Perception & Psychophysics
1974, Vol. 15, No. 2, 391-398

Effect of practice on the identification of auditory sequences*

ULRIC NEISSER and WILLIAM HIRST
Cornell University, Ithaca, New York 14850

Using a variant of the up-and-down method to establish duration thresholds for correct report of the order of sound sequences, six experiments were run with a practiced crew of three Ss. All showed striking improvement with practice, but substantial individual differences were observed. Two Ss found spaced sequences harder than simple cyclic ones and single presentations hardest of all, had increased difficulty when the component frequencies were brought closer together, and were disturbed by irregular timing. These variables made no difference to the third and most sensitive S, whose threshold reached 23 msec per component. All Ss displayed poorer performance if noise, additional signals, or a distrating activity occurred between the stimulus sequence and the response.

In the experiments of Warren, Obusek, Farmer, and Warren (1969), naive listeners were presented with a cyclic sequence of four meaningless sounds, and asked to report the order in which they occurred. This proved to be impossible at durations of 200 msec per component or less. The reasons for the difficulty are not entirely clear. Warren et al suggested that the nature of the sounds themselves was responsible, a view which has been amplified by others. For example, Bregman and his associates (Bregman & Campbell, 1971; Bregman, 1972) have suggested that sounds which differ widely in their frequency composition cannot be ordered with respect to one another because they fall in different "streams." However, it is known that performance in Warren's task improves with practice; both D. A. Norman (personal communication) and Nickerson and Freeman[1] have found enormous improvements in single Ss run for prolonged periods. Our own pilot studies (Neisser, 1972) showed similar, though smaller, improvements. Thus, performance is not limited only by the nature of the signals, but by the skills which the listener brings to the task. What is the nature of these skills?

To report the order of a sequence of stimuli, a S must produce a correctly ordered series of responses. Since he typically produces his report only after the stimulus has terminated, it must be based on a memory representation of the stimulus. A representation which itself consists of an ordered sequence of parts, corresponding to the response sequence it will mediate, may conveniently be called a *string*. One strategy for a listener would be to construct a string as he listens–to say something like "hiss, scratch, high, low" in synchrony with the corresponding sounds. Although this is by no means the only way temporal order can be perceived, it is almost the only possibility for a naive S on his first trial (Warren, 1972a, has expressed a similar view). Since such a strategy would require a series of choice reactions, each made in 200 msec or less, we need not be surprised that it usually fails.

To succeed with more rapid sequences, a S must use another type of representation. An indication that this is possible comes from studies using a "same/different" method. Wilcox, Neisser, and Roberts (1972) found that people can judge whether or not two temporal sequences were identical at substantially shorter durations. Their Ss were correct about 84% of the time in comparing two one-shot (i.e., not cyclic) sequences separated by a 100-msec silent interval. Warren (1972a, 1973b) has subsequently reported comparable results in a "same/different" paradigm. With such a procedure, the S need never form a *string* at all, in the sense defined above. It is enough if his representation of the first sequence (which he maintains until the second appears) preserves just enough information to distinguish the two. We may call such partial representations *analogs*, bearing in mind that they need not be "templates" (Warren, 1973a) of the entire acoustic pattern.

There must be many possible analogs of every stimulus sequence, many different sorts of information which can be picked up and stored by the listener. Moreover, analogs may be used not only for same/different judgments, but for explicit reports of order as well. For example, a listener who has succeeded in forming a string during the stimulus presentation itself might reason, "The low tone was surely first . . . there was a high-hiss transition . . . the scratch certainly wasn't *before* the hiss . . . so it must have been low, hiss, high, scratch." The overt response string may really be a complex post hoc construction based on minimal cues available in an analog form. It seems plausible to suppose that different listeners might construct different types of analogs in such a task, and also that the nature of the analogs might change with practice.

*This research was supported by Grant MH20082 from the National Institute of Mental Health and by Grant GB-28712 from the National Science Foundation. Requests for reprints should be sent to Ulric Neisser, Department of Psychology, Cornell University, Ithaca, New York 14850. We wish to thank Gordon W. Wilcox for his help and advice in all phases of the study, William Hemsath for his invaluable technical assistance, and Kurt Lichtmann for his help both in programming and in running subjects for some of the sessions.

For these reasons, we conducted an empirical study of the effects of extended practice in Warren's task, using three Ss, with a method devised to permit continuous tracking of a "threshold" for temporal order. A number of variables which are important with naive Ss were studied to determine whether their effects would survive prolonged training and whether they would continue to affect each S in the same way. We were specifically concerned with the following questions:

(1) Warren et al (1969) found that their Ss could more easily identify the order of a four-component sequence presented only once (a "one-shot") than the order of a cyclically repeating sequence. They attributed this to the greater salience of the first and last elements in the one-shot case. There is some reason to wonder whether this result would hold for practiced Ss, however, since the critical information is available many times in the cyclic case and only once in a one-shot.

(2) The "stream" hypothesis of Bregman and Campbell (1971) suggests that sequences will become more difficult to identify as their components are given more widely separated frequencies. However, a recent negative result (Nickerson & Freeman[1]) suggests that this may not be universally true.

(3) Warren (1972) has shown that naive Ss can identify a cyclic sequence more readily if silent pauses are interpolated between repetitions of the basic cycle. We expected this result to hold for practiced listeners as well, but it did not appear in our pilot experiments (Neisser, 1972).

(4) The pilot experiments had suggested that accurate order judgments of a cyclic sequence are possible even when silent pauses are introduced irregularly into a cyclic sequence to disturb its rhythm; this counterintuitive result seemed to require further study.

(5) Casual observation indicates that Ss often hesitate for several seconds after the end of a sequence before responding. By our hypothesis, this is the period when the stored analog is being used for the construction of a suitable response string. If this is true, it should be possible to impair performance by interfering with the memory or the constructive process or both. Thus, a preliminary exploration of various types of "backward masks" seemed warranted.

(6) Finally, it seemed desirable to order the 24 possible one-shot sequences in terms of the ease with which they could be identified. Changes in the ordering of difficulty with practice or among Ss might provide clues to the types of analogs being used.

METHOD

The four signals used throughout most of the study were a *high* tone (3,400 Hz), a *low* tone (500 Hz), a *hiss* (broad-band white noise), and a *scratch*, made by passing white noise through a Schmitt trigger to leave only the peaks, which were then amplified. Their amplitudes were adjusted to yield approximately equal loudness at the S's binaural headphones, where the SPL was approximately 80 dB. Each signal was fed to a separate electronic gate; the opening and closing of the gates was controlled by a PDP-8/E computer. The gates provided rise and fall times of about 10 msec; as one signal decayed, its successor was already rising.

The S sat in an experimental room adjacent to that in which the computer and other apparatus were housed. The signals were presented over headphones; lights on a display panel indicated when a signal would occur, when a response was permitted, and whether or not it was correct. Response was made on a typewriter keyboard; the numerals 1, 2, 3, and 4 corresponded to *hiss, scratch, low*, and *high*. After a given sequence had been presented, the S pressed these keys in the order in which he thought the component signals had occurred. He could change his mind; keypresses were accepted until he indicated that he was finished by striking the *return* button. The sequence of the last four keys pressed was taken to be his response.

The durations of the stimuli were under the control of a computer program. On each trial, the program opened the signal gates one after the other, each for a specified duration, T, in an order which was internally determined as a random permutation of the digits 1-4. In a one-shot mode, it then stopped and awaited the S's response; in a cyclic mode, it repeated the same sequence of gates 10 times before stopping. The durations, the stimulus sequence, and the response were then typed out at the control console and also stored internally. If the response was correct, the program decreased the stimulus durations by a fixed decrement, D, i.e., set $T = (T - D)$ for the next trial; if incorrect, the durations were increased to $(T + D)$. A new permutation was then selected, and the next trial presented. The permutations were selected without replacement from the $4! = 24$ possible orderings, so that in 24 trials every sequence occurred exactly once. A typical run consisted of 48 trials, though sometimes 54 were used, with the first 6 being treated as practice and not analyzed. At the beginning of each run, the E initialized the program by indicating the desired condition (cyclic, one-shot, and others to be described below), prescribing an initial duration, T (usually the S's threshold value from the preceding session), and specifying the increment, D (10 msec in the early weeks of the experiment, 5 msec beginning on Day 19). At the end of a run, the S rested briefly while the stored data were transferred to magnetic tape for later analysis; then another run began. This procedure insured that the series of durations actually presented during a run remained near the value at which 50% could be identified correctly. The mean of durations during a run is thus an estimate of the S's threshold.

Most sessions consisted of 4-6 such runs, depending on the conditions being tested, and lasted from 60 to 80 min. Each of the three Ss was run in one session a day, 4 days a week, for about 7 weeks. The Ss, paid for their services, regarded the experiment as a part-time job. Two of them (one male, D.B., and one female, S.D.) were Cornell students; the third (male, E.S.) was a high school graduate and amateur musician. Tests with a Grason-Stadler audiometer indicated that all had normal hearing in both ears. On the Wing Standardized Test of Musical Aptitude, E.S. and D.B. scored "A," while S.D. scored "C."

Six distinguishable experiments were conducted during the 27 days of the study.

EXPERIMENT I (DAYS 1-7, 21)

This principal experiment compared thresholds in three conditions: (a) *one-shots*; (b) *cyclic* series, consisting of 10 full cycles; (c) *spaced* series, in which a silent interval, also of Duration T, occurred between successive cycles. (Thus, the duration of a cycle was 4T in *cyclic* and 5T in *spaced*.) Each condition was presented in two different runs on each day, using an ABCCBA order.

Figure 1 shows the very substantial effects of practice in these three conditions for S D.B. On the first day, he could identify *one-shots* only if each component lasted nearly 200 msec; by Day 7, his threshold was down to 86 msec; and on Day 21, it was down to 56 msec. The corresponding drop for *cyclic* and *spaced* series was from about 100 msec to under 50 msec. (Even the figures for Day 1, of course, represent a good deal of practice; we have no way of estimating the threshold on the initial trial.) In every stage of practice, however, *one-shots* remained more difficult than either *cyclic* or *spaced* sequences. Contrary to expectation, the silent intervals in the *spaced* condition did not make it easier; *spaced* thresholds were generally higher than the *cyclic* ones in the early part of the study, and never consistently lower.

Figure 2 shows a similar pattern of results for S S.D. Starting with even longer threshold durations (over a third of a second for *one-shots*), she eventually attained approximately the same levels of performance as D.B. For her, too, *one-shots* were the most difficult condition. More consistently than D.B., she found that *spaced* sequences were harder than *cyclic* ones.

S E.S., however, shows quite a different pattern (Fig. 3). Though his performance on the first day wascomparable to that of the other Ss, he soon outstripped them. By Day 6, he was below 50 msec in every condition; by Day 21, he was down to about 30 msec. And, although he, too, began by finding *one-shots* more difficult than the other conditions, his curves soon converge; all conditions became equally easy for him after the first few days.

None of the Ss allowed the stimulus durations to move very far above or below their threshold values. On Day 21, with a step size D of 5 msec, the standard deviations of the duration (averaged across six conditions) were 10.3, 13.0, and 7.2 msec for D.B., S.D., and E.S., respectively.

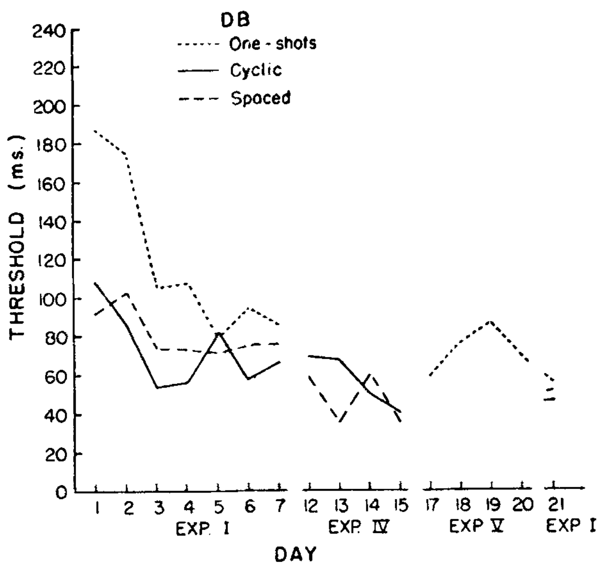

Fig. 1. Mean T (see text) each day in the three conditions of Experiment I and in other experiments where the same conditions were presented. S D.B.

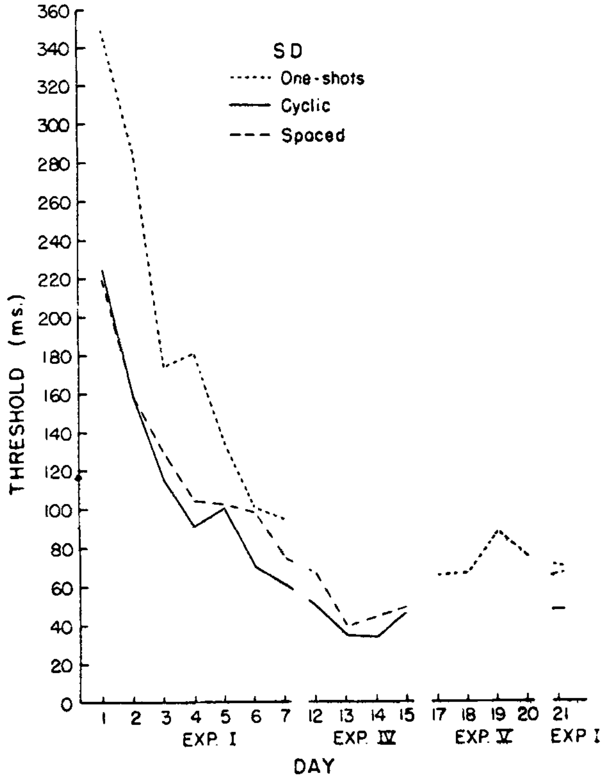

Fig. 2. Mean T (see text) each day in the three conditions of Experiment I and in other experiments where the same conditions were presented. S S.D.

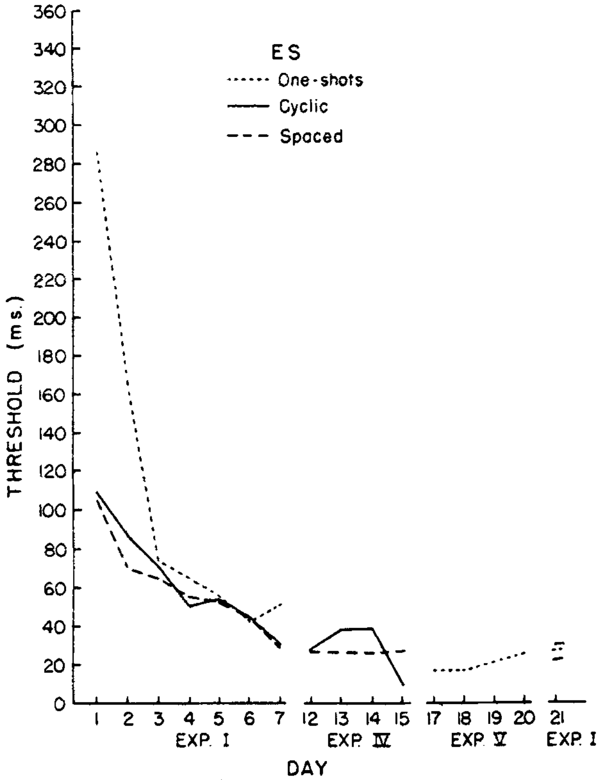

Fig. 3. Mean T (see text) each day in the three conditions of Experiment I and in other experiments where the same conditions were presented. S E.S.

These results raise a number of questions. Why are one-shots harder than either of the other two conditions? Why does spacing make cyclic series more difficult when it matters at all? What is the nature of the practice effect? Why are E.S.'s data so different from those of the other two Ss?

The Ss' own introspective comments did not shed much light on these issues, but a number of hypotheses suggest themselves. The greater difficulty of one-shots is not surprising, even though Warren et al (1969) had made the opposite finding with naive Ss; in a cyclic series, the listener has 10 times as many opportunities to pick up the relevant information. The adverse effect of spacing was more surprising; even our pilot studies had not prepared us for it. Hindsight, however, suggested a possible explanation. A cyclic series, ABCDABCDABCD . . ., can actually be organized in four different ways, each based on a different permutation: ABCD, BCDA, CDAB, or DABC. In general, these four will not be equally easy to identify: both Ss' introspections and casual inspection of the data suggested that some one-shots were more difficult than others. In the *cyclic* condition, a sophisticated listener might be able to choose which of the four to use as the basis of his own organization and response (though naive Ss like those of Warren, 1972b, probably would not possess this skill). In the *spaced* condition, however, he would be restricted to a single one of these permutations, often not the easiest.

These considerations indicated the need to determine whether the various permutations, taken as one-shots, were indeed of unequal difficulty. This determination was made in Experiment II. It was also hoped that the patterns of difficulty might shed some light on the individual differences among our Ss. Moreover, data on the relative difficulty of the permutations would permit an explicit test of the foregoing hypothesis about *spaced* series: their relative difficulty should be increased if they are constructed only of hard one-shots and decreased if they are made up of easy ones. Such a test is reported below as Experiment VI.

EXPERIMENT II (DAYS 8, 22)

This experiment studied the individual difficulty of the 24 possible one-shot sequences. The up-and-down method of the other experiments was not employed. On both days, the S was simply presented with six runs of 48 one-shots each (two random passes through the 4! sequences), all the same duration T. This duration was chosen, on the basis of his performance on the previous day, to be one at which approximately 50% correct responses would result.

The number of correct responses (out of 12 trials) for each permutation is given separately by S and by day in Table 1. It is evident that some sequences are more difficult than others. The standard deviations in each column—representing the variability of the number correct across the 24 permutations—are all higher than the chance value for binomial samples of Size 12 (1.73 if $p = \frac{1}{2}$ and less otherwise).

However, the six columns of Table 1 (three Ss each run twice) do not give the *same* ordering of difficulty. Only E.S. (the best S) produced an ordering that was stable across both days (rho = 0.80). Both of his orderings were *negatively* correlated with those of the other two Ss, whose agreement with each other was significant on Day 8 (rho = 0.52) but not on Day 22 (rho = 0.19).

Even in the case of E.S., it is not easy to determine what principles govern the ordering, though there was apparently a tendency for one-shots beginning with the low tone to be difficult. He himself could not verbalize any principle which might distinguish hard sequences from easy ones; indeed, he could not even accurately report which sequences fell in either category.

EXPERIMENT III (DAYS 9-11, 23-25)

In this experiment, we tried to compare the difficulty of one-shots made up of various types of components. Four conditions were used: *standard*, in which each one-shot consisted of the hiss, scratch, 500-Hz, and 3,400-Hz signals used in the rest of the study; *pink*, in which a "pink noise" weighted toward lower frequencies was substituted for the "hiss"; 500/590, in which the original "hiss" was used, but the 3,400-Hz tone was replaced with one of 590 Hz, judged to be more nearly in the same "pitch stream" in Bregman's sense; 500/510, in which it was replaced with one of 510 Hz, discriminable from the "low" tone but very similar to it. Days 9, 10, and 23 were devoted to an attempt to determine which of the two noises could be more easily discriminated from the "scratch," and which of the three higher tones could be more easily discriminated from the low one. (It was thought that sequences containing more easily discriminable pairs would be relatively easier to order.) However, this attempt was unsuccessful; we were unable to make the five discriminations difficult enough to distinguish among them. Data will be reported here only from Days 11 and 24-25. On these days, two runs of each of the four types of one-shots described above were presented in ABCDDCBA order.

The results appear in Table 2. It is apparent that the 500/590 sequences are *not* necessarily easier than the *standard* (500/3400) ones. For D.B., they were actually harder on 2 days out of 3, and the *standard* threshold for the remaining day (Day 11) is probably invalid: it is far higher than the threshold for the same condition 4 days earlier in Experiment I (see Fig. 1). For S.D., the 500/590 sequences were harder on all 3 days. Only for E.S. were they easier, and the differences are slight. As for 500/510, it was much harder than 500/590 for both D.B. and S.D., though E.S. is once more the exceptional case.

Table 1
Results of Experiment II

	S D.B.		S S.D.		S E.S.	
Permutation	Day 8 70 msec	Day 22 50 msec	Day 8 85 msec	Day 22 60 msec	Day 8 35 msec	Day 22 25 msec
HSUD	10	8	8	8	7	9
HSDU	9	7	10	7	7	11
HUSD	4	5	6	3	5	8
HUDS	8	11	3	4	8	9
HDSU	6	4	5	8	9	10
HDUS	4	10	5	7	4	5
SHUD	3	10*	4	4	8	11
SHDU	3	9*	4	3	10	11
SUHD	4	10	3	4	10	8
SUDH	4	6	2	6	4	6
SDHU	6	12	7	5	4	7
SDUH	3	4	4	6	6	9
UHSD	9	9	6	11	8	9
UHDS	7	0	7	4	7	11
USHD	7	7	3	5	8	9
USDH	5	3	7	4	2	6
UDHS	11	11	8	5	5	3
UDSH	7	8*	8	7	4	2
DHSU	6	6	7	9	4	3
DHUS	8	3	8	0	0	3
DSHU	6	6	2	5	5	7
DSUH	6	2	10	3	0	3
DUHS	6	7	8	9	4	3
DUSH	7	12	7	7	4	3
Mean	6.2	7.2	5.9	5.6	5.5	6.9
s	2.2	3.4	2.4	2.5	2.8	3.1
rho	0.12		0.14		0.80	

Note–Each entry is the number of one shots correctly identified out of 12 presented (11 presented in entries marked ""). Components: H, hiss; S, scratch; U, high tone; D, low tone. The last row indicates the reliability of Ss' ordering of difficulty with Spearman rank correlations; only E.S.'s is significantly greater than zero ($p < .01$).*

The greater difficulty of 500/590 for D.B. and S.D. may have been due to its lesser familiarity. Certainly, sequences containing the unfamiliar *pink* noise were harder than the *standard* sequences for all Ss. It seems that great frequency separations do not automatically correspond to greater difficulty. Bregman has pointed out to us, however (personal communication), that our use of one-shots rather than cyclic sequences in this experiment may be responsible for our negative result. Segregated "streams" may require a number of cycles to become established.

Table 2
Results of Experiment III

Condition	Day 11	Day 24	Day 25
		S D.B.	
Pink	91	62	50
Standard	113	39	40
500/590	88	64	57
500/510	121	82	82
		S S.D.	
Pink	88	68	75
Standard	70	70	62
500/590	98	76	69
500/510	96	152	98
		S E.S.	
Pink	52	36	20
Standard	42	24	17
500/590	32	20	16
500/510	37	20	17

Note–Each entry is the mean threshold (milliseconds) from two runs of 48 trials each.

EXPERIMENT IV (DAYS 12-15)

This experiment studied the effect of introducing temporal irregularities into cyclic sequences. In the standard *cyclic* condition, a permutation of the four basic components was cycled 10 times uninterruptedly, each component having Duration T. In the *spaced* condition, as in Experiment I, a silent interval of duration T was interpolated at the end of each cycle. In *irregularly spaced*, a silent interval was introduced into each cycle at a randomly determined point (after the first, second, third, or fourth component with equal probability) and thus occurred at randomly different places in the 10 successive repetitions; the result was an oddly irregular rhythm. In a *variable duration* condition, the value T (initialized by the E and increased or decreased according to the successive responses) was made the mean of a random variable t, which could with equal probability assume any value in the range of 0.5T

Table 3
Results of Experiment IV

Condition	S D.G. Day 12	13	14	15	S S.D. Day 12	13	14	15	S E.S. Day 12	13	14	15
Irregular	76	65	66	50	84	76	44	32	55	33	26	14
Variable	89	79	60	56	91	61	46	55	42	29	31	24
Mean of Irregular and Variable	83	72	63	53	88	69	45	44	49	31	29	19
Mean of Cyclic and Spaced	64	50	55	38	59	38	39	47	28	38	32	22
Cyclic	69	64*	50*	40*	51	36*	34*	45*	28*	38*	39*	16*
Spaced	59	36*	60*	35*	67	40*	44*	48*	27*	–	25*	27*

Note–Entries are thresholds (milliseconds) on two runs of 48 unless marked "" (based on one run) or unless marked "mean."*

< t < 1.5T. A new value of t was chosen for each cycle. Thus, the first cycle through the four components was at one duration, the second at a different duration, and so on through the 10 cycles of a trial. Each of the four conditions was presented twice in an ABCDDCBA order on Day 12; this took so long that on the three remaining days *cyclic* and *spaced* were presented only once and the others twice.

The results appear in Table 3. Both types of temporal irregularity caused difficulty when they were first presented. Both benefited from practice, but for D.B. they remained harder than *cyclic* or *spaced* throughout the 4-day experiment; this was roughly the case for S.D. as well. The two types of irregular sequences did not differ consistently from one another for those two Ss. For E.S., none of the four conditions posed any noticeable difficulty at all after the first day. He is therefore the only S whose behavior was consistent with the pilot data reported earlier by Neisser (1972).

EXPERIMENT V (DAYS 15-20)

This equipment studied the effect of various kinds of postsequence masking on the identification of one-shots. There were five conditions: *standard*, comparable to the one-shot condition of several other experiments; *wait*, in which the S was not permitted to respond until 2 sec after the end of the one-shot sequence; *noise*, in which a loud "pink noise," easily distinguishable from the hiss, occurred during the mandatory 2-sec waiting period; *count*, in which the S had to count aloud from 1 to 4 before responding (his counting was monitored over an intercom by the E); and *extra sequence*, in which a second one-shot sequence, randomly chosen, was presented after the first and before the response (the S was instructed to ignore it). Each condition was presented once on each experimental day, after a short extra practice run in the control condition.

Apparatus difficulties and procedural changes which occurred during the course of Experiment V invalidate some of the data. In the first session (Day 16), the signal intensities were abnormally weak due to a calibration difficulty which was corrected to some extent on Days 17 and 18 but not definitely understood until Day 19. Thus, data from Day 16 will not be presented. Two other changes were made on Day 19: the increment step, D, was reduced from 10 to 5 msec to permit more accurate measurement, and the intensity of the pink masking noise in the *noise* condition, which had been only 20 dB louder than signal hiss, was increased to be 30 dB louder instead. The loud noise greatly disturbed S E.S. on Day 19, affecting his performance not only in the *noise* condition but in several conditions which he encountered afterwards; at his suggestion, we have discarded all his data for that day.

Table 4 shows the results. If the *standard* condition is regarded as a baseline, the enforced 2-sec delay in *wait* affected the Ss differently. It was helpful to D.B., made little difference to S.D. (the atypically high value on Day 20 is probably artifactual), and had a negative effect

Table 4
Results of Experiment V

Condition	S D.B. Day 17	18	19	20	S S.D. Day 17	18	19	20	S E.S. Day 17	18	19	20
Standard	59	77	85	69	66	67	88	76	18	18	–	25
Wait	65	44	54	63	66	66	97	132	29	22	–	38
Noise*	77	64	95*	83*	99	120	133*	88*	44	50	–	32*
Count	72	55	70	77	306	157	132	180	57	32	–	26
Extra Sequence	118	125	142	136	245	179	215	219	116	49	–	55

*Note–Each entry is a threshold (milliseconds) based on two runs of 48 trials each. (*The pink masking noise was 10 dB louder on Days 19-20 than on Days 17-18.)*

Table 5
Easy and Hard Permutations Selected for Use in Experiment VI Based on Day 22 of Experiment II

S D.B.				S S.D.				S E.S.			
Easy		Hard		Easy		Hard		Easy		Hard	
UDHS	11	SUDH	6	DHSU	9	UDHS	5	HSUD	9	DHSU	3
UHSD	9	SDUH	4	UHSD	11	SDUH	6	HSDU	11	DUHS	3
SDHU	12	USDH	3	SDHU	5	DHUS	0	HUSD	8	DHUS	3
HUDS	11	DSHU	6	UDSH	7	HUDS	4	SHUD	11	UDSH	2
SUHD	10	UHDS	0	HDSU	8	DSUH	3	UHDS	11	DSUH	3
DUSH	12	USHD	7	DUSH	7	SHDU	3	SHDU	11	DUSH	3
Mean	10.8		3.3		7.8		3.5		10:2		2.8
Difference of Means	7.5					4.3				7.4	

Note–Components: H, hiss; S, scratch; U, high tone; D, low tone. The number of times each S had been correct on each permutation (out of 12) on Day 22 is also given, with mean number correct for each category.

for E.S. The three masking conditions, *noise, count*, and *extra sequence*, had marked negative effects for all the Ss. Introducing the pink noise during the 2 sec raised D.B.'s threshold by about 20 msec compared to *wait*, had an even larger effect on S.D. for the first 3 days, and raised E.S.'s threshold on 2 of the 3 days for which data are available. Being made to *count* aloud was not quite as disturbing as the extraneous noise for D.B. or E.S., but far more so to S.D. The introduction of an extra and irrelevant sequence was the most disturbing condition of all; it raised thresholds by a factor of 2 or more in nearly every case.

EXPERIMENT VI (DAYS 26-27)

It had become apparent as Experiment I progressed that the *spaced* condition did not produce better performance than simple cyclic repetition; indeed, for S.D. it was consistently worse. One possible reason for the difficulty of *spaced* sequences, as noted above, is that they deprive the listener of an option available in the unsegmented *cyclic* case. Hearing ABCDABCDABCDABC . . ., the listener may focus on ABCD, BCDA, CDAB, or DABC as he pleases; in the spaced condition, he is limited to ABCD ABCD ABCD . . . . This hypothesis predicts that *spaced* sequences made up of particularly easy permutations (as established in Experiment II) should *not* be harder than corresponding *cyclic* sequences, while *spaced* sequences made up of particularly difficult permutations should be much harder than *cyclic* ones. Although the complete logic of the hypothesis applies only to S.D., who alone continued to show a *spaced/cyclic* difference throughout the study, it was tested for all three Ss on Days 26 and 27. For each S, we selected six easy and six hard one-shots, using primarily data from Day 22 of Experiment II. Table 5 shows which ones were selected for each S. In the *easy spaced* condition, every stimulus consisted of 10 spaced presentations of one of the six easy permutations; in *easy cyclic*, there were no spaces, but the same six basic permutations were used; *hard spaced* and *hard cyclic* were formed analogously. Each condition was presented once, for 60 trials, on each day.

Table 6
Results of Experiment IV

Condition	S D.B. Day 26	S D.B. Day 27	S S.D. Day 26	S S.D. Day 27	S E.S. Day 26	S E.S. Day 27
Easy Cyclic	28	32	41	46	16	14
Easy Spaced	28	47	38	61	14	13
Hard Spaced	44	43	71	82	26	33
Hard Cyclic	34	36	38	41	34	25

Note–Each entry is a threshold (milliseconds) based on one run of 60 trials.

Table 6 shows the results. The selection of easy and hard permutations was evidently effective. *Hard spaced* sequences produced higher thresholds than *easy spaced* ones on both days for S.D. and E.S., as well as on the first day for D.B. In fact, *hard cyclic* was also generally harder than *easy cyclic*. The difference between *spaced* and *cyclic* was exaggerated when both were *hard*, as had been predicted, not only for S.D. but for D.B. as well. Also as predicted, there were no consistent differences between *spaced* and *cyclic* when both were built up of *easy* permutations. Thus, although Experiment VI did not continue long enough to establish fully reliable results, it seems likely that our explanation of the difficulty of spaced sequences is correct.

CONCLUSIONS

It is evident that success in naming the order of a rapid sequence of signals depends to a great extent on the skills and strategies that the listener can bring to bear on the task. All of our Ss showed spectacular improvement over the course of Experiment I, and ended with thresholds far below 100 msec per component. The best S, E.S., was at 23 ± 7 msec on the last day of Experiment I.

E.S. was remarkable not only for his low threshold, but because many of the conditions which affected the performance of D.B. and S.D. left his unimpaired. They found *one-shots* to be harder than repetitive sequences; he did not. They (especially S.D.) found *spaced* sequences harder than *cyclic* ones, evidently because spacing forced them into a particular analysis of the repetitive stimulus; he did not. For them, replacing the highest (3,400-Hz) tone with one of 590 or 510 Hz (thus bringing it closer to the 500-Hz low tone) made the sequences more difficult; it made no difference to him. The introduction of temporal and rhythmic irregularity in Experiment IV made the task harder for S.D. and especially for D.B., but not for E.S. When the difficulty of individual one-shot sequences was studied, his behavior was again unlike theirs. On the first day of Experiment II, D.B. and S.D. were in considerable agreement in finding certain sequences hard and others easy, but E.S.'s ordering was negatively correlated with theirs; while his ordering remained consistent from Day 8 to Day 22, theirs did not. Apparently E.S. developed sophisticated and powerful skills for coping with the experimental task rather early and used them throughout, while the other two Ss continued to change and improve approaches which were never as effective. He attributed his success to his musical background, and often used concepts and terminology taken from music to discuss the task. Unfortunately, we have not succeeded in using either his comments or his data to construct a satisfactory theoretical account of his methods.

Apart from the effects of practice and of individual difference, our most consistent finding concerns the effects of "backward masking" in Experiment V. Forcing Ss to wait 2 sec before reporting the order of a one-shot sequence does not necessarily impair their performance, but the introduction of noise, enforced activity, or an additional signal during the 2 sec definitely does. This supports our notion that the response string is not fully available at first; the S must construct it gradually out of a less articulate analog. Either the stored analog or the process of construction must be vulnerable to the interfering manipulations we used.

In summary, we can conclude that the order of a temporal succession of sounds is not given to listeners "directly," any more than any other property of the perceivable world. Information about the order must be extracted, and an appropriate response constructed, by means of skills and strategies that require time to execute. Ss acquire these skills relatively slowly and in unequal measure. Little can be predicted on the basis of stimulus variables alone, and results obtained with naive Ss cannot safely be generalized to sophisticated ones. Whether the components are close together in frequency or far apart, whether the temporal spacing is regular or arhythmic, whether the sequence occurs once or repeatedly—the importance of such factors is drastically dependent on the listener and his skills.

REFERENCES

Bregman, A. S. Effects of stream segregation on the perception of order. Paper presented to the American Psychological Association, Honolulu, September 1972.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 1971, 89, 244-249.

Neisser, U. On the perception of auditory sequences. Paper presented to the American Psychological Association, Honolulu, September 1972.

Warren, R. M. Temporal resolution of auditory events. Paper presented at the American Psychological Association, Honolulu, September 1972a.

Warren, R. M. Perception of temporal order: Special rules for the initial and terminal sounds of sequences. Journal of the Acoustical Society of America, 1972b, 52, 167 (abstract).

Warren, R. M. Temporal order discrimination: Recognition without identification by untrained subjects. Journal of the Acoustical Society of America, 1973a, 53, 316 (abstract).

Warren, R. M. Temporal order discrimination: Identification and recognition by trained listeners. Journal of the Acoustical Society of America, 1973b, 53, 316 (abstract).

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusion of patterns other than speech or music. Science, 1969, 164, 586-587.

Wilcox, G. W., Neisser, U., & Roberts, J. Recognition of auditory temporal order. Paper presented to the Eastern Psychological Association, Boston, Mass., 1972.

NOTE

1. Nickerson, R. S., & Freeman, B. Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation, tone-to-interval ratio, and extensive practice. In preparation, Bolt Beranek & Newman Inc., Cambridge, Massachusetts, 1973.

(Received for publication September 4, 1973; revision received December 10, 1973.)