UNIVERSITY OF ILLINOIS PRESS

Council for Research in Music Education

Musical vs. Psychoacoustical Variables and Their Influence on the Perception of Musical Intervals
Author(s): Gerald J. Balzano
Source: *Bulletin of the Council for Research in Music Education*, No. 70 (Spring, 1982), pp. 1-11
Published by: University of Illinois Press on behalf of the Council for Research in Music Education
Stable URL: http://www.jstor.org/stable/40317707
Accessed: 17/01/2011 17:19

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=illinois.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

University of Illinois Press and *Council for Research in Music Education* are collaborating with JSTOR to digitize, preserve and extend access to *Bulletin of the Council for Research in Music Education.*

http://www.jstor.org

article

Musical vs. Psychoacoustical Variables and their Influence on the Perception of Musical Intervals

Gerald J. Balzano, University of California, San Diego

Psychoacoustics & Music

The term psychoacoustician is probably a misnomer for those who are seriously interested in studying the perception of musical objects and events. Psychoacoustics, properly speaking, is a branch of psychophysics, the latter being a study of those elementary sensations that can be put into one-to-one correspondence with simple physical variables. It is within the scope of psychoacoustics to investigate subjective pitch as a function of objective frequency, and such experiments have been performed (Stevens, Volkmann, & Neumann, 1937; Stevens & Volkmann, 1940). But if we examine any of the music-theoretic literature, pitch is seen as a much richer domain, giving rise to a host of purely musical functions and effects, and not reducible to—perhaps even independent of—a specific mapping of the frequency domain. If the perceiver of music can be said to perceive anything properly musical at all, the simple unidimensional mapping proffered by psychoacoustics must be gone beyond and the richer domain of music-theoretic pitch relations must be confronted directly.

This directive is stated with some urgency, as there is already a considerable body of literature on the perception of unidimensional pitch, some of it even purporting to study such fundamentally musical objects as the 12 basic musical intervals. A relatively recent example is the experiment by Plomp, Wagenaar and Mimpen (1973), where identification errors in the perception of briefly sounded musical intervals are broken down and analyzed in such a way as to lend credence to the unidimensionality thesis. To be sure, identification errors occur more frequently to interval categories that are close in width to a sounded interval—we would certainly be surprised if this weren't true—but there are other patterns in the results of Plomp *et. al.* that are either glossed over as "slight exceptions" (p. 106) or not mentioned at all. These patterns are of two basic types (see Table 1 if a guide to the nomenclature is needed):

Table 1

Summary information on the 12 basic intervals

Interval	*Frequency Ratio*	*Number of Semitones*
minor 2nd (M12,m2)	16:15	1
major 2nd (MA2;M2)	9:8	2
minor 3rd (MI3,m3)	6:5	3
major 3rd (MA3,M3)	5:4	4
perfect 4th (PE4,p4)	4:3	5
tritone (TRI,t)	7:5	6
perfect 5th (PE5,P5)	3:2	7
minor 6th (MI6,m6)	8:5	8
major 6th (MA6,M6)	5:3	9
minor 7th (MI7,m7)	16:9	10
major 7th (MA7,M7)	15:8	11
octave (OCT,p8)	2:1	12

(al) Despite the fact that the intervals p4 and p5 are two semitones (*s*) apart, and each only one semitone from the tritone (6*s*), the data clearly show p4 and p5 to be confused with one another more than either is with the tritone.

(a2) Identification errors to m2 stimuli (1*s*) occur more often in the M7 (11 *s*) category than any other category except M2.

(b) The interval M2 is more often confused with m2 than with m3, though equidistant from both. The m3 is more often confused with M3 than with M2. In general, 2nds are more confusable with other 2nds, 3rds with other 3rds, and so on.

In (al) and (a2) the pairs p4-p5 and m2-M7 bear a special musical relation. Musicians would say each is a pair of inverses; we will say they are *chroma-equivalent* pairs, borrowing Revesz' (1954) term chroma for the cyclic aspect of musical pitch, displayed in Figure 1. As for (b), we will describe the relationship between intervals like m2 and M2 as scale-step equivalence. That scale-step equivalence significantly affects the apparent similarity of intervals is never alluded to by Plomp *et.al.*, but its effects are clearly present in every relevant comparison in the data.

While chroma and scale-step are musically both interesting and important properties of intervals, they are *not* psychophysical properties. Yet it would be a serious error to assert that they are any less real because of this; their reality is amply attested to by the data.

In the remainder of this paper a more complete empirical documentation of the perceptual reality of chroma and scale-step is provided, as they influence the recognizability and similarity of musical intervals. The paradigm employed differs substantially from the error/confusion methodology of Plomp *et.al.* Instead of inducing errors, the strategy was to monitor more closely the time course of the process of recognizing a musical interval, measuring response latencies in several simple tasks.

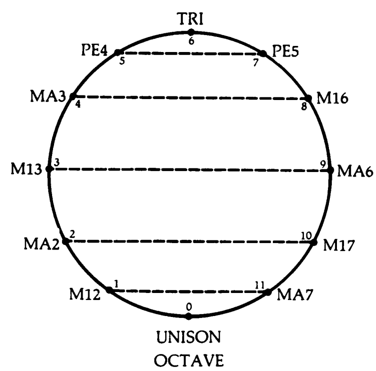

Figure 1. Chroma circle of musical intervals. Chroma-equivalent intervals (inverses) are connected with dashed lines.

Summary of Methodology

In assessing the perceived similarity of, say, a major 2nd and a minor 3rd, we are interested in how soon a listener in the aural presence of the first knows (s)he is *not* hearing the second, and vice versa. Put another way, how early in the process of encoding a heard major 2nd is the information potentially available that it is *not* a minor 3rd? To get at this, a question was put to listeners at the start of each trial, in the form of a visually presented musical interval name such as MA2. This was understood as a shorthand for "Is the interval you are about to hear a MA2?" Then, 1.6 seconds after the onset of the question, a musical interval stimulus was played from a tape, and the listener's task was to respond either affirmatively or negatively by pressing one of two keys as soon as the answer to the question was evident. Figure 2 depicts the events of an experimental trial. For the most part this response could be performed quite rapidly by our listeners, requiring for most subjects about 500-900 milliseconds from sound onset to keypress. For example, in the first experiment the mean latency or reaction time (RT) was 797 msec., with a 13 percent error rate. Subjects' comments following the errors suggested that most could have been eliminated had they waited longer to respond. Subjects were nonetheless encouraged to respond briskly, as this renders the latency measure more sensitive than it would be if subjects were asked to wait until they were absolutely sure. In general, it appeared that those question-stimulus combinations leading to errors were also those that would have led to long latencies in any case: latency-error rate correlations were uniformly high in all experiments.

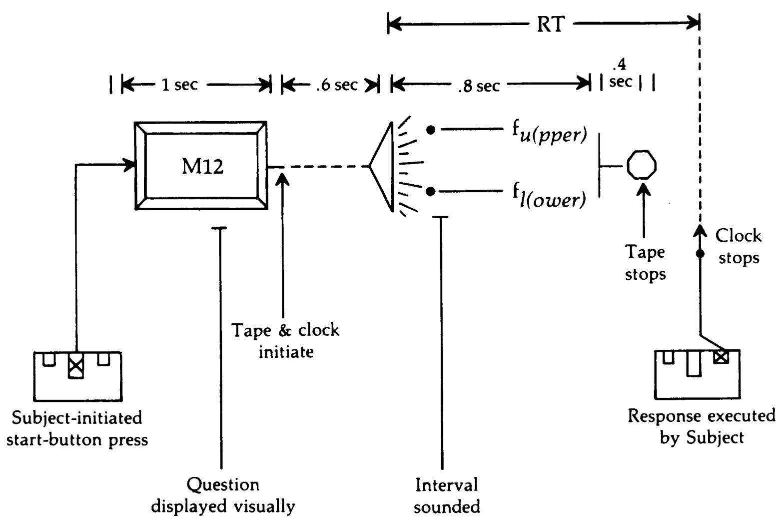

Figure 2. The events of an experimental trial.

In all of the experiments the stimuli were harmonic intervals composed of pure tones. For the first experiment, both questions and stimuli were drawn from the set of 12 basic musical intervals from m2 to p8. The 144 question-stimulus combinations are displayed in matrix form in Table 2. Elements on the diagonal correspond to cases where question and stimulus "match" (Same trials), off-diagonal elements correspond to mismatches (Different trials). Both types occurred equally often in the first experiment, through interest was primarily in Different trials.

Multidimensional Scaling and Chroma

A data matrix such as Table 2 potentially contains much more structure than can be revealed from direct inspection. The tech-

(St_j)

		M12	MA2	M13	MA3	PE4	TRI	PE5	M16	MA6	M17	MA7	OCT
	M12												
	MA2												
	M13												
	MA3												
	PE4				$M_{i,j}$								
(Q_i)	TRI												
	PE5												
	M16												
	MA6												
	M17												
	MA7												
	OCT												

Table 2

A 12 x 12, Question(Q)-by-Stimulus(St), interval x interval matrix

nique of nonmetric multidimensional scaling (Kruskal, 1964 a,b; Shepard, 1962 a,b) displays this structure in a more accessible form by translating the objects under study (here, the 12 intervals) into points in space and by translating the Different RT values into distances, using only the rank order of the RT values to determine relative point placements. Thus, if a long time is required to distinguish a major 2nd from a minor 3rd, the M2 and m3 points would be placed close together in multidimensional space to indicate their perceptual similarity or closeness to one another.

That reference is made to a multidimensional space in no way implies that the dimensionality of the space is known in advance. On the contrary, this is very much an empirical question: some patterns in similarity matrices will admit of a representation with all points lying on a line relative to one another without undue violence to the data. If musical intervals are indeed unidimensional percepts, then we ought to expect such a pattern emerging from the similarity matrix. Within the multidimensional scaling algorithm it is possible to obtain solutions where the points are free to vary in as many or as few dimensions as we please. For each solution, a statistic known as stress is computed, the value of which informs us just how well the pattern of similarity data is captured in a space of a particular number of dimensions. Of course, stress will always decrease as the points are allowed to vary in more dimensions; ultimately n points can always be fit perfectly (stress = .000) in a space of n-1 dimensions. But in general, adding dimensions will eventually fail to capture significant additional structure (produce trivial decreases in stress), and the point where this occurs is the best estimate of the true dimensionality of the space systematically implied by the data. Usually stress values above about .15 or .20 indicate unacceptable fits.

A word about subjects: 18 different musically inclined listeners from Stanford University supplied a total of 35 sessions to this experiment. Each session consisted of 264 trials, and was usually completed within 50 minutes.

The best-fitting undimensional solution for the data is given in Figure 3. It is clear that the sequence of points does not even remotely resemble the ordering that would be anticipated on the basis of semitone width. Rather, the data more closely resemble an ordering based on ratio simplicity (see Table 1), with the perfect intervals of the octave, fifth and fourth at one extreme and the seconds, sevenths, and tritone at the other.

m2 M2 M7 t m7 m6 m3 M6 p4 M3 p5 p8

Figure 3. Best-fitting unidimensional configuration for the 12 intervals.

It is possible to preload the multidimensional scaling program with a suggested configuration. An attempt to force a semitone-width ordering in this way had the following effects: (a) the algorithm rearranged several of the points before finding a (local) minimum stress value, disrupting the inputted semitone ordering considerably, (b) the obtained solution, with its permuted points had an unacceptably high stress value of .363. The upshot is that a unidimensional scale based on semitone width is not even a good first approximation to the data. And while the ratio simplicity ordering, recovered without preloading or forcing, gave a considerably better fit, its stress value (.269) was still too large to be acceptable. As a consequence, higher dimensional solutions were sought to account for the results.

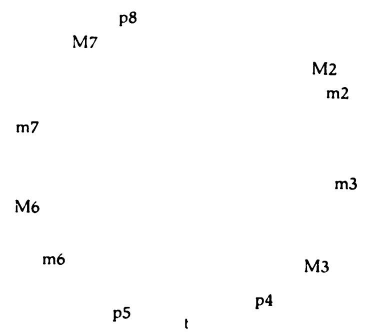

Figure 4. Best-fitting two-dimensional configuration for the 12 intervals.

The two-dimensional solution is displayed in Figure 4. Stress has been considerably reduced, to a value of .160, but even more significant is the arrangement of the points. There is no mistaking the resemblance of this configuration to what has been called the chroma circle by such writers as Shepard (1964), Bachem (1950), and Revesz (1954). It is the cyclic, and not the rectilinear, aspect of pitch structure that is most clearly revealed in the way our musical listeners perceived (and misperceived) the set of musical intervals. This type of pattern was remarkably robust, appearing in analyses of both error rates and latency data, and across several experiments besides the one reported here (see Balzano, 1977 for a more complete discussion)[1]. Musical intervals may be realizable by variation of a single physical parameter, but this does not mean that they are unidimensional percepts.

Scale-Step Categories

A further characteristic of the 12 intervals, one that we have referred to earlier, is that they fall into eight distinct higher-level categories, a feature directly reflected in the nomenclature. Thus the intervals that are one and two semitones wide are both referred to as seconds, those of three and four semitones are thirds, and so on up to the 8th high-level category given in the name octave. This suggests that not all semitone distances are truly equal, since some of them reflect a distinction at two levels (as between major 2nd and a minor 3rd) while others differ only at the lower level (such as minor 3rd vs. major 3rd). This possibility is of course outside the scope of the unidimensional hypothesis without the introduction of specifically musical factors the defenders of this hypothesis had hoped to avoid.

Table 3

A comparison of Different RTs to interval pairs that are
(a) scale-step equivalent
(b) one semitone apart but not scale-step equivalent
(c) more than one semitone apart

	a	b	c
RT in msec:	1014	758	687

Nonetheless the distinction appears to have considerable perceptual reality, as can be seen in Table 3. The dependent variable is the amount of time required to respond "Different" and all of the intervals pairs contributing to the two values on the left differ by a single semitone. There is a 256 msec. difference ($p < .001$) between pairs that share the same value of scale-step and pairs that straddle scale-step boundaries; a difference of this magnitude is really quite large relative to the sorts of effects usually found in reaction-time experiments. It took considerably longer for our listeners to distinguish between intervals possessing this higher-order equivalence than intervals that did not[2]. The third value is the average of all other off-diagonal entries in the data matrix. The interval pairs contributing to this value all differ in width by more than a semitone, yet the 71 msec. difference with the 2nd column is less than one-third the difference between the first two columns. Once the effect of scale-step is taken out, there is really very little "width" effect left.

This result was so convincing that it led to yet another experiment, where the questions were sampled from both low-level and high-level categories. On some trials subjects were to respond "Same" if the interval they heard was either (say) a major or minor 2nd (high-level), while on other trials they might be asked to respond Same to a minor 2nd, and Different otherwise, as in the initial experiments (low-level). Four different high-level questions—2nd, 3rd, 6th, and 7th—were examined, along with the corresponding eight low-level counterparts (m2, M2, m3, M3, m6, M6, m7, M7). A subset of 10 subjects participated in this experiment.

The intuitive prediction most psychologists (and many musicians) would (and did) make is that the high-level questions ought to take longer, on the notion that it is harder to listen for two things (a m3 and a M3) than for just one (a m3). This notion has certainly received much support from research on recognition of characters (Sternberg, 1966), forms (Cavanaugh, 1972), and even tones (Clifton and Cruse, 1977); the longer the list of things one is looking (or listening) for, the greater the time required to respond.

Yet this did not occur in the present experiment. Figure 5 shows results for each interval class separately, but the trends are all in the same direction. Listeners are actually faster ($F(1,9) = 8.14$, $p < .025$) to respond to a minor 3rd when they are listening for either a minor or major 3rd than when they are listening only for a minor 3rd! Error rates showed the same pattern, $F = 9.36$ ($p < .02$).

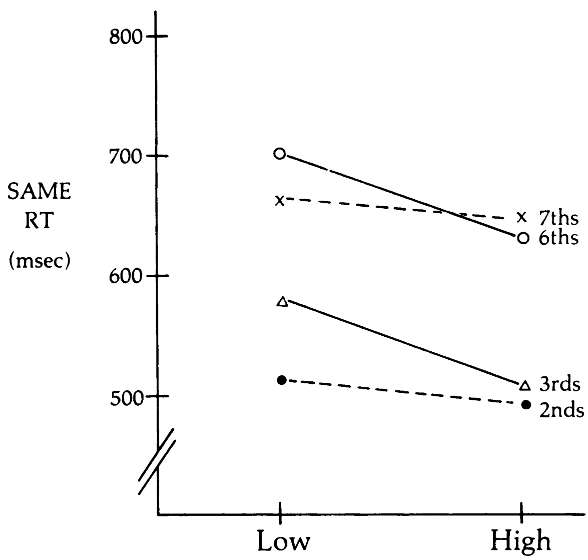

Figure 5. The effect of high- vs. low-level questions on Same RT to recognize 2nd, 3rds, 6ths, and 7ths.

A partial account of this seemingly paradoxical result comes from a consideration of ideas and experiments presented by Rosch and her colleagues (Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976). They speak of a basic level of categorization, one that is perceptually prior to other levels. For example, it is possible to conceive of what I am sitting on as "furniture," as "a chair" or "a swivel-back chair," but when I recognize one, I primarily recognize it as a chair. To be sure, the basic level differs for different people, but the message of the present experiments seems to be that, for trained musicians at least, the basic level for musical intervals is the scale-step, and not the semitone level. If the process these musicians were going through upon hearing a minor 3rd were something like "that's a minor 3rd, therefore it must be a third," there is no way the high-level question could have led to shorter latencies. These people must, therefore, have been recognizing intervals at the scale-step level directly.

There is no element of self-fulfilling prophecy in this result. Subjects were asked to predict the results of this experiment in advance on a questionnaire and nearly everyone indicated the semitone level as perceptually prior, to varying degrees, for all intervals. The correlation between predicted and observed levels effects over subjects in this experiment was negative ($r = -.32$) and no subject's predicted values correlated significantly with his or her latency data.

Conclusions

In this brief report I have tried to show that the time it takes to answer questions about heard musical intervals is a rather sensitive measure of the ease of recognition and categorization of these intervals. Valuable structure-detecting data-analysis techniques such as nonmetric multidimensional scaling can be applied to such latency measures, and provide interpretable, meaningful results. The central finding uncovered in this way was that musical intervals are perceived, not as one-dimensional objects varying only in width, but in a space of at least two dimensions where the cyclic nature of musical pitch is preserved. This documents the perceptual reality of that aspect of pitch called chroma by several authors.

Perhaps even more strikingly revealed in these experiments was the perceptual salience of higher-order scale-step properties of musical intervals. Pairs of intervals that are scale-step equivalent are perceived as more similar to one another than equally distant pairs of intervals that are not scale-step equivalent. In fact, the scale-step level of definition appears to have perceptual primacy, as revealed by the decreased recognition latencies and error rates to high-level questions over low-level questions. This last finding in particular suggests that students learning to identify intervals be taught to recognize scale-step qualities such as 2nd first, and to differentiate these into minor 2nds and major 2nds only later on. The outcome of such a teaching procedure would clearly be of considerable interest.

NOTES

[1]The ratio-simplicity dimension, missing from the Figure 4 representation, was recovered as a 3rd dimension only with the use of more sophisticated metric scaling procedures (INDSCAL —individual differences scaling), and was not a consistent feature of all the experiments performed. To some extent, it is implicit in the chroma representation, with 2nds and 7ths projecting near one another and similarly for 3rds, 4ths, 5ths, and 6ths. The octave (a simple ratio) and the tritone (presumably a more complex ratio) do not fit this pattern so well, however.

[2]It might be argued that the added confusability of intervals possessing scale-step equivalence is due to the similarity in the names of the intervals. On closer inspection, however, this hypothesis does not appear intelligible. We do not confuse dogs with logs, though we may mistake a dog for a wolf. The point is that similarity of names seems irrelevant unless it is accompanied by perceptual similarity. And given perceptual similarity, this appears to override any similarity of a purely nominal sort. In any case, the next experiment to be reported will demonstrate the perceptual similarity of scale-step equivalent intervals directly, and it will be seen that no naming hypothesis will even apply.

REFERENCES

Bachem, A. Tone height and tone chroma as two different pitch qualities. *Acta Psychologica*, 1950, *7*, 80-88.

Balzano, G. J. Chronometric studies of the musical interval sense. Ph.D. Dissertation, Stanford University, 1977.

Cavanaugh, J. P. Relation between the immediate memory span and the memory search rate. *Psychological Review*, 1972, *79*, 525-530.

Clifton, C., Jr., and Cruse, D. Time to recognize tones: Memory scanning or memory strength? *Quarterly Journal of Experimental Psychology*, Nov., 1977, vol. 29(4), 709-726.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964a, *29*, 1-27.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964b, *29*, 115-129.

Plomp, R., Wagenaar, W. A., and Mimpen, A. M. Musical interval recognition with simultaneous tones. *Acustica*, 1973, *29*, 101-109.

Revesz, G. *Introduction to the Psychology of Music*. Norman, Oklahoma: University of Oklahoma Press, 1954.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. Basic objects in natural categories. *Cognitive Psychology*, 1976, *8*, 382-439.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 1962a, *27*, 125-139. II. *Psychometrika*, 1962b, *27*, 219-246.

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

Sternberg, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception and Psychophysics*, 1967, *2*, 45-53.

Stevens, S. S. and Volkmann, J. The relation of pitch to frequency: A revised scale. *American Journal of Psychology*, 1940, *53*, 329-353.

Stevens, S. S., Volkmann, J. and Newmann, E. B. A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, 1937, *8*, 185-190.

(The time interval between receipt of this article and its publication was excessively long. The editors appreciate Professor Balzano's understanding.)