# On interpretation of interactions

GEOFFREY R. LOFTUS
*University of Washington, Seattle, Washington 98195*

The principle focus of this paper is on interpretation of interactions that are obtained when response probability is used as a dependent variable. It is argued that results obtained with probability (or any dependent variable) are only interesting insofar ˜s they reflect something about a corresponding theoretical component. It follows that the functional mapping of response probability (which is measured) onto the state of a theoretical component (which is inferred) must be somehow specified if conclusions are to be meaningful. Depending on the nature of such a mapping, various types of results, particularly results involving interactions, may or may not be interpretable.

The purpose of this paper is to provide a new look at an old caveat: that certain types of interactions make sense only if a particular scale is assumed. This point has been made before in the measurement literature (e.g., Krantz & Tversky, 1971), in the developmental psychology literature (e.g., Bogartz, 1976), as well as in design textbooks (e.g., Winer, 1971). It occasionally makes furtive appearances in the memory literature in the guise of cautions about "floor" and "ceiling" effects. Unhappily, however, it is a point that often continues to be ignored.

The present attack on this issue rests on the assumption that the primary purpose of experimentation in psychological research is to test theories, models, hypotheses, and speculations (hereafter collectively designated "theories"). Given this assumption, three topics will be discussed. First, I present a representation of the relationship between theory and data in memory research. This representation is not particularly profound, but it provides a convenient framework within which subsequent arguments will be couched. Second, I consider a problem in data interpretation that frequently occurs in experimental reports. This problem involves the issue of what type of conclusions may be drawn from the results of an experiment in which response probability has been used as a dependent variable. Finally, some modest beginnings of a solution to this problem will be offered.

## THEORY AND EXPERIMENTATION IN MEMORY RESEARCH

Figure 1 is a representation of some of the relation-

ships between theory and data in memory research. The center section of Figure 1 lists some of the major components of interest in the theory of memory. This list is not, of course, exhaustive, but is meant to represent theory at a fairly general level.

To test a theory or some component of a theory, one does experiments. An experiment consists of manipulating some independent variable (top panel of Figure 1) which is assumed to cause variation in the state of one or more of the components of the memory

INDEPENDENT VARIABLES

Manipulation of independent variables is assumed to cause variation in one or more components of a theory of memory.

THEORY OF MEMORY

A. Characterization of input information.
B. Characterization of stored information.
C. Identification of processes that operate on stored and input information.
D. Identification of how long these processes take.
E. How the results of A – D map into responses that a subject is permitted to make in an experiment.

DEPENDENT VARIABLES

Assumed to measure variations in one or more components of a theory
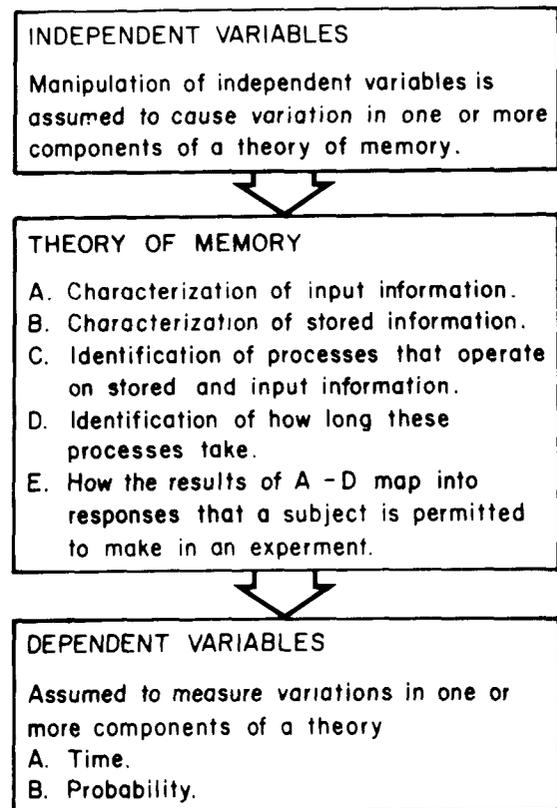A. Time.
B. Probability.

Figure 1. A general view of the relationship between theory and experimentation in memory research.

theory (hereafter, theoretical components). Variation in the state of a theoretical component is, in turn, assumed to be reflected by an associated variation in some dependent variable that is being measured in the experiment.

Memory research has produced a proliferation both of theoretical components and of the independent variables that are assumed to cause variations in the states of these components. In contrast, however, memory research is woefully impoverished in terms of the dependent variables that are assumed to measure such variations. The vast majority of experiments designed to investigate memory use but two dependent variables. time and probability.[1] Wickelgren (Note 1) has enumerated reasons for exercising caution when interpreting data based on reaction time. The arguments to be sketched in this paper will similarly identify precautions that should be observed when interpreting data based on response probability.

Time (e.g., reaction time), when used as a dependent variable, is generally assumed to be a direct (i.e., linear) reflection of time as a theoretical component[2] (e.g., Donders, 1969; Sternberg, 1969). When response probability is used as a dependent variable, however, the situation is somewhat different, in the sense that one cannot generally assume a linear function mapping the state of some theoretical component onto response probability. Rather, this function may be linear, but it may also be negatively accelerated, ogival, or something else, depending on the theory and the specific theoretical component under consideration. Often, the function is not well specified but is merely assumed to be monotonic. To the extent that the nature of this function is not considered, certain types of data may be overinterpreted.

## CONCLUSIONS MADE FROM EXPERIMENTS USING RESPONSE PROBABILITY

The fact that the function relating some theoretical component to response probability is (1) often not well specified and (2) not necessarily linear means that there are limits on the sorts of conclusions that can be based on response probability data. A hypothetical example will serve to illustrate what is meant by this.

Suppose an experiment is done on forgetting in the Brown-Peterson paradigm.[3] Two retention intervals are used: 5 sec and 20 sec. Additionally, a second, two-level, independent variable is factorially combined with retention interval. This variable could be any one of a number of things (e.g., stimulus meaningfulness, number of initial repetitions, type of instructions, etc.). I will refer to the second variable as A and to the two levels as $A_1$ and $A_2$.

Figure 2 (top panel) shows a possible outcome of this experiment. (For ease of subsequent discourse, the four data points have been labeled a, b, c, and d.)
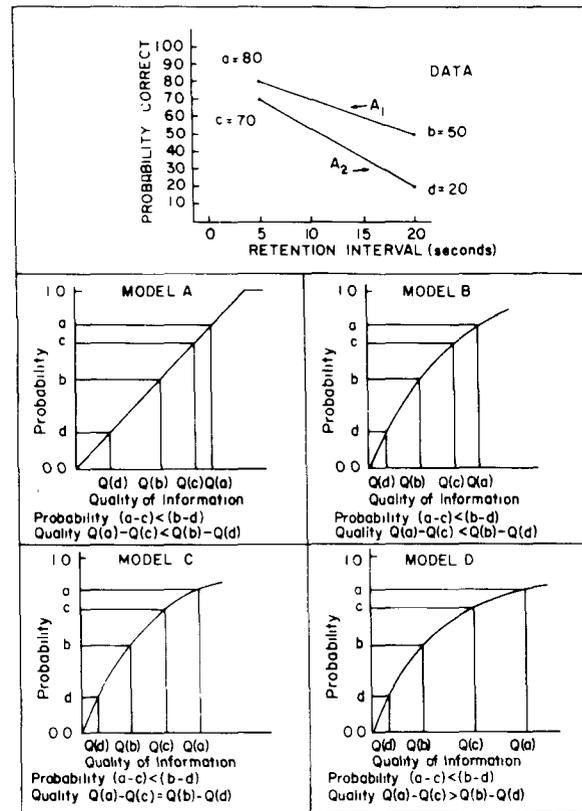


Figure 2. A hypothetical outcome in the Brown-Peterson paradigm. Of four possible models relating "quality of information" to response probability, two preserve the original interaction, whereas the other two do not.

Assume that statistical analysis has revealed all simple main effects and the interaction to be significant. What sorts of conclusions might be made on the basis of these results? Based on the significant main effects, a research report would probably claim something such as, "Condition $A_1$ leads to better overall memory performance than does Condition $A_2$ and overall memory performance decreases over retention interval." The report might then go on to claim, based on the significant interaction, that "forgetting is faster in Condition $A_2$ than in Condition $A_1$." It is the third conclusion that raises problems.

### Model-Dependent Statements

All three conclusions imply some model of memory and of forgetting. At a very general level, the following assumptions must underlie the conclusions:

(1) When a correct response is made, the subject must have been basing that response on information "stored" at some time. This information must be information "about" the to-be-remembered stimulus. (What is meant by "stored" and "about" is left vague at this general level and can only be defined in the context of some more specific model.)

(2) The greater the "quality" of this information,

the more often the subject is able to respond correctly. (Again, the precise meaning of "quality" can only be defined within the context of a more specific model.)

Within the framework of this very general model, the three conclusions based on the hypothetical data now translate into the following:

Conclusion 1. The quality of information is greater in Condition $A_1$ than in Condition $A_2$.

Conclusion 2. The quality of information declines over retention intervals.

Conclusion 3. The decline of quality is faster in Condition $A_2$ than in Condition $A_1$.

Based on the data and the general model, Conclusions 1 and 2 follow unequivocally. However, Conclusion 3 does not follow unequivocally. Whether Conclusion 3 is true or false depends entirely on how "quality" is defined and how quality is functionally related to response probability. In short, Conclusion 3 is a conclusion that can be made only within the context of a more specific model than the one described above.

**Four Definitions of "Quality"**

To illustrate why Conclusion 3 is model dependent, the bottom panels of Figure 2 depict four different models of "quality," where a model is characterized here simply in terms of the function that maps quality onto response probability.

Model A depicts a linear function. Here, Conclusion 3 is true in the sense that the independent variables interact in terms of quality in the same way that they interact in terms of response probability. More formally stated, the four data points (probability values) a, b, c, and d map into the four quality values Q(a), Q(b), Q(c), and Q(d), respectively. In terms of probability, the interaction may be represented by the fact that $a - c < b - d$. The equivalence of the interaction in terms of quality is then reflected by the fact that $Q(a) - Q(c) < Q(b) - Q(d)$. Models B, C, and D depict negatively accelerated functions relating quality to response probability. In Model B, the claimed interaction still exists in terms of quality, since it is still the case that $Q(a) - Q(c) < Q(b) - Q(d)$. Therefore, within the contexts of Models A and B, Conclusion 3 above is still valid: Forgetting is faster in Condition $A_2$ than in Condition $A_1$.

However, the situation changes rather dramatically when Models C and D are considered. In Model C, the interaction has been transformed away; that is, in terms of quality, $Q(a) - Q(c) = Q(b) - Q(d)$. And finally, in Model D, the interaction in terms of quality is reversed relative to what it was in terms of response probability; that is, with Model D, $Q(a) - Q(c) > Q(b) - Q(d)$.

Hence, working within the context of Model C would yield the conclusion that the rate of forgetting is equal in Conditions $A_1$ and $A_2$. With quality as defined by Model D, the original conclusion would reverse: It would be that the rate of forgetting is slower in Condition $A_2$ than in Condition $A_1$.

The point of this example is to illustrate that, when a negatively accelerated function maps some theoretical component—in this case, quality—onto response probability, the sort of interaction depicted in the top panel of Figure 1 is uninterpretable. That is to say, one cannot tell whether the interaction will be the same, will be transformed away, or will reverse itself in terms of the theoretical component. Which of these three outcomes will obtain depends entirely on the exact quantitative form of the mapping function.

**CLASSES OF THEORETICAL COMPONENTS**

The remainder of this paper will generalize the above illustration to other kinds of mapping functions and other sorts of interactions. The strategy will be to enumerate classes of theoretical components, where a "class" is defined in terms of the general type of function that maps the state of the component onto response probability. Having characterized such classes, it is then a fairly straightforward exercise to determine what types of conclusions about a particular class of components can be made from various sorts of data. Specifically, five classes of theoretical components will be defined, corresponding to functions that are monotonic, ogival, negatively accelerated, inverse, and linear. For each class, I will (1) give an example of a member of that class drawn from the memory literature and (2) specify what types of data based on response probability, particularly data involving interactions, are interpretable in terms of the theoretical component of interest. An interaction (or lack of interaction) in terms of response probability is defined here to be interpretable if the ordering of the response probability data and the form of the mapping function under consideration unambiguously specify the form of the interaction that will exist in terms of the theoretical component.[4] Formal mathematical proofs will not be provided, as this paper is a short tutorial exercise. The reader is invited to convince himself or herself of the validity of the assertions to be made about the interpretability of various kinds of interactions by making up examples similar to those depicted in Figure 1.

Examples will be given of data patterns that form interpretable vs. uninterpretable interactions for the various classes of components. These examples will be restricted to designs involving only two levels of each independent variable. This is because 2 by 2 designs are simple, and sufficient for illustrating the major arguments to be made. Designs with more than two levels per independent variable will be discussed briefly in a later section.

**MONOTONIC FUNCTIONS**

In the vast majority of theories about memory, the function relating the state of some theoretical
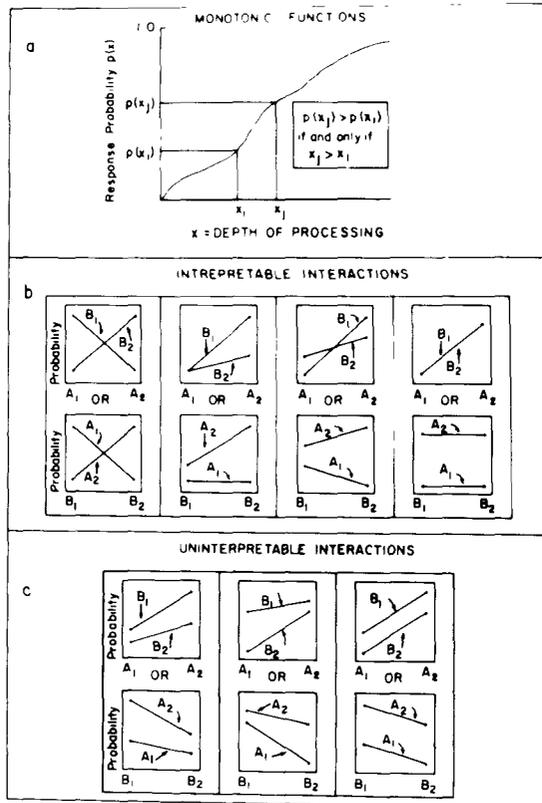
**Figure 3.** Interpretable and uninterpretable interactions when monotonicity is the only assumption made about the function mapping a theoretical component into response probability.

component to response probability is not quantitatively specified; rather, only the weak assumption is made (usually implicitly) that the function is monotonic.[5] An example of such a situation comes from the "depth-of-processing" literature (e.g., Craik & Lockhart, 1972). As illustrated in Figure 3a, this framework assumes that the greater the depth to which some item is processed, the higher will be the probability of correctly responding to that item.

### Conclusions Based on Main Effects

For a function $p(x)$ to be monotonic, the following must be true: Given two states $x_i$ and $x_j$ of a theoretical component, and their corresponding probabilities, $p(x_i)$ and $p(x_j)$, $x_i > x_j$ if and only if $p(x_i) > p(x_j)$.

Concluding that there exists a main effect in terms of probability would be equivalent to the statement that $p(x_i) > p(x_j)$. Therefore, whenever a monotonic relationship exists between $x$ and $p(x)$, a main effect in terms of probability would imply a main effect in terms of depth of processing, since probability is assumed to be monotonically related to depth of processing.

The rest of the classes of components discussed in this paper all involve monotonic functions.[7,8] It will

therefore be assumed that main effects obtained with probability are always interpretable.

### Conclusions Based on Interactions

The bottom panels of Figure 3 illustrate what sorts of interactions are interpretable and what sorts of interactions are not interpretable when monotonicity is the only assumption made about the mapping function. Note that the data patterns are arranged in seven pairs: The two members of each pair represent the same outcome but with the abscissa and curve parameters reversed (i.e., the top member of each pair plots Factor A on the abscissa with Factor B as the curve parameter, whereas the reverse is true for the bottom member of each pair).

Any interaction that is a crossover interaction is interpretable in the context of a monotonic function (Figure 3b). A crossover interaction (also referred to as ordinal nonindependence), of course, exists when the ordering of the data points corresponding to the levels of one independent variable depends on the level of the other independent variable (cf. Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971). Graphically, a crossover interaction exists when the two curves touch each other when either of the two independent variables is plotted on the abscissa. Lack of an interaction is uninterpretable when both independent variables show a main effect; however, when one or both of the independent variables fail to show a main effect, this lack of interaction is uninterpretable. Any interaction that is not a crossover interaction is not interpretable (Figure 3c).

## POTENTIALLY QUANTIFIABLE FUNCTIONS

The remainder of the classes of theoretical components are assumed to have the capability of being quantitatively related to response probability. Each such class breaks into two subclasses. In the first subclass, the exact function and its parameters are assumed to be specified; that is, given any probability, $p(x)$, the numerical value of the corresponding $x$ may be determined. For this subclass, the arguments made in this paper are superfluous, because the probability that is initially measured in an experiment can be transformed into, and analyses can be performed directly on the state of the component of interest. This is done, for example, when response (hit and false alarm) probabilities measured in a recognition memory experiment are transformed into d' scores.

The second subclass of components involves instances in which the general form of the function is known, or intuited (e.g., the function is negatively accelerated), but for one reason or another, the exact form and/or the parameters of the function are not determined (e.g., it is not known whether the function is logarithmic or exponential). When the theoretical component
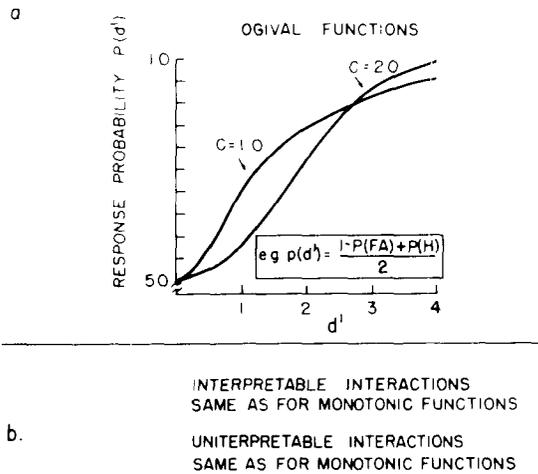
a.

OGIVAL FUNCTIONS



**Figure 4.** Ogival functions are no better than monotonic functions in terms of interpreting interactions.

being tested falls into the second subclass, response probabilities are the only data generally presented in a research report, and the arguments presented in this paper do apply.

## Ogival Functions

Many variables in the world are assumed to be normally distributed. With an underlying normal distribution, a theoretical component of interest is often measured by a z score or by something closely akin to a z score. The measure $d'$ from the theory of signal detection is an example of such a theoretical component. Figure 4a, which assumes a yes-no recognition memory experiment and the standard assumptions of signal-detection theory, depicts the function mapping $d'$ into probability correct for two criterion values.

In terms of interpreting interactions, one is in no better shape with an ogival function than with a function that is merely assumed to be monotonic. Parts of the ogival function are negatively accelerated, whereas other parts are positively accelerated. This again means that only crossover interactions are interpretable.

## Negatively Accelerated Functions

A second class of functions that map some theoretical component into response probability is the class of negatively accelerated functions. For example, as shown in Figure 5a, Atkinson and Shiffrin (1968) assume long-term memory strength of an item to be exponentially related to the number of trials that the item spent in the rehearsal buffer. When such a function obtains, most types of interactions are interpretable, as depicted in Figure 5b. In addition to the crossover interactions, any interaction involving curves that diverge toward higher values of response probability or any lack of interaction must similarly result in a diverging

interaction when transformed to the theoretical component. The only type that is not interpretable is shown in Figure 5c. Such an interaction occurs when the curves converge toward higher levels of response probability. An example demonstrating why this type of interaction is uninterpretable within the context of a negatively accelerated mapping function has been presented earlier in this paper.

## Inverse Functions

When the time to perform some operation is the theoretical component of interest, an inverse function of one sort or another often maps time into response probability. That is, the longer the time it takes to carry out some process, the lower is the probability of successfully accomplishing whatever the task at hand. An example of such a situation is drawn from an experiment reported by Sperling, Budiansky, Spivak, and Johnson (1971). In this experiment, a series of letter arrays was presented to a subject in rapid succession. Embedded somewhere in one of the arrays was a digit, and the dependent variable in the experiment was the probability of correctly reporting
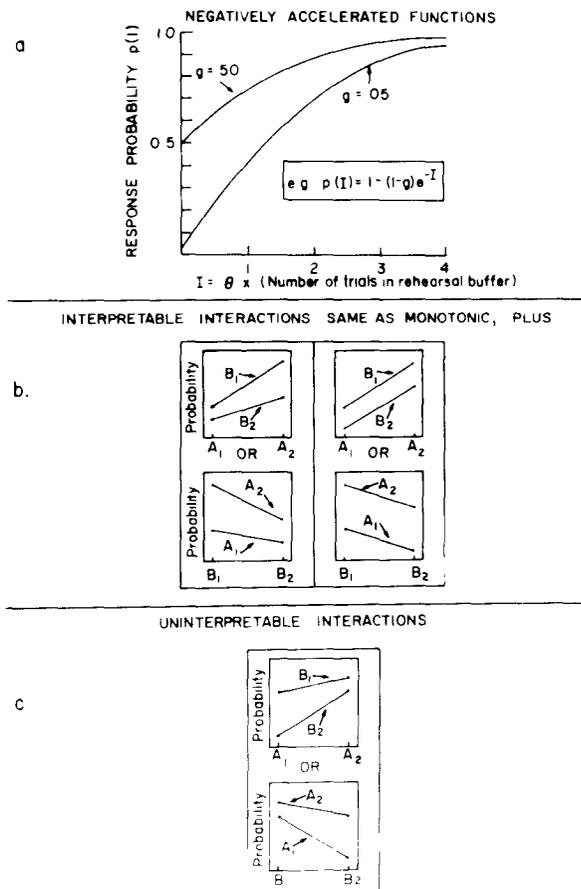


**Figure 5.** Interpretable and uninterpretable interactions when a negatively accelerated function is assumed to map some theoretical component into response probability.

INVERSE FUNCTIONS

$p(t) = \frac{g(1-g) + gn\,t}{nt}$

e.g.

g = 05
l = 100
n = 25

g = 05
l = 50
n = 10

t = Time (m sec) to scan one array location

INTERPRETABLE INTERACTIONS SAME AS MONOTONIC PLUS
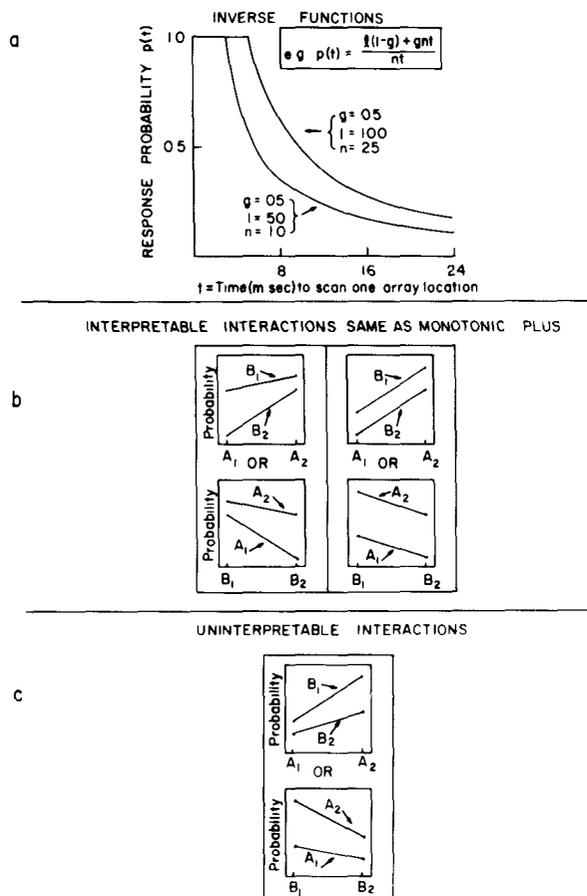
UNINTERPRETABLE INTERACTIONS

Figure 6. Interpretable and uninterpretable interactions when an inverse function is assumed to map some theoretical component into response probability.

the digit's location. Sperling et al. assumed that if the subject had scanned the location of the digit, he would respond correctly; otherwise, he would guess and be correct with some guessing probability g.

Designating 1 as the length of time each array is shown and n as the number of symbols per array, the function relating response probability to the time to scan a location is shown in Figure 6a. Because it is positively decelerated, this sort of function represents, in a sense, the converse of the negatively accelerated

function shown in Figure 5a. Any interaction involving curves that converge toward higher levels of probability or any lack of interaction must result in a similar converging interaction in terms of the theoretical component. The only uninterpretable interactions are ones involving curves that diverge toward higher levels of probability.[7]

**Linear Functions**

The final class of functions to be considered in this paper is the class of functions that linearly relate some theoretical component of interest to response probability. An example of such a situation may be drawn from the Sperling et al. (1971) experiment described in the previous section. Suppose that in this paradigm, the theoretical component of interest is the number of locations scanned per array (as opposed to the time to scan a single location). The function relating number of scanned locations to response probability is linear, as shown in Figure 7a. For this class of theoretical components, all interactions in terms of response probability are interpretable.[7]

## INTERACTIONS IN DESIGNS INVOLVING MORE THAN TWO LEVELS OF AN INDEPENDENT VARIABLE

As noted above, 2 by 2 designs have been used as examples in this paper because they are the simplest for illustrating the paper's basic point. However, the same general arguments made above obtain for designs involving more than two levels of each independent variable. Suppose, for example, that the design in the top panel of Figure 1 had included two other values, 10 and 15 sec, of the retention interval factor, and suppose that the same diverging curves had emerged. The same conclusion, that forgetting is faster for Condition $A_2$ than for Condition $A_1$, might easily be made. And this conclusion would be an invalid one by the same reasoning that applied to the 2 by 2 design.

**Conjoint Measurement**

The general topic of nonlinear scale transformations is a crucial one in the conjoint measurement literature

Table 1
Examples of Hypothetical Data Stemming from 3 by 3 Designs

| | Example 1 | | | | | Example 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Factor 1 | | | | | Factor 2 | | | |
| Factor 2 | a | b | c | | Factor 2 | a | b | c | |
| r | ra .50 | rb .70 | rc .90 | rb = .70 > qc = .65 qa = .45 > pb = .35 | r | ra .50 | rb .70 | rc .90 | rb = .70 > qc = .65 qa = .45 > pb = .35 |
| q | qa .45 | qb .50 | qc .65 | and ra = .50 > pc = .40 | q | qa .45 | qb .50 | qc .65 | but ra = .50 < pc = .55 |
| p | pa .10 | pb .35 | pc .40 | Double cancellation is satisfied. | p | pa .10 | pb .35 | pc .55 | Double cancellation is not satisfied. |

Note—In both examples, ordinal independence is satisfied. However, double cancellation is satisfied only in Example 1.

(e.g., Krantz et al., 1971; Krantz & Tversky, 1971). One of the major goals of the conjoint measurement technique is to enumerate necessary and sufficient conditions for the existence of transformations that will remove interactions. Consider a set of data arising from a multifactor design. The question may be raised, do there exist monotonic transformations on the independent and dependent variables that will remove any existing interaction in the data? For any 2 by n design (n ⩾ 2), lack of ordinal independence in the data is, as noted in the above arguments, sufficient to reject the existence of such transformations. However, for designs involving three or more levels of both independent variables, a second characteristic of the data is also sufficient to reject the existence of such transformations. This characteristic involves any failure of what is referred to as double cancellation. A thorough discussion of double cancellation is beyond the scope of this paper; however, for completeness, the description provided by Krantz and Tversky (1971) is briefly recapitulated below. Krantz and Tversky as well as Krantz et al. (1971) provide excellent and much more complete coverage on the topic.

Consider the 3 by 3 design depicted in Table 1. The notation is similar to that used by Krantz and Tversky (1971). The three levels of Factor 1 are labeled a, b, and c; and similarly, the three levels of Factor 2 are labeled p, q, and r. Accordingly, the nine cells are labeled ap, aq, and so on.

Now consider the relationships of Cell rb to qc and of Cell qa to pb. Satisfaction of double cancellation may be stated as follows: If

$$rb \geqslant qc \qquad (1)$$

and

$$qa \geqslant pb, \qquad (2)$$

then

$$ra \geqslant pc.$$

The term "double cancellation" stems from the fact that the bs on the left-hand side of Inequation 1 and on the right-hand side of Inequation 2 may be "cancelled," as may the qs on the right-hand side of Inequation 1 and on the left-hand side of Inequation 2. Such cancelling leaves the resulting ra on the left side and pc on the right side of the inequation.

Finally, consider the difference between the two examples depicted in Table 1. In both of these examples, ordinal independence is satisfied (i.e., the ordering of the dependent variable is the same for each row and is the same for each column). In Example 1, double cancellation is also satisfied. Thus, the interaction existing in these data could be transformed away.
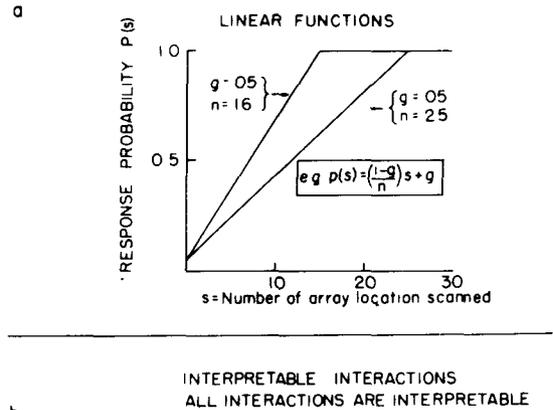


Figure 7. All interactions are interpretable when a linear function is assumed to map some theoretical component into response probability.

However, double cancellation fails in Example 2. This means that the interaction is a "true" one: No monotonic transformation exists that will remove it.

It should be noted that with designs involving more than three levels of either of the independent variables, there is more than one test of double cancellation. In general, in an M by N design (M and N both ⩾ 3) there are $\binom{M}{3} \times \binom{N}{3}$ tests of double cancellation. If any of these tests fails, then the existing interaction may not be transformed away.

## CONCLUSIONS

The arguments made in this paper are not, of course, new. Psychologists are generally familiar with the fact that some interactions can be made to appear and disappear at will by performing a nonlinear transformation on the data. The crucial point, however, is that a dependent variable such as response probability is not intrinsically interesting. Rather, it is only interesting because of what it reflects about a component of some theory. Getting from the dependent variable back to the theoretical component may well involve a transformation on the dependent variable. If this transformation is nonlinear, then one is limited in fairly specifiable ways when interpreting data involving the original dependent variable.

### REFERENCE NOTE

1. Wickelgren, W. Speed-accuracy tradeoff and information processing dynamics. Paper presented at the annual meeting of the Psychonomic Society, Boston, November 1974.

### REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation (Vol. II). New York: Academic Press, 1968.

BOGARTZ, R. S. On the meaning of statistical interactions. *Journal of Experimental Child Psychology*, 1976, **22**, 178-183.

CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 761-784.

DONDERS, F. C. [On the speed of mental processes.] In W. G. Koster (Ed. and trans.), *Attention and performance II.* Amsterdam: North-Holland, 1969. (Reprinted from *Acta Psychologica*, 1969, **30**.)

KRANTZ, D. H., LUCE, R. D., SUPPES, P., & TVERSKY, A. *Foundations of measurement.* New York: Academic Press, 1971.

KRANTZ, D. H., & TVERSKY, A. Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 1971, **78**, 151-169.

SPERLING, G., BUDIANSKY, J., SPIVAK, J., & JOHNSON, M. C. Extremely rapid visual search: The maximum rate of scanning letters for the presence of a numeral. *Science*, 1971, **174**, 307-311.

STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II.* (Reprinted from *Acta Psychologica*, 1969, **30**, 276-315.)

WINER, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

## NOTES

1. Both time and probability can, of course, be broken into subcategories. "Time" would include, for example, reaction time to a tachistoscopic presentation, time to do a long sorting task, duration of an eye fixation, and so on. Likewise, "probability" would include probability of recalling an item, probability of choosing one item as opposed to another, probability of choosing a given confidence rating, and so on. The dependent variable, "number of items," is a disguised form of probability, since it is simply probability multiplied by a constant.

2. In some experiments, however, time as a dependent variable is assumed to reflect something other than time as a theoretical component. For example, reaction time is sometimes used as a measure of memory strength. These experiments are subject to some criticisms (presented below) of experiments that use probability as a dependent variable.

3. The Brown-Peterson paradigm was chosen arbitrarily. The example could have been made within the context of any of a number of memory paradigms.

4. This definition ignores problems having to do with error variance in the data. Sample means, obtained from the data, may form an interpretable interaction, whereas the population means being estimated may form an uninterpretable interaction or vice versa. This, of course, dictates additional caution in data interpretation.

5. The to-be-presented arguments actually apply only to "strictly monotonic" functions. For ease of discourse, the less cumbersome term "monotonic" will be used.

6. This is not, of course, to say that all functions relating a theoretical component to some measure of probability are monotonic. For instance, in yes-no recognition experiments, criterion value ($\beta$) is nonmonotonically related to probability correct.

7. Actually, neither inverse nor linear functions are strictly monotonic, since they run into an abrupt ceiling at $p(x) = 1.0$ and/or an abrupt floor at $p(x) = 0.0$. Ceiling and floor effects will not be dealt with extensively in this paper; suffice it to say that the arguments assume that, in an experiment, none of the probabilities being estimated is 1.0 or 0.0. The arguments could, however, be extended to cover the more general case that does not depend on this assumption.