

EFFECT OF REPETITION OF STANDARD AND COMPARISON TONES ON RECOGNITION MEMORY FOR PITCH¹

DIANA DEUTSCH²

Center for Human Information Processing, University of California, San Diego

In all conditions, pitch recognition was required after a 4-sec. interval during which four other tones were played. The effects of including in the interpolated sequence a tone at the same pitch as the standard (S) or comparison (C) tone were investigated. When the S and C tones were identical in pitch, inserting in the interpolated sequence a tone at that pitch caused a reduction in errors. When the S and C tones differed in pitch, inserting a tone at the pitch of the S tone produced a reduction in errors and inserting a tone at the pitch of the C tone produced an increase in errors. Inserting one tone at the pitch of the S tone and another at the pitch of the C tone produced no significant change in errors, compared with the condition in which no tone at either pitch was included in the interpolated sequence.

Several studies have been made of recognition memory for pitch when the tones to be compared are separated by a retention interval. Koester (1945) has demonstrated that memory for both loudness and pitch deteriorates with time, using the differential threshold (DL) as a measure of memory. Bachem (1954) found a continuing deterioration of pitch recognition up to a week, yet showing that we can maintain some sort of memory for an auditory stimulus over very long periods. Harris (1953) measured the DL with various interstimulus intervals (ISIs) under two conditions. When the standard tone was fixed at 1,000 cps, the DL was found to remain constant at around 4 cps for ISIs of up to 3.5 sec., with a decline of only .8 cps with an ISI of 15 sec. With a roving standard tone varying in frequency from 950 to 1050 cps, no increase in the

DL occurred up to a 1.0-sec. ISI. Then a linear increase in the DL was found up to 15-sec. ISI where the DL was increased by 3.7 cps. It can be concluded that memory for pitch decays spontaneously with time. However, the amount of deterioration manifest in 15 sec. is small.

Until recently, the effects of interpolated stimuli on pitch recognition have not been investigated. Wickelgren (1966, 1969) has found that a tone incorporated between the standard (S) and comparison (C) tones produces memory deterioration which increases with increased duration of the intervening tone. In an experiment designed to study the effects of increased interference in a retention interval of constant duration, the present author presented Ss with a set of S and C tones, each separated by a $4\frac{1}{4}$ -sec. retention interval. A sequence of tones was interpolated during this retention interval, always beginning at 300 msec. after the end of the S tone and ending 2 sec. before the start of the C tone. The interpolated tonal sequence consisted of either four, six, or eight tones spaced regularly within the remaining time span. A statistically significant increment in errors was obtained, both when the number of tones interpolated during the retention interval was increased from four to six and also when it was increased from six to eight. It can be concluded that memory for tonal pitch is subject to interference by other tones even

¹ The work reported in this paper was performed in partial fulfillment of the requirement for a PhD thesis at the Department of Psychology, University of California at San Diego. I thank my advisor, G. Mandler, and the members of my thesis committee, A. Garsia, D. M. Green, I. Jacobs, and D. E. Rumelhart, for their advice and encouragement. I also thank D. A. Norman for use of computer facilities supported by United States Public Health Service Training Grant NS-07454-06. The work was supported in part by United States Public Health Service Training Grant MH-10835 and Grant MH-15828-01.

² Requests for reprints should be sent to Diana Deutsch, Department of Psychology, P.O. Box 109, University of California, San Diego, La Jolla, California 92037.

when the retention interval is held constant. Massaro (1970) has also reported that increasing the number of tones interpolated during an interval of constant duration increases the number of errors in pitch recognition. In this experiment, Massaro varied the number of interpolated tones between one and four.

The above findings show clearly that memory for pitch does not simply deteriorate with time but is subject to some interference effect. A further experiment (Deutsch, 1970b) has shown that this interference cannot be explained in such general terms as a distraction of attention or a limitation in information storage capacity. In this study it was shown that the incorporation during the retention interval of spoken numbers, which Ss are required later to recall, produces only a minimal decrement in the same pitch-recognition task as is severely disrupted by the interpolation of other tones. Interactive effects must therefore take place within the pitch memory store itself. The following experiment represents a further investigation into the pitch memory store. It was theorized that there must be a representation in memory both of the item to be remembered, and also of the time or order in which this item occurred. One can obtain information about these representations by repeating the same item at a different point in time, since here item and order information are not simultaneously covaried. This procedure was used in the experiment described here. A preliminary investigation of this nature is reported in Deutsch (1970a).

METHOD

Procedure.—In all experimental conditions, Ss listened to an S tone which was followed by a sequence of four interpolated tones, and later, after a pause, by a C tone. The Ss were instructed to try to remember the S tone, ignore the four interpolated tones if they wished, and then to judge whether the C tone was or was not the same in pitch as the S tone. They then indicated their judgments by writing "same" or "different" on paper. All the tones were of equal loudness, and 200 msec. in duration. The interval between the S tone and the first interpolated tone was 300 msec. and the interpolated tones were also spaced 300 msec. apart. A 2-sec.

pause was incorporated between the last interpolated tone and the C tone.

Conditions.—The different conditions of the experiment are shown in Fig. 1. It can be seen that the presence or absence in the interpolated sequence of a tone at the pitch of the S tone was systematically varied; as was the presence or absence of a tone at the pitch of the C tone. Further, the placement of these critical tones in the interpolated sequence was systematically varied between the second and third serial positions. The entire tape consisted of 96 sequences. These were presented in 12 groups of eight. Sequences within each group of eight were separated by 10-sec. pauses; and there were 2-min. pauses between the groups. The sequences were arranged in random order with no separation by condition except that each group of eight contained four sequences in which the S and C tones were the same in pitch, and four in which they were different, to minimize the development of response biases. Also, no group of sequences contained more than one example of each S and C tone pitch (see below) so that pitch repetition effects for these critical tones were minimized. The Ss listened to the entire tape on two different occasions and the results were averaged.

S and C tones.—S and C tone pitches were taken from an equal-tempered scale (international pitch: A = 435) ranging from the C# a semitone above middle C to the C an octave above. The frequencies employed were: C# = 274, D = 290, D# = 308, E = 326, F = 345, F# = 366, G = 388, G# = 411, A = 435, A# = 461, B = 488, C = 517. Each of these tonal pitches was employed equally often within each condition, either as an S tone, or as a C tone, or both. When the S and C tones differed in pitch, in half of the sequences the S tone was higher than the C tone and in the other half it was lower.

Interpolated tones.—The interpolated tonal pitches were also taken from the equal-tempered scale (A = 435) and ranged from middle C to the C# an octave above. The frequencies employed were: C = 259, C# = 274, D = 290, D# = 308, E = 326, F = 345, F# = 366, G = 388, G# = 411, A = 435, A# = 461, B = 488, C = 517, C# = 561. The tones in the intervening sequences were chosen randomly from this set, except that no sequence contained repeated tones unless specified by the experimental condition. This restriction also covered repetition of tones separated by exactly an octave.

Subjects.—Fifteen Ss participated in this experiment. These were 14 University of California at San Diego undergraduates and 1 housewife and they were paid for their participation. The Ss were selected on the basis of obtaining a score of at least 75% correct on a small tape containing similar sequences. About one out of five applicants was accepted by this procedure.

Apparatus.—The tones were generated by a Wavetek oscillator controlled by a PDP-9 computer. The output was recorded on high-fidelity tape. The tape was played to Ss on a high-quality tape recorder through loudspeakers with the loudness at a comfortable listening level.

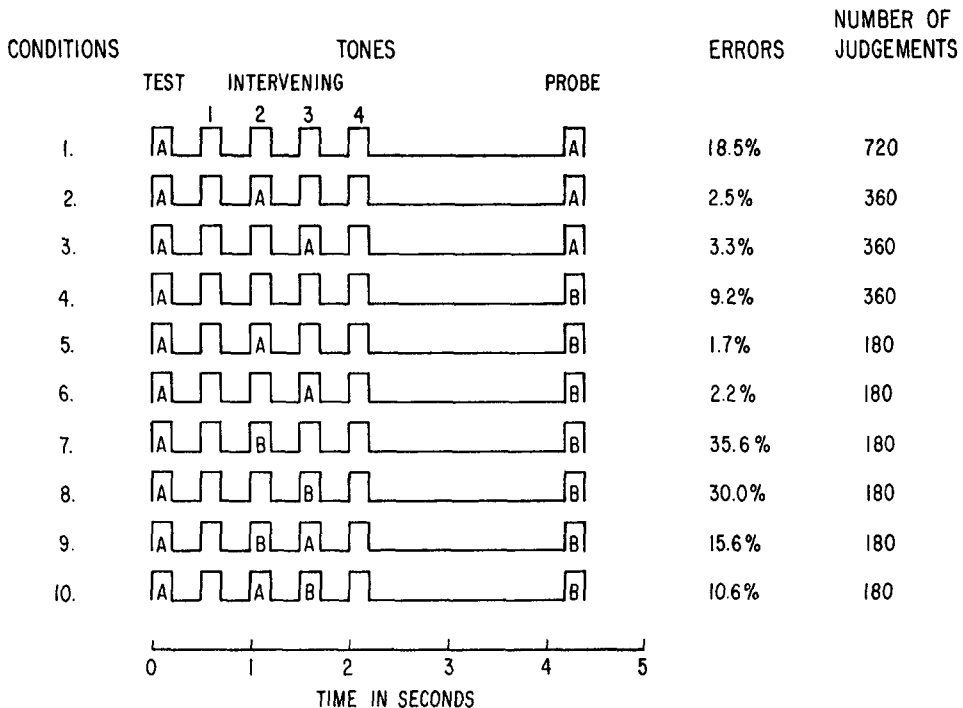


FIG. 1. The different conditions of the experiment with number of judgments and percentage of errors made in each. (Each condition is diagrammed by a set of rectangular elevations, each of which indicates the presentation of a tone. The scale at the base of the figure indicates the time during which each tonal sequence is played. The initial elevation in each sequence indicates that the S tone occurred. The next four elevations labeled 1, 2, 3, and 4 indicate that the interpolated tones occurred, and show their serial position. The final elevation indicates that the C tone occurred. The initial elevations in all the sequences are labeled A and indicate the S tone. When the symbol A occurs elsewhere in a sequence, this indicates that a tone at the same pitch as the S tone occurred at that time. Those final elevations which are not labeled A are labeled B. This indicates that the C tone differed in pitch from the S tone. An elevation in the intervening sequence labeled B indicates that a tone at the pitch of the C tone (when this differed from the S tone) occurred in the indicated position of the intervening sequence.)

RESULTS

The results of this experiment are shown in Fig. 1.

Effects of the Interpolated Tones

The effect of inserting a tone at the pitch of the S tone, when the S and C tones are identical in pitch is shown by comparisons of Cond. 1 with Cond. 2 and 3. A considerable reduction in errors is produced. For both comparisons, this improvement is significant on sign tests beyond the .001 level.

The effect of inserting a tone at the pitch of the S tone when the S and C tones differ in pitch is shown by comparisons of

Cond. 4 with Cond. 5 and 6. An improvement in detection of a difference between the S and C tones is produced. This effect is significant for both comparisons on Wilcoxon tests beyond the .01 level (two-tailed).

The effect of inserting a tone at the pitch of the C tone, when the S and C tones differ in pitch is shown by comparisons of Cond. 4 with Cond. 7 and 8. It can be seen that a striking increase in errors occur. For both comparisons, this effect is significant on sign tests beyond the .001 level.

The effect of inserting two tones, one at the pitch of the S tone and the other at the pitch of the C tone when the S and C tones

differ in pitch, is shown by comparison of Cond. 4 with Cond. 9 and 10. Here there are no significant differences in the number of errors, on Wilcoxon tests ($p > .05$, two-tailed). Thus, the dramatic increase in errors produced when the C-tone pitch alone is included in the intervening sequence does not persist when the S-tone pitch is also included in the same sequence. In fact, the reduction in errors produced both when Cond. 9 is compared with Cond. 7, and also when Cond. 10 is compared with Cond. 8 is significant on Wilcoxon tests beyond the .01 level (two-tailed). Thus the benefit to discrimination produced by inserting the S-tone pitch in the intervening sequence effectively counteracts the damage produced by inserting the C-tone pitch in the same sequence.

Serial Position Effects

Including in the intervening sequence a tone at the pitch of the S tone produced a greater reduction in errors when the critical tone was in the second serial position than in the third. This was true both when the S and C tones were identical (Cond. 2 compared with Cond. 3) and also when the S and C tones differed in pitch (Cond. 5 compared with Cond. 6).

Including in the intervening sequence a tone at the pitch of the C tone, when the S and C tones differed in pitch, produced a greater increase in errors when the critical tone was included in the second serial position than in the third (Cond. 7 compared with Cond. 8).

When tones at both the pitches of the S tone and the C tone were included in the same intervening sequence, errors were greater when a tone at the pitch of the C tone was included in the second serial position and a tone at the pitch of the S tone in the third, than when the serial positions of these critical tones were reversed (Cond. 9 compared with Cond. 10). This difference is significant on a Wilcoxon test at the .05 level (two-tailed).

Although only one of these effects reached statistical significance, all the trends described are consistent with the following generalization: When the included tone was

closer in serial position to the S tone, there was an increased tendency to identify its pitch as that of the S tone.

Effect of Lower Compared with Higher S- and C-Tone Pitches

With the S- and C-tone pitches in the lower half of the range employed in this study (274–366 cps), errors were greater than with the S- and C-tone pitches in the upper half of the range (388–517 cps). Errors in these two conditions were at 15.5% and 9.7%, respectively. This effect is significant on a Wilcoxon test beyond the .01 level (two-tailed). It may be explained by considering that the equal-tempered scale is a ratio scale, and so absolute differences in terms of cps increase as the scale is ascended; yet the DL for pitch over the range employed here remains roughly constant in absolute terms as the scale is ascended (Shower & Biddulph, 1931). Discrimination would therefore be easier at the higher frequencies since here the absolute-pitch differences between S and C tones are greater. In this experiment, however, all S- and C-tone pitches occurred equally often in all conditions.

DISCUSSION

An attempt at rigorous theoretical quantification in explaining these results appears premature; however, a theory can be suggested which seems to be most in accord with the findings.

It is proposed that memory for the pitch of a tone is laid down simultaneously both on a pitch continuum and also on a temporal continuum. This may be represented as in the three-dimensional diagram in Fig. 2. A tone is here represented as laid down in memory in the form of a three-dimensional bell-shaped distribution. As time proceeds, this distribution spreads in both directions, but particularly along the temporal continuum. Also with the passage of time, each distribution shrinks in height, and so the more recent the tone the larger the distribution representing it. When the distributions underlying two tones overlap, the overlapping portions sum. In this way, summation between distributions underlying tones presented at different points in time will occur, due to the spread of these distributions along the temporal continuum.

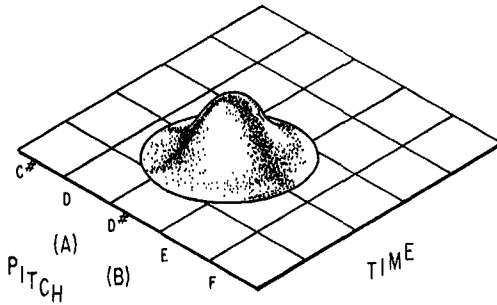


FIG. 2. Proposed distribution underlying memory for the pitch of a tone.

It is further proposed that the decision as to whether a given tone presented at one point in time (in this experiment, the C tone) is or is not the same in pitch as another tone presented at a different point in time (in this experiment, the S tone) is made in the following way. Assume that *S* divides the pitch continuum arbitrarily into sections according to his knowledge of the quantal steps employed in the experiment. He similarly divides the temporal continuum into sections. He is also able to estimate the volume of distribution within each three-dimensional section (see Fig. 2). Now when the C tone sounds, *S* examines the distribution in the section which corresponds to the same pitch as that of the C tone, and to the same temporal region as that of the S tone. He then compares the volume of distribution in this section with those in sections which are adjacent along the pitch continuum, but in the identical temporal region. If the volume of distribution in this section is much larger than those in neighboring sections in the same temporal region, *S* concludes that the S and C tones were the same in pitch. Conversely, if the volume of distribution in this section is much smaller than in a neighboring section in the same temporal region, *S* concludes that the S and C tones were different in pitch. As the volume of distribution in the appropriate section becomes more similar to that in a neighboring section in the same temporal region, the decision as to whether the S or C tones were the same or different in pitch becomes increasingly more difficult and so errors more numerous. Thus, the degree of accuracy of discrimination is determined by the size of a ratio between two distribution volumes.

The diagrams in Fig. 3 represent cross sections of the model on Fig. 2 taken at the point along the temporal continuum representing the presentation of the S tone. The two adjacent

sections are arbitrarily labeled A and B to conform to the terminology in Fig. 1 and represent a semitone difference in pitch. The S tone is always A and the C tone is either A or B.

In Fig. 3, a, no tone either at pitch A or at pitch B occurs in the intervening sequence. The distribution at this point in time is therefore almost entirely due to presentation of the S tone A. (There will be some spread due to presentation of other tones at least a whole tone removed from A but for the purpose of simplicity this is not considered at present: In this experiment, such tones are randomly presented.) There is therefore a substantial difference in the amount of distribution in the A section compared with the B section; however, errors will still be made due to the distribution in the B section.

In Fig. 3, b, a tone at the pitch of the S tone (A) is included in the intervening sequence. The distribution underlying this included tone will spread along the temporal continuum and sum with that already existing due to presentation of the S tone A. The result shown in Fig. 3, b, will be an increased volume of distribution most specifically in the A section of the pitch continuum. The proportional difference between the volume of distribution in the A section and neighboring sections will therefore be increased; and discrimination will become more accurate. This would explain the decrease in errors produced by inclusion of the S-tone pitch in the intervening sequence, both when the S and C tones are identical in pitch (Cond. 2 and 3 compared with Cond. 1) and also when the S and C tones are different in pitch (Cond. 5 and 6 compared with Cond. 4). Further, the closer in time this included tone is to the S tone, the more will its distribution overlap with that of the S tone and the greater will be the resultant improvement in discrimination. This would explain the greater improvement in performance found when the included tone A is presented in the second serial position of the intervening sequence (Cond. 2 and 5) compared with the third serial position (Cond. 3 and 6).

In Fig. 3, c, a tone at the pitch of the C tone B is included in the intervening sequence. The distribution representing this included tone will spread along the temporal continuum and sum with that already existing due to presentation of the S tone A. The distribution existing at the point along the temporal continuum representing the presentation of the S tone A will therefore look somewhat as in Fig. 3, c. It can be seen here that the proportional difference between the volume of

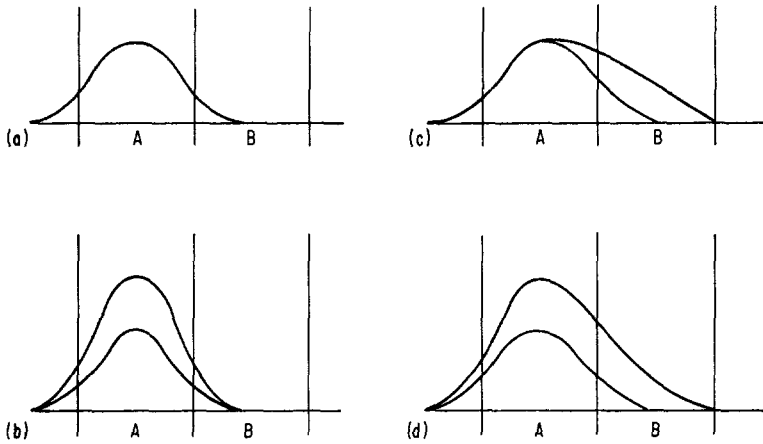


FIG. 3. Cross sections of the distributions underlying memory for the pitch of a tone under different experimental conditions.

distribution in Sections A and B is now decreased; so discrimination would become more difficult and errors more numerous. This would explain the increase in errors in Cond. 7 and 8 compared with Cond. 4. Further, the closer in time the included tone B is to the S tone A, the more will its distribution overlap with that of the S tone, and so the greater will be the decrement in performance. This would explain the greater increase in errors found when the included tone B is presented in the second serial position of the intervening sequence (Cond. 7) compared with the third serial position (Cond. 8). One should also expect from this theory that accuracy of performance in Cond. 1 (S and C tones identical in pitch, with this pitch not included in the intervening sequence) should depend very much on whether or not a tone a semitone removed from the S tone is included in the intervening sequence. Such an inclusion should produce a substantial increase in errors. Although there is no condition in this experiment which formally tests this prediction, it was found that in 8 of the sequences in Cond. 1, there was no such inclusion; and in 16 of the sequences, a tone a semitone removed from the S tone did indeed occur in the intervening sequence. Analysis of errors on this basis showed that where there was no such inclusion, errors were only 12.1% and where there was such an inclusion errors increased to 21.7%. Such a finding would be expected on the present theory.

In Fig. 3, d, tones both at the pitch of the S tone and also at the pitch of the C tone are included in the intervening sequence. This

causes both the effects shown in 3, b, and 3, c, to occur. Thus, the volumes of distribution in both Sections A and B are increased; and so a considerable difference between them is retained. Discrimination should therefore be relatively easy. This would account for the nullification of the damage to discrimination produced by inclusion of a tone at the pitch of the C tone in the intervening sequence (Cond. 7 and 8) by the inclusion also of a tone at the pitch of the S tone (Cond. 9 and 10). Further, a powerful serial position effect should be expected in this condition, since here the serial positions of the included tones at both Pitch A and at Pitch B are simultaneously reversed. And a statistically significant serial position effect is indeed found when Cond. 9 and 10 are compared. One should also expect from this theory that accuracy of performance in those sequences in which the S and C tones were the same in pitch, with a tone at that pitch included in the intervening sequence (Cond. 2 and 3), should depend on whether or not a tone a semitone removed from the S-tone pitch is also included in the intervening sequence. Although there is no condition which formally tests this prediction it was found that in 15 out of the 24 sequences in Cond. 2 and 3 there was no such inclusion and in 9 of the sequences such an inclusion did occur. Analysis of errors on this basis showed that where there was no such inclusion, errors were 2.4%, and where such a tone was included, errors rose to 3.7%. This increment in errors was in the direction predicted by the theory.

It is apparent from this study, as well as

from that described in Deutsch (1970b), that tonal pitch deteriorates rapidly in the presence of other tones. This finding has important bearing on theories of musical information storage. It has been suggested by several theorists (Cohen, 1962) that musical information (with particular reference to pitch information) is retained in its absolute form in an information store of limited channel capacity. However, the rapid deterioration of pitch information found in these experiments demonstrates that most of us are incapable of short-term storage of more than a minimal amount of absolute-pitch information, in a situation where other tones are also presented. Further, even when highly selected Ss are used, so that the average level of performance on this task is very high, it is demonstrated in the present experiment that the presence of a critical tone in the middle of a sequence can have a highly deleterious effect on the recognition of the first tone in that sequence. We may recognize that a certain tone had occurred in a sequence without being able to tell whether it was or was not the first tone in that sequence.

It must be concluded from these findings that we cannot store transitional probabilities between tones with any success. It is suggested instead that information concerning absolute pitch is rapidly discarded, and that we store pitch information in an abstracted or recoded form. This conclusion is quite consistent with everyday knowledge of musical information processing. We recognize a tune much more readily than we recognize what key it was played in. The same argument holds for recognition of harmonic se-

quences; since not only are these transposable, but also their component chords are invertible. A theory of how pitch information is abstracted has been described by Deutsch (1969).

REFERENCES

- BACHEM, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, **26**, 751-3.
- COHEN, J. E. Information theory and music. *Behavioral Science*, 1962, **7**, 137-163.
- DEUTSCH, D. Music recognition. *Psychological Review*, 1969, **76**, 300-307.
- DEUTSCH, D. Dislocation of tones in a musical sequence: A memory illusion. *Nature*, 1970, No. 5242, 226. (a)
- DEUTSCH, D. Tones and numbers: Specificity of interference in short term memory. *Science*, 1970, **168**, 1604-1605. (b)
- HARRIS, J. D. The decline of pitch discrimination with time. *Journal of Experimental Psychology*, 1952, **43**, 96-99.
- KOESTER, T. The time error in pitch and loudness discrimination as a function of time interval and stimulus level. *Archives of Psychology, New York*, 1945, No. 297.
- MASSARO, D. W. Retroactive interference in short-term recognition memory of pitch. *Journal of Experimental Psychology*, 1970, **83**, 32-39.
- SHOWER, E. G., & BIDDULPH, R. Differential pitch sensitivity of the ear. *Journal of the Acoustical Society of America*, 1931, **3**, 275-287.
- WICKELGREN, W. A. Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1966, **72**, 250-259.
- WICKELGREN, W. A. Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, 1969, **6**, 13-61.

(Received August 23, 1971)