Signal recognition as influenced by presentation schedules

T. A. TANNER, JR., *AMES RESEARCH CENTER, NASA*
R. W. HALLER AND R. C. ATKINSON, *STANFORD UNIVERSITY*

Performance in a recognition task involving two amplitudes of the same tone was investigated over a wide range of presentation schedules. The task was arranged so that there was no trial-to-trial feedback or other information regarding the relative frequencies of the two tones. The hit and false alarm rates (the proportion of "loud" responses to loud and soft stimuli, respectively) on any given trial were strongly influenced by the stimulus and response on the preceding trial. In general, Ss tended to repeat the last response and were more accurate after a stimulus alternation than after a stimulus repetition. In addition, hit and false alarm rates were inversely related to the presentation probability of the loud tone, in contrast to the direct relation typically found in signal detection experiments and in recognition experiments with trial-to-trial feedback. A mathematical model incorporating three processes (memory, comparison, and decision) was shown to give a good account of these data.

The typical signal recognition experiment consists of a series of discrete trials. A presentation set is defined which includes at least two stimuli (signals); on each trial one member of the set is presented. The observer's (O's)[1] task on each trial is to identify which stimulus was in fact presented. Of particular interest are those experiments in which the signals in the presentation set are highly similar, so that O has great difficulty in making identifications. The task is further complicated by the fact that only one member of the stimulus set is presented on each trial. Therefore, O is forced to rely on information about the stimuli gained from previous presentations, which must be remembered from trial to trial, in order to compare and identify the individual members of the set.

This memory factor for previously presented stimuli distinguishes signal recognition from signal detection experiments. In the signal detection task the background noise from which the signal must be discriminated is present and available for comparison whenever the signal occurs. We can conceive of pure detection and pure recognition experiments as falling at opposite ends of a continuum. The position of any given identification experiment on this continuum is determined by the extent to which O's memory for previous stimuli is involved in determining his behavior. At one end of the continuum of experimental tasks we may study decision making as it interfaces with the sensory process, at the other end, decision making as it interacts with both memory and sensory processes.

In the experiment to be considered here, the stimulus-presentation set included two tones of equal frequency and duration, one slightly louder than the other. On each of a series of trials one of the two tones was presented and O was required to judge whether the loud or soft tone had occurred. The following notation will be used in describing the experiment:

$S_{1,n}$ = presentation of the loud tone on trial n,

$S_{0,n}$ = presentation of the soft tone on trial n,

$A_{1,n}$ = O's response identifying the signal on trial n as loud,

$A_{0,n}$ = O's response identifying the signal on trial n as soft.

Thus, on each trial either S_1 or S_0 was presented and O made either an A_1 or an A_0 response. The major independent variable was the stimulus presentation schedule. The schedules were binomial sequences of S_1 and S_0 events. We will let Y denote the probability of presenting S_1; in the present experiment Y took on values ranging from 0.1 to 0.9. The major dependent variable was the probability of an A_j response given that S_i occurred on the same trial, where i and j can take on the values 1 or 0. The four possible outcomes on trial n can be represented by the *performance matrix*:

$$
\begin{array}{c}
S_{1,n} \\
S_{0,n}
\end{array}
\begin{array}{cc}
A_{1,n} & A_{0,n} \\
\left[\begin{array}{cc}
Pr(A_{1,n}|S_{1,n}) & Pr(A_{0,n}|S_{1,n}) \\
Pr(A_{1,n}|S_{0,n}) & Pr(A_{0,n}|S_{0,n})
\end{array}\right]
\end{array}
\quad (1)
$$

We shall refer to the entries in this matrix as the first-order statistics.

In keeping with the literature on signal detectability theory (SDT), an A_1 response made to an S_1 signal will be called a *hit*, and an A_1 response made to S_0 will be called a *false alarm*; thus,

$$
\begin{aligned}
Pr(\text{hit}) &= Pr(A_{1,n}|S_{1,n}) \\
Pr(\text{false alarm}) &= Pr(A_{1,n}|S_{0,n})
\end{aligned}
\quad (2)
$$

Fixing the hit and false alarm probabilities completely specifies the performance matrix since each row of the matrix sums to 1.0. Other probabilities of interest can be defined in terms of hits and false alarms. The probability of an A_1 response on trial n, independent of the signal event, is simply

$$
Pr(A_{1,n}) = Pr(A_{1,n}|S_{1,n})Y + Pr(A_{1,n}|S_{0,n})(1-Y) \quad (3)
$$

and the probability of a correct response is

$$
\begin{aligned}
Pr(C_n) &= Pr(A_{1,n}|S_{1,n})Y \\
&\quad + [1 - Pr(A_{1,n}|S_{0,n})](1-Y)
\end{aligned}
\quad (4)
$$

Also of interest are sequential (trial-to-trial) fluctuations in hits and false alarms. The sequential effects of particular interest concern the stimulus and response events on trial n-1 as they influence the response on trial n; specifically,

$$Pr(A_{1,n} | S_{1,n} A_{j,n-1} S_{i,n-1}) \qquad (5)$$

$$Pr(A_{1,n} | S_{0,n} A_{j,n-1} S_{i,n-1}) \qquad (6)$$

Expression 5 denotes the probability of a hit, given that on the preceding trial an S_i was presented and an A_j occurred. Expression 6 denotes the probability of a false alarm, given that on the preceding trial S_i was followed by A_j. For each of the four outcomes on trial n-1 ($A_1 S_1$, $A_1 S_0$, $A_0 S_1$, $A_0 S_0$) we can define a hit and false alarm probability on the current trial n. Hereafter, asymptotic performance values will be indicated by omitting the trial subscripts (n and n-1) from both the sequential and the first-order statistics. At all times the temporal order of events in the sequential statistics should be interpreted as corresponding to that in Expressions 5 and 6.

Signal recognition has been investigated much less thoroughly than signal detection. However, attempts have been made to compare performance in the two tasks (Kinchla, 1966; Luce, 1963a; Tanner, 1956). Kinchla contrasted results obtained from a recognition experiment with results typically obtained in detection experiments. The presentation set consisted of two tonal amplitudes. When trial-to-trial feedback about the correctness of O's response was omitted, variations in γ (from 0.25 to 0.75) had a smaller influence on hits and false alarms than that usually obtained in signal detection studies. In addition, Kinchla found that the sequential effects were dramatically greater than those typically found in detection experiments. These deviations from the pattern of performance found in detection experiments were shown by Kinchla to be greater for those Os who were not given trial-to-trial feedback than for those who were given such information.

In signal detection experiments O usually is told the proportions of signal and no-signal trials at the start of each session. In addition, O frequently is given trial-to-trial feedback concerning the presence or absence of the signal immediately after his response on each trial. Green and Swets (1966, p. 395) cite several studies in support of the statement that, "It [feedback] helps to bring about a rapid approach to asymptotic performance, but its effect is relatively small when the signal is previewed prior to each block of trials." Thus it often has been assumed by proponents of SDT that an O would approach approximately the same asymptote with trial-to-trial feedback as he would without it, but that he would approach this asymptote at a faster rate when given feedback. Theorists favoring a threshold approach (Atkinson &

Kinchla, 1965; Luce, 1963b) have taken a similar view, namely, that feedback and knowledge of the relative frequencies of the signals allow the subject to bias his reports (usually in the sense of the number of nondetections that are reported as detections) in order to maximize his gain.

Swets et al (1961) have shown that subjects are remarkably good at producing, on demand, various hit and false alarm rates in a detection task or, in terms of SDT, at adjusting their criteria to ideal values. Mere knowledge of the presentation schedule should be sufficient to produce proper performance under either an SDT or a threshold interpretation. Kinchla's (1966) results, however, suggest that feedback plays some role other than merely helping the subject to adjust his bias or criterion. Kinchla found that subjects given trial-to-trial feedback did indeed match their A_1 response probability to the stimulus probability, whereas subjects who were told only the value of γ did not. Instead, their A_1 response probabilities regressed toward 0.50. In addition, large sequential effects were observed. In general, subjects tended to repeat the last response and to be more accurate after a stimulus alternation than after a repetition.

Parducci and Sandusky (1965) reported a study in which subjects were required to discriminate two different points in space (indicated by turning on one of two lights). Using presentation probabilities of 0.2, 0.5, and 0.8 and without giving trial-to-trial feedback or indicating the stimulus presentation schedule, they found a consistent but nonsignificant trend indicating an inverse relation between γ and hit and false alarm rates, viz, that hit and false alarm rates decreased as γ increased. Sequential effects were observed that were essentially the same as those reported by Kinchla (1966) in the no-feedback condition of his experiment.

The purpose of the present experiment was to determine whether the omission of both trial-to-trial feedback and information about the value of γ in an auditory recognition task would yield results like those of Parducci and Sandusky (1965) for visual recognition. In order to investigate the relationships in detail, a more extensive set of γ values was employed than in the two previous studies. Additional data points were obtained by using a four point confidence rating scale (rather than a simple binary choice) for O's responses.

METHOD

The Os were six male college freshmen (O_1 to O_6) and six female housewives (O_7 to O_{12}), ranging in age from 20 to 25 and 30 to 50, respectively. All Os reported having normal hearing. They were paid at the rate of $2.50 per h. At the beginning of the experiment they all were naive with respect to the experimental conditions. During testing each O sat

in a separate, sound-attenuated booth (Industrial Acoustics 400-A).

The task consisted in O's judging, on each of a series of trials, which of two auditory amplitudes had occurred, and indicating whether he was certain or uncertain of his judgment. The sequence of events on each trial was as follows: a 1-sec ready period, designated by the illumination of a small white light on a panel before O; the presentation of one of the two signals for 0.1 sec; a 1.9-sec. response period, designated by O's response buttons (also in the panel) being illuminated. A 2-sec interval followed each trial, making a total of 5 sec between signal presentations.

The tones were 1000 Hz sinusoids, presented through earphones (Permoflux PDR-8) for a duration of 100 msec. They were generated by a Hewlett-Packard, Model 201CR Audio Oscillator; timed by a Grason-Stadler, Model 471-1 Interval Timer; and gated with 10 msec rise and decay times by a Grason-Stadler, Model 620D Electronic Switch. No external noise was presented. The amplitude of S_1, the louder signal, was constant throughout the experiment at a sound pressure level of 70 dB above 0.0002 dyne per cm^2. The amplitude of S_0, the soft signal, was adjusted individually for each O, contingent on his performance during three practice sessions. The adjustment was made after each block of 50 trials so that by the end of the third practice session O was responding correctly on about 70 percent of the trials. At this time the amplitude settings of S_0 were as follows for the 12 Os (O_1 to O_{12}): 66.8, 67.2, 67.7, 67.4, 67.5, 67.2, 67.9, 67.9, 67.7, 67.1, 67.9, and 68.0 (mean = 67.5). These amplitudes for S_0 were held constant for the remainder of the experiment. During the practice sessions γ was set equal to 0.5.

Each session consisted of 400 trials. Within a session the sequence of S_1 and S_0 trials was determined by one of five schedules: $\gamma = 0.1, 0.3, 0.5, 0.7$, and 0.9. In successive five-day blocks O performed for one day on each of the five schedules. Within each five-day block the order of schedules was randomly determined. The experiment involved 15 sessions (exclusive of the practice sessions); therefore, each schedule was run on three separate days. Within each session S_1 and S_0 were randomized within successive blocks of 50 trials. In order to discount possible warmup effects, the first 50 trials of each session were eliminated from the data analysis.

The Os were given no information concerning the presentation schedules and no trial-to-trial feedback at any time during the experiment.

The responses were recorded by O's pressing one of four buttons, which were arranged on the panel before him in a horizontal array. For half the Os, the buttons were labeled from left to right: *soft-certain, soft-uncertain, loud-uncertain,* and *loud-certain.* For the other half of the Os, the ordering of the buttons was reversed. We shall use the following notation for these rating responses: A_1^{**} = loud-certain, A_1^* = loud-uncertain, A_0^* = soft-uncertain, A_0^{**} = soft-certain.

RESULTS

Table 1 presents hit rates, $Pr(A_1|S_1)$, false alarm rates, $Pr(A_1|S_0)$, and the unconditional A_1 response probability, $Pr(A_1)$, for individual Os. In Fig. 1 these statistics, averaged over the 12 Os, are plotted against γ. There was a general bias for reporting the occurrence of S_1, i.e., $Pr(A_1)$ was above 0.50 for all values of γ. $Pr(A_1)$ increased monotonically with γ, although the observed increase of A_1 (0.52 to 0.60) was quite small compared to that of γ (0.1 to 0.9). This small influence of γ on $Pr(A_1)$ is similar to that found by Parducci and Sandusky (1965). However, Kinchla (1966) found that $Pr(A_1)$ tended to match γ when his subjects received trial-to-trial feedback. When he omitted
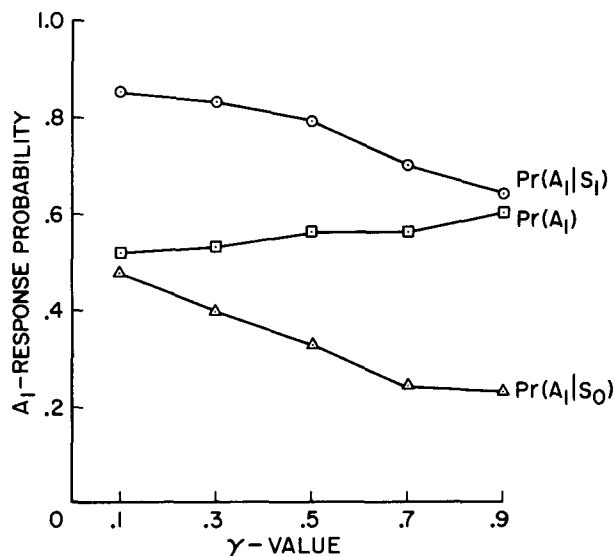

Fig. 1. Probability of hits, false alarms, and the A_1 response averaged over observers.

Table 1. Probability of hits, false alarms, and the unconditional probability of A_1 for each of the 12 observers

Observer	$Pr(A_1\|S_1)$ γ-Value					$Pr(A_1\|S_0)$ γ-Value					$Pr(A_1)$ γ-Value				
	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
1	.99	.96	.99	.94	.91	.90	.74	.69	.55	.38	.91	.81	.84	.82	.86
2	.89	.90	.83	.66	.68	.37	.29	.13	.13	.14	.42	.47	.48	.50	.63
3	.85	.76	.69	.64	.54	.41	.31	.28	.22	.23	.45	.44	.48	.51	.51
4	.88	.81	.74	.64	.53	.43	.36	.26	.14	.20	.48	.50	.50	.49	.50
5	.81	.67	.70	.63	.52	.47	.37	.35	.32	.28	.50	.46	.52	.54	.50
6	.70	.75	.72	.53	.51	.36	.30	.24	.13	.08	.39	.44	.48	.41	.47
7	.93	.83	.74	.68	.60	.45	.31	.24	.17	.13	.50	.47	.49	.53	.55
8	.95	.95	.88	.84	.75	.72	.60	.45	.32	.27	.74	.70	.66	.68	.70
9	.75	.68	.69	.51	.48	.22	.16	.18	.11	.05	.27	.32	.44	.39	.44
10	.89	.95	.93	.86	.84	.46	.47	.38	.29	.45	.50	.61	.66	.69	.80
11	.68	.80	.71	.61	.62	.56	.54	.52	.46	.57	.57	.62	.62	.56	.62
12	.93	.88	.86	.81	.70	.42	.29	.24	.10	.04	.47	.47	.55	.60	.63

feedback, but told subjects the value of Υ, $Pr(A_1)$ regressed toward 0.50, but to a much lesser degree than in the present study. Figure 1 also shows that both the hit and false alarm rates decreased as Υ increased.

Figure 2 presents hit and false alarm rates plotted on a receiver-operating-characteristic (ROC) graph, with Υ as the parameter. The five points appear to lie on a curve similar to that generated in signal detection experiments. When these data are plotted on normal-normal coordinates, a straight line with a slope of 1.0 provides a fairly good fit.[2] This is consistent with the predictions of SDT for performance in a signal detection task of the "yes-no" type (Green & Swets, 1966, Ch. 4). However, the ordering of the points along the ROC curve is the reverse of that typically obtained in detection experiments, where the probabilities of both hits and false alarms increase as Υ is increased.

The results of the present study, viz, that hits and false alarms decreased as Υ increased, are the same as those obtained by Parducci and Sandusky (1965) for recognition of visual displacement. However, the extent of the effect, as indicated by the spread of the points in the ROC space, was somewhat greater in the present study for comparable variations in Υ. The data in Table 1 verify that the group averages presented in Fig. 1 are typical of the individual Os.

If the data displayed in Fig. 2 did not represent asymptotic performance, then the relationship between Υ and the hit and false alarm rates should change over time. Figure 3 presents the hit and false alarm probabilities for each of the three experimental ses-
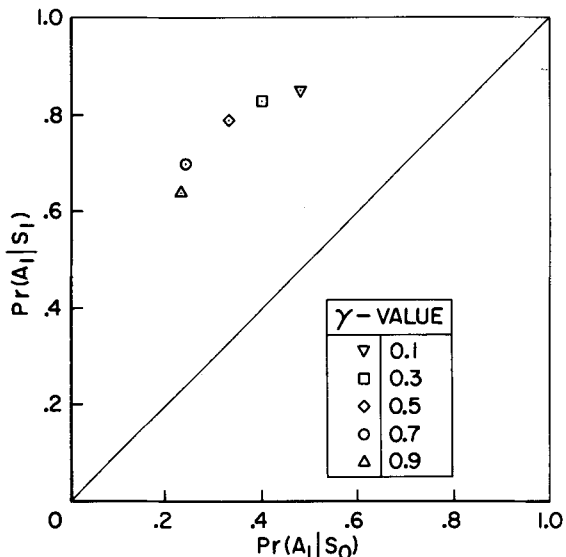

Fig. 2. ROC graph of hit and false alarm probabilities for each of the five Υ values.

sions during which a given γ value was in effect (upper graphs) and for three consecutive blocks of trials averaged over all sessions for a given value of γ (lower graphs). The symbols for the five γ conditions are the same as those used in Fig. 2. The three upper graphs indicate that the relationship between γ and the hit and false alarm rates was consistent over sessions. The three lower graphs show that performance was consistent within sessions. This evidence suggests that the relationship in Fig. 2 is representative of performance throughout the experiment and that there was no trend toward the opposite relationship observed in detection studies.

Figure 4 presents an ROC graph of the data generated by the confidence ratings, A_0^{**}, A_0^*, A_1^*, and A_1^{**}. These data provide a measure of the consistency of the relationship between γ and performance as measured at different points along the rating scale. The points on the graph were generated using the method described by Green and Swets (1966, pp. 101-103). The confidence ratings were considered as a scale ranging from A_0^{**} to A_1^{**}. This scale was compressed into binary forced-choice data by considering successively each of the confidence boundaries as defining a criterion for reporting the occurrence of S_1. Thus, 15 points were generated, one for each intersection of the three confidence boundaries and the five values of γ. The open points are the same as those plotted in Fig. 2 since they were generated by dichotomizing the ratings between A_1 and A_0. For the half filled points only A_1^{**} was considered as an A_1, and for the solid points $A_1^{**} + A_1^* + A_0^*$ were considered together as an A_1 response. Figure 4 shows that, regardless of where on the rating scale A_1 and A_0 are dichotomized, the relationship demonstrated in Fig. 2 holds, i.e., hit and false alarm rates decreased as γ increased.

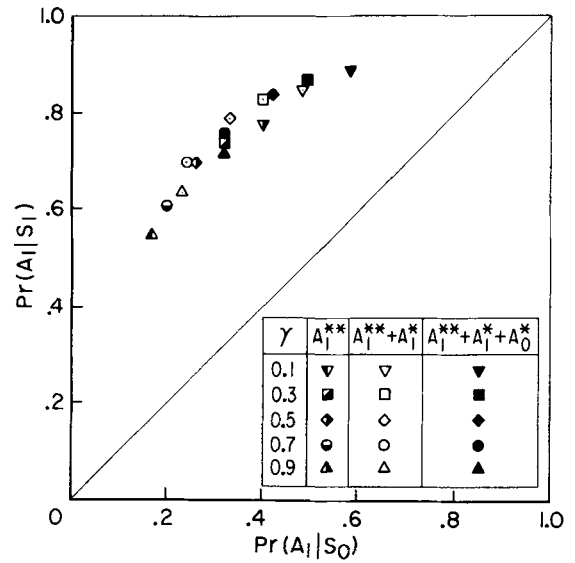| γ	A_1^{**}	$A_1^{**} + A_1^*$	$A_1^{**} + A_1^* + A_0^*$
0.1	▼	▽	▼
0.3	◪	□	■
0.5	◆	◇	◆
0.7	◐	○	●
0.9	◭	△	▲

Fig. 4. ROC graph of the rating scale data.

Table 2 gives estimates of $Pr(A_{1,n} | S_{1,n} A_{j,n-1} S_{i,n-1})$ and $Pr(A_{1,n} | S_{0,n} A_{j,n-1} S_{i,n-1})$ for the individual Os. The blank (—) entries are due to the fact that some signal-response contingencies did not occur. These blank entries and most of the probability estimates of 0 and 1 (which resulted when the number of observations on which the probabilities were based was quite small) occurred when γ was either 0.1 or 0.9.

Figure 5 shows ROC graphs of these sequential statistics averaged over the 12 OS. Each of the four points plotted for a given value of γ (the circles and squares) represents the hits and false alarms based on one of the four possible outcomes on trial n-1. The relative ordering of the four points within the
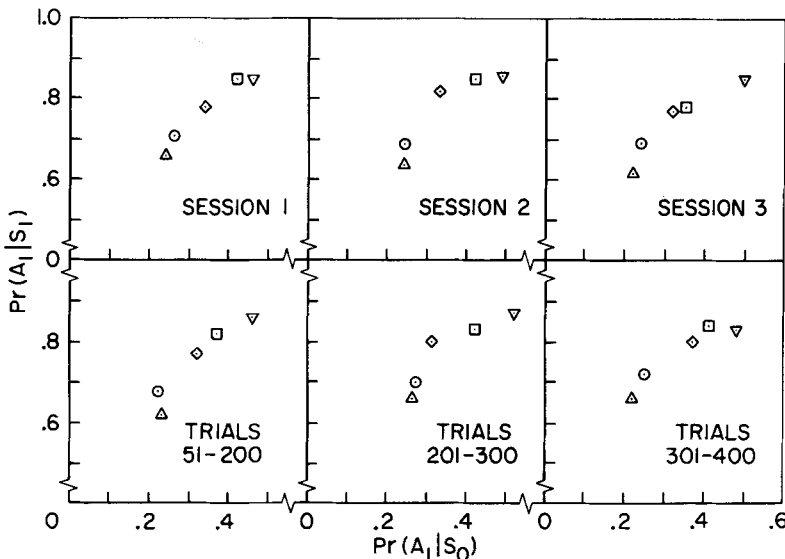
Fig. 3. ROC graphs across sessions (upper panels) and within an average session (lower panels).

Table 2. Hit (H) and false alarm (F) probabilities on trial n, as a function of $A_j S_i$ on trial n-1, for each of the 12 observers

Observer		$\gamma = 0.1$				$\gamma = 0.3$				$\gamma = 0.5$				$\gamma = 0.7$				$\gamma = 0.9$			
		A_1S_1	A_0S_1	A_1S_0	A_0S_0	A_1S_1	A_0S_1	A_1S_0	A_0S_0	A_1S_1	A_0S_1	A_1S_0	A_0S_0	A_1S_1	A_0S_1	A_1S_0	A_0S_0	A_1S_1	A_0S_1	A_1S_0	A_0S_0
1	H	1.00*	--	.99	1.00*	.96	.50†	.99	.90	.98	1.00†	1.00	.98	.96	.87	.96	.83	.94	.79	.97	.87
	F	.89	.00†	.94	.73	.71	.20*	.83	.57	.67	.67†	.75	.59	.55	.36*	.67	.39	.43	.00*	--	.14*
2	H	1.00†	1.00†	.94	.82	.98	.89*	.99	.81	.94	.74	.95	.72	.81	.44	.91	.54	.84	.41	.75*	.43
	F	.69	.06*	.67	.12	.50	.14	.40	.13	.22	.03	.24	.05	.22	.04	.20*	.05	.22	.00	.00†	.00*
3	H	1.00†	1.00†	.91	.77	.83	.71	.84	.69	.75	.57	.76	.66	.67	.61	.78	.57	.61	.45	.64	.53
	F	.53	.00*	.50	.34	.30	.15	.40	.29	.29	.22	.33	.24	.24	.10	.33	.24	.22	.12	.00†	.38*
4	H	.83*	1.00†	.84	.88	.76	.55*	.85	.85	.70	.66	.84	.73	.64	.56	.77	.74	.53	.54	.68*	.43
	F	.26	.35*	.45	.47	.25	.16	.48	.39	.25	.19	.27	.31	.13	.05	.30*	.19	.19	.16	1.00†	.33*
5	H	.60*	.00†	.82	.86	.65	.39	.78	.68	.68	.70	.80	.69	.70	.55	.68	.61	.60	.42	.58	.45
	F	.41	.46*	.53	.42	.32	.22	.53	.33	.33	.33	.38	.36	.32	.24	.39	.41	.30	.25	1.00†	.00*
6	H	.75*	.00†	.91	.64	.82	.52*	.89	.70	.81	.70	.91	.63	.58	.44	.80	.57	.59	.45	.75*	.47
	F	.33	.11	.46	.31	.34	.10	.40	.26	.26	.20	.36	.18	.14	.12	.31*	.10	.09	.07	--	.25*
7	H	.80*	--	.95	.88	.83	.71*	.90	.85	.77	.58	.93	.71	.69	.56	.84	.72	.66	.48	.89*	.64
	F	.44	.17*	.49	.44	.31	.25	.37	.31	.21	.13	.31	.30	.17	.11	.10*	.25	.13	.10	--	.75†
8	H	1.00*	--	1.00	.71*	.96	1.00*	.99	.86	.89	.81	.97	.84	.88	.70	.86	.79	.83	.59	.89	.63
	F	.55	.40*	.80	.64	.60	.22*	.68	.48	.44	.24	.54	.39	.33	.10	.42	.36	.31	.14	.00†	.33*
9	H	1.00†	.67†	.93	.65	.92	.59	.97	.59	.88	.57	.96	.59	.77	.30	.80	.39	.75	.30	.75†	.34
	F	.36	.08	.56	.13	.18	.06	.45	.11	.21	.15	.44	.12	.23	.03	.50*	.00	.11	.00	1.00	.00*
10	H	1.00*	--	.99	.78	.96	.83*	.95	.94	.95	.74	.99	.90	.89	.71	.92	.82	.85	.79	.92	.87
	F	.53	.11*	.60	.33	.45	.36*	.58	.42	.41	.25	.52	.29	.28	.03	.44	.27	.49	.71*	.43†	.00†
11	H	1.00†	1.00*	.44	.89	.69	.64*	.78	.95	.57	.87	.66	.88	.46	.72	.52	.90	.55	.74	.51	.65
	F	.55	.75	.38	.80	.37	.79	.40	.81	.40	.76*	.37	.75	.30	.64	.34	.65	.56	.70	.33	.75†
12	H	1.00*	1.00†	.96	.92	.87	.69*	.87	.90	.85	.89	.80	.87	.78	.86	.86	.82	.70	.64	1.00†	.86
	F	.41	.33*	.43	.43	.29	.35	.25	.31	.22	.16	.30	.24	.06	.14	.17*	.18	.02	.06	--	.00†

* Based on six to twenty occurrences of $S_{i,n}A_{j,n-1}S_{i,n-1}$

† Based on one to five occurrences of $S_{i,n}A_{j,n-1}S_{i,n-1}$

-- No occurrences of $S_{i,n}A_{j,n-1}S_{i,n-1}$

ROC space is independent of γ; i.e., for all values of γ the hit and false alarm rates were highest when an A_1 was made to an S_0 on the preceding trial, lowest when an A_0 was made to an S_1, and higher when an A_1 was made to an S_1 than when an A_0 was made to an S_0. These trial-to-trial sequential effects are the same as those reported by Kinchla (1966). Kinchla also showed that the sequential effects were much stronger without trial-to-trial feedback than they were with feedback. In view of these results we had hypothesized that in the present experiment, where information about the value of γ was omitted, the sequential effects might be still greater. However, this does not appear to be the case; the magnitude of the sequential effects shown in Fig. 5 is comparable to, but no greater than, those obtained by Kinchla when his subjects were given no trial-to-trial feedback.

Parducci and Sandusky (1965) did not report sequential data in the form presented in Table 2 and Fig. 5. However, they have supplied us with the necessary statistics, and the same ordering of points as that shown in Fig. 5 was found. Thus, this relative ordering of the sequential statistics $Pr(A_1 | S_1 A_j S_i)$ and $Pr(A_1 | S_0 A_j S_i)$ is observed in the Parducci-Sandusky study, the Kinchla study, and in the present study, suggesting that the order does not depend on the presentation schedule, knowledge of presentation probabilities, or the sensory modality.

DISCUSSION

A theory of signal recognition formulated by Haller and Atkinson (1967) may be used to interpret the findings of this study. The theory assumes three processes: a *memory process* that stores an image of the signal presented on the preceding trial, a *comparison process* that calculates a difference function on the stored image and the incoming signal, and a *decision process* that selects a response on the basis of the comparison process. Except for the explicit inclusion of the memory process, the theory is similar to SDT and makes comparable types of predictions.

As noted above, we assume that the O has in memory an image of the signal presented on the immediately preceding trial. This stored image will be referred to as the *trace*. Due to errors of rehearsal and the influence of other noise sources, O's trace of the signal S_i will take on different values from trial to trial and is best described as a random variable T_i. It will be assumed that T_i is normally distributed with mean t_i and variance σ_T^2. The trace distributions for the signals S_1 and S_0 have different means, t_1 and t_0, but a common variance.[3]

On each trial of the experiment, O processes both the image of the presented signal and the trace of the last signal. We shall call the image associated with signal S_i the *input*, a random variable denoted as I_i. It is also assumed that this random variable is normally distributed with mean s_i and variance σ_I^2. Thus, the two signals S_1 and S_0 are characterized by two input distributions with means s_1 and s_0 but a common variance.

According to the theory, on each trial of the experiment O compares the trace from the preceding signal with the input of the current signal. He then calculates the algebraic difference between the trace and the input on the relevant dimension (the dimension on which O is asked to base his judgment). If signal S_k was presented on the preceding trial and signal S_i is presented on the current trial, then the difference score d_{ik} is distributed as a random variable D_{ik} that is specified by the equation

$$D_{ik} = I_i - T_k \qquad (7)$$

To avoid confusion, it should be noted that whereas the trace on trial n is determined by the stimulus input on trial n-1, the input on trial n is assumed to be independent of the trace active on trial n. Thus D_{ik} is normally distributed with mean $s_i - t_k$ and variance $\sigma_D^2 = \sigma_T^2 + \sigma_I^2$.

The decision process uses the output of the comparison process to generate a response as follows:

$$\text{If} \left\{ \begin{array}{l} d_{ik} > \delta_0 \\ d_{ik} < \delta_1 \\ \text{otherwise} \end{array} \right\} \quad \text{then} \quad \left\{ \begin{array}{l} \text{make response } A_1 \\ \text{make response } A_0 \\ \text{repeat response} \\ \text{made on the} \\ \text{preceding trial} \end{array} \right\} \quad (8)$$

where $\delta_0 \geq \delta_1$. If the difference between the input and the trace is greater than some criterion value δ_0, then an A_1 response occurs; if the difference is less than some criterion value δ_1, than an A_0 occurs; if the difference does not exceed either the lower or the upper criterion, then the response made on the preceding trial is repeated. In essence, when O subtracts the trace of the last signal from the image of the current signal and obtains a ''large'' positive difference, he calls the current signal loud; when he obtains a ''large'' negative difference, he calls the current signal soft; and when he obtains little or no difference, he identifies the current signal as a repetition of the last one and repeats his last response.

The derivations and resulting equations for quantities like $\Pr(A_1|S_1)$, $\Pr(A_1|S_0)$, $\Pr(A_1)$, and $\Pr(C)$ are fairly complex in this theory. However, certain sequential statistics take on extremely simple forms. Let $\Phi(x)$ be the integral of the unit normal density function from $-\infty$ to x, i.e.,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(1/2)y^2} dy \qquad (9)$$

Then, from the assumptions presented above, it can be shown that

$$\Pr(A_1|S_iA_jS_k) = \Phi\left(\frac{s_i - t_k - \delta_j}{\sigma_D}\right) \qquad (10)$$

where i, j, and k can take on the values 0 or 1.

In evaluating predictions that may be made from this theory, we shall consider first those that do not depend on parameter estimates. It can be shown that the point $[\Pr(A_1|S_0A_jS_i), \Pr(A_1|S_1A_jS_i)]$ will fall on an ROC curve of the form predicted by SDT, assum-

TRIAL n-1   ○ $A_1 S_0$ □ $A_0 S_0$
            ● $A_1 S_1$ ■ $A_0 S_1$

$\gamma = 0.1$ $\gamma = 0.3$ $\gamma = 0.5$
$\gamma = 0.7$ $\gamma = 0.9$ COMPOSITE

y-axis: $\Pr(A_{1,n}|S_{1,n}\,A_{j,n-1}\,S_{i,n-1})$
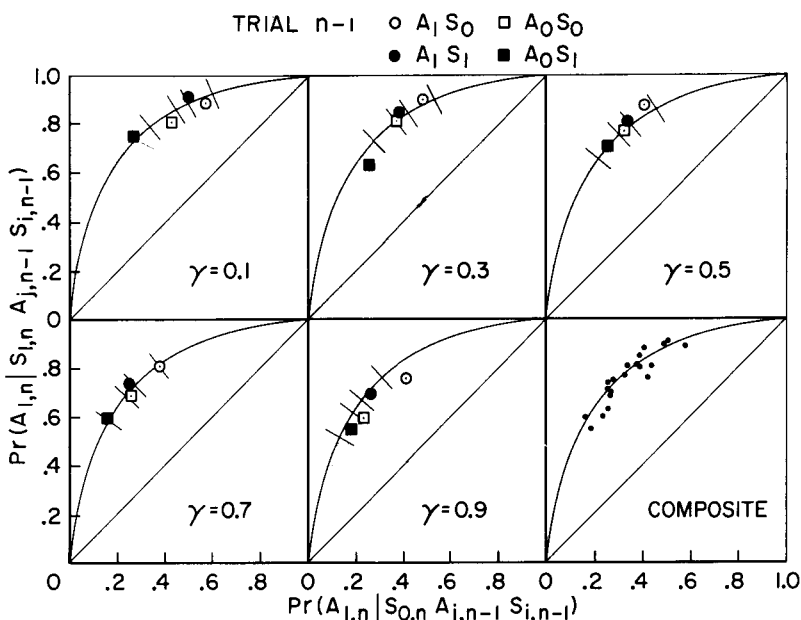x-axis: $\Pr(A_{1,n}|S_{0,n}\,A_{j,n-1}\,S_{i,n-1})$

Fig. 5. ROC graphs of sequential statistics with predicted values marked on the theoretical curve.

ing underlying normal distributions with equal variance (Green & Swets, 1966, p. 88f).[4] Thus, the four observed points (circles and squares) in the graphs for each of the five Y conditions in Fig. 5 should fall on a bow-shaped curve that is symmetric about the negative diagonal and passes through the points (0,0) and (1,1). This appears to be the case, as is shown by the theoretical curve that has been drawn through the observed points. This predicted ROC curve depends on the parameter σ_D and we assume that σ_D is not influenced by the presentation schedule (Y value). In general, the observed points for each Y value are reasonably well fit by the theoretical curve. The accuracy of the fit is further indicated in the Composite Graph of Fig. 5, where the curve has been drawn through the 20 points taken from the other five graphs.

An additional way to test the above predictions would be to plot the 20 points from the Composite Graph of Fig. 5 on double-probability paper. All the points in such a plot should fall on a straight line with a slope of 1.0 and an intercept that is a function of σ_D. To evaluate the predictions for individual subjects, such a plot was made for the data in Table 2. These graphs require too much space to be reproduced here, but inspection of them indicates that a straight line provides a fairly accurate fit for individual Os. For two Os there is a suggestion that the slope of the best fit ROC curve is greater than 1.0, but for the others a straight line with unit slope is satisfactory.

In order to make more detailed predictions with this theory, it is necessary to estimate parameters. Fortunately, there are fewer parameters to be estimated than may appear at first. As noted previously, we assume that σ_D is independent of Y. Similarly s_1 and s_0 are viewed as measures of the signals, and are assumed to be constant over the five values of Y. Furthermore, s_1 and s_0 may be regarded as scaling parameters and arbitrarily set to any values. They might be set equal to the actual amplitudes of the two signals, but for mathematical convenience we let $s_1 = 1$ and $s_0 = 0$. The ROC curve for the sequential statistics shown in Fig. 5, when plotted on double-probability paper, will have slope 1 and an intercept which is simply $\Phi[(s_1 - s_0)/\sigma_D]$. Since we have let $s_0 = 0$ and $s_1 = 1$, the intercept becomes $\Phi(1/\sigma_D)$.

In the present experiment both the hit and false alarm rates vary as a function of Y. To account for this the theory must assume that variations in Y affect either the decision criteria δ_1 and δ_0 or the means of the trace distributions, t_1 and t_0. For reasons elaborated more fully in Haller and Atkinson (1967), it is assumed that δ_1 and δ_0 are independent of Y, at least in experiments in which O receives no information regarding the relative proportions of S_1 and S_0 events. There is some justification for this assumption in the present experiment since none of the Os reported an awareness of the fact that Y was being varied from session to session.[5] Therefore, we will assume that t_1 and t_0 depend on Y. The postulated relationship is linear

and is specified by a single parameter α as follows:

$$t_1 = \alpha + (1 - \alpha) Y \qquad (11)$$

$$t_0 = (1 - \alpha) Y \qquad (12)$$

The reasons for this assumption are discussed in Haller and Atkinson (1967) and will not be reviewed here. However, note that a signal that appears with high frequency tends to be remembered with less distortion than a signal that occurs only rarely. For example, when Y is close to 1, the S_1 signal is occurring with high probability and $t_1 \approx 1$ and $t_0 \approx 1 - \alpha$; when Y is close to 0, then $t_1 \approx \alpha$ and $t_0 \approx 0$. Thus the more frequent signal is remembered with more accuracy, while recall of the less frequent signal is less accurate. When the two signals occur with equal frequency, both t_1 and t_0 move away from the limiting values of 1 and 0 by an equal amount, i.e., $t_0 = 1 - t_1$. Equations 11 and 12 are consistent with assumptions about adaptation levels corresponding roughly to those proposed by Helson (1959) and Parducci and Sandusky (1965). Kinchla's (1967) random walk model might also be applied to justify the assumed relationships, if the walk parameter and the presentation parameter, Y, were functionally related.

Given these constraints on the general theory, we need to estimate σ_D, δ_1, δ_0, and α. We shall not describe the techniques used to estimate these parameters here; the interested reader should consult Haller and Atkinson (1967). However, note that σ_D can be estimated by fitting a straight line to the 20 points in the Composite Graph of Fig. 5, when they are plotted on double-probability paper. Identical procedures can be used to estimate δ_1, δ_0, and α, as is indicated in Haller and Atkinson (1967). Estimates of the parameters for the group data are as follows: $\sigma_D = 0.83$; $\delta_1 = -0.29$; $\delta_0 = 0.04$; $\alpha = 0.22$.

The predicted and observed values for $Pr(A_1|S_1)$, $Pr(A_1|S_0)$, $Pr(A_1)$, $Pr(C)$, and the sequential hit and false alarms $Pr(A_1|S_1A_jS_i)$ and $Pr(A_1|S_0A_jS_i)$ are shown in Table 3. The predicted values of the sequential statistics are indicated in Fig. 5 by the points at which the four short lines in each graph are drawn perpendicularly through the theoretical ROC curve. For each value of Y the four predicted points are in the same order along the curve as that of the corresponding observed points. With four points predicted for each of the five Y values, a total of 20 points have been fit, each with two degrees of freedom. Thus, there are a total of 40 degrees of freedom in Fig. 5 and only four have been used to estimate parameters. Given this consideration, the fits are remarkably good.

The theoretical relation between $Pr(A_1|S_1)$ and $Pr(A_1|S_0)$ does not take a simple mathematical form such as the ROC curve for the sequential statistics. However, note that for any value of Y the point $[Pr(A_1|S_0), Pr(A_1|S_1)]$ in Fig. 2 can be obtained by a weighted average of the corresponding four sequen-

tial points in Fig. 5. Thus, according to the theory, the points $[Pr(A_1|S_0), Pr(A_1|S_1)]$ for the five γ conditions of Fig. 2 should fall on a curve that closely approximates the ROC curve of Fig. 5, but that is slightly below it. This is shown in Fig. 6, where the observed values from Fig. 2 (open points) are plotted along with the corresponding predicted values (solid points); also shown is the theoretical ROC curve from Fig. 5.

CONCLUSION

The most significant finding of the present study was the marked influence on hits and false alarms of the response and signal events on the preceding trial. Independent of γ both the hit and false alarm rates were ordered as follows for the four possible outcomes on trial n-1: $Pr(A_1|S_iA_1S_0) > Pr(A_1|S_iA_1S_1) > Pr(A_1|S_iA_0S_0) > Pr(A_1|S_iA_0S_1)$ for $i = 1$ or 0. This ordering and the spacing of the four points in the ROC space are similar to the results reported by Parducci and Sandusky (1965) and by Kinchla (1966). On the basis of this evidence it seems reasonable to conclude that strong sequential effects occur consistently in recognition tasks.

The results of the present study and those of the other recognition studies reviewed here indicate that the influence of the presentation schedule on hits and false alarms depends on whether or not the subject is given information about the schedule. If, as in the present study (or that of Parducci and Sandusky), no information is given, then hits and false alarms decrease as γ increases. If, as in Kinchla's (1966) study, such information is provided, then the relation is reversed; the hit and false alarm rates increase as γ increases.

We have demonstrated that the results of the present study can be predicted by the Haller-Atkinson Model. Kinchla's (1966) results also can be accounted for by the model. When a subject is told the schedule for a session but not given trial-to-trial feedback, we would expect him to adjust his δ_i values in order to bias responses in favor of the more frequent stimulus. This would cause hits and false alarms to be appropriately ordered, whereas the ordering and spacing of the sequential statistics would not be greatly affected. When trial-to-trial feedback is added, the subject should still adjust δ_i to favor the more frequent stimulus and, in addition, when the trace and the input are perceived to be the same, he should make the response that was designated as correct on the last trial, rather than simply repeating the last response. Thus we would expect to find that the effect of the pre-

Table 3. Observed and predicted values for sequential and first-order statistics

Statistic	$\gamma = 0.1$		$\gamma = 0.3$		$\gamma = 0.5$		$\gamma = 0.7$		$\gamma = 0.9$	
	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.
$Pr(A_1|S_1A_1S_1)$	.91	.88	.85	.84	.81	.79	.74	.73	.70	.67
$Pr(A_1|S_0A_1S_1)$	.50	.50	.38	.43	.33	.35	.25	.28	.26	.22
$Pr(A_1|S_1A_1S_0)$	.89	.93	.90	.90	.88	.86	.81	.81	.76	.76
$Pr(A_1|S_0A_1S_0)$	.57	.60	.48	.53	.40	.45	.38	.38	.41	.31
$Pr(A_1|S_1A_0S_1)$	.75	.79	.63	.73	.71	.66	.60	.59	.55	.52
$Pr(A_1|S_0A_0S_1)$	.27	.34	.25	.28·	.25	.22	.16	.17	.18	.12
$Pr(A_1|S_1A_0S_0)$	.81	.86	.81	.81	.77	.75	.69	.69	.60	.62
$Pr(A_1|S_0A_0S_0)$	.43	.44	.37	.37	.32	.30	.26	.24	.23	.18
$Pr(A_1|S_1)$	.85	.89	.83	.84	.79	.79	.70	.70	.64	.61
$Pr(A_1|S_0)$	.48	.52	.40	.43	.33	.34	.24	.25	.23	.19
$Pr(A_1)$	.52	.56	.53	.55	.56	.56	.56	.56	.60	.57
$Pr(C)$	.55	.52	.67	.65	.73	.73	.72	.72	.65	.63

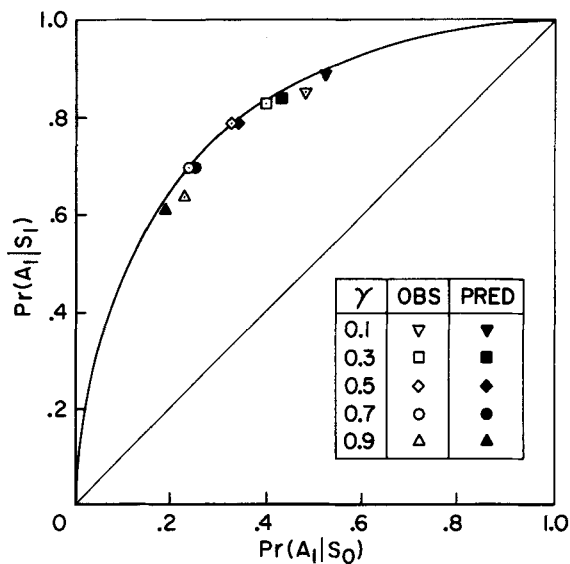| γ	OBS	PRED
0.1	▽	▼
0.3	□	■
0.5	◇	◆
0.7	○	●
0.9	△	▲

Fig. 6. ROC graph of observed and predicted values for hits and false alarms.

ceding response on hits and false alarms is greatly reduced, a result that was reported by Kinchla (1966).

As noted in the introduction, memory factors play a more important role in recognition experiments than in detection experiments. In a recognition task the Haller–Atkinson Model specifies that a comparison is made between the current signal and the signal that occurred on the preceding trial. In a detection task, the signal is compared with a noise background that is always present when the signal occurs; thus the comparison does not necessarily involve memory for previous signals. As we have suggested, differences in performance between the two tasks can be altered by the introduction of information that identifies the signals. Giving the subject information about the presentation schedule in a recognition task allows him to adjust his response bias accordingly. When trial-to-trial feedback is added, O's accuracy on a given trial is enhanced since he then has a correct label for the trace of the preceding stimulus against which he must compare the current one. Thus, the

results of the present study, as well as the other studies discussed here, suggest that memory factors play a much more important role in signal recognition tasks than has been assumed in the literature.

References

Atkinson, R. C., & Kinchla, R. A. A learning model for forced-choice detection experiments. Brit. J. math. statist. Psychol., 1965, 18, 183-206.

Green, D. M., & Swets, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.

Helson, H. Adaptation-level theory. In S. Koch (Ed.), Psychology: A study of a science. New York: McGraw-Hill, Vol. 1, 1959. Pp. 565-621.

Haller, R. W., & Atkinson, R. C. A memory-detection model for signal recognition. Institute for Mathematical Studies in the Social Sciences, Stanford University, Tech. Rep., 110, 1967.

Kinchla, R. A. A comparison of sequential effects in detection and recognition. Experimental Psychology Series, Psychology Department, New York University, Tech. Rep., 1, 1966.

Kinchla, R. A., & Smyzer, F. A diffusion model of perceptual memory. Percept. & Psychophys., 1967, 2, 219-229.

Luce, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology. New York: Wiley, 1963a. Pp. 103-189.

Luce, R. D. A threshold theory for simple detection experiments. Psychol. Rev., 1963b, 70, 61-79.

Parducci, A., & Sandusky, A. Distribution and sequence effects in judgment. J. exp. Psychol., 1965, 69, 450-459.

Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. G. Decision processes in perception. Psychol. Rev., 1961, 68, 301-340.

Tanner, W. P., Jr. Theory of recognition. J. Acoust. Soc. Amer., 1956, 28, 882-888.

Notes

1. The terms observer and subject and the symbol O are used interchangeably throughout this paper.

2. Normal-normal paper (also known as double-probability paper) is graph paper in which the x and y coordinates have been transformed so that normal deviates are linearly spaced.

3. In reference to the assumption of normal trace distributions it may be noted that Kinchla (1967) has developed a random-walk model for a task in which the subject is asked to report whether the second of two tones is the same as or lower in amplitude than the first. The actual trace distributions generated by such a process are binomial and approach normality in the limit.

4. Throughout this paper, a point in the ROC space will be denoted as an ordered pair (x, y) where the first member denotes the value of the abscissa, and the second, the ordinate.

5. Although none of the Os were aware of changes in y, some did report that they thought the difficulty of the task was being varied from day to day.