THE DETECTION OF MARKOVIAN SEQUENCES OF SIGNALS
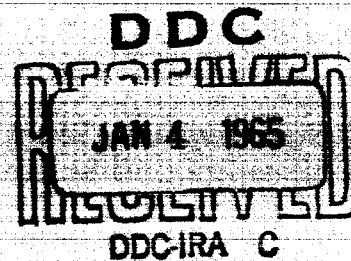
Morton P. Friedman and Edward C. Carterette

Technical Report No. 23

15 July, 1964

UNIVERSITY OF CALIFORNIA, LOS ANGELES

DEPARTMENT OF PSYCHOLOGY

HUMAN COMMUNICATION LABORATORY

THE DETECTION OF MARKOVIAN SEQUENCES OF SIGNALS

Morton P. Friedman and Edward C. Carterette

Technical Report No. 23

15 July, 1964

UNIVERSITY OF CALIFORNIA, LOS ANGELES

DEPARTMENT OF PSYCHOLOGY

HUMAN COMMUNICATION LABORATORY

The Detection of Markovian Sequences of Signals

Morton P. Friedman and Edward C. Carterette

Human Communication Laboratory, Department of Psychology

University of California, Los Angeles

ABSTRACT

2134 0

The influence of constrained stimulus sequences on detection was studied in a two-alternative temporal forced-choice task with feedback. Three observers listened to a weak pure tone embedded in noise whose probabilities of occurrence and repetition in an interval were governed by a first-order Markov process. Each observer listened to examples of each of nine different Markov chains. Results were: (a) A single function relating detections to false alarms fitted individual sets of data well, in agreement with the theory of signal detectability, except (b) that detection was higher for more extreme repetition probabilities; (c) Responses depended strongly on the previous stimulus with (d) the dependence being peculiar to a given chain; (e) Detection probabilities increased during runs of signals in the same interval, yet (f) probability of detection on the first trial of a run in a given interval did not depend on the length of the preceding run in the other interval.

Author 7

The Detection of Markovian Sequences of Signals

It is well known that the judgments of practiced observers in psychophysical experiments depend in an important way on both the previous sequence of stimuli and the observer's response to these stimuli. As long ago as 1860, Fechner[1] pointed out that there were substantial sequential response effects in the data of psychophysical experiments. For the most part, however, researchers interested in psychophysical problems have not concerned themselves with sequential effects.

Recent years have seen a rebirth of interest in the analysis of sequential effects in psychophysical judgments. This interest stems from two sources: first, the development of analytic techniques such as information theory and finite state methods; second, the emergence of theoretical interpretations of psychophysical experiments for which sequential effects are regarded as important aspects of the data.

The use of information measures in the analysis of sequential dependencies is illustrated by McGill's[2] experiment on serial effects in auditory judgments. On each of series of trials, one of four tones of equal intensity but differing in frequency was presented against a background of noise. The observer's task was to identify the tone presented on each trial. The difficulty of the judgment was varied by varying the noise level. Using uncertainty measures, McGill analyzed his data to determine the joint effect of previous responses (called presponse effects) and the stimulus in determining

judgments.   McGill found a marked presponse effect for S/N ratios
lower than 22 db.   This corresponded to judgments at or below the
threshold.   As the S/N ratio increased, the presponse influence
decreased.   The nature of the presponse effect was the tendency
of observers to avoid repeating responses on successive trials.
In one series of trials run by McGill, no tones were ever presented.
This can be regarded as a $-\infty$ S/N ratio.   McGill argued that
there was a continuum of presponse effects, running from the $-\infty$
S/N ratio condition in which observers simply guessed on each
trial (although they were unaware that no tones were presented) to
high S/N ratios in which the judgment was determined solely by
the present stimulus.   In a signal detection task, Speeth and
Mathews[3] obtained results on presponse effects which were es-
sentially comparable to McGill's.   Speeth and Mathews applied an
interesting finite state approach to their data which promises to be
quite useful.   An excellent discussion of research on sequential
response effects using uncertainty measures is contained in a recent
book by Garner.[4]

The second recent source of interest in sequential effects has
been the development of new theories of detection such as those of
Atkinson,[5] LaBerge,[6] Luce,[7] and Restle,[8] some of which regard
sequential response statistics as important aspects of detection res-
ponding.   Like the theory of signal detectability,[9] these new theories
are judgmental theories in the sense that they deal explicitly with
psychophysical and situational variables.   But whereas the main

emphasis in theory of signal detectability has been on psychophysical variables, the major concern of these competing theories has been the influences of situational variables such as feedback and signal probability on detection responding.

These newer theories may be considered as an outgrowth of research on discrimination and probability learning in mathematical learning theory. In the probability learning or guessing experiment (reviewed by Estes[10] and Anderson[11]) the observer's task is simply to predict on each of series of trials which one of a set of events will occur. Following his response, he is told which event actually occurred on that trial. In the simplest sort of probability learning experiment, termed the "non-contingent" case, the sequence of events which occur on each trial is programmed independently of the observer's responses according to some probabilistic schedule. The non-contingent probability learning experiment is strikingly like the $-\infty$ S/N ratio condition (in which no tones were presented) in the McGill[2] experiment discussed earlier. The probability learning experiment may be said to represent an important limiting case of judgment, in which responding is completely determined by situational variables. Sequential statistics have played a major role in both the descriptive and theoretical analysis of the probability learning experiment. Assuming a continuity between the probability learning experiment and the usual psychophysical experiment in which the judgment is jointly determined by situational and psychophysical variables, it might be assumed that the analysis of sequential

statistics will prove of equal importance in understanding psycho-physical experiments. The research in this laboratory on sequential effects in judgment is motivated by these considerations.

In the present experiment, we focused attention on the role of stimulus sequence in determining the sequence of judgments To do this, we employed Markovian schedules of signal presentation, and compared the effects on detection performance of a number of different first-order Markov chain generators of trial sequences in a forced choice task.

METHOD

A conventional two-alternative temporal-forced-choice detection task was employed. On each trial, a gated 1000 c/s sinusoid appeared in one of two temporal observation intervals against a continuous background of wide-band Gaussian noise. The observer's task was to indicate which interval he thought contained the signal.

The temporal intervals were indicated to the observer by a display of pilot lamps. Each trial began with a one second warning interval followed by the two .75 second observation intervals. The observer made his response by pressing one of two pushbutton switches. As soon as the observer responded, one of two feedback lamps came on for one second to indicate the interval which contained the signal on that trial. After a 1.5 second intertrial interval, the next trial began.

On each trial, an electronic switch gated the signal on for 100 ms. (whose rise and fall times were equal to 25 ms.) in the middle

TABLE 1

Transition Matrices for the Markov chains used to
generate stimulus sequences and examples of sequences.

Friedman and Carterette
"Detection of Markovian Sequences of Signals"
Acoustical Society of America, New York City,
May 8, 1964

Chain A   P(S₁) = 0.5

Transition matrix (Trial n → Trial n+1):

	S₁	S₂
S₁	.5	.5
S₂	.5	.5

```
2122122122122121211
2222111221122222211221
1221212122212111121111
1121111111121121111121
212222221121222211111122
2111221212
```

Chain B   P(S₁) = .8

	S₁	S₂
S₁	.8	.2
S₂	.8	.2

```
1122112111211112211111
111111122121111111111
1111122212111112111221
1121111111211121111121
2111221111121212221
121111111111112222112211
1121111211
```

Chain C   P(S₁)= .2

	S₁	S₂
S₁	.2	.8
S₂	.2	.8

```
1222222222222222211222
2222222212222212122222121
22222221222221212122222221
1222222222222222222212222
122222222222222211222222
1222121212222222221222211
2222112222
```

Chain D   P(S₁) = .5

	S₁	S₂
S₁	.8	.2
S₂	.2	.8

```
12222221222211122111111
11111112222222211222222
2222222222222222222112
2222111111111111111111
2221111111221122211111111
1112111111122211222222
2222222221
```

Chain E   P(S₁) = .5

	S₁	S₂
S₁	.2	.8
S₂	.8	.2

```
12221221112222211121211
21122221112221122121122
22221112121211121211221
21212111221121121211212
21212121121212121211212
121212121
```

Chain F   P(S₁) = .71

	S₁	S₂
S₁	.8	.2
S₂	.5	.5

```
1221221211122111211122
11111221111111112111
112111212111112211122
1112111111111111122222
1211111211211111111111
2111111111
```

Chain G   P(S₁) = .38

	S₁	S₂
S₁	.2	.8
S₂	.5	.5

```
2212222222122112212212212
1222222212122222122122121222
212212121212221212121211
121221212122222221121212121
2121221222222221112121211
21112122212
```

Chain H   P(S₁) = .62

	S₁	S₂
S₁	.5	.5
S₂	.8	.2

```
1111112121121211112111
12222212112112211122221
1222112212121211212122
11112211212121212112122
1112111221
```

Chain I   P(S₁) = .29

	S₁	S₂
S₁	.5	.5
S₂	.2	.8

```
12121121111112222121212222
222122222222222112121222
222222222212222211121222222
22111222221122222111122222
21112222222211111122221
1112222211
```

of one of the observation intervals. The noise was not filtered. Observers listened binaurally over PDR-8 earphones in an anechoic chamber. The same signal and noise levels were used for all three observers. The calculated E/No value was about 5.2.

Nine different first-order Markov chains were used to generate the sequence of intervals in which the signal appeared on successive trials. The chains differed in both the overall (apriori) probabilities of occurrence of the signal in the first interval, and in the probabilities of repetition of signal in a given interval on a pair of successive trials. Transition matrices for the nine Markov chains and examples of the sequences generated with them are shown in Table 1. A cell entry shows the probability that the signal

Table I follows

on the present [or $(\underline{n+1})$st] trial is in interval 1 or 2 — the columns — given that the signal was in interval 1 or 2 on the preceding (or $\underline{n}$th) trial — the rows. As an example consider Chain B. The first entry in the first row is .8, the second .2, which means that if the signal was in interval 1 on the last trial, it will be in interval 1 on this trial with probability .8, and in interval 2 on this trial with probability .2. The column entries for the second row are identical with those for the first row. This means that the transitions do not depend upon trials, that is to say, successive trials are independent. Furthermore, the overall probability of a signal in the first interval, indicated by $P(S_1)$, is equal to .8. Chains A, B, and C were all independent trial processes with $P(S_1)$ equal to .5, .8, and .2, respectively.

Chain D generated sequences with $P(S_1) = .5$ but with the probability of a repetition of the signal in the same interval on successive trials being .8. The sample sequences from this generator have longer homogeneous runs of signals in a given interval, compared with those from generator A, even though both A and D have $P(S_1)$ equal to .5. Chain E sequences had $P(S_1)$ of .5 but the probability of the signal interval alternating on successive trials was equal to .8. Thus, the sample sequence for E shows shorter homogeneous runs of signals than A or D with the same apriori probabilities. In Chain F, if the signal occurred in the first interval in trial n, it repeated in that interval only with probability .5. Chain H had the same structure as F with intervals reversed. In Chains G and I, one signal tended to repeat in an interval with probability only .2, while the other repeated with probability .5.

Three paid observers were used in the study, and all had at least 10,000 trials of practice on the various schedules before the actual data were taken. Each daily experimental session consisted of four 164-trial blocks. There was a five minute rest period between blocks in which observers left the experimental chamber. The first four trials of each block were "memory" trials, in which the signal level was increased 10 db. and the signal was presented twice in each observation interval in an 1221 order. Only the last 150 trials of each block were used in the analysis of the data. In a given daily session, only sequences generated by one of the nine chains were used. The nine chains were run through in random

order in successive experimental sessions, then were replicated
in a new random order. For each of the nine chains, eight 160-
trial sequences were constructed, and they were presented to each
observer in a different random order. Additional sequences were
constructed for use in the preliminary practice sessions. The
sequences were generated by a computer program, and only those
sequences were used which a χ^2 first-order Markov chain goodness-
of-fit test could not reject at the .05 level.

The practice sessions which preceded the experiment also
consisted of four 164-trial blocks, but sequences from chains A,
B, C, D, and E were presented in a haphazard order in each session.

Observers were not told anything about the purpose of the
experiment, were not informed about the various Markov chains
used to generate the trial sequences, and were never told when
conditions were shifted.

RESULTS AND DISCUSSION

Before pooling the data over blocks of trials, χ^2 tests for homo-
geneity of proportions were performed for each observer on the pro-
portion of correct responses in each interval for the 8 blocks of
trials for each chain. In only 3 of the 54 tests could the hypothesis
of homogeneity be rejected at the .05 level, and there was no apparent
pattern to the magnitude of the observed χ^2's. It appears that under
the conditions of this experiment, using feedback on each trial and
shifting schedules during the preliminary sessions, observers come
to adjust rather quickly to the context of a given type of sequence
of trials. In all analyses, the data are pooled over blocks and

replications for each chain.

Mean detection rates on the various schedules for the three observers are shown on ROC plots in Figure 1. For illustrative

Figure 1 follows

purposes, we have fitted by eye functions derived from the theory of signal detectability to the data. There are a few things to note about these results.

First, although strong biases were induced by the various schedules, the single functions give a reasonable approximation to the data. Second, the pattern of biases induced by the various schedules are similar for the three observers. For the most part, the position of a given chain on the curve is determined by the over-all apriori probabilities, irrespective of the stimulus dependencies. Thus, chains with high apriori probabilities of the signal being in the first interval are the highest points on the function, and chains with the lowest apriori probabilites of the signal in the first interval are the lowest points on the functions. The third point to be made about the data in Figure 1 is that there is some suggestion of greater sensitivity on schedule D, in which signals tend to repeat and long runs of signals occur in the same interval. Sequences from Chain D yield the largest estimate of $\underline{d}'$ for all observers. This apparent increase in sensitivity may simply be due to a decrease in the observer's uncertainty as to the onset time of the stimulus during a run of trials in which the signal recurred in the same interval. Egan, Greenberg, and Schulman[12] have shown signal uncertainty to
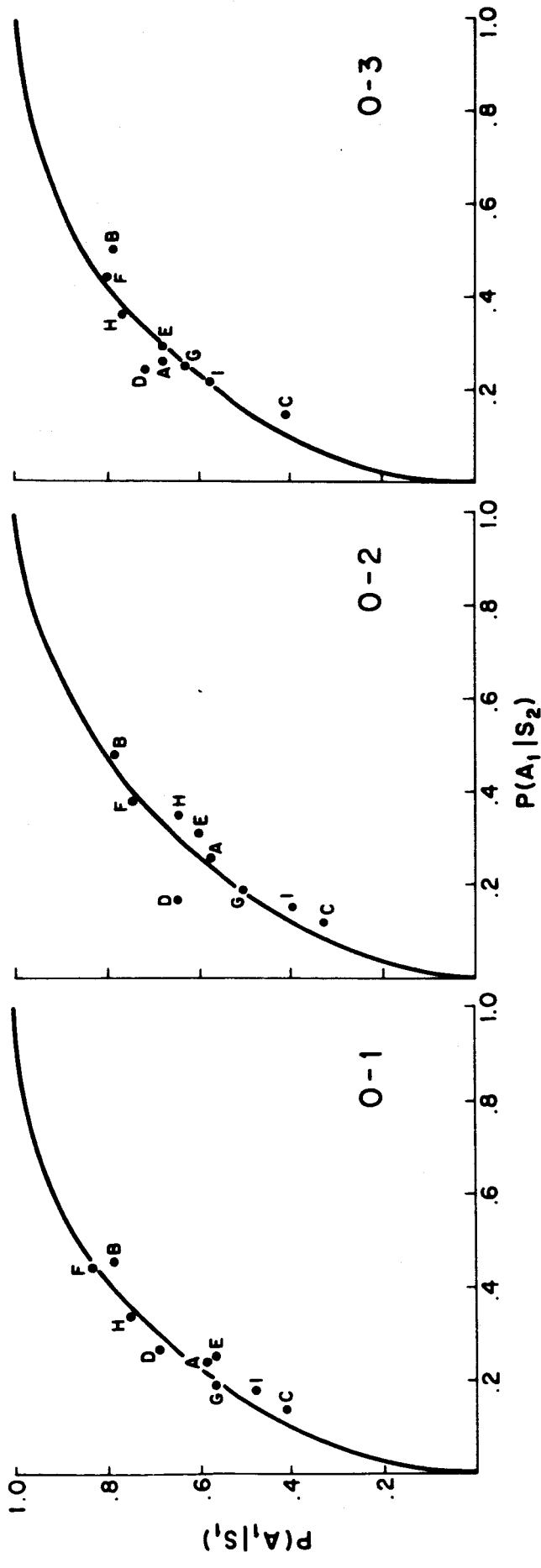
be a potent determiner of detection performance and Green[9] has

discussed theoretical implications of this for the theory of signal

detectability. It is reasonable to assume that, in responding to

sequences generated by Chain D which tended to have long runs

of trials in which the signal recurred in the same interval, ob-

servers' uncertainty as to the onset time of the signal would

decrease during a run and detection would be enhanced.

Of greater interest to the study of stimulus dependencies is

responding during runs of trials on which the signal recurs in the

Figure 2 follows

same interval. Figure 2 shows the proportion of correct responses

to the various chains during such runs. Since the data from the

three observers were similar, we have averaged over the three

observers. And, whenever possible, we have combined data with-

in and between symetrical chains, such as B and C, and G and I.

Thus, we have combined frequencies of correct responding for both

intervals in Chain A. Likewise, the points labeled $B-S_1$ are from

the combined tabulations of runs of signals in the first interval for

Chain B and runs of signals in the second interval for Chain C. Data

are presented only for those points for which there are at least 125

observations. Because the number of possible observations of runs

depends on the structure of the chain, the number of points for each

chain differ. Points for runs of length k include observations for

the kth trial of runs longer than k. The gist of Figure 2 seems to be

as follows: For schedules in which successive trials are independent,
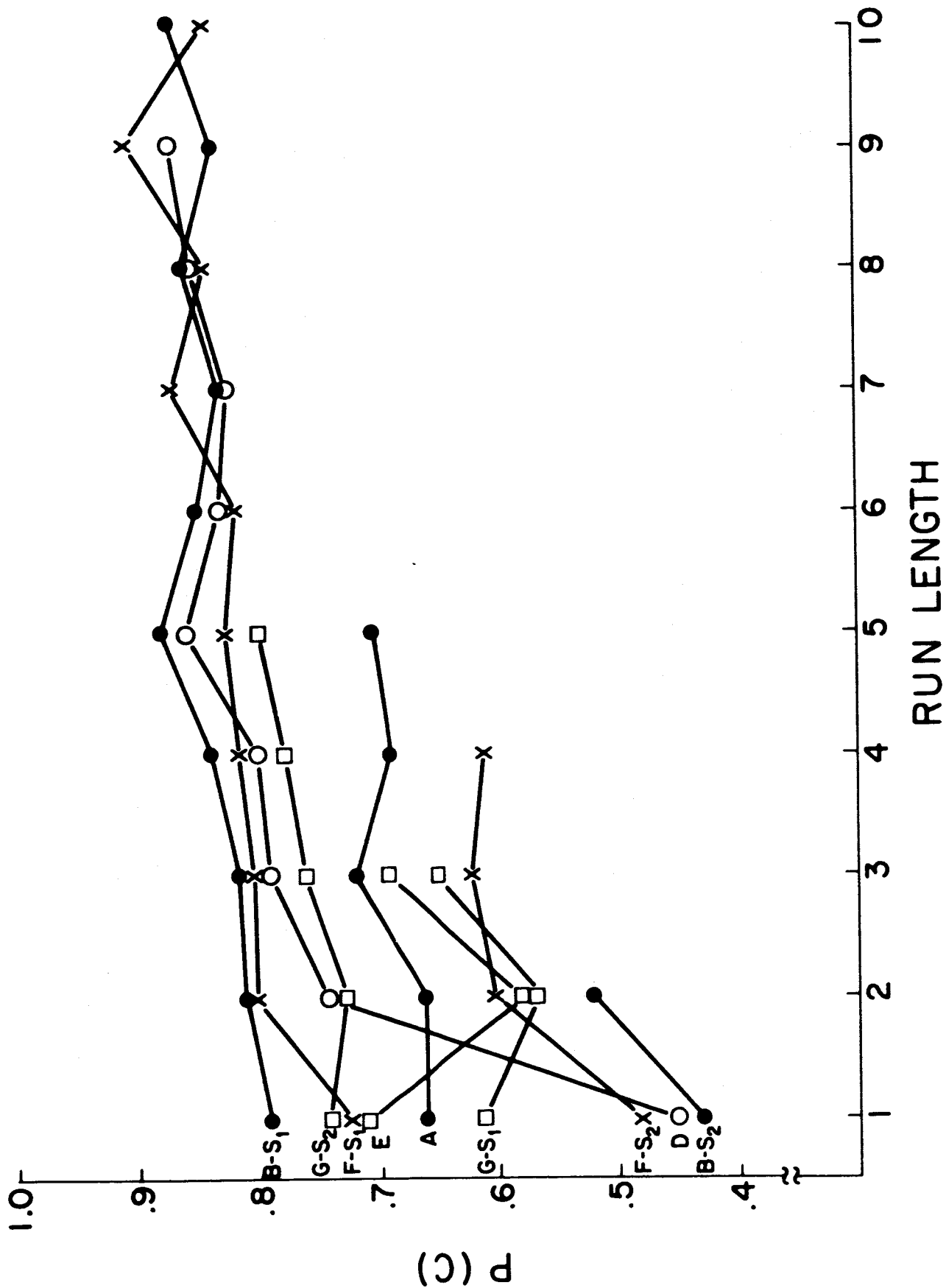
there is a slight increase in correct responding during a run. Thus

for Chain A, with apriori probabilities of .5, and independent trials,

the mean detection rate for the first trial of a run is .66, and in-

creases to .70 on the fifth trial of a run of signals in the same inter-

val. On Chain B, with an apriori probability of signal occurrence

in the first interval equal to .8, the curve labeled $B-S_1$ shows the

detection rate during runs of trials in which the signal recurred

in the first interval. On the first trial of a run, the mean detection

rate was .79, and on the 11th trial of the run, the mean detection

rate was .84.

For chains in which there were dependencies in the stimulus

sequences, detection responding mirrored these dependencies

markedly. For example, consider Chain D, in which $P(S_1)$ was

.5, but signals recurred in the same interval with probability .8

on successive trials. Figure 2 shows that on the first trial of a

run, the detection rate was .45, but that it increased rapidly to

about .85. Responding to signals in the first interval for Chain F,

(the points labeled $F-S_1$) show the same effect to a lesser extent.

For Chain E, where the signal tended to alternate in the two inter-

vals on successive trials, the detection rate on the first trial of a

run was .70, and dropped to .58 on the second trial of a run, again

mirroring the dependencies in the signal sequence. The same effect

occurs in responding to runs of signals in the first interval for

Chain G (the points labeled $G-S_1$). These results make it clear that

signal events on preceding trials are important determiners of trial-

to-trial response biases.

Another statistic of interest in gauging the influence of stimulus dependencies on responding is the detection rate on those trials which break a run of signals in a given interval. Figure 3

Figure 3 follows

shows the detection rate on the first trial of a run of signals in a given interval as a function of the length of the preceding run of signals in the other interval. The points are labeled and the data were tabulated in the same way as the run data in Figure 2. The result here is quite clear: There is little or no effect of the length of the preceding run on detection responding, excepting a slight effect for chains which tend to alternate. These data indicate that when the signal switches intervals, the immediately preceding trial is the important determiner of bias.

We have discussed only the effects of stimulus contingencies on responding. The observed dependence of responses on preceding responses is quite small compared to the effects of preceding stimuli. This is probably because of our use of feedback on every trial. Research in this laboratory [13] and elsewhere [14] indicates that when feedback is employed, it is the important determiner of sequential responding; when there is no feedback, then previous responses appear to be important. The nature of the dependence when there is no feedback is a tendency to alternate responses, as in McGill's [2] study, or to repeat responses, as in a recent study by Parducci and Sandusky. [14] Much clarification is needed of the variables
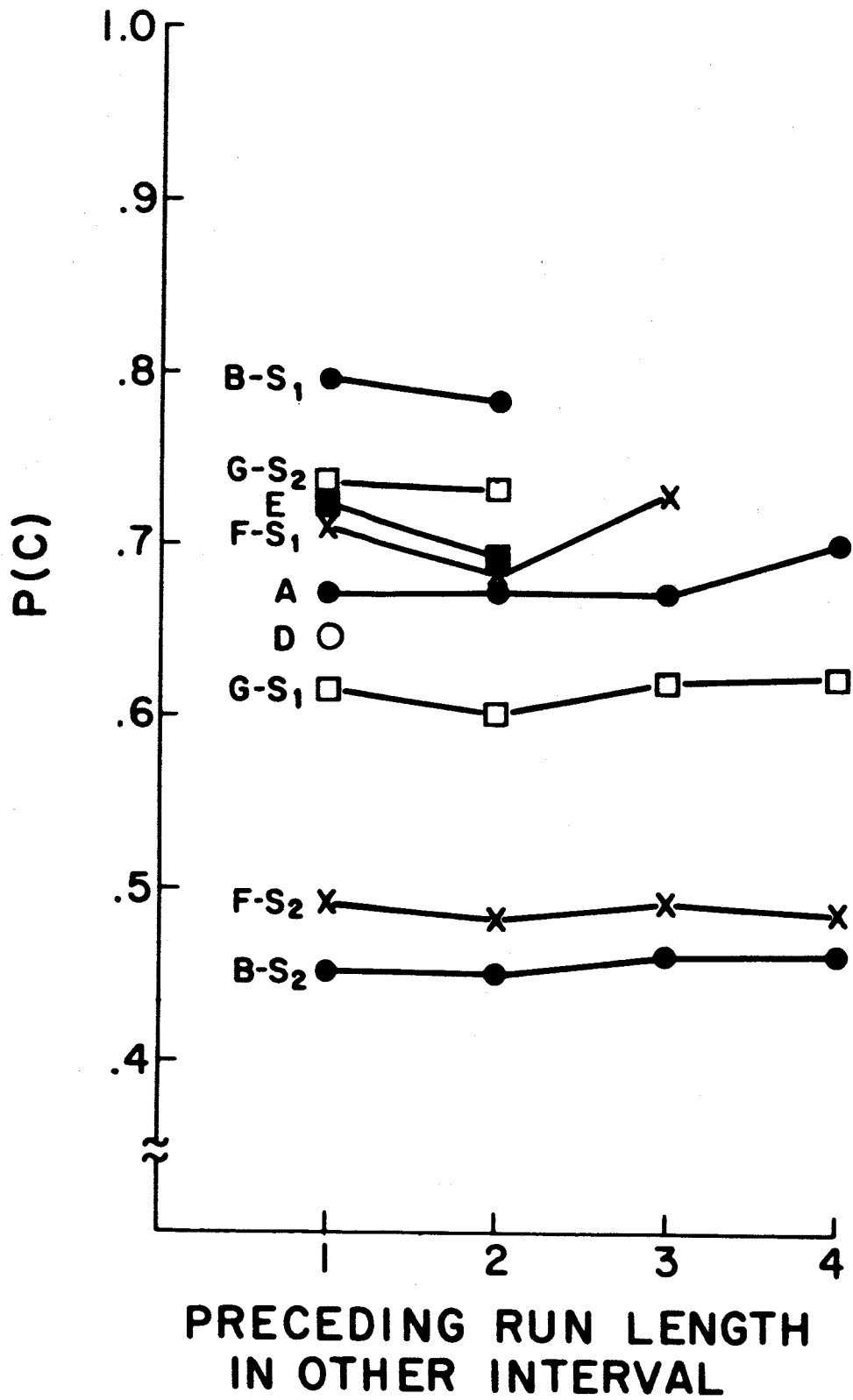
which control responding in the absence of feedback.

The theory of signal detectability has not yet been extended to deal with sequential response effects, but Atkinson's[5] detection model has been developed in enough detail to allow exact quantitative predictions of sequential statistics. It is essentially a threshold theory of the sort proposed by Blackwell.[15] Atkinson formulated his theory in the language of stimulus sampling theory, but we shall discuss it in more general terms here. As applied to the yes-no experiment with feedback, it is assumed that on every trial, the observer either correctly detects the presence or absence of the signal, or he is uncertain whether the signal was presented. If he is uncertain, he guesses. According to Atkinson's model, sequential effects are due to the fact that the guessing probabilities change in a trial-to-trial fashion as a function of the feedback. If a signal is presented on a given trial, then the probability of guessing yes associated with the uncertain state increases; if the feedback indicates that no signal was presented, then the probability of guessing yes decreases. Atkinson's theory has had success in accounting for sequential effects as well as mean response proportions in both feedback and non-feedback visual[16] and auditory[17] detection experiments.

However, Atkinson's theory is not consistent with our results. The main difficulty seems to be with predictions concerning responding during homogeneous runs of signals in the same interval. Atkinson's prediction is, that during homogeneous runs, detection

rate should continue to increase, and assuming that sensitivity is constant, responding on all schedules should approach the same asymptote. This is clearly not the case in Figure 2. Similarly, Atkinson's theory predicts that the probability of a correct detection on the first trial of a run of signals in an interval should be a decreasing function of the preceding run length in the other interval. The data presented in Figure 3 indicates that responding on the first trial of a run is independent of the length of the preceding run in the other interval.

We do not regard these negative results in themselves as particularly damaging to Atkinson's model. For one thing, it is clear that if Atkinson's theory about changes in guessing probability when in the uncertain state were modified to put greater weight on the outcome of the preceding trial, then predictions would be more in line with our results. Also, sequential analyses of data from other experiments in this laboratory[13,17] have yielded results which are at least in qualitative agreement with Atkinson's model. The main difference between this study and the others seems to be that in the present study, observers were exposed to a much wider variety of probabilistic schedules. Our impression is that under the conditions of this experiment, observers quickly adjusted to the context of a new schedule, and the observed sequential effects are determined by a few short term strategies that the observers are using. The quantification of a theory incorporating these notions is difficult, but the finite state methods used by Speeth and Mathews[3] seem to offer a reasonable approach.

References

1. G. T. Fechner, _Elemente der psychophysik,_ (Breitkopf and Hartel, Leipzig, 1860)

2. W. McGill, "Serial Effects in Auditory Threshold Judgments, " J. exp. Psychol. 53: 297-303 (1957).

3. S. D. Speeth and M. V. Mathews, "Sequential Effects in the Signal-Detection Situation, " J. Acoust. Soc. Amer. 33: 1046-1054 (1961).

4. W. R. Garner, _Uncertainty and Structure as Psychological Concepts._ (Wiley, New York, 1962)

5. R. C. Atkinson, "A Variable Sensitivity Theory of Signal Detection, " Psychol. Rev. 70: 91-106 (1963)

6. D. LaBerge, "A Recruitment Theory of Simple Behavior, " Psychometrika. 27: 375-396 (1962)

7. R. D. Luce, "A Threshold Theory for Simple Detection Experiments, " Psychol. Rev. 70: 61-79 (1963)

8. F. Restle, _Psychology of Judgment and Choice_ (Wiley, New York, 1961).

9. D. M. Green, "Psychoacoustics and Detection Theory, " J. Acoust. Soc. Amer. 32: 1189-1203 (1960)

10. W. K. Estes, "Probability Learning" in _Categories of Human Learning,_ edited by A. Melton, (Academic Press, New York, 1964.)

11. N. H. Anderson, "An Evaluation of Stimulus Sampling Theory" in _Categories of Human Learning,_ edited by A. Melton, Academic Press, New York, 1964)

12. J. P. Egan, G. Z. Greenberg, and A. Schulman, "Interval of Time Uncertainty in Auditory Detection, " J. Acoust. Soc. Amer. 33: 771-778 (1961).

13. E. C. Carterette and R. C. Atkinson, "Sequential Phenomena in Signal Detection, " J. Acoust. Soc. Amer. 34: 172 A (1962).

14. A. Parducci and A. Sandusky, "Distribution and Sequence Effects in Judgment. " J. exp. Psychol. in press (1964).

15. H. R. Blackwell, "Psychophysical Thresholds: Experimental Studies of Methods of Measurement. " Bull. Engng. Res. Inst. Ann Arbor, No. 36 (1953).

16. R. A. Kinchla "A Learning Factor in Visual Discrimination, " in Studies in Mathematical Psychology, edited by R. C. Atkinson (Stanford Univ. Press, Stanford, 1963).

17. R. C. Atkinson, E. C. Carterette and R. A. Kinchla, "Sequential Phenomena in Psychophysical Judgments: A Theoretical Analysis" IRE Transactions on Information Theory, IT-8, 5155-62 (1962).

Figure Captions

Figure 1.   ROC plots for the three observers.

Proportion of correct responses in Interval 1 [$P(A_1 S_1)$], are

plotted against the proportion of incorrect responses in

Interval 2 [$P(A_1 S_2)$].   Points labeled with letters are from

different Markov chain generators.   The smooth curves are

derived from the theory of signal detectability.   The $\underline{d}'$ values

for the three observers were 1.08 for $\underline{0}$-1 and $\underline{0}$-3, and .94

for $\underline{0}$-2.

Figure 2.

Proportion of correct responses [$P(C)$] averaged over the

three observers as a function of run length for the various

Markov chain generators.

Figure 3.

Proportion of correct responses [$P(C)$] on the first trial of

a run as a function of the preceding run length in the other

interval shown for the various Markov chain generators.

Data are averaged over the three observers.