

Memory strategies in absolute identification of "circular" pitch

ALAN COSTALL, STEVE PLATT, and ALEX MACRAE
University of Southampton, Southampton SO9 5NH, United Kingdom

Shepard (1964) has devised a sequence of synthesized tones that appear to constitute a "circular" pitch dimension; each tone in the sequence appears to be higher in pitch than the preceding one, and yet the first tone also seems higher than the last. Two series of such stimuli (a six-alternative and an eight-alternative set) were used to investigate memory processes in absolute identification performance when end stimuli are unavailable as possible anchors. Clear evidence was obtained for short-term anchoring strategies in this task, but none was obtained for the stable retention of those stimuli corresponding to the ends of the response continuum. The results are examined in the light of Eriksen and Hake's (1957) subjective-standard hypothesis.

The original development of the method of absolute judgment (or single stimuli) as a rating procedure for measuring discrimination thresholds was based on the assumption that people are able to remember the values of the judgmental categories. This memory assumption seemed to be justified, since difference thresholds determined by the method of absolute judgment were in close agreement with those obtained from the other psychophysical procedures (Wever & Zener, 1928; Woodworth, 1950, pp. 425-426). However, it should be noted that the stimuli in these early experiments were very closely spaced, so perceptual resolution was a major limiting factor in performance (cf. Pynn, Braida, & Durlach, 1972).

Since the late 1940s, the task has been used primarily to study identification performance. The stimulus alternatives are assigned unique numerical labels in order of magnitude or level and are spaced so widely that they would be impossible to confuse in a simple discrimination task. Nevertheless, if the stimulus alternatives vary along a single sensory dimension, errors occur when seven or even fewer alternatives are presented (Miller, 1956). Clearly, in these experiments, performance is no longer constrained by limits of sensory discrimination, yet, despite the explicit assumption underlying the introduction of the absolute judgment procedure, memory processes were largely neglected as a possible explanation of the low limit on identification performance (but see Pollack, 1954, 1956, 1959).

Recently, Siegel (1972) has reported some studies of pitch identification with feedback that demonstrate the importance of short-term memory processes in this task. He found that identification of a tone

was much improved if that same stimulus had occurred shortly before in the sequence of trials, and he concluded that people make their judgments by trying to match the value of the presented stimulus to the unstable memory traces of the preceding stimuli.

It seems likely, however, that additional anchoring strategies are involved. For example, in Siegel's experiment, performance did not fall to chance level even when several trials had intervened since the previous occurrence of the presented tone, and, in an earlier study by Hartman (1954), also on pitch identification, a burst of noise presented before each test tone to "wash out" the traces of preceding tones had only a limited disruptive effect.

The possible existence of longer term memory processes in absolute identification was recognized some time ago by Eriksen and Hake (1957) in their subjective-standard hypothesis. They suggested that the observer attempts to retain the values of a few stimuli in the series as standards for judging the remainder and thus, in effect, transforms the task into one of relative judgment. More specifically, they proposed that people tend to adopt the two extreme stimuli in the series as anchors and hence make their judgments by interpolation between the remembered values of these subjective standards (see also Durlach & Braida, 1969; Lockhead, 1973). The "end effect" in absolute identification performance would seem to provide good evidence for such an anchoring strategy; identification of the extreme stimuli in the series is usually markedly enhanced. There are, however, several alternative explanations for the end effect, although two of these can be readily eliminated. The end effect does not result from a tendency to over-employ the end response categories, since it is found even when (theoretically defined) bias-free measures of resolution are used (Braida & Durlach, 1972). Nor

Requests for reprints should be sent to A.P. Costall, Department of Psychology, The University, Southampton, SO9 5NH, United Kingdom.

can the end effect be attributed to a failure of Weber's law for the extreme stimuli; it appears to be a true context effect, since the confusability of adjacent stimuli is a function of their position in the set rather than of their levels *per se* (Hartman, 1954).

Two additional rivals to Eriksen and Hake's account of the end effect cannot be dismissed so easily. The first, suggested by Garner (1953), is that the end stimuli are better identified simply because errors in identifying these can occur in only one direction (cf. John, 1973). The second possibility is that the end stimuli are in some way perceptually distinct.

These alternative explanations are difficult to assess if we consider only the identification of stimuli varying along bipolar sensory dimensions such as loudness or length. But, some sensory dimensions, such as the "color circle," have no obvious ends, in which case the stimuli can be labeled from any point we choose. Then, Eriksen and Hake argued, the observer will adopt as subjective standards those stimuli associated with the ends of the response continuum, so an end effect should be found even for circular stimulus dimensions. On the other hand, if the end effect in judgments of bipolar dimensions results from the greater distinctiveness of the end stimuli or because errors in identifying them can occur in only one direction, then no end effect should be found for judgments along a circular dimension.

Eriksen and Hake investigated absolute identification of a series of stimuli whose color varied only in hue and which were spaced evenly around the color circle. They instructed observers to use the numbers 1 through 20 to identify the colors along the series, but, for one group, numbering began in the red region of the color circle, while, for the other group, it began in the blue-green region. For the first group, an end effect was found, the stimulus associated with the beginning of the response series being better identified than the others, but the results of the second group were not so clear, and no obvious end effect occurred.

The interpretation of Eriksen and Hake's results is complicated by the presence of especially distinctive hues along the color circle; furthermore, although few of us possess "perfect pitch," we all have "perfect color." What is needed for a clearer test is another circular dimension, but there are few, and the most convenient, angular orientation, is ruled out by the existence of distinctive positions along the continuum (the horizontal and vertical). Fortunately, Shepard (1964) has devised an interesting sequence of complex tones that has the property of circularity. Each of the tones consists of 10 frequency components spaced by octaves, the energy of each tone being concentrated within the middle components. Each successive tone is obtained by shifting the base frequency slightly higher and then adding the additional components. The loudness levels of

the components are adjusted to conform to a spectral envelope, identical for all of the tones in the set. Eventually, a tone is obtained whose base frequency corresponds to the second lowest component of the very first tone in the sequence but whose highest component is an octave above that of the first tone. Now we can replace this tone with the first, since the two differ only in their extreme components, which are of low intensity and hence virtually inaudible. Thus, a circle of tones spanning an octave is obtained. Every tone in the sequence sounds higher than the one before (although similar in quality and loudness), and yet the first tone in the sequence also sounds higher than the last. The resulting sequence is therefore analogous to the never-ending staircases in Escher's pictures, which appear to be continuous although every step seems higher than the one before.

Such a sequence of "circular pitch" stimuli was used in the present experiment to reexamine Eriksen and Hake's explanation of the end effect in terms of their subjective-standard hypothesis.

METHOD

The circular pitch stimuli were generated by means of a PDP-11/20 computer, which was also used to control the experiment and to record the responses. Each tone consisted of 10 frequency components spaced by octaves, the loudness levels of the components being set according to the spectral envelope specified by Shepard (1964). The base frequency of the first tone in the sequence was 6.8 Hz.

The waveforms of the tones were computed prior to the experiment, and a cycle of each successive tone was stored as a sequence of digital values on disk. These digital values specified displacement as a function of subsequent presentation time at a rate of 14,200 samples/sec via a digital-to-analog converter. Sampling noise was eliminated by means of a fifth-order Butterworth filter, with a 3-dB cutoff at 4.4 kHz and a slope of 30 dB/octave. A circuit that produced a 20-msec exponential rise and fall of the presented tones ensured click-free switching.

The tones were presented to both ears through AKG-K60 moving-coil earphones, the loudness of the tones being set at a comfortable level for the listener. A total of 24 undergraduate students were tested individually in a quiet cubicle. The response panel consisted of a horizontal line of push buttons, numbered in order from left to right, with a feedback light beside each button.

One group of 13 observers judged a sequence of six tones spaced evenly around the pitch circle by equal logarithmic steps in frequency. Each person took part for two sessions run on consecutive days. A session consisted of two blocks separated by a rest period of not less than 5 min. The order of the 90 trials within each block was determined by random selection without replacement from a set consisting of 15 of each of the 6 tones.

Two labeling conditions were employed, the numbers 1 to 6 being assigned to the tones in order of pitch. For six of the observers, numbering of the tones started from the first tone in the sequence on the first session but from the fourth tone in the sequence on the second session. The remainder of the observers were presented with the two labeling conditions in the reverse order. They were not informed that the labeling of the tones had been changed on the second session.

The six-tone series was consistent with the tuning of Western musical scales of equal temperament and, in fact, constituted a whole-tone scale. An additional group of 11 observers judged a sequence of eight circular pitch stimuli not consistent with

Western musical scales but, again, spaced by equal logarithmic steps. Two labeling conditions were again employed, five of the listeners being given the condition in which numbering began from the first tone of the generated sequence on the first session. Each of the tones occurred 12 times within each block. In other respects, the procedure was identical to that used for the six-tone set.

Before each block of judgments, the full set of tones was presented twice, in order of labeling, with a pause of 4 sec between each presentation. Each of the tones lasted for 300 msec, and the interstimulus interval was 1 sec. After a delay of 2 sec, the identification task began. Stimulus presentation lasted for 300 msec, and the listener could record a response only after the tone had ended. Feedback was given immediately by means of a light beside the correct push button, and then, 2 sec later, the next tone was presented for identification. The task was thus self-paced but with a minimum interstimulus interval of 2 sec.

The observers were told that the tones were equally probable and were presented in a random order, and they were also informed of the unusual (i.e., circular) nature of the pitch series.

RESULTS AND DISCUSSION

Ordinal Position Effects

Identification of the six-tone set will be considered first. Figure 1 represents performance as a function of the position of the tone along the response continuum. The measure of performance is that used by Eriksen and Hake (1957), the percentage of responses correctly assigned. No end effect is apparent, accuracy being no higher for the tones labeled 1 and 6 than for the intermediate tones. Figure 1A shows the effect of ordinal position for the first group when numbering began from the first tone in the generated sequence on the first session (filled circles) and from

the fourth tone in the sequence on the second session, a day later (open circles). Figure 1B shows ordinal position effects for the second group of observers, who received the labeling conditions in the reverse order, for the first session (filled circles) and the second session (open circles). Performance did improve from the first session to the next, but there is no evidence that ordinal position effects developed with practice.

A three-factor repeated measures ANOVA revealed that none of the main sources of variance (ordinal position, labeling condition, and order in which the labeling conditions were given) was reliable. There were two reliable interaction effects. The first, the interaction between labeling condition and ordinal position, indicates a slight but reliable effect due, presumably, to an unevenness in quality of the tones [$F(5,120) = 2.76, p < .05$]. The second, the interaction between labeling condition and the order in which it was presented, corresponds to the effect of practice on overall performance [$F(1,120) = 31.26, p < .05$].

Similar results were obtained when the number of alternatives was increased to eight tones. Figures 1C and 1D show performance under the two labeling conditions and across the two sessions. Again, no end effect occurred, and the only reliable effect was that corresponding to the improvement in the overall level of performance from the first to the second sessions [$F(1,135) = 24.46, p < .01$].

The results for both the six- and the eight-tone series indicate, therefore, that no end anchoring to the extremes of the response continuum occurs when people make absolute judgments of circular pitch stimuli. It should be noted that the manner in which the tone series was initially presented, as well as the linear layout of the response panel, favored anchoring to those tones associated with the ends of the response continuum. Nevertheless, although a number of the observers reported that at the outset of the experiment they regarded Tone 1 as the "tonic" of the set, they no longer thought of this tone as in any way special once the "circular" nature of the series had been experienced.

Short-Term Anchoring Strategies

Analysis of sequential effects indicated the existence of strong short-term anchoring strategies in this task. Figures 2A and 2B represent performance as a function of "retention interval" (i.e., the number of items intervening since the previous occurrence of the presented stimulus). The measure of performance makes allowance for the influence of preceding stimuli on subsequent response tendencies; the accuracy associated with each retention interval is expressed as the ratio of the number of times each stimulus was correctly identified to the number of times the corresponding response had been used (both correctly and

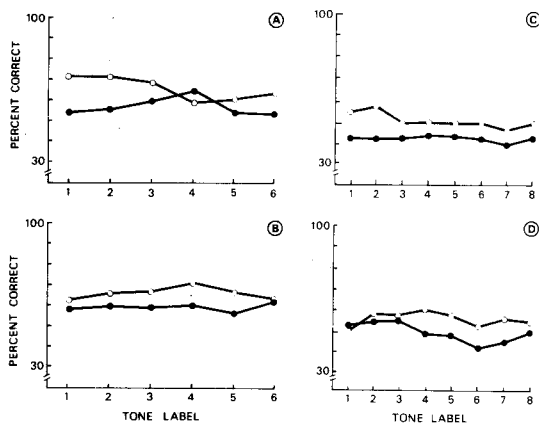


Figure 1. Identification performance as a function of tone label, both for the six-tone set, A and B, and for the eight-tone set, C and D. The upper figures show the results for the observers for whom numbering began from the first tone in the sequence on the first session (filled circles) and from the fourth tone in the sequence on the second session (open circles). The lower figures show the results of the other observers, who received the labeling conditions in the reverse order, for the first session (filled circles) and the second session (open circles). The measure of performance is the percentage of responses correctly assigned.

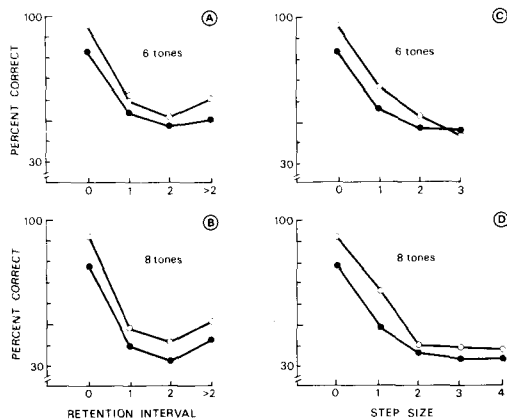


Figure 2. Retention interval effects in the six-tone condition (A) and in the eight-tone condition (B). Identification accuracy as a function of extrapolation interval in the six-tone condition (C) and in the eight-tone condition (D). The filled circles represent performance on the first session, and the open circles represent performance on the second session. The measure is the percentage of responses correctly assigned.

incorrectly) after the same number of intervening items. Figure 2A represents the averaged results of the individuals in the six-stimulus condition, and Figure 2B represents the results for the eight-stimulus condition; retention interval curves for both the first session (filled circles) and for the second session (open circles) are given. Identification was very accurate when the presented tone had also occurred on the immediately preceding trial, but retention appears to have been equally poor for all other retention intervals, and the pattern of these retention interval effects was relatively stable over the two sessions.

The reports of the observers concerning their performance suggested that they had used the remembered value of the immediately preceding stimulus not only on repetition trials (by the matching process proposed by Siegel, 1972), but also on nonrepetition trials by means of extrapolation; they had tried to estimate the number of category steps between the presented tone and the one presented just before. The existence of such an extrapolation strategy was confirmed by an analysis of identification accuracy as a function of the difference in category steps between the presented and immediately preceding stimulus (Figures 2C and 2D). (Again, the measure of performance is the percentage of responses correctly assigned.) The results for both the first session (filled circles) and for the second session (open circles) show that identification accuracy was high when the immediately preceding tone was close in value to the presented tone and decreased as a function of the extrapolation interval.

The Subjective-Standard Hypothesis

The implication of Eriksen and Hake's subjective-standard hypothesis is that the choice of the ends of

a bipolar stimulus continuum as a reference for judgment is optional; when end stimuli are not available, the prediction is that those stimuli associated with the ends of the response continuum will serve as anchors for judgment. The present experiment indicates that such response anchoring does not occur in judgments of circular pitch. Furthermore, the results suggest that the observers relied heavily on short-term anchoring for their judgments.

Nevertheless, there can be little doubt that longer term memory strategies can occur in absolute judgment performance. In a study of loudness judgments by Ward and Lockhead (1970), the entire stimulus set was shifted up or down in level from day to day; although the observers were not informed of the shifts in loudness, constant errors occurred in the direction of the previous scale values. Clearly, a relatively stable impression of the stimulus range was retained from one day to the next. Further evidence of long-term anchoring comes from Cuddy's research on the training of pitch judgments (Cuddy, 1968, 1970; Cuddy, Pinn, & Simons, 1973). Initially, the middle tone in the set was presented much more frequently than the other alternatives, but, as training progressed, its frequency of presentation was gradually decreased until all of the stimuli were equally probable; even when the tones were presented equally often, identification of the middle and neighboring tones was improved.

In conclusion, long-term anchoring can occur in absolute judgments and may even account (in part at least) for the end effect in judgments of bipolar sensory continua. What the results of the present experiment do indicate, however, is that the choice of subjective standards is not as "subjective" as Eriksen and Hake's hypothesis seems to imply. Long-term anchoring occurs only to a stimulus that is intrinsically distinctive in some way, such as an end stimulus, perhaps, or a more frequently presented item.

REFERENCES

- BRAIDA, L. D., & DURLACH, N. I. Intensity resolution II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, 1972, **51**, 483-502.
- CUDDY, L. L. Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America*, 1968, **43**, 1069-1076.
- CUDDY, L. L. Training the absolute identification of pitch. *Perception & Psychophysics*, 1970, **8**, 265-269.
- CUDDY, L. L., PINN, J., & SIMONS, E. Anchor effects with biased probability in absolute judgment of pitch. *Journal of Experimental Psychology*, 1973, **100**, 218-220.
- DURLACH, N. I., & BRAIDA, L. D. Intensity perception I: Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, 1969, **46**, 372-383.
- ERIKSEN, C. W., & HAKE, H. W. Anchor effects in absolute judgments. *Journal of Experimental Psychology*, 1957, **53**, 132-138.
- GARNER, W. R. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, 1953, **46**, 373-380.

HARTMAN, E. B. The influence of practice and pitch-distance between tones on the absolute judgment of pitch. *American Journal of Psychology*, 1954, **67**, 1-14.

JOHN, I. D. Sequential effects in absolute judgments of loudness without feedback. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.

LOCKHEAD, G. R. Choosing a response. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.

POLLACK, I. Intensity discrimination under several psychophysical procedures. *Journal of the Acoustical Society of America*, 1954, **26**, 1056-1059.

POLLACK, I. Identification of sound level and 'matching to sample.' *Journal of the Acoustical Society of America*, 1956, **28**, 412-415.

POLLACK, I. Identification of elementary auditory displays and the method of recognition memory. *Journal of the Acoustical Society of America*, 1959, **31**, 1126-1128.

PYNN, C. T., BRAIDA, L. D., & DURLACH, N. I. Intensity perception III: Resolution in small-range identification. *Journal of the Acoustical Society of America*, 1972, **51**, 559-566.

SHEPARD, R. N. Circularity in judgment of relative pitch. *Journal of the Acoustical Society of America*, 1964, **36**, 2346-2353.

SIEGEL, W. Memory effects in the method of absolute judgment. *Journal of Experimental Psychology*, 1972, **94**, 121-131.

WARD, L. M., & LOCKHEAD, G. R. Sequential effects and memory in category judgments. *Journal of Experimental Psychology*, 1970, **8**, 27-34.

WEVER, E. G., & ZENER, K. E. The method of absolute judgment

in psychophysics. *Psychological Review*, 1928, **35**, 466-493.

WOODWORTH, R. S. *Experimental psychology*. London: Methuen, 1950.

Appendix
The Complete Stimulus-Response Confusion Matrices
for Both the Six- and Eight-Tone Sets

	S1	S2	S3	S4	S5	S6	S7	S8	
R1	457	103	41	23	39	87			750
R2	94	504	99	51	25	38			811
R3	35	88	488	103	46	23			783
R4	33	29	85	483	101	36			767
R5	57	20	36	90	473	124			800
R6	104	36	31	30	96	472			769
	780	780	780	780	780	780			4680
R1	251	86	34	14	10	15	33	66	509
R2	64	258	72	24	7	11	17	27	480
R3	37	75	272	71	30	16	20	14	535
R4	19	31	85	283	89	28	18	9	562
R5	9	15	28	77	262	97	31	21	540
R6	18	18	16	29	84	247	86	36	534
R7	38	17	8	15	32	80	250	86	526
R8	92	28	13	15	14	34	73	269	538
	528	528	528	528	528	528	528	528	4224

(Received for publication November 20, 1980;
revision accepted March 26, 1981.)