Temporal order psychometric functions based on confidence-rating data

LORRAINE G. ALLAN
McMaster University, Hamilton, Ontario, Canada

Psychophysical tasks involving confidence judgments allow the simultaneous generation of a family of psychometric functions. Sternberg, Knoll, and Mallows (1975) have demonstrated the power of the multiple-function approach in evaluating models concerned with specifying the source of errors in judgments of simultaneity and temporal order. In the present paper, data from a temporal order task requiring confidence ratings are examined, and a number of models for successiveness and order judgments evaluated.

Sternberg, Knoll, and Mallows (1975) have extended their general model for order judgment, the independent-channels model (Sternberg & Knoll, 1973), to confidence judgments. Experimental tasks involving confidence ratings allow the simultaneous generation of a family of psychometric functions. Sternberg et al. discuss the theoretical usefulness of multiple psychometric functions, and, in particular, they demonstrate the power of the multiple-function approach in evaluating models concerned with specifying the source of errors in judgments of simultaneity and temporal order.

The independent-channels model is the general form for all published models for order judgments. In essence, the model states that a decision function converts the difference in central arrival times of two inputs into an order judgment. Sternberg and Knoll (1973) showed that the psychometric function relating the probability of a particular order response to the temporal separation between the two inputs (t), can be expressed as the convolution of the decision function with the distribution of arrival time differences, $f_t(I)$. One important contribution of the Sternberg et al. paper is the specification of the conditions under which the family of psychometric functions, generated by a confidence-rating procedure, are parallel—that is, differ only by translation along the horizontal t axis.

In the present paper, data are reported from a temporal order task in which the subject was required to rate his confidence about his order response. The family of psychometric functions generated from the data are systematically nonparallel. On the basis of the shape of these functions, a number of specific models for order discrimination can be shown to be incorrect.

METHOD

Six paid subjects participated in the experiment. The subject was required to indicate the order of occurrence of two offsets (of a light and of a tone), and to rate his confidence about his order response.

Each subject was seated in a dimly illuminated, IAC sound-attenuated auditory chamber, approximately 66 cm from the visual display. A glow modulator bulb (Sylvania R1131C), driven by an Iconix power supply (Model 6195-4), was used to generate a 50-fL light signal. The bulb was enclosed in a metal box with an aperture of 4 mm in diam subtending a visual angle of approximately 21 min. The auditory signal was a 810-Hz, 72-dB pure tone produced by a Wavetek function generator, and presented to the observer through earphones. The presentation and timing of the signals and the recording of the responses were under the control of a PDP-8/E computer.

On each trial, one light-tone pair was presented. The onsets of the light and the tone were always physically simultaneous. On one-third of the trials, the offsets were simultaneous (an S_0 pattern), on one-third of the trials, the light terminated before the tone (an S_1 pattern), and on one-third of the trials, the tone before the light (an S_2 pattern). When an S_0 pattern was presented, the duration of both the light and the tone was 2,000 msec. When an S_1 pattern was presented, the duration of the light was 2,000 msec and the duration of the tone was 2,025, 2,050, 2,075, or 2,100 msec. That is, t (the temporal interval between the offset of the light and the offset of the tone) was 25, 50, 75, or 100 msec. The four values of t were equally likely. When an S_2 pattern was presented, the duration of the tone was 2,000 msec and the duration of the light was 2,025, 2,050, 2,075, and 2,100 msec. Again, the four values of t, -25, -50, -75, and -100 msec were equally likely. (A negative value of t indicates that tone offset occurred prior to light offset).

An experimental session consisted of three blocks of 96 trials each, with a 1-min rest between blocks. During each block, 32 S_0 patterns, 32 S_1 patterns, and 32 S_2 patterns were presented in a random order. Of the 32 S_1 patterns, there were 8 at each of the four levels of t; similarly for the S_2 patterns.

The response period began 1.5 sec after the termination of the stimulus pattern, and was marked by the on period of a visual cue. The subject was given 2 sec to make one of four response: he was certain the light offset was first (R_1C), he was certain the tone offset was first (R_2C), he was uncertain that the light offset occurred first (R_1U), or he was uncertain that the tone offset occurred first (R_2U). The interval between the end of the response period and the onset of the next stimulus pattern was 1.5 sec.

At no time during the experiment was the subject given any information regarding his performance, other than that related to consistency and stability. The subjects were practiced until fairly consistent performance was obtained. The actual number of data sessions is shown in Table 1 for each subject.

Table 1
Summary Statistics for Each Subject

t	$P(R_{1C})$	$P(R_1)$	$1 - P(R_{2C})$	$P(R_{1C})$	$P(R_1)$	$1 - P(R_{2C})$
	Subject V.R. (6)			Subject L.M. (6)		
-100	.049	.083	.167	.021	.035	.078
-75	.029	.076	.264	.077	.126	.210
-50	.111	.229	.465	.098	.239	.437
-25	.153	.430	.743	.237	.453	.748
0	.212	.594	.891	.249	.611	.877
25	.245	.748	.951	.259	.727	.892
50	.444	.854	.944	.352	.789	.915
75	.688	.908	.964	.514	.803	.951
100	.909	.972	.972	.773	.936	.979
	Subject V.V. (6)			Subject B.M. (7)		
-100	.125	.146	.146	.126	.235	.319
-75	.174	.201	.236	.132	.246	.401
-50	.292	.396	.472	.115	.382	.618
-25	.347	.444	.653	.161	.530	.857
0	.378	.570	.685	.141	.576	.899
25	.458	.625	.778	.132	.557	.922
50	.430	.618	.757	.161	.631	.946
75	.566	.685	.790	.268	.696	.928
100	.799	.833	.889	.442	.693	.902
	Subject N.C. (4)			Subject A.T. (6)		
-100	.000	.000	.021	.049	.201	.340
-75	.000	.000	.010	.133	.399	.588
-50	.000	.000	.083	.175	.490	.776
-25	.021	.117	.447	.189	.517	.755
0	.116	.490	.938	.274	.613	.836
25	.188	.708	.969	.271	.653	.833
50	.510	.844	.969	.278	.618	.875
75	.854	.979	1.000	.399	.741	.881
100	1.000	1.000	1.000	.437	.704	.852

Note—The number of data sessions is shown in parentheses.

RESULTS AND DISCUSSION

In Table 1, the summary statistics of interest, $P(R_{1C})$, $P(R_1)$, and $1 - P(R_{2C})$, are presented for each subject, and in Figure 1, these probability values are plotted as a function of t. Overall, the three psychometric functions are decidedly nonparallel. In

Figure 1. $P(R_{1C})$, $P(R_1)$, and $P(R_{2C})$ as a function of t for each subject.

general, $P(R_1)$ is fairly symmetric, $P(R_{1C})$ is negatively skewed, and $1 - P(R_{2C})$ is positively skewed. This is more clearly demonstrated by assuming that the distribution of arrival time differences, $f_t(l)$, is normally distributed and transforming the probability values to standard scores, $Z(R_{1C})$, $Z(R_1)$, and $Z(R_{2C})$. If the three psychometric functions were simply horizontally displaced, the three transformed functions would be parallel. In Figure 2a, the standard scores, averaged over the six subjects, are shown as a function of t. It is clear that the three functions are not parallel.

Sternberg et al. have shown that regardless of the shape of the $f_t(l)$ distribution, parallel decision functions on the arrival time difference dimension, on the l axis, require the psychometric functions to be distinguished by horizontal displacement only on the t axis. An example of a decision model which postulates parallel decision functions is one which states that the subject adopts three boundary or criterion values of I, $C_2 \leqslant C \leqslant C_1$. If the difference in arrival times between the offset of the light and the offset of the tone is larger than C_1 ($I > C_1$), he reports that he is certain that the light terminated first, an R_{1C} response; if $I < C_2$, he responds R_{2C}. He is uncertain about the order of occurrence of the offsets for the intermediate I values, and responds R_{IU} if $C \leqslant I \leqslant C_1$ and R_{2U} if
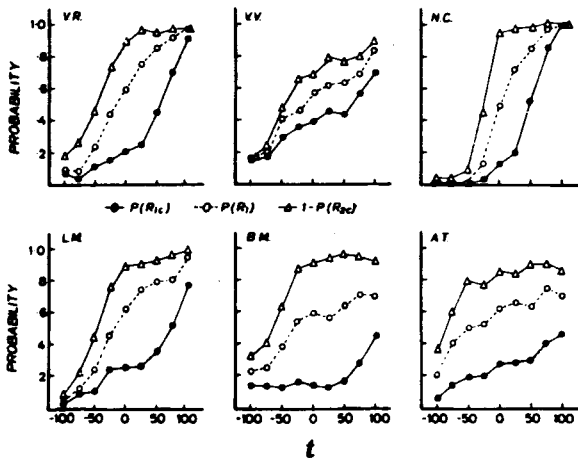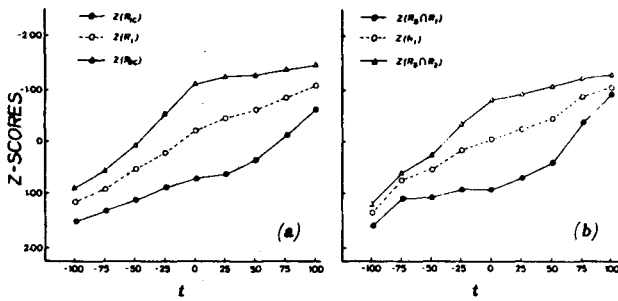
Figure 2. (a) Standard scores as a function of t averaged over the six subjects. (b) Standard scores as a function of t averaged over the three subjects in Allan's successiveness/order task.

$C_2 \leqslant 1 < C$. This type of nonthreshold decision process specifies that the three psychometric functions in the present rating experiment should be parallel.

A somewhat different model of the decision process specifies that there is a range of values of I, $C_2 < 1 < C_1$, within which order information is not available. Within this threshold range, the subject makes an R_1U response with probability β (and an R_2U response with probability $1 - \beta$). If $1 \geqslant C_1$, he responds R_1C, and if $1 \leqslant C_2$, he responds R_2C. This type of threshold decision process requires two of the psychometric functions, $P(R_1C)$ and $1 - P(R_2C)$, to be distinguished by horizontal displacement only.

Sternberg et al. have also shown that for the above two models of the decision process, even if the criteria fluctuate from trial to trial, the psychometric functions will be parallel as long as the criterion distributions differ in location only.

In sum, the data from the present rating experiment deviate systematically from the predictions of models of the decision process which postulate a set of decision functions which are parallel on the arrival time difference dimension. This is the case whether the criteria are fixed or variable, as long as the criterion distributions differ only in location.

In a recent paper, Allan (1975) also presented data which were used to generate a family of psychometric functions. Under one condition of the experiment, the subject made a successiveness decision followed by an order decision on each trial. He judged whether the offsets of a light and a tone were successive, R_S, or simultaneous, R_0, and then whether the light offset preceded the tone offset, R_1, or the tone offset preceded the light offset, R_2. Thus, as the temporal interval between the light offset and the tone offset (t) was varied, three psychometric functions, $P(R_S \cap R_1)$, $P(R_1)$, and $1 - P(R_S \cap R_2)$, were generated. As in the present experiment, the three psychometric functions were not parallel. In Figure 2b, the Z transformations of the psychometric functions are shown averaged over the three subjects in Allan's experiment. This set of functions is remarkably similar to the set generated in the present

experiment (Figure 2a). If, in the joint successiveness/order task, the subject bases his two decisions on the same internal observation, a value of I, then Allan's successiveness/order data, in agreement with the data from the present experiment, provide grounds for rejecting models which incorporate decision functions that are parallel on the I axis.

Allan (1975) extended Kristofferson's (1967) attention-switching model to the joint successiveness/order task, and found that the model was able to predict the shapes of the obtained psychometric functions. The important feature of the attention-switching model, for the present purposes, is the postulated set of decision functions. These functions, which are not parallel on the I axis, are reflected in a family of nonparallel psychometric functions. Thus, there is one specific model, the attention-switching model, which predicts a family of psychometric functions which reflect the shapes of the empirical functions obtained under two different rating procedures.

While the nonparallel decision functions of the attention-switching model can account for the data, Sternberg et al. discuss an alternative decision model which also specifies nonparallel decision functions which could yield appropriately shaped psychometric functions. A basic assumption of the attention-switching model is that the perception of successiveness or nonsimultaneity is a necessary and sufficient condition for the perception of the correct temporal order (Allan, 1975). The alternative suggested by Sternberg et al. is that the perception of order is based on information that is different and to some extent independent of the information upon which the perception of successiveness is based, and that the perception of successiveness does not always result in the correct perception of order. Somewhat surprisingly, these two very different views can predict similar-order psychometric functions.

Psychometric functions generated from a successiveness/simultaneity rating task may be useful in discriminating between the attention-switching model and models which specify that the perception of successiveness is not sufficient for the perception of order. Nonparallel successiveness psychometric functions would be consistent with the attention-switching model. Sternberg et al. do not specifically consider the decision functions for the successiveness task. However, from their discussion, it appears that the decision functions would be parallel on the I dimension, resulting in parallel successiveness psychometric functions. To my knowledge, appropriate empirical successiveness functions do not exist, and we are currently undertaking a study to produce them.

The empirical psychometric functions from the present experiment and from Allan's (1975) earlier experiment provide grounds for rejecting decision models with fluctuating criteria whose distributions

differ only in location. However, Sternberg et al. have shown that they are consistent with decision models which specify criterion variability, if certain constraints are placed on the shapes of the criterion distributions. In particular, a symmetrical distribution of C values, a distribution of C_1 values that is skewed toward small values of I, and a distribution of C_2 values that is skewed toward the large I values, will result in a family of psychometric functions which reflect the shape of the empirical functions.

Sternberg et al. suggest that these constraints on the shape of the criterion distributions are plausible given the symmetrical nature of Allan's experimental tasks. It would be of interest to study asymmetrical tasks and to examine the psychometric functions resulting from such tasks.

CONCLUSIONS

The family of psychometric functions generated from rating data in the present experiment and in Allan's earlier experiment provide grounds for rejecting a number of specific models for temporal order and successiveness judgments. The attention-switching model yields functions which reflect the shapes of the empirical functions. However, so do two alternative models which are discussed by Sternberg et al.

The multiple-function approach has proven useful in rejecting a number of plausible models for judgments of order and successiveness. Research is currently being considered to determine whether the multiple-function approach will provide a means for discriminating between the remaining models.

REFERENCES

ALLAN, L. G. The relationship between the perception of successiveness and the perception of order. *Perception & Psychophysics*, 1975, 18, 29-36.

KRISTOFFERSON, A. B. Attention and psychophysical time. *Acta Psychologica*, 1967, 27, 93-100.

STERNBERG, S., & KNOLL, R. L. The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973. Pp. 629-685.

STERNBERG, S., KNOLL, R. L., & MALLOWS, C. L. Conditions for parallel psychometric functions based on rating-scale data: Application to temporal-order judgments. *Perception & Psychophysics*, 1975, in press.