Canadian Acoustics /
Acoustique canadienne
1988, 16 (3):3-13

OCTAVE DISCRIMINATION: TEMPORAL AND CONTEXTUAL EFFECTS

Lola L. Cuddy and Peter A. Dobbins

Department of Psychology
Queen's University at Kingston
Kingston, Ontario, Canada K7L 3N6

ABSTRACT

Discrimination of the musical octave was studied using the psychophysical method of constant stimuli. Stimuli were two successive pure tones whose ratio varied in discrete steps from 100 cents below to 100 cents above the physical octave of 1200 cents. Listeners judged whether each pair was flat or sharp with respect to a correctly tuned octave. Two measures were estimated for each of ten listeners. The first was a measure of sensitivity, the difference limen. The second was the subjective criterion for the octave, or that tuning judged equally likely to be sharp or flat. For two tones in immediate succession the results were in accord with previous investigations: the discrepancy between the subjective octave and the physical octave was typically in favour of a "stretched" subjective octave about 20 cents wider than the physical octave. However, the magnitude of the stretch decreased, and sensitivity of discrimination increased, when the two tones were separated either by silence or by two musically related tones—in this case, the notes of the equal-tempered major triad. The results suggest that the criterion for the subjective octave is influenced by context and by the listener's strategy. Thus, in music, a flexible choice of tuning criteria may be desirable.

SOMMAIRE

La discrimination de l'octave musicale a fait l'objet d'une étude psychophysique au moyen de la méthode des stimuli constants. Les stimuli étaient constitués de deux sons purs successifs dont le rapport variait par pas discontinus de 100 cents en-deça à 100 cents au-delà de l'octave physique de 1200 cents. Les sujets jugeaient si chaque paire était bémol ou dièse en fonction d'une octave correctement accordée. Deux mesures ont été obtenues auprès de chacun des dix sujets. La première était une mesure de sensibilité, soit le seuil différentiel. La seconde était le critère subjectif pour l'octave ou l'accord jugé de façon équi-probable comme étant dièse ou bémol. Pour deux sons se suivant immédiatement, les résultats étaient en accord avec les études antérieures: la différence entre l'octave subjective et l'octave physique était généralement en faveur d'une octave subjective "étendue" d'à peu près 20 cents par rapport à l'octave physique. Cependant, l'ampleur de l'extension diminuait et la sensibilité de la discrimination augmentait lorsque les deux sons étaient séparés soit par un silence soit par des sons interreliés musicalement—dans ce cas, les notes de la triade majeure également tempérée. Ces résultats suggèrent que le critère de l'octave subjective est partiallement influencé par le contexte et par la strategie de l'auditeur. Ainsi, dans la musique, un choix flexible du critère accordable peut être désirable.

INTRODUCTION

Experiments have shown that the subjective musical octave is an interval somewhat wider than the physical octave of frequency ratio 2 to 1. That the subjective octave is stretched with respect to the physical octave is a finding that holds across variations in fundamental frequency, timbre and intensity (Sundberg & Lindqvist, l973; Ward, 1954), across levels of musical experience, and different methods of measurement (Dobbins & Cuddy, l982), and across cultures (Dowling & Harwood, 1986, Ch. 4). The extent of the stretch—about one-third of a semitone on the average—is nevertheless quite variable among listeners. In a previous report (Dobbins & Cuddy, l982), we noted a dispersion of individual estimates over more than a semitone, and commented upon the difficulties that would attend music performance if individual tuning preferences had to be taken into account. Obviously, the setting of tuning standards for performance involves a fair degree of compromise.

Even within a single listener, preferences vary, and the selection of a given musical intonation may reflect a compromise between conflicting influences (e.g., Hall & Hess, 1984; Terhardt, 1984). Makeig and Balzano (1982) suggest a number of possible factors influencing preference. One is that tuning preferences reflect the intentional effort of the performer to produce a variety of aesthetic effects. Another is musical context. Terhardt and Zick (1975) found that, for melodic passages, listeners preferred a physical intonation that was stretched as opposed to normal (in this case, equal-tempered) or contracted intonation. For harmonically rich passages involving complex spectral patterns, however, stretched tuning was judged to be the poorest of the three types studied. To account for these results, Terhardt and Zick suggested that interaction patterns between simultaneous harmonic components produce a subjective stretch (a concept included in later theoretical developments, e.g., Terhardt, 1979). Thus, adding a physical stretch would result in a percept that was "out-of-tune".

Specifically with respect to the octave, Ward (1954) reported that simultaneous (harmonic) presentation of component tones yielded a narrower estimate of the subjective octave than did successive (melodic) presentation of tones. Terhardt (1978) also noted this effect, and recent unpublished results in our laboratory were consistent: the harmonic frequency ratio 2:1 was judged a reasonably "correct" octave by listeners, while the melodic presentation of the same ratio sounded "flat" by comparison.

The musical implication is that there may be two standards for the subjective octave, both capable of exerting an influence on preferred tuning in musical contexts. When melodic judgments are made with reference to harmonic standards, a preference for stretched tuning will result. A wider melodic interval is needed to match, perceptually, a harmonic interval of the same physical ratio. If, on the other hand, melodic judgments are made with reference to melodic standards, they will be stretched to a lesser extent, or not at all. It is difficult, of course, to predict which standard will prevail in any given instance, but it seems reasonable to expect that the choice will be influenced by the availability of specific harmonic or melodic cues in the musical context itself.

The present experiment was an exploration of some of these musical notions within the experimental control and abstraction of a psychoacoustical paradigm. The purpose of the present experiment was to examine octave judgment in the presence of musically related tones--the tones of the well-tempered major triad. The melodic triad is a salient cue to tonal structure (Cuddy & Badertscher, 1987). It was thought that these tones, presented melodically, might therefore cue a melodic standard of judgment for the octave.

The psychophysical method was the method of constant stimuli and the method of parameter estimation was adapted from a solution proposed by Olson & Ogilvie (l972) and verified for the octave discrimination paradigm (Dobbins & Cuddy, l982). Two performance measures were estimated: the difference threshold, reflecting sensitivity of octave resolution; and the magnitude of the subjective octave, or that tuning equally likely to be judged sharp or flat. The Olson and Ogilvie model incorporates the traditional assumptions of signal detection theory as well as the assumption that the subjective metric is linearly related to the physical scale. The psychophysical function can be used to estimate the difference threshold and point of subjective equality in physical units (Dobbins & Cuddy, 1982). Another point made in our previous report is that a good fit of the model to the data implies a continuous function relating psychophysical and physical dimensions; thus a categorical model of musical interval perception (e.g., Siegel & Siegel, 1977) was not supported.

There were four experimental conditions. On each trial for each condition two test tones approximating the physical octave were presented, and the listener was asked to rate the sharpness or flatness of tuning. In the first condition, the two tones were presented in immediate succession as in our earlier study. In the second condition, the two test tones were separated by two interpolated tones. The interpolated tones, along with the first test tone, formed an equal-tempered major triad. The third and fourth conditions were experimental controls. In the third condition, the test tones were separated by a silent delay equal in duration to that of the interpolated tones of the second condition. In the fourth condition, the interpolated tones occurring between the test tones were mistuned with respect to the equal-tempered scale.

An example of the sequence of events for each of the four conditions is given in musical notation in Figure l. The example is given with the octave C5 to C6 as the test octave in all conditions (one octave below the actual frequency range used in the experiment). The first row of notation represents the events of a trial for Condition 1. The second row shows the interpolation of a descending major triad, in its first inversion, between test tones (Condition 2). The third row shows a silent delay, or musical rests, between test tones (Condition 3). The fourth row shows mistuned auditory material between test tones (Condition 4). Mistuning is represented by crosses over the notes.

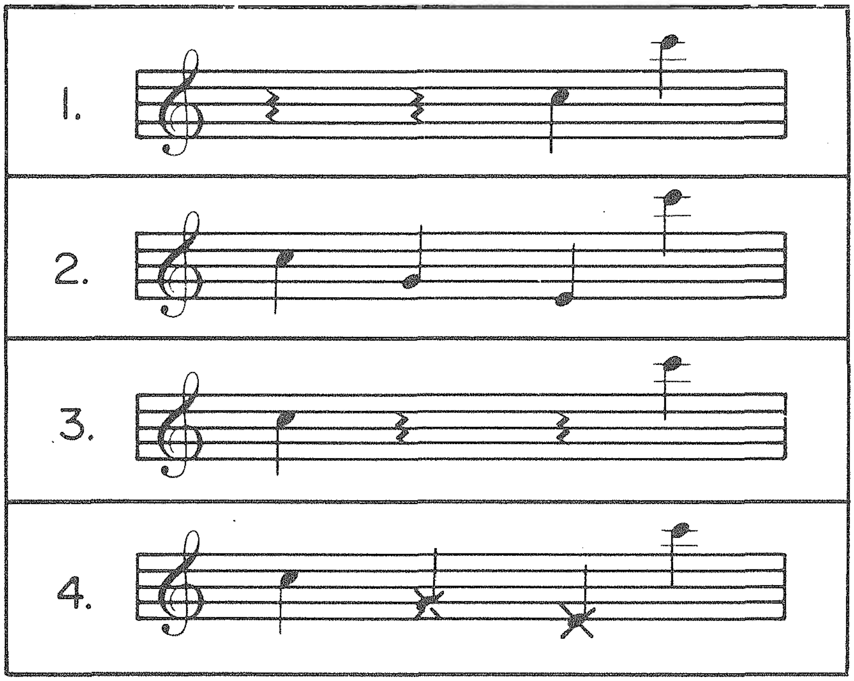


FIGURE 1. The sequence of events for each of the four experimental conditions. The musical notation is given one octave below the actual frequency range used.

METHOD

1. Listeners

The listeners were ten undergraduate students ranging from 20 to 24 years of age. All were either Bachelor of Music students or had taken university-level music courses, and all had achieved the Grade VI level of piano performance for the Royal Conservatory of Music, Toronto. Most were also equally proficient in at least one other instrument. Listeners were paid for their voluntary participation.

2. Apparatus

Stimulus tones were produced by a General Radio 1161-A coherent decade frequency synthesizer under the control of a Digital PDP8/I computer. All tones were sinusoidal with linear rise and fall times of 30 msec. The output signal was attenuated to 50 dB SPL by a Hewlett-Packard 350B attenuator set, and in turn fed to a Soundcraftsman 20-12 audio frequency equalizer, where levels were adjusted so that all tones were of equal subjective loudness.

Listeners were tested individually in a sound-isolated booth, and heard the tones binaurally through MB-300 headphones. Listeners made their responses by closing one of four touch-sensitive illuminated switches, which were labelled "very flat", "flat", "sharp", and "very sharp" in their left-to-right order on the response panel.

3. Stimuli

Each trial contained a 1-sec standard tone (S) and a 1-sec comparison tone (C). With the offset of C, a response light was illuminated on the switch panel, at which point listeners had 3 sec to respond before the next trial began.

For each trial, S was assigned one of three frequencies: 987.8 Hz, 1046.5 Hz, and 1108.7 Hz. These values correspond to B5, C6, and C#6, respectively, according to equal temperament rules and with A4 at 440 Hz. Within each block of trials, each value of S was paired with each of 11 different C-values, ranging from 1100 to 1300 cents above S in 20-cent gradations. (A cent is the 100th logarithmic division of an equal-tempered semitone.) Thus, S-C intervals ranged from one semitone below to one semitone above the physical octave (PO).

S-C intervals were presented within four contextual conditions as described in the text above and in Figure 1. The duration of each interpolated tone (Conditions 2 and 4) was 1 sec and the silent delay of Condition 3 was 2 sec. The two interpolated tones of Condition 2 formed along with S an equal-tempered descending interval of five semitones followed by an equal-tempered descending interval of three semitones. The two interpolated tones of Condition 4 were tuned 130 cents sharp and 160 cents flat from their respective counterparts in Condition 2.

A visual metronome was employed to eliminate temporal uncertainty on the part of the listeners. During each trial, a 0.5-sec light on the response panel flashed four times at 1-sec intervals, so that for Conditions 2, 3, and 4 the onsets of S and C would coincide with the onsets of the first and fourth flashes respectively. However, in Condition 1 trials, S and C onsets coincided with the third and fourth flashes, so that listeners would expect C immediately to follow S when the first two flashes were not accompanied by tones.

Each of the 33 S-C pairs (3 S frequencies by 11 S-C intervals) was presented three times under each of the four contextual conditions for a total of l32 trials per block. A basic block was permuted to produce three blocks of randomly ordered trials.

4. Procedure

All listeners were instructed to evaluate each S-C interval with respect to an octave, selecting the most appropriate response from the four available alternatives. No "correct" or reference octave was demonstrated to the listeners before or during the experiment.

Each session began with a practice block of 24 trials. This block contained all four contextual conditions, all three S frequencies, and six (1100, 1140, 1180, 1220, 1260, and 1300 cents) of the 11 S-C intervals. After their assurance that the instructions were understood, listeners heard the three experimental blocks in an order randomized for each listener, with 5-min breaks between blocks.

RESULTS

The data for each condition within each block for each listener were entered into a 11 x 4 confusion matrix (11 C-values by 4 response alternatives). A maximum-likelihood solution for parameter estimation was derived from a model proposed by Olson and Ogilvie (I972) for the method of constant stimuli with two or more response categories and implemented through the program MALCOS. The solution provides two theoretically independent measures of each listener's performance, reported in physical units of cents. The measures are the difference limen (DL), based upon one-half of the interquartile range of the response distribution, and the subjective octave (SO), or that tuning judged equally likely to be flat or sharp.

The data averaged across blocks and across listeners are given in Table 1. Table 1 shows for each contextual condition the DL in cents, the discrepancy in cents between the subjective octave and the physical octave, and the between-subjects standard error for each measure.

TABLE I. Mean DL and SO-PO in cents and between-subject standard errors (SE) for both means.

	Contextual Condition			
	1	2	3	4
Mean DL	42.1	29.3	32.5	44.7
Mean SO-PO	22.2	1.9	4.4	22.3
SE between-subjects for DL	4.6	3.9	5.0	4.4
SE between-subjects for SO-PO	8.4	4.6	6.2	9.1

The data show similar results for Conditions 1 and 4, and similar results for Conditions 2 and 3. Analysis of variance confirmed this observation. The main effect of condition was significant for both the DL ($F(3,24) = 8.17$, $p < .001$) and SO ($F(3,24) = 4.57$, $p < .025$) measures. Orthogonal contrasts within the effect of contextual condition indicated that the conditions in which S and C tones were presented in immediate succession or were separated by two mistempered tones did not differ significantly, nor was there a significant difference between conditions where well-tempered tones or a 2-sec delay separated S and C tones. But the latter two conditions, compared with the former, produced greater sensitivity of discrimination and less bias in favour of a stretched octave (for DL, $F(1,6) = 24.43$, $p < .005$; for SO, $F(1,6) = 15.57$, $p < .01$). Between-subjects standard error for the SO measure was also reduced in the latter conditions compared to the former. The main effect of blocks was significant, indicating a slight but significant linear trend towards greater sensitivity and reduced bias (for DL, $F(1,6) = 28.87$, $p < .005$; for SO, $F(1,6) = 6.32$, $p < .05$). There was no significant interaction, however, between blocks and conditions; the differences attributable to contextual condition held for each block of trials.

Condition 1 used the temporal parameters of earlier studies. The mean discrepancy of 22.2 cents in favour of a stretched SO in Condition 1 is close to the estimate of 22.5 cents for musically trained listeners with the method of constant stimuli (Dobbins and Cuddy, I982) and the estimate of 21.4 cents reported by Ward (I954) for a fixed S tone at 1180 Hz with the method of adjustment. In Condition 1, eight of the ten listeners were consistently biased toward a stretched SO ($p < .05$), and a ninth listener was similarly biased on two of the three experimental blocks.

Of the 120 matrices computed, 110 were found to fit the Olson and Ogilvie (I972) model according to a chi-square test of the difference between expected and observed response proportions, with p set at .01. Only three pairs of DL and SO estimates could not be obtained, due to response inconsistencies that occurred in the more difficult conditions. (Because of these missing values, degrees of freedom for the statistical tests reported above were reduced accordingly.) Inspection of the rejected confusion matrices revealed that listeners occasionally made reversal errors along the decision axis, e.g., would use the "very sharp" category for a comparison that was physically very flat. Such errors could be evidence of categorical perception: the listener detects the mistuning but the direction of mistuning is not discriminated (Siegel & Siegel, 1977). However, the reversal errors were relatively infrequent and could just as easily be attributed to lapses of attention.

ADDITIONAL EXPERIMENTS

Additional tests yielded the following supplementary information:

1. Three highly practiced listeners (the musically trained listeners from Dobbins and Cuddy (1982)) were also tested under all conditions of the main experiment. Discrimination was slightly more sensitive than the main experiment (average DL = 32.1 cents) and showed less variation with condition. The SO-PO discrepancy was 31.2 cents for the average of Conditions 1 and 4, but only 11.4 cents for the average of for Conditions 2 and 3. This pattern of results for SO-PO discrepancy replicates the pattern of the main experiment.

2. Four highly practiced listeners (different from the above) were tested under conditions where the S-C delay varied between 0 and 8 sec. No conditions involving interpolated tones were included. For three listeners, there was some evidence of improved sensitivity or reduced bias with delays of 1 or 2 sec. The effects were smaller and less reliable than in the main experiment, however. Ward (1954) earlier reported no effect from interpolation of a 1-sec silent delay between tones. These results suggest that the magnitude of the delay effect in the main experiment was due, at least in part, to the embedding of the delay condition with the condition containing the major triad.

   In a subsequent group discussion, the listeners from the main experiment agreed that the presence of the major triad seemed to make judgment easier. Many suggested that during the silent delay, therefore, they attempted to recreate a triadic structure from memory. In the presence of mistuned tones it was difficult to produce this auditory image. What the listeners emphasized was that the silent delay was accompanied by an active attempt to retrieve a musical cue.

3. Two listeners, highly trained both in music performance and with the psychophysical task, were asked to tune the subjective melodic octave with the psychophysical method of adjustment. The physical parameters were as in the main experiment, but here, on each trial, the listener adjusted a variable comparison tone to be one octave above the standard. On preliminary cuing trials, listeners heard a standard PO presented either harmonically or melodically and attempted to tune a melodic octave to match the standard. The octave to be tuned was located six semitones above the standard. Then, on half the experimental trials, listeners were asked to use the memory for a harmonic octave as the standard for tuning the melodic octave; on the remaining half, to use the memory for a melodic octave. For the first listener, the mean SO-PO discrepancy was 36.7 cents and 15.1 cents for harmonic and melodic standards respectively. For the second listener, the mean SO-PO discrepancy was 16.5 cents and -5.7 cents for harmonic and melodic standards respectively. Although the absolute values of the

discrepancy were quite different for the two listeners, the difference between the two conditions was about 20 cents for both listeners.

DISCUSSION

Our results suggest that listeners are capable of adopting two different criteria for the subjective melodic octave, criteria that differ by about 20 cents. The main experiment showed that for two temporally contiguous tones in the frequency region studied, the subjective octave was stretched from the physical octave. However, octave discrimination was enhanced and bias in favour of a stretched octave was reduced when the octave was presented in a tone sequence forming a well-tempered major triad. Listeners also tended to be in greater agreement when the triad was present, as evidenced by the lowered between-subject variability. These results suggest that it was the pattern of the major triad that influenced judgment. Had listeners merely tried to judge the interval between the third and fourth note presented in Condition 2, the preferred stretch for an interval of 20 semitones should have been obtained.

Convergence to the physical octave was also obtained in the main experiment when the two tones of the octave were separated by a 2-sec silent gap. The alteration in performance cannot be attributed merely to the temporal separation of the octave tones; with the gap filled with mistuned tones, there was no change in discrimination or bias. The results for the silent-delay condition, in conjunction with listeners' self-reports and the results of a supplementary experiment, suggest that listeners adopted a strategy of rehearsing the major triad during the silent interval. Such rehearsal was prevented, however, by the presence of mistuned interpolated tones.

In the main experiment, therefore, the equal-tempered major triad appears to have cued a reference standard for the subjective octave that was applied locally on specific trials (those generated from Conditions 2 and 3). Supplementary experiments suggested that such a standard was unlikely to be adopted, even during a silent delay, unless the listener was especially instructed to try to match a remembered melodic octave. It appears that a stretched subjective octave is the natural standard to adopt. However, it may be replaced by other standards in a rather flexible manner.

The good fit of the Olson and Ogilvie (1972) model again implies a continuous function for octave discrimination. Consequently, we reject categorical models that suggest musicians are unable to distinguish flat tuning from sharp. However, the concept of an octave category is useful where it defines a range of physical values that are acceptable members of the category. For individual listeners, there is probably a best or prototypic member of the category, exemplified by natural tuning preferences. (See Krumhansl (1983) for a similar description applied to the the tonal organization of musical pitches.)

Two contemporary models of pitch perception have incorporated the phenomenon of octave enlargement. Both models locate the origin of the stretch in the auditory system itself, not in familiarity with physically stretched intervals. For Terhardt (1978, 1979) octave enlargement is a by-product of speech perception. Due to nonlinearities in the auditory system, the harmonics present in voiced speech sounds are stretched, and these stretched percepts become the standards for judging musical intervals, including the octave. For Ohgushi (1983) the stretched octave is the result of systematic bias in the temporal pattern of firing in auditory neurons. Of the two, Terhardt explicitly attributes the stretch to the auditory processing of simultaneous harmonic components, and argues that successive tones provide an unaffected, direct, measure of pitch (Terhardt, 1978). Thus, though the stretch may arise through primary auditory processes, there is also a plausible basis for a second reference standard for the subjective octave.

The musical implications of our data must be approached with caution, because we need to know much more about how perception of isolated intervals applies to larger musical contexts of phrases and other musical structures. Nevertheless, some interesting ideas emerge. First, our data, and those of other investigators, suggest that a natural preference for a stretched interval will usually prevail. But, second, these preferences may be modified by context or by strategy. Certain musical texts may elicit these secondary preferences, as Terhardt and Zick (1975) have suggested. There are, as well, musical instances where maintaining consistent intonation may require special strategies. For example, suppose a melody in an unaccompanied soprano voice is followed by a voice in a different register taking its pitch cue from the soprano. Singers may prefer to preserve a stretched scale, but the idiosyncratic nature of the stretch will create difficulties with overall consistency. Therefore, a choirmaster/organist of our acquaintance, John Gallienne of St. George's Cathedral, Kingston, insists on training choristers to adopt a standard for tuning the subjective octave that, in his words, is "hard, cold and narrow". This percept arises from the melodic presentation of the frequency ratio 2:1, and Gallienne argues that it produces a more reliable standard than the naturally preferred stretch.

In conclusion, listeners are capable of modifying judgment of the subjective octave in the direction of the physical octave under appropriate cuing or rehearsal. Under such conditions, agreement among listeners is increased. Where convergence of intonation is a musical goal, it may make sense to emphasize listening to, and matching, melodic intervals. Generally, the results support the musical notion that, where feasible, a flexible but systematic choice of tuning criteria is desirable.

ACKNOWLEDGMENTS

This research was supported by grants to the first author from the Natural Science and Engineering Research Council of Canada and the Advisory Research Council of Queen's University. We thank Geoffrey Anderson, Jamie North, Tom Mawhinney, and Karen Smith for technical assistance, and John Gallienne and M.G. Wiebe for valuable discussions.

REFERENCES

Cuddy, L.L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: some comparisons across age and levels of musical experience. *Perception and Psychophysics*, *41*, 609-620.

Dobbins, P.A., & Cuddy, L.L. (1982). Octave discrimination: An experimental confirmation of the "stretched" subjective octave. *Journal of the Acoustical Society of America*, *72*, 411-415.

Dowling, W.J., & Harwood, D.L. (1986). *Music Cognition.* New York: Academic.

Hall, D.E., & Hess, J.T. (1984). Perception of musical interval tuning. *Music Perception*, *2*, 166-195.

Krumhansl, C.L. (1983). Perceptual structures for tonal music. *Music Perception*, *1*, 28-62.

Makeig, S., & Balzano, G. (1982). Octave tuning—two modes of perception. *Proceedings of the Research Symposium on the Psychology and Acoustics of Music*. Lawrence: U. of Kansas.

Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon. *Journal of the Acoustical Society of America*, *73*, 1694-1700.

Olson, C.L., & Ogilvie, J.C. (1972). The method of constant stimuli with two or more categories of response. *Journal of Mathematical Psychology*, *9*, 320-338.

Siegel, J.A., & Siegel W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics*, *21*, 399-407.

Sundberg, J., & Lindqvist, J. (1973). Musical octaves and pitch. *Journal of the Acoustical Society of America*, *54*, 922-929.

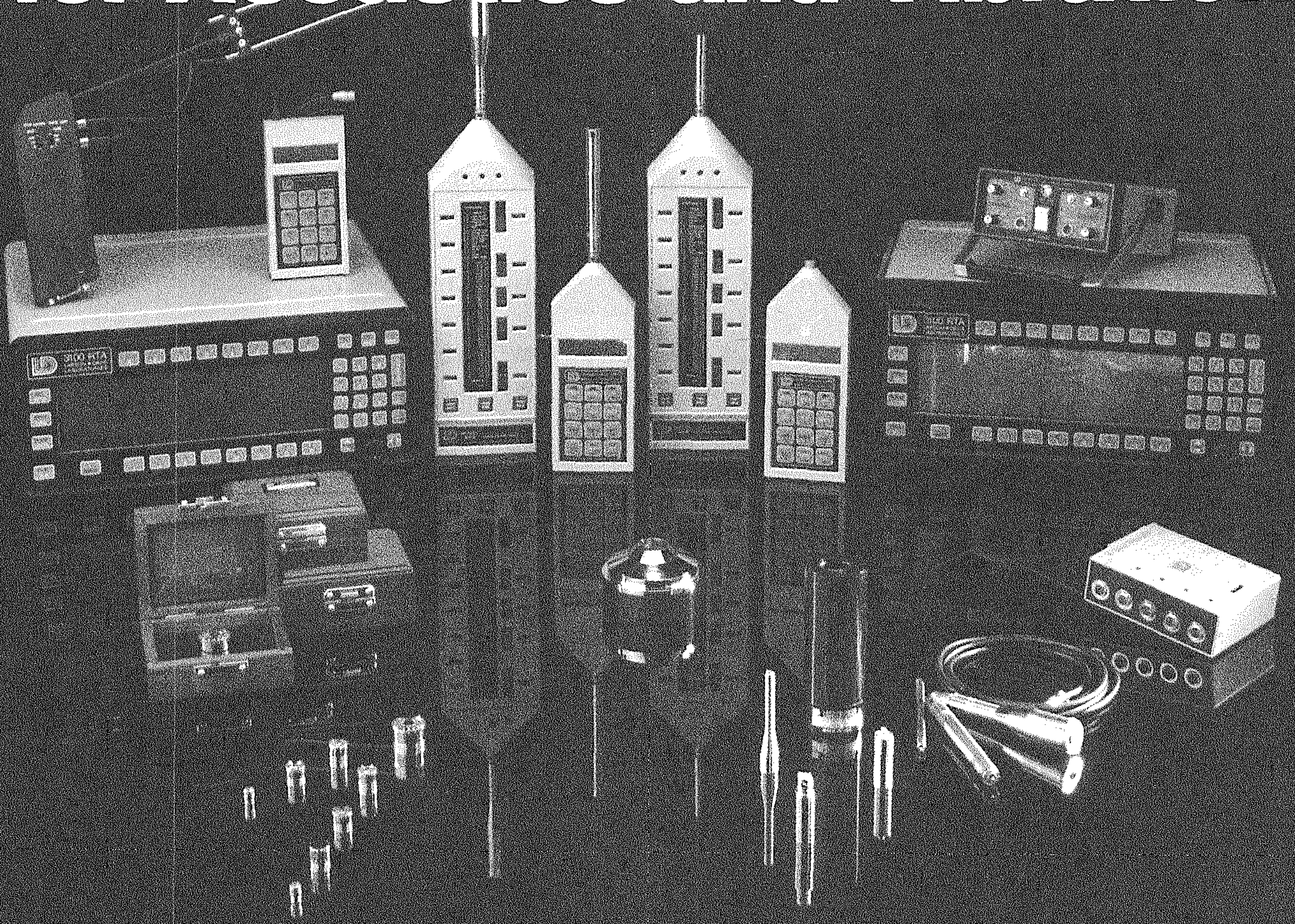
Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception and Psychophysics*, *23*,, 483-492.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research*, *1*, 155-182.

Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception*, *1*, 276-295.

Terhardt, E., & Zick, M. (1975). Evaluation of the tempered tone scale in normal, stretched, and contracted intonation. *Acoustica*, *32*, 268-274.

Ward, W.D. (1954). Subjective musical pitch. *Journal of the Acoustical Society of America*, *26*, 369-380.



3100 RTA
3100 RTA