
- (3) Over an infinite number of independent presentations of the same stimulus, i , the distribution of ψ_i values is normal with mean μ_i and variance σ_i^2 .
- (4) If on any judgment occasion, any stimulus pair, i and j , is presented then the subject will judge i to be greater on the relevant attribute than j if and only if $\psi_i > \psi_j$ (where ψ_i and ψ_j are the discriminial processes produced by i and j on that occasion).
- (5) For any stimulus, i , the values of μ_i and σ_i^2 are constant across subjects.
- (6) For all pairs of stimuli, i and j , $\sigma_i^2 = \sigma_j^2 = \sigma^2$.
- (7) For all pairs of stimuli, i and j , the product moment correlation coefficient between ψ_i and ψ_j , ρ_{ij} , is constant.

Proposition (1) is one of the less plausible features of Thurstone's theory in this context. There is no compelling evidence that attributes like the favorability of attitude statements are quantitative and, if they are not, it does not seem likely that they would be represented psychologically in a quantitative way. But Thurstone's theory needs an underlying quantity in order to explain the $p(x > y)$ values and so it must be introduced. Of course, this need is only relative to his conviction that the processes underlying pair comparison judgments of this kind are analogous to those underlying pair comparisons when the relevant attribute is clearly quantitative.

Proposition (2) has its plausible side. If we are presented with a stimulus, like an attitude statement, and we are comparing its favorability with that of another statement then we must make an assessment of its favorability. This will be its ψ_i value. Of course, we do not seem to be aware that such ψ_i values are quantitative

but that need not count against Thurstone's theory. They may be quantitative without us being aware of the fact.

Proposition (3) says that over different occasions the ψ_i values produced by i need not all be the same. That is, for example, our assessment of how favorable an attitude statement is need not always be the same. Indeed, proposition (3) asserts that if the occasions are independent then such assessments will in all probability vary. By "independent" here is meant that the ψ_i value (or discriminial process) produced on any one occasion is not influenced by the value produced on any other occasion. In stating what happens over an infinite number of occasions proposition (3) is clearly hypothetical and it is this hypothesis that gives much of the mathematical strength to Thurstone's theory. It has the consequence that differences between judgments involving the same stimuli are a random, or chance, phenomenon. Coupled with proposition (5) this means that not only differences within the same subject but also differences between subjects are so, as well. Although this is plausible within the context of psychophysics where small differences on some perceptually obvious physical variable are involved, it is much less plausible in the purely psychological areas (like attitude measurement) to which Thurstone's theory has been applied. In that context it is at least possible that differences between subjects in their interests, values, and beliefs contribute to differences between them in pair comparison judgments on attitude statements. This is an issue that could easily be experimentally investigated.

Historically, however, this hypothesis was a virtue for Thurstone's theory because it related it to the statistical theory of normally distributed variables. This is an area of mathematics that psychologists have some familiarity with because of its use in statistical inference. So Thurstone's choice of the normal distribution, from the infinite number that could have been chosen, was largely pragmatic.

Proposition (3) means that the ψ_i value produced by a stimulus, i , on any occasion may be regarded as μ_i plus an error component, ϵ_i , that varies randomly across occasions. That is, $\psi_i = \mu_i + \epsilon_i$ (where ϵ_i is a normally distributed random variable with a mean equal to zero and a variance of σ_i^2). The aim of Thurstone's scaling procedure is to estimate the value of μ_i for each i from the pair comparison judgments.

Proposition (4) gives the principle whereby subjects make pair comparison judgments and the principle whereby the μ_i values may be estimated. It says that a pair comparison judgment between i and j is based on the order relation between ψ_i and ψ_j . Over many independent judgments $p(i > j)$ will be related to the proportion of times ψ_i is greater than ψ_j and this in turn must be related to the order relation between μ_i and μ_j , the magnitude of the difference between μ_i and μ_j and the sizes of σ_i^2 and σ_j^2 .

Propositions (1) to (4) present the core of Thurstone's theory. Propositions (5) to (7) are introduced to enable the estimation of μ_i and μ_j . Proposition (5) allows observations from different subjects to be combined and propositions (6) and (7) simplify the equations used, reducing the number of unknown values to a workable few. In some applications of Thurstone's theory all of (5), (6), and (7) are

not accepted, other propositions being substituted for them. These differences give rise to the various cases of the "law of comparative judgment" (see Thurstone, 1927a, and Torgerson, 1958).

This so-called "law" (it is really nothing more than a hypothesis) relates the probability of judging stimulus i as greater than stimulus j to the characteristics of the probability distributions associated with i and j . The form of this relationship may be understood as follows. Consider the three hypothetical cases illustrated in Fig. 5.1. Here μ_i and μ_j are separated by increasingly larger amounts: in (a) $\mu_j - \mu_i = \sigma/2$; in (b) $\mu_j - \mu_i = \sigma$; and in (c) $\mu_j - \mu_i = 2\sigma$. Now imagine taking two values, ψ_i and ψ_j , at random from the i and j distributions in each of these three figures. In each case the probability that $\psi_j > \psi_i$ exceeds the probability that $\psi_i > \psi_j$ simply because $\mu_j > \mu_i$. Additionally, moving from (a) to (c), the probability that $\psi_j > \psi_i$ increases and the probability that $\psi_i > \psi_j$ decreases because $\mu_j - \mu_i$ increases. In this way one can easily see an important feature of this version of Thurstone's theory: $\mu_j - \mu_i$ and the probability that $\psi_j > \psi_i$ are monotonically related. Thurstone's theory may be used to predict the precise form of this monotonic relationship.

On any occasion when the subject is presented with i and j , j is judged greater than i if and only if $\psi_j - \psi_i > 0$. Because ψ_i and ψ_j fluctuate randomly over occasions so must $\psi_j - \psi_i$ and its probability distribution is also normal with a mean $\mu_j - \mu_i$ and variance $\sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_i\sigma_j$. For example, if $\mu_j - \mu_i$ equals one standard deviation of the distribution of $\psi_j - \psi_i$ then the situation is as illustrated in Fig. 5.2. Then the probability that $\psi_j > \psi_i$ equals .84 and the probability that $\psi_i > \psi_j$ equals .16.

In this case, if only the probability that $\psi_j > \psi_i$ is known then the value of $\mu_j - \mu_i$ can be derived. The point that cuts the distribution into two parts, with 16% in the lower part, and 84% in the upper, is the z score (or normal deviate) corresponding to this probability. From the table of the normal curve this is found to be -1.0 . This point is the zero point in Fig. 5.2 and because its z score is -1.0 , it must be one standard deviation below the mean. That is,

$$z_o = \frac{0 - (\mu_j - \mu_i)}{(\sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_i\sigma_j)^{1/2}} = -1$$

and therefore

$$\mu_j - \mu_i = -z_o (\sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_i\sigma_j)^{1/2}$$

In more general terms,

$$\mu_j - \mu_i = z_{ji} (\sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_i\sigma_j)^{1/2}$$

(where z_{ji} is the normal deviate corresponding to the value of the probability that $\psi_j > \psi_i$). From propositions (6) and (7) of Thurstone's theory it follows that

$$\begin{aligned} \sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_i\sigma_j &= 2\sigma^2 - 2\rho\sigma^2 \\ &= 2\sigma^2 (1-\rho) \end{aligned}$$

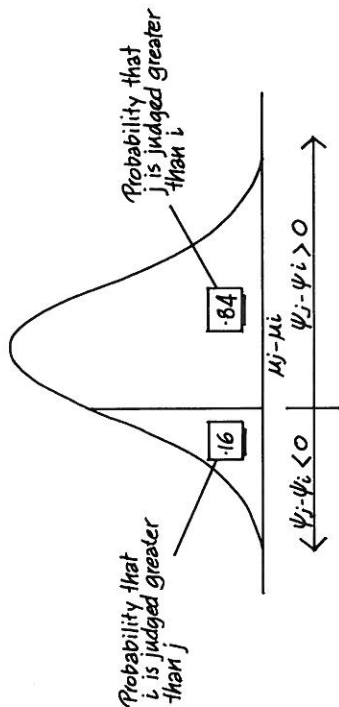


FIG. 5.2 Hypothetical distribution of $\psi_j - \psi_i$.

and so

$$(\sigma_j^2 + \sigma_i^2 - 2\rho_{ij}\sigma_j\sigma_i)^{1/2} = \sigma(2(1-\rho))^{1/2}$$

As this value is constant across all stimuli and subjects let it be called k . Then,

$$\mu_j - \mu_i = kz_{ji}.$$

By then fixing the unit of measurement at k (and because the unit is always arbitrary it may be fixed at any value) the following result is reached:

$$\mu_j - \mu_i = z_{ji}.$$

So if the probability that $\psi_j > \psi_i$ is known then $\mu_j - \mu_i$ may be deduced because z_{ji} is a direct function of this probability and may be conveniently obtained from the table of the normal curve. That this probability is never known must, of course, be obvious. However this is not a serious impediment. What may be observed is the proportion of times j is judged greater than i over a finite number of independent trials, $p(i > j)$, and this may provide a good estimate of the above mentioned probability. This is the basis of Thurstone's scaling procedure. Torgerson (1958) discussed the analytical procedures for deriving least squares estimates of the values from these sample proportions. If Thurstone's theory is true, then psychological measurement is possible. But this is a big "if."

THURSTONE'S THEORY AND CONJOINT MEASUREMENT

The central equation of Thurstone's theory, from his point of view, was that $\mu_j - \mu_i = z_{ji}$ because that enabled the values of μ_j and μ_i to be assessed. From the point of view of conjoint measurement it is more interesting to remember that z_{ji} and the probability that $\psi_j > \psi_i$ are monotonically related and, therefore, so is

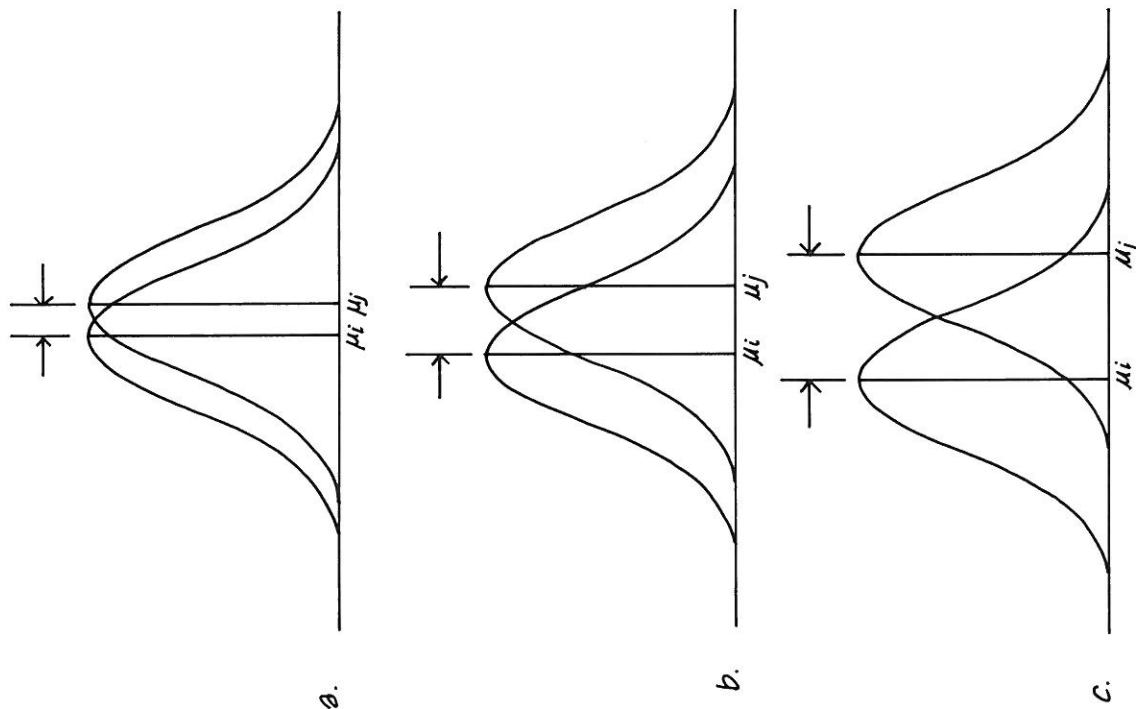


FIG. 5.1 Hypothetical distributions of ψ_i and ψ_j . When making a comparative judgment between stimuli i and j , a ψ_i and a ψ_j value are randomly sampled from these distributions. As can be seen, the probability that $\psi_j > \psi_i$ increases as $\mu_j - \mu_i$ increases.

that probability and $\mu_j - \mu_i$. It is, thus, a consequence of Thurstone's theory that for any stimuli h, i, j , and k ,

$$\mu_h - \mu_i \geq \mu_j - \mu_k \text{ if and only if } P(h > i) \geq P(j > k)$$

(where $P(h > i)$ and $P(j > k)$ are the probability that $\psi_h > \psi_i$ and the probability that $\psi_j > \psi_k$ respectively). It is in view of this monotonic relationship that Thurstone's theory is often known as a "monotone theory for pair comparisons" (see Suppes & Zinnes, 1963).

Now if, in terms of the symbolism introduced last chapter, $\mu_h, \mu_i, \mu_j, \mu_k, \dots$, are all values of A , $A = X$, and $P = A - A = A + (-A)$, then we have a basis for applying the theory of conjoint measurement to Thurstone's theory. This is because the order of the probabilities, $P(h > i), P(j > k), \dots$, is the order of $P = A - A$ and the proportions, $p(h > i), p(j > k), \dots$, are, according to Thurstone's theory, estimates of these probabilities.

It is customary to represent pair comparisons judgments in what is called a *pair comparisons matrix*. This is a table that has a row and a column for each stimulus and $p(i > j)$ is the entry in row i column j . Such a matrix is schematically illustrated in Table 5.1.

For any stimulus pair, i and j , $p(i > j)$ is set at $1 - p(j > i)$ (i.e., these two proportions are not derived from independent judgments), only the top half of the pair comparisons matrix need be inspected.

What the theory of conjoint measurement requires of all monotone theories of pair comparisons is that the probabilities (of which the proportions in the pair comparisons matrix are sample estimates) must satisfy the hierarchy of cancellation conditions. Because the prediction applies to the probabilities and the data are the corresponding proportions, it is necessary to use the methods of statistical inference to test these predictions. Using these methods a series of decisions may be made about the orders of the probabilities corresponding to the proportions. However, we are usually only dealing with the top half of the matrix and it must be 4×4 before independence can be tested, 6×6 before double cancellation can be tested and so on.

So the application of the theory of conjoint measurement to Thurstone's theory of comparative judgment provides a feasible way whereby it may be tested. In fact, as indicated, it provides not only a test of Thurstone's theory but also a test of a much wider class of theories: the class of all monotone theories of pair comparisons. As there are other such theories (e.g., Luce's, 1959, theory of choice), it is quite a useful connection to make.

It is interesting to note how monotone theories differ from one another. Thurstone's hypothesis that for any stimulus, i , the distribution of discriminial processes (ψ_i values) is normal (proposition (3) of previous section) leads to the conclusion that for any stimulus pair, i and j , the distribution of $\psi_i - \psi_j$ is also normal. Hence, the relationship between $P(i > j)$ and $\mu_i - \mu_j$ is that of the cumulative normal function, as illustrated in Fig. 5.3. This function, of course, is monotonic and it is in virtue of this property that Thurstone's theory is a

TABLE 5.1.
A Schematic Representation of a Pair Comparison Matrix

	...	h	i	j	k	...
i		$\vdots$	$\vdots$	$\vdots$	$\vdots$	
h		.	$p(i > h)$	$p(j > h)$	$p(k > h)$	...
i			.	$p(j > i)$	$p(k > i)$	...
j				.	$p(k > j)$	...
k					.	...
i						

monotone theory of pair comparisons. Any theory claiming that the relationship between $P(i > j)$ and $\mu_i - \mu_j$ increases from left to right is such a monotone theory, and it is with respect to the shape of this increase that such theories may differ from one another.

Now the possibility that conjoint measurement may be used to measure the differences, $\mu_i - \mu_j$, raises the prospect that the form of this function may be discoverable and so this facet of Thurstone's theory is testable.

The application of conjoint measurement theory to Thurstone's theory illustrates an important aspect of theory testing in experimental science. This is that the failure of the conditions for conjoint measurement (i.e., failure of some of the cancellation conditions to prove true) would not entail that the hypothesized psychological variable involved is not quantitative. Failure could come about for other reasons, for example, the falsity of propositions (5) or (6). If μ_i and σ_i^2 are not constant across subjects then proportions based on the judgments of more than one subject will not necessarily conform to the cancellation conditions. Or,

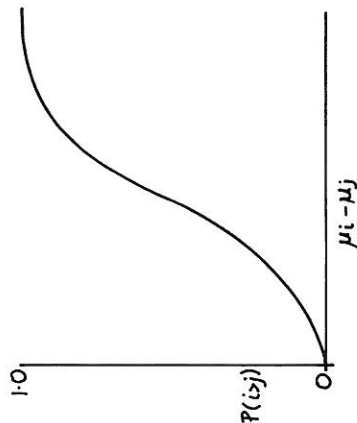


FIG. 5.3 The monotonic relationship between $P(i > j)$ and $\mu_i - \mu_j$.

if the discriminational processes have different variances then for each stimulus pair, $P(i > j)$ is no longer necessarily related to $\mu_i - \mu_j$ by the same monotonic function. Hence, $P(i > j) > P(h > k)$ and $\mu_i - \mu_j < \mu_h - \mu_k$ may both be true. So once again violations of the cancellation conditions would be possible. In complex theories like Thurstone's it is difficult to test the different propositions of the theory independently of one another. One must be prepared to settle for the kind of qualified testing permitted here. In this kind of situation, failure to satisfy the cancellation conditions means that, in the absence of more direct evidence, the hypothesis that the variable involved is quantitative is not proven and may well be false.

AN EXAMPLE OF CONJOINT MEASUREMENT

The application of the theory of conjoint measurement to Thurstone's theory may be illustrated by considering some data collected by Thurstone (1927d) himself. At the University of Chicago, 266 students made all 171 pair comparison judgments between 19 crimes, the crimes being judged with respect to their seriousness. Thurstone used this data to derive a scale for the "measurement" of the seriousness of these crimes. Our interest, however, is in whether or not his data satisfies the conjoint measurement cancellation conditions. Only a very small subset of Thurstone's data will be considered: a single 3×3 matrix involving the crimes of bootlegging, counterfeiting, libel, perjury, abortion, and kidnapping. This enables tests of single and double cancellations to be illustrated.

The relevant submatrix of Thurstone's data is shown in Table 5.2 below. The cell entries in this table are the proportion of times that the column crime was judged to be more serious than the row crime by Thurstone's subjects. The theory of conjoint measurement states that an order, $\geq$, on $P = A \times X$ satisfies single cancellation (or independence) if and only if

- (i) for any a, b in A and x in X , if $ax \geq bx$ then for every y in X , $ay \geq by$; and
- (ii) for any x, y in X and a in A , if $ax \geq ay$ then for every b in A , $bx \geq by$.

TABLE 5.2
A 3×3 Submatrix From Thurstone's (1927d) Pair Comparison Matrix Reporting Comparative Judgements about the Seriousness of Crimes

	Perjury	Abortion	Kidnapping
Counterfeiting	.532	.756	.804
Libel	.702	.809	.894
Bootlegging	.728	.872	.924

The meaning of this condition is illustrated in the two 2×2 matrices in Fig. 5.4.

In these matrices the single line arrows represent the antecedent order (the "if" part) and the double line arrows represent the consequent order (the "then" part). Matrix (i) illustrates part (i) of single cancellation and matrix (ii), part (ii). As is evident from Fig. 5.4, single cancellation implies that all rows must be ordered alike and that all columns must be ordered alike. This may be generalized to matrices larger than 2×2 . If the rows and columns of a data matrix are permuted so that the cells of the first row and the first column are ordered from least to greatest, (say, from left to right and from top to bottom) then the cells in the other rows and columns must also be ordered from least to greatest in the same way.

In testing Thurstone's theory it is the matrix of probabilities that must satisfy this condition. This hypothesis about the probabilities may be tested statistically via the proportions given in Table 5.2, for if Thurstone's theory is correct, these

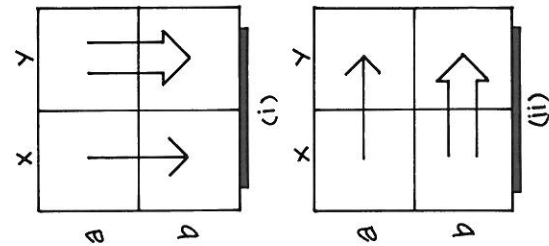


FIG. 5.4 A schematic representation of single cancellation (independence). The single line arrow is the antecedent and the double line arrow, the consequent.

proportions are estimates of those probabilities. For pairs of adjacent cells in the 3×3 matrix the hypothesis that the two probabilities were equal was tested with type one error rate set at .05. The results are represented in Fig. 5.5. Here, an arrow from one cell to another signifies that the probability in the one cell exceeds that of the other and an equals sign between cells means that the two probabilities are equal.

Obviously, these order relations violate single cancellation and, consequently, this prediction from Thurstone's theory, via the theory of conjoint measurement, is falsified. The extent of the violation of single cancellation may be assessed by counting the number of alterations to the order shown in Fig. 5.5 necessary to satisfy single cancellation. If $CA > LA$, (rather than $CA = LA$) and $LA = BA$ (rather than $LA > BA$) then single cancellation would be satisfied. Hence, Fig. 5.5 deviates from the predicted result with respect to two of the order relations.

As pointed out in the previous chapter, there are eight different versions of double cancellation that might be tested in any 3×3 matrix. However, these are not all independent of transitivity, or single cancellation, and if these two conditions are taken as true then only one version of double cancellation needs to be tested. This is the version given there in the section on "Double Cancellation." As shown in that section there are six logically independent tests of that version in any 3×3 matrix. However, again only one of these is independent of single cancellation. This test is illustrated in Fig. 5.6, where it is assumed that the rows are ordered from least to greatest, from left to right, and the columns are ordered from least to greatest, from top to bottom. Here, once again, the single line arrows represent the antecedent inequalities and the double line arrows represent the consequent inequalities. Double cancellation is violated if and only if the antecedent orders hold while the consequent ones do not, so the data patterns leading to a rejection of double cancellation are those illustrated in Fig. 5.7. Anything else confirms double cancellation. Thus, if $\geq$ on P satisfies transitivity and single cancellation, then double cancellation may be tested by searching all 3×3 matrices for one or other of the data patterns illustrated in Fig. 5.7. If either of these do not occur then double cancellation may be accepted.

However, what if single cancellation does not hold within some 3×3 matrix,

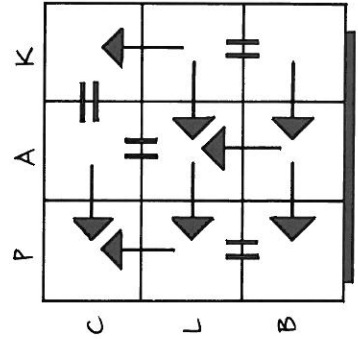


FIG. 5.5 Ordinal relations between the probabilities corresponding to the proportions in Table 5.2. An arrow between two cells means that the probability at the arrow's tail exceeds that at its head; double lines between two cells mean that the probabilities are equal. These ordinal relations are the ones considered in testing single cancellation (independence).

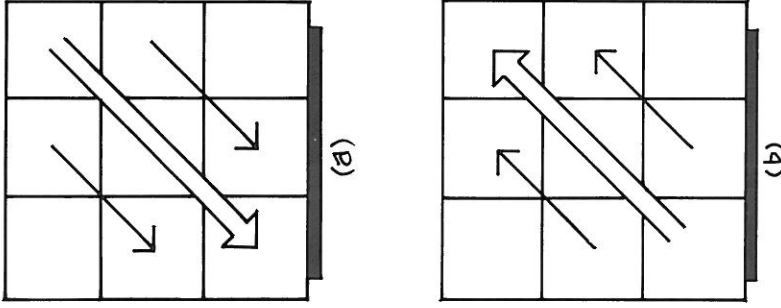


FIG. 5.6 If the rows are ordered from left to right and the columns from top to bottom then the single necessary test of double cancellation is represented in (a) and (b).

as in Fig. 5.5? One could proceed to test double cancellation in all its many versions if one was so inclined. Generally what one wants to know is how far the data under consideration departs from confirming the existence of a conjoint, noninteractive, quantitative structure. In the case of a 3×3 matrix this may be seen by considering how far the data departs from satisfaction of single cancellation, and considering over and above that how far it departs from satisfying that single test of double cancellation that is independent of single cancellation. This may be assessed by finding the minimum number of alterations to the ordering on the cells necessary to satisfy single and double cancellation (without at the same time violating transitivity). So even if independence is violated, one only need consider that single test of double cancellation illustrated in Fig. 5.6.

The details of this test are as follows. Separating the relation, $\geq$, into its components, $>$ and $=$, the relation between any pair of cells may be $>$, $=$ or $<$. Considering the three pairs of cells joined by arrows in Fig. 5.6 the problem is to identify those patterns or orders that satisfy double cancellation and those that do not. Labelling the rows and columns of the 3×3 matrix as a , b , c and x , y and z (as before), and calling the order relation between bx and ay "1", that between cy and bz "2", and that between cx and az "3" (as in Fig. 5.8) the logically possible patterns may be indicated as in Table 5.3.

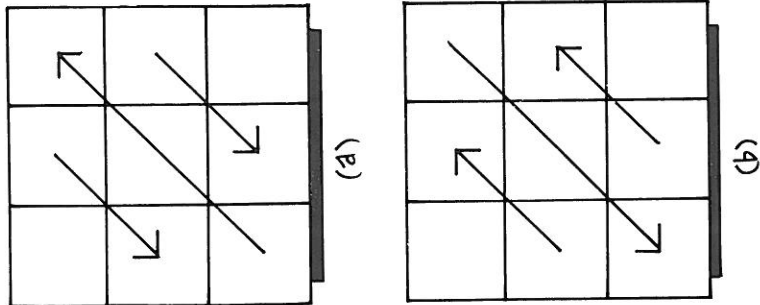


FIG. 5.7 In the test represented in Fig. 5.6, double cancellation is violated if and only if either of the data patterns (a) or (b) occurs.

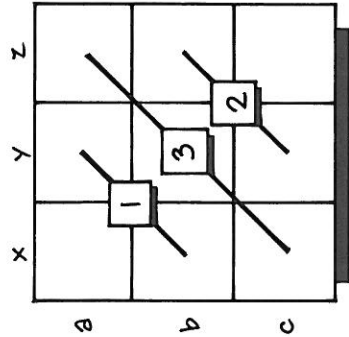


FIG. 5.8 The three order relations to be considered in the test of double cancellation represented in Fig. 5.6. Each relation may be either $<$, $=$ or $>$.

As there are three order relations to consider (1, 2, and 3) and each of these may be $>$, $=$ or $<$, the number of possible patterns is $3^3 = 27$. These are all shown in Table 5.3: 13 patterns satisfy double cancellation and 14 violate it. Despite the claims to the contrary made by Krantz & Tversky (1971), Levelt, et al. (1972), Perline et al. (1979) and van der Ven (1980), none of these 27 possible patterns are indifferent to double cancellation (see also Michell, 1988).

The pattern obtained from the proportions in Table 5.2 after the necessary

TABLE 5.3
Each Relation Identified in Fig. 5.3 is Assigned a Row of This Table (Column R). Each Such Relation May Be Either $>$, $=$, or $<$. All 27 Logical Possibilities are Shown Here: 13 Satisfy Double Cancellation and 14 Violate It. (Adapted from Michell (1988, Table 1). Adapted by permission of Academic Press.)

R	1	2	3
PATTERNS SATISFYING DOUBLE CANCELLATION	$<$	$<$	$<$
1	$<$	$<$	$<$
2	$<$	$<$	$<$
3	$<$	$<$	$<$
4	$<$	$<$	$<$
5	$<$	$<$	$<$
6	$<$	$<$	$<$
7	$<$	$<$	$<$
8	$<$	$<$	$<$
9	$<$	$<$	$<$
10	$<$	$<$	$<$
11	$<$	$<$	$<$
12	$<$	$<$	$<$
13	$<$	$<$	$<$
PATTERNS VIOLATING DOUBLE CANCELLATION	$<$	$<$	$<$
14	$<$	$<$	$<$
15	$<$	$<$	$<$
16	$<$	$<$	$<$
17	$<$	$<$	$<$
18	$<$	$<$	$<$
19	$<$	$<$	$<$
20	$<$	$<$	$<$
21	$<$	$<$	$<$
22	$<$	$<$	$<$
23	$<$	$<$	$<$
24	$<$	$<$	$<$
25	$<$	$<$	$<$
26	$<$	$<$	$<$
27	$<$	$<$	$<$

statistical tests is shown in Fig. 5.9 and that is pattern 2 in Table 5.3. This satisfies double cancellation.

If the orders in Figs. 5.5 and 5.9 are then combined, the result is shown in Fig. 5.10. This contains the violations of single cancellation mentioned above together with a violation of transitivity: $BA = LK$ and $LK = BK$, by transitivity implies that $BA = BK$, however, in fact, $BA > BK$. The minimum number of alterations necessary for the order on the cells to satisfy both single and double cancellation (together with transitivity) is three (Figs. 5.11a and b). This illustrates how single and double cancellation may be tested in a 3×3 matrix. In this instance single cancellation and transitivity (a special case of double cancellation) were violated, suggesting that either seriousness of crimes is not a quantitative variable or else some other part of Thurstone's theory of comparative judgment is false.

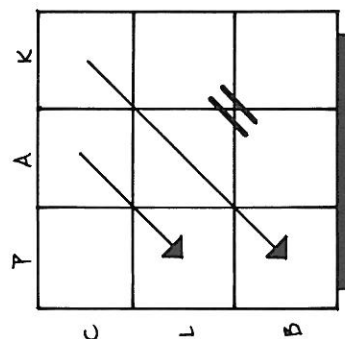


FIG. 5.9 Ordinal relations between the probabilities corresponding to the proportions in Table 5.2.

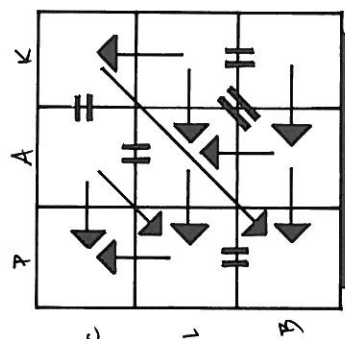


FIG. 5.10 The ordinal relations represented in Figs. 5.5 and 5.9 combined.

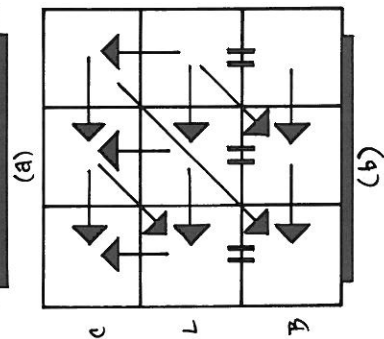
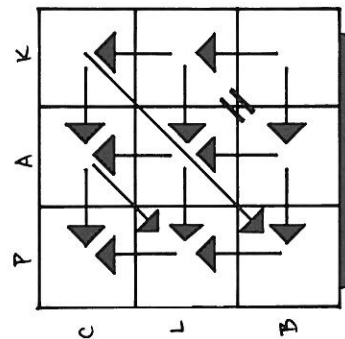


FIG. 5.11 Two sets of ordinal relations in a 3×3 matrix which satisfy both single and double cancellation. The relations represented in Fig. 5.10 approximate these.