Auditory Illusions Demonstrating That Tones Are Assimilated to an Internalized Musical Scale

Abstract. *Judgments of relations between tones in a sequence derived from the major diatonic scale* (do, re, mi, . . .) *by systematic distortions (either equalization or stretching of intervals) reveal that Western listeners interpret the tones relative to that musical scale even when they try to judge the purely physical relations between the tones.*

Diverse phenomena of perception and memory indicate that our experience is determined by internal schemata (*1*) as well as by external inputs. Thus musical tones, though physically variable along a continuum of frequency, tend to be interpreted categorically (*2*) as the discrete notes (named *do, re, mi, fa, sol, la, ti, do*) of an internalized musical scale (*3, 4*). We suggest that the internal schema may act as a template that, when brought into register with a tonal input (Fig. 1a), maps the unequally spaced physical tones into the discrete steps of the schema, with a resulting unique conferral of tonal stabilities (indicated in Fig. 1 by the lengths of the bars labeled *Tonic, Dominant*, and so forth) on the tones (*4*).

We sought more direct evidence concerning the operation of the proposed tonal schema by presenting listeners with sequences of eight sinusoidal tones realizing each of two different distortions of the major diatonic scale: (i) an equalized scale in which each of the seven intervals between successive tones within an octave spanned exactly the same difference in log frequency and (ii) a stretched scale in which the unequal half- and whole-tone spacing of the standard diatonic scale was preserved but uniformly stretched (by the factor 13/12) to extend one half-tone beyond an octave (*5*).

Listeners trying to judge the physical relations between the tones (for example, on the log frequency continuum) might unwittingly interpret the tones according to the diatonic schema. In the case of the equalized scale (Fig. 1b), the third and seventh successive intervals, though physically equal to the others, might then be judged larger because those two intervals would be wide in relation to the narrower gaps "expected" by the input template. In the case of the stretched ascending scale (Fig. 1c), each successively higher tone, being slightly sharp, might be accommodated by a small upward adjustment of the whole diatonic schema. Upon accommodation to the final tone, the schema would be shifted a half-tone upward (Fig. 1d), corresponding to a modulation from the key of C to that of C#. Listeners might then mistakenly accept, as the original starting tone, not that tone, C, but the C# immediately above it—even though that particular C# had never been presented. A similar distortion in the opposite direction should follow a stretched descending scale.

We tested the equalized scale with two groups of undergraduate students who had just studied the physical scale of log frequency of tones in a course on perception. We twice played the sequence of eight equally spaced sinusoidal tones ascending from middle C to the C an octave above to give the listeners two opportunities to judge each interval. The tones were 500 msec long and separated by 50-msec intervals. The 43 listeners in group 1 indicated whether they judged the difference in log frequency between each tone and the next to be larger, smaller, or the same in comparison with the other intervals by marking a +, −, or 0, respectively, above the line segments between eight equally spaced circles corresponding to the successively higher tones. The 41 listeners in group 2 made a similar rating of the size of each successive interval, but, each time, in comparison with the immediately preceding interval only (with no rating for the interval between the first two tones). We particularly emphasized to this second group that comparisons were to be made solely on the basis of the equality or nonequality of the intervals with respect to the physical scale of log frequency (rather than on any musical basis).

The obtained percentages of "larger" ratings (Fig. 2a) and "smaller" ratings (Fig. 2b) confirm our prediction for both sets of procedures and instructions. The interval between the third and fourth tones was judged larger by 33 and 59 percent of listeners in groups 1 and 2 and smaller by only 12 and 2 percent, respectively; the interval between the seventh and last tones was judged larger by 72 and 63 percent in groups 1 and 2 and smaller by only 7 percent in both groups. Possibly the effect was strongest for the last interval because the diatonic schema tends to be most fully instantiated upon completion of the scale (with its transition from the leading tone back to the most stable tone, the tonic). This predicted tendency to rate the third and seventh intervals of the scale as larger than the other intervals was statistically significant ($P < 0.001$ for each of the two groups) (*6*).

We presented the stretched diatonic scale, also, as a sequence of 500-msec

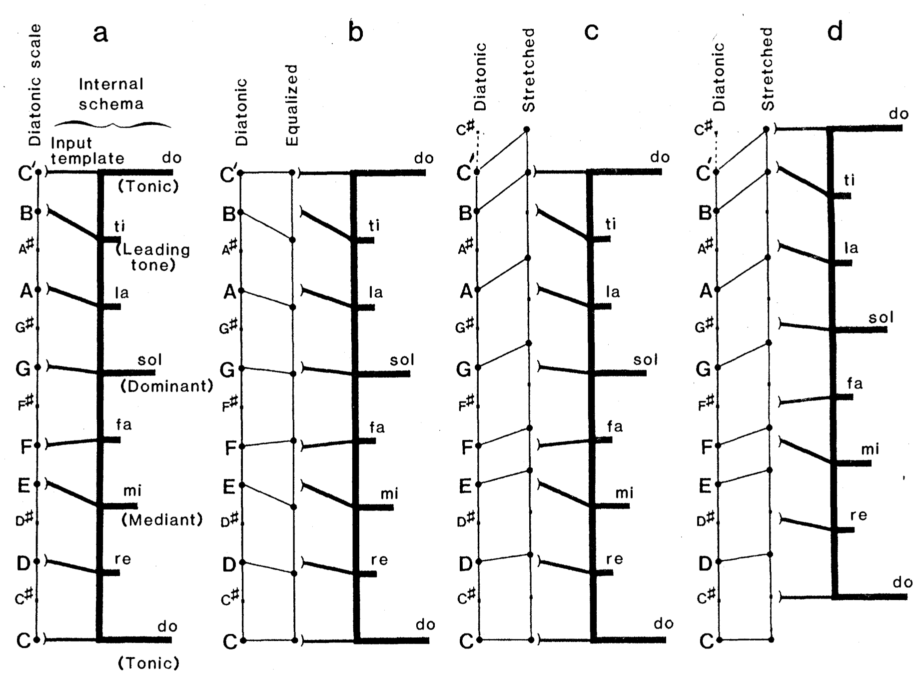

Fig. 1. (a) Illustration (for the key of C major) of how the input template of an internalized diatonic schema would map the unequally spaced tones of a major musical scale into an internal representation of the discrete scale steps (*do, re, mi, . . .*) with an associated hierarchy of tonal stabilities (*Tonic, Dominant, Mediant*, and so forth). (b) Illustration of why intervals 3 and 7 of an equalized scale might sound too large. (c and d) Illustration of how an internal schema initially aligned with C might be shifted up to C# as a result of playing the stretched scale.

sinusoidal tones separated by 50-msec silent intervals. We followed each of four presentations of the stretched scale by a single test tone that was either the original starting tone or a tone a half step removed from that original tone. The 57 listeners, again undergraduate students, were asked to judge whether each test tone was definitely lower (−2), probably lower (−1), about the same (0), probably higher (+1), or definitely higher (+2) than the original starting tone. We twice presented the stretched ascending scale from middle C to the C# in the next octave above (C#′)—once followed by the original C as the test tone, once followed by the never presented C# a half step above that starting tone. We also twice presented the stretched descending scale from C#′ to middle C—once followed by the original C#′, once followed by the C′ a half step below that starting tone.

Judgments after the stretched ascending scale (Fig. 2c) supported our predictions (Fig. 1d). Of all listeners, 86 percent rated the repetition of the original starting middle C as lower than that tone (with 48 percent even rating it definitely lower), while only 9 percent correctly rated it about the same. The departure of the (negative) mean rating (−1.23) from zero was statistically significant [$t(56) = -9.62$, $P < 0.0001$]. At the same time, 63 percent rated the never presented C# a half step above the starting tone as about the same as that tone, while only 11 percent correctly rated it as higher.

Similarly, after the descending scale (Fig. 2d), 68 percent of the listeners rated the repetition of the starting tone (C#′) as higher than that tone, while only 25 percent correctly rated it about the same. The departure of the (positive) mean rating (+0.88) from zero was statistically significant [$t(56) = 6.99$, $P < 0.0001$]. In contrast, 49 percent rated the tone (C′) a half step below the starting tone about the same as that tone, while only 15 percent correctly rated it as lower. The predicted pattern may have emerged somewhat more strongly after the ascending scale because scales are more commonly played or sung, as scales, in ascending order (with the half-tone transition from the leading tone back to the tonic occurring at the end).

Evidently, during the equalized scale, the physically equal differences between successive tones were perceived as larger for those steps for which the input template of the diatonic schema expects a smaller interval; after a scale stretched by a half-tone, the tone judged to be the same as the starting tone was systematically displaced in pitch as if the entire input template had been shifted by approximately a half-tone. In subsequent experiments, in which we obtained ratings of how well test tones corresponding to all notes of the scale fit in with the presented scale (compare with *4*), we have obtained evidence (i) corroborating this stretch-induced shift, (ii) suggesting that after the descending scale, the schema, in addition to being shifted, is itself somewhat stretched, and (iii) disconfirming the possible alternative explanation that the listeners in experiment 2 were merely judging the (octave) relation of the test tone to the last tone of the stretched scale (*7*).

ROGER N. SHEPARD
DANIEL S. JORDAN

Department of Psychology, Stanford University, Stanford, California 94305

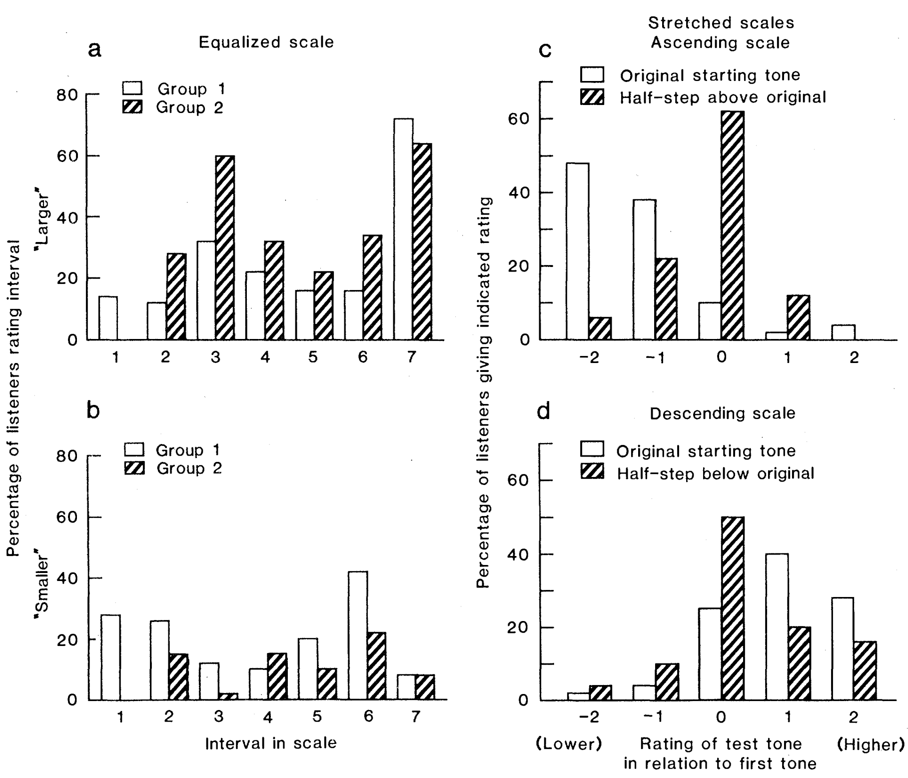

Fig. 2. (a) Percentages of listeners rating each interval between successively sounded, equally spaced, ascending tones as larger than the other intervals (unshaded bars) or as larger than the immediately preceding interval (shaded bars). (b) Percentages of listeners rating each interval as smaller. As predicted (Fig. 1b), intervals 3 and 7, although equal, were rated larger. (c and d) Percentages of listeners rating a test tone as lower (−2 or −1), equal (0), or higher (1 or 2) with respect to the starting tone of a preceding stretched diatonic scale, (c) when the stretched scale was ascending and the test tone was equal (unshaded bars) or a half-tone higher (shaded bars), and (d) when the stretched scale was descending and the test tone was equal (unshaded) or a half-tone lower (shaded). As predicted (Fig. 1, c and d), the tone most accepted as the starting tone was one appropriately shifted from it by a half-tone.

References and Notes

1. F. C. Bartlett, *Remembering* (Cambridge Univ. Press, Cambridge, 1932); J. Piaget, *The Origins of Intelligence in Children* (International Universities Press, New York, 1952); R. S. Woodworth, *Experimental Psychology* (Holt, New York, 1938).
2. E. M. Burns and W. D. Ward, *J. Acoust. Soc. Am.*, **63**, 456 (1978); J. A. Siegel and W. Siegel, *Percept. Psychophys.* **21**, 143 (1977); *ibid.*, p. 399; R. S. Zatorre and A. R. Halpern, *ibid.*, **26**, 384 (1979).
3. W. J. Dowling, *Psychol. Rev.* **85**, 341 (1978).
4. C. L. Krumhansl and R. N. Shepard, *J. Exp. Psychol. Hum. Percept. Perform.* **5**, 579 (1979).
5. The tone sequences were generated by a Mountain Computer digital synthesizer, controlled by an Apple II computer, with software provided by the Alpha-Syntauri music system and by E. J. Kessler.
6. We tested the statistical significance of the predicted pattern in the average ratings of each of the two groups of listeners by means of contrasts [B. J. Winer, *Statistical Principles in Experimental Design* (McGraw-Hill, New York, ed. 2, 1971), chapter 4]. With weights for the contrasts of (−1, −1, 2.5, −1, −1, −1, 2.5) for the seven intervals rated by the first group and (−1, 2, −1, −1, −1, 2) for the last six of these intervals rated by the second group, reflecting in both cases the prediction that the third and seventh of the seven equal intervals would be judged larger, the F ratios (using residual error terms from a repeated measures design) for a Sheffé test were $F(6, 252) = 44.7$ and $F(5, 200) = 24.1$, respectively. Significant F ratios were also obtained with contrasts reflecting the further expectation that the seventh interval would be more strongly affected than the third. Thus, with weights of (−1, −1, 2, −1, −1, −1, 3) and (−1, 1.5, −1, −1, −1, 2.5) for the first and second groups, $F(6, 252) = 51.1$ and $F(5, 200) = 34.2$, respectively ($P < 0.001$, in both cases).
7. We reported some of that further evidence, along with the present evidence, at the annual meeting of the Psychonomic Society, Minneapolis, 11 November 1982.
8. Supported by NSF grant BNS80-05517. We thank J. Chowning, C. Krumhansl, and B. Wandell for help and suggestions, and E. Kessler for extensive assistance in generating the tone sequences and in conducting the experiments.

9 April 1984; accepted 2 August 1984