UNIVERSITY OF CALIFORNIA PRESS
JOURNALS + DIGITAL PUBLISHING

An Experimental Study of Internal Interval Standards in Javanese and Western Musicians
Author(s): Marc Perlman and Carol L. Krumhansl
Source: *Music Perception: An Interdisciplinary Journal*, Vol. 14, No. 2 (Winter, 1996), pp. 95-116
Published by: University of California Press
Stable URL: http://www.jstor.org/stable/40285714
Accessed: 04/04/2013 15:41

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

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 California Press is collaborating with JSTOR to digitize, preserve and extend access to *Music Perception: An Interdisciplinary Journal.*

http://www.jstor.org

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

Music Perception
Winter 1996, Vol. 14, No. 2, 95–116

© 1997 BY THE REGENTS OF THE
UNIVERSITY OF CALIFORNIA

An Experimental Study of Internal Interval Standards in Javanese and Western Musicians

MARC PERLMAN
Cornell University & Brown University

CAROL L. KRUMHANSL
Cornell University

Six Javanese and six Western musicians performed a magnitude-estimation task using 36 melodic intervals ranging from 60 to 760 cents at 20-cent increments. Several musicians displayed well-defined regions of confusion in which a range of intervals was assigned approximately equal magnitude estimates. The results suggest that these listeners assimilate the intervals to a set of internal interval standards. No evidence for assimilation was found for other musicians in both groups, some of whom made highly accurate estimates. For the Javanese musicians who showed assimilation to internal interval standards, the regions corresponded to the two Javanese tuning systems, *sléndro* and *pélog*. For the Western musicians, the regions corresponded to the equal-tempered scale. The relatively wider regions of confusion for the Javanese musicians may reflect the greater variability of intonation in Java. In addition, the Javanese musicians seemed able to choose between internal interval standards based on the two tuning systems.

In perceiving, remembering, and reproducing melodies, individuals use cognitive representations, or schemas, of musical intervals. Evidence for the existence of cognitive representations is found in patterns of errors in judging unfamiliar intervals (Dowling, Lung, & Herrbold, 1987; Francès, 1988, pp. 38-39; Shepard & Jordan, 1984) and also in identification and discrimination performance in studies of categorical perception (e.g., Acker, Pastore, & Hall, 1995; Burns & Campbell, 1994; Burns & Ward, 1978; Siegel & Siegel, 1977b; Wapnick, Bourassa, & Sampson, 1982; Zatorre & Halpern, 1979). To the extent that individuals acquire these interval standards through learning (Burns & Ward, 1982), we would expect to find

Address correspondence to Marc Perlman, Department of Music, Box 1924, Brown University, Providence, RI 02912 (e-mail: Marc_Perlman@Brown.edu) or Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. (e-mail: clk4@cornell.edu)

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

cross-cultural differences. Psychological measures of internal interval standards can be compared with the intervals in musical practice and the conceptualization of interval usage within a culture's explicitly formulated music theory, if one exists. The experiment reported here investigated interval perception by Javanese musicians expert in the Surakarta style of Central Javanese gamelan music and a group of expert Western musicians.

Acoustic phenomena, auditory physiology, and cognitive processes may impose constraints on the class of possible musical scales. Nevertheless, a great variety of scales can be found among the world's musical traditions. This fact, established by Alexander Ellis more than a century ago (Ellis, 1885), has been confirmed by subsequent ethnomusicological research (e.g., Berliner, 1978, pp. 63–68; Hood, 1966; Morton, 1976, pp. 232–237; Tracey, 1970, p. 125). The most obvious dimensions of difference across cultures are the number of scale steps within an octave and the size of those steps. A third dimension of difference is the amount of variance in tuning that is found—that is, the degree of standardization. Javanese music, considered in relation to Western music, illustrates all three.

The fixed-pitch instruments of a Javanese gamelan orchestra are tuned to two distinct scales, called *sléndro* and *pélog,* each of which uses intervals that do not occur in Western music. *Sléndro* has five pitches per octave (although the singers and flexible-pitch string instruments occasionally use additional "slanted" [*miring*] tones in performance). *Pélog* makes seven pitches available in each octave, but compositions in *pélog* are almost always pentatonic—two of the scale tones are "substitute" tones, which can temporarily replace their neighbors. With very rare exceptions, compositions do not mix tones from the two tuning systems, nor are *sléndro* and *pélog* thought to be derived from any single superordinate pitch collection.

Figure 1 indicates the type of relationship that can be found between *sléndro* and *pélog*. It shows tone measurements of two instruments (one *pélog* and one *sléndro*) made by one of the authors (MP).[1] Five of the *pélog* pitch names are identical to *sléndro* pitch names, but the actual pitches are not necessarily considered identical. However, in this particular ensemble the tone called *nem* is considered to be identical in *sléndro* and *pélog*, and their tunings are very similar. Figure 1 also gives the number of cents between successive tones, illustrating the difference between typical Javanese intervals and those of Western music. *Sléndro* uses intervals between major seconds and minor thirds, whereas *pélog* uses intervals between minor and major seconds.

The tuning pattern presented in Figure 1 is unique to this gamelan; other gamelan can differ substantially. For example, in an ongoing study (by

1. The tones were measured by a Sanderson Accu-Tuner. The integers after the plus or minus sign give the deviation in cents from the equal-tempered reference pitches based on A_4 = 440 Hz.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

Pélog

panunggul	*gulu*	*dhadha*	*pélog*	*lima*	*nem*	*barang*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
D_3-26	$E\flat_3$-10	F_3-45	G_3+24	$G\sharp_3$+43	A_3+43	B_3+35

Intervals	116	165	269	119	100	192

Sléndro

gulu	*dhadha*	*lima*	*nem*	*barang*
(2)	(3)	(5)	(6)	(1)
D_3+28	F_3-35	G_3+17	A_3+40	C_4-10

Intervals	237	252	223	250

Fig. 1. Tone measurements of the *pélog* and *sléndro slenthem* of the gamelan at the Sri Wedhari theater auditorium in Solo, expressed in cents deviations from tones of the Western equal-tempered scale (based on A_4 = 440 Hz). For each tuning system, the top line gives the names of the tones; immediately underneath, in parentheses, are the numerals used to represent the tones in Javanese cipher notation. The next lower line contains the measurements obtained by one of the authors (MP) using a Sanderson Accu-Tuner. The bottom line shows the sizes of the intervals between successive tones in cents.

MP) of 22 fixed-pitch instruments in the *sléndro* tuning system, no interval was found to vary less than 30 cents, and some varied as much as 75 cents. Surjodiningrat, Sudarjana, and Susanto (1972) found similar variability in tunings of 28 gamelan sets. This variability is much greater than that found among fixed-pitch instruments in Western music. Although considerable tuning variations appear in Western performances by voice and certain instruments (e.g., unfretted strings; cf. Burns & Ward, 1982, p. 258), they are apparently conditioned by local melodic context (Rakowski, 1990, 1991); similar contextual effects are also found in North India (Jairazbhoy & Stone, 1963; Levy, 1982) and Thailand (Morton, 1974). These variations seem qualitatively different from the context-independent variability of fixed-pitch instrumental tunings in Java and certain other cultures (e.g., the harp and *amadinda* xylophone music of Uganda; cf. Wachsmann, 1950, 1967).

The effect of these differences in intonational practice on musicians' internal interval standards is largely unknown. Almost all of the available data pertain to the Western tradition and generally show conformance with the equal-tempered chromatic scale. However, one study (Burns & Campbell, 1994) finds a tendency to compress small intervals and to stretch large ones. Several studies indicate that the location and width of internal standards vary between musicians (Acker, Pastore, & Hall, 1995; Burns & Campbell, 1994; Burns & Ward, 1978; Siegel & Siegel, 1977b). Finally, an effect of tonal stability (Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979) has been found on accuracy of identifying intervals (Miyazaki, 1995).

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

These and other effects may be found cross-culturally. In a study of Ugandan tuning practices, Cooke (1992), drawing on Siegel and Siegel (1977b), suggested that a relationship may exist between size of scale step, breadth of interval category, and variability of tuning. For example, quasi-equidistant pentatonic scales (whose steps approximate 240 cents each), such as those found in Uganda and Java, may give rise to broad, tolerant interval categories. From a perceptual point of view, this may mean that intervals differing by as much as 100 cents (say, 190 cents and 290 cents) are assimilated to the same internal interval standard. Comparative data of this sort are interesting because they yield information concerning how interval standards are internalized and applied in musical practice.

Most research on internal interval standards uses the experimental paradigms associated with categorical perception: identification (labeling by interval name) and discrimination (typically an ABX design). Various considerations suggest that these may need to be modified or selectively applied when non-Western music is studied. For example, the identification task assumes the participant's mastery of standardized labels for frequency ratios. This ability is valued in the Western art music tradition and is inculcated in musicians through a regimen of ear-training, but this is not true of all musical cultures. In some societies, many or all musical intervals have no labels. Even when such labels exist, they may not refer primarily to frequency ratios but to instrumental techniques (i.e., how intervals are produced rather than how they sound). Finally, even if a complete set of interval labels is present in a culture, it may pertain more to speculative music theory than to practical musical training.

In addition, some identification and discrimination tasks constantly vary the base tone of the stimulus intervals (e.g., Burns & Ward, 1978; Siegel & Siegel, 1977b; Zatorre & Halpern, 1979, Experiment 2). This assumes that the participant can judge the size of intervals independently of their absolute pitch. This ability, too, is more highly valued in Western art music than in some other cultures. Western musicians learn to judge the sizes of intervals relative to constantly shifting reference points, both the local shifts of a harmonic progression and the larger shifts of tonal framework produced by modulation between keys. In other music cultures, such shifts are rare or nonexistent, and the ability to recognize an interval under transposition is irrelevant.

Previous studies that used identification and discrimination tasks with non-Western musicians have sometimes encountered problems related to these factors. Burns (1977), for example, tested North Indian musicians to see if they could identify and discriminate more than 12 interval categories per octave. Indian music theory divides the octave into 22 microtonal steps (called *sruti*). In the identification task, the musicians responded both in terms of Indian solfège note labels (*svaras*) and *sruti*. However, the responses

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

in terms of *sruti* were highly inconsistent (p. 108). This result is compatible with the view, commonly held by ethnomusicologists, that the microtonal scale of 22 *srutis* is a speculative theoretical construct, of dubious relevance to modern musical practice (cf. Powers, 1980, p. 98). The musicians' performance was also impaired, although to a lesser degree, by the constantly shifting reference tones of the experiment. Because Indian musical practice uses an immutable tonic, Burns suspected that his participants might perform more consistently if the base tone was held constant. A subsequent experiment confirmed this prediction (Burns, 1977, p. 110).

Another, more recent, study (Keefe, Burns & Nguyen, 1991) presented an identification task to a Vietnamese musician who is also a musicologist and thus quite familiar with Western pitch nomenclature. Response categories were taken from a previous experiment that measured the musician's tunings of various Vietnamese scales and were expressed in terms of Western pitch names modified by plus and minus signs. Some well-defined interval categories were found, but for certain interval regions, the musician was unable to respond consistently and the boundaries between some interval categories did not match measurements of his own tuning behavior. The authors concluded that the "Western concept of interval does not exist in Vietnamese music" (p. 463), and in a subsequent experiment, they dispensed with pitch-labels.

The musical tradition we chose to investigate presents a similar challenge to the use of identification tasks. Only two Javanese interval labels are commonly known, and they seem to refer primarily to instrumental technique, not tonal distance. A discrimination task would be more appropriate because it does not require labels. However, another consideration weighed against the use of this method. We felt it was desirable to use tones recorded from gamelan instruments. We wanted to use stimulus timbres that would evoke a perceptual set appropriate to Javanese music (and not Western music with which the Javanese musicians were also familiar). Because the tones were recorded, listeners might be successful in a same/different, ABX, or oddity task because of distinctive characteristics of the particular tones used in the experiment. For example, they might be successful in these tasks by paying attention to peculiarities of amplitude variations or tone onsets that have nothing to do with pitch per se.

These considerations weighed in favor of a magnitude estimation type of task, as was used by Siegel and Siegel (1977a, Experiments 3 and 4; 1977b). In their experiments, the musicians were asked to give numerical responses. We decided to use instead the seemingly simpler nonnumerical response scale of a horizontal line that they marked to indicate the interval size. In terms of this response measure, evidence of internal interval standards may take the form of a range of intervals all of which receive approximately the same magnitude estimates. If listeners assimilate the stimuli to their inter-

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

nal interval standards, they would underestimate the difference between the stimulus intervals that are mapped to the same internal interval standard. Such assimilation would produce regions of confusion, that is, ranges of intervals for which the participants gave similar responses. To assess this, the data analysis identifies (for each participant) intervals with responses that are not different by a test of statistical significance. Finally, various considerations led to the selection of stimulus intervals and the participants for the experiment. We used a relatively large range of intervals (60–760 cents) compared with most studies of musical interval perception (e.g., Burns & Ward, 1978, 250–550 cents; Siegel & Siegel, 1977b and Wapnick et al., 1982, 480–720 cents; Zatorre & Halpern, 1979 and Zatorre, 1983, 300–400 cents). This was because some of the Javanese musicians might relate the stimulus materials to the *sléndro* scale. The steps of *sléndro* are quite wide, and a range of 700 cents is needed to cover three of them. The stimulus intervals varied at increments of 20 cents in order to locate fairly precisely the boundaries between possible regions of confusion and identify corresponding intervals in the tuning systems.

Although we would have preferred to work with Javanese subjects who had no exposure to the scales of Western music, this was not feasible. The Indonesian national anthem, all patriotic school-songs, and virtually all popular music use the Western diatonic scale. We worked instead with Javanese who had maximum exposure to and familiarity with Javanese music: expert gamelan musicians. We were mainly interested in the internal interval standards of Javanese musicians, but for purposes of comparison we collected responses from Western listeners who performed exactly the same task with exactly the same stimulus materials. We therefore used two groups of expert musicians, one with expertise in Javanese music and the other with expertise in Western music.

Method

PARTICIPANTS

The participants were six Javanese musicians and six Western musicians. All the Javanese participants were professional musicians currently (or recently) active in the two cultural centers of Central Java, Surakarta and Yogyakarta. All but one were (or had been) performers or teachers at the major musical and educational institutions in Java: the Indonesian College of the Arts and the Gamelan Studio of the local station of Indonesian National Radio. The one exception was a successful free-lance performer. All the Western musicians were members of the Cornell University community. Two were degree candidates in musicology, and the others were teachers or graduate students of Western music performance. All of the Western musicians had extensive performance experience, and many had made commercial recordings. Some specialized in historical performance practice. Table 1 gives more information about the individual participants.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

TABLE 1
Participants in the Magnitude Estimation Task

	Javanese Musicians
S1	Gamelan teacher and instrumentalist at Indonesian National Radio, performs regularly at one of the courts, avocation of tuning gamelan instruments
S2	Singer and instrumentalist on staff of conservatory in Java
S3	Freelance instrumentalist in Java
S4	Instrumentalist on staff of conservatory in Java
S5	Instrumentalist, former instructor at conservatory in Java
S6	Instrumentalist on staff of conservatory in Java, extensive experience tuning gamelan instruments
	Western Musicians
S7	Singer, one semester's experience learning gamelan at Cornell University
S8	String player, extensive experience with historical performance practice
S9	Keyboard player and teacher, extensive experience tuning harpsichords, historical performance practice
S10	Wind-instrument player, extensive experience with historical performance practice
S11	Singer and voice teacher
S12	String player, absolute pitch

APPARATUS AND MATERIALS

The stimulus materials consisted of 36 ascending melodic intervals varying in size between 60 and 760 cents, at increments of 20 cents. The lowest tone was always G$\sharp_4$. Tone durations were approximately 2.5 sec; the onset of the second tone immediately followed the decay of the first tone. The tones were recorded from a *gendèr barung,* which consists of forged bronze bars suspended over tuned resonators, played with a padded beater. Spectral analysis showed the tones are quite harmonic and have a strong fundamental frequency, a weaker second harmonic, and a very weak fourth harmonic. Five tones (G$\sharp_4$, A$\sharp_4$+60, C_5+40, E_5+20, and F_5+80, where the integer following the plus sign indicates the number of cents above the equal-tempered tone) were sampled by a Macintosh Powerbook. They were processed by a program that transposes in units of semitones. Thus, the upper tone of the intervals of 100, 200, 300, 400, 500, 600, and 700 cents were based on G$\sharp_4$; the upper tone of the intervals of 120, 220, 320, 420, 520, 620, and 720 cents were based on E_5+20, the tone 20 cents higher than an equal-tempered tone, and so on.

Ten blocks of trials presented the 36 melodic intervals in different random orders. An announcement of the number of the trial preceded each interval. Each block began with three example intervals: an octave, the smallest interval in the experiment (60 cents), and the largest interval (760 cents). The tapes were recorded and played back on a Sony WMD6C Professional Walkman tape recorder; the participants listened through headphones. The response sheet for each block of trials consisted of three sheets, each with 12 numbered horizontal lines 180 mm long.

PROCEDURE

The experiment was completed in either one or two sessions at the option of the participant. The total time varied from 2-1/2 hr to 4 hr. The experiment began with as many practice trials as the participants wished. The 10 blocks of trials were then presented in different random orders. Participants were free to rewind the tape if they wished to hear an interval again, although few did so.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

The instructions asked participants to judge the size of each melodic interval by marking the corresponding line on the response sheet. The instructions, which were spoken in Indonesian to the Javanese participants, included the following:

> This is an attempt to investigate your "aural feeling" (*rasa pendengaran*). Every person hears differently—every musician's ears have their own tendencies in listening. This procedure will give us a sort of "earprint" (*tapak telinga*), a description of how you hear tones. . . . On these tapes you will hear a series of pairs of tones, each pair forming a different interval (*jangkah nada*). The first (lower) tone will always be the same. The second (higher) tone will sometimes be low (*besar,* lit. 'big'), and so close to the first tone; sometimes it will be high (*kecil,* lit. 'small'). . . . The timbre (*warna suara*) of these tones is that of a *gendèr.* Some of these intervals may sound like *sléndro,* some like *pélog,* and some neither *sléndro* nor *pélog*. . . . For each interval, please make a mark on the sheet, on the line with the corresponding number. If you wish, you may think of the whole line as representing an octave (*gembyang*). Let the leftmost end of the line represent the lower tone of the pair, and make a mark showing where the upper tone would be. When the two tones are close [sing example], make the mark close to the leftmost end of the line; when the tones are far [sing example], make the mark farther from the left end.

Parts of the instructions were modified for the Western participants. The reference to Javanese tuning systems was changed to "some of the intervals they would hear would be in tune according to the tempered scale, but others would not."

The first participant, a Javanese musician, requested that the experimenter add evenly spaced vertical guidelines to the response sheets. We decided to allow this kind of strategy if the participant spontaneously adopted it. As it turned out, half of the participants in each group of listeners added evenly spaced vertical guidelines, or tick marks, as noted in the results section. Effectively, these participants modified the experiment to bring it into line with Siegel and Siegel's (1977b, p. 402) experiment, in which participants were asked to assign numbers to intervals. Without exception, they assigned constant numerical increments to semitone intervals, analogous to the vertical guidelines or tick marks used by some of our participants.

Results

The distance between the left end of the response line and the listener's mark was measured for each trial. Initial inspection of these values showed that the judgments of each listener increased regularly with interval size. As shown in Table 2, the average judgments for each interval correlated very strongly with the actual interval size. These correlations averaged r (df = 34) = .991, $p < .0001$ and were on average only very slightly lower for Javanese musicians, r (df = 34) = .987, than for Western musicians, r (df = 34) = .994. Thus, all listeners clearly understood the idea that interval size could be represented on the response scale.

The responses were also reasonably reliable; Table 2 shows for each listener the standard deviation across the 10 blocks of trials averaged over the intervals. As a group, the Western listeners gave somewhat more reliable responses than the Javanese listeners. Considerable overlap was ap-

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

TABLE 2
Responses in Magnitude Estimation Task

	r with cents	*SD* (mm)	Est. *d′* for 20 cents	Est. *d′* for 40 cents	Modification of Response Sheet
			Javanese Musicians		
S1	.992	4.85	.585	1.186	Guidelines, labeled with tones of *pélog*
S2	.987	13.33	.231	0.487	No guidelines or tick marks
S3	.981	13.19	.195	0.380	No guidelines or tick marks
S4	.983	5.89	.511	0.997	Guidelines, labeled with tones of *sléndro*
S5	.992	11.19	.159	0.322	No guidelines or tick marks
S6	.988	4.76	.526	1.035	Tick marks, labeled with tones of *sléndro*
Average	.987	8.91	.368	0.735	
			Western Musicians		
S7	.977	4.96	.529	1.055	No guidelines or tick marks
S8	.974	8.69	.486	0.985	Tick marks, labeled with diatonic tones
S9	.991	3.36	.848	1.719	Tick marks, labeled with diatonic tones
S10	.958	4.76	.342	0.705	No guidelines or tick marks
S11	.970	5.29	.489	0.999	No guidelines or tick marks
S12	.991	3.12	.916	1.872	Tick marks, labeled with diatonic tones
Average	.994	5.70	.602	1.222	
			All Musicians		
Average	.991	7.30	.485	0.978	

parent between the groups, however. Two of the Javanese subjects (S1, S6) gave more reliable responses than all but two (S9, S12) of the Western listeners.

Table 2 also shows estimates for each listener of *d′* for intervals that differ by 20 and 40 cents. These were computed as follows. The mean response for each interval was computed. The estimated *d′* for 20-cent differences was the average difference between the means for all intervals differing by 20 cents divided by the average standard deviation. The estimated *d′* for 40-cent differences was the average difference between the means for all intervals differing by 40 cents divided by the average standard deviation. Again, although the Javanese values were somewhat lower on average, the two groups overlapped. Thus, in terms of the linearity of

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

the magnitude estimates, the degree of reliability, and estimated d' values, the two groups of listeners were quite comparable.

Figures 2 and 3 show the results for the Javanese and Western musicians, respectively. For each musician, we present two graphs. The graph on the left shows the position of the marks made in response to each interval in the 10 blocks of trials. The positions of any vertical guidelines or tick marks added to the response sheets, as noted in Table 2, are indicated by the horizontal dashed lines. The graph on the right shows which intervals were confused with which other intervals. It summarizes the results of an analysis of variance done for each individual using block of trials as the repeated measure. A post-hoc test, the Fisher PLST (Statview (c) II, v. 1.03), tested which intervals were significantly different from one another (at $p <$ .05). The nonsignificant pairs of intervals are indicated by dots.[2] Specifically, the graph shows for each interval (on the x-axis) all of the intervals larger than it (on the y-axis) that were not significantly different. For example, a dot for $x = 60$ cents and $y = 80$ cents means that the responses for these intervals were not significantly different from one another. Thus, triangular areas containing dots correspond to regions of confusion in which similar responses were given.

Two general patterns were found for the Javanese musicians (Figure 2). One pattern is shown by S2, S3, and S5. The left graph shows that the judged interval size increased quite linearly with the actual interval size. However, the graphs exhibit quite a bit of variability in the responses for each interval. These listeners had the largest average standard deviations (see Table 2). As a consequence, a fairly large number of interval pairs were not significantly different from one another, as can be seen in the graphs on the right. None of these listeners added tick marks or guidelines to the response sheet. No clearly defined regions of interval confusion were obtained, although S2 exhibited a slight grouping that distinguishes small from large intervals. S2 claimed to be comparing the intervals to the *sléndro* tuning system, with the base tone of the intervals corresponding to *dhadha* (3). He also claimed to be using additional tones called *miring*. Neither S3 nor S5 reported trying to relate the intervals to tones in either tuning system.

The other general pattern is found for S1, S4, and S6. All three exhibited plateaus in their response functions, shown in the graphs on the left. The responses tended to cluster around the guidelines or tick marks that they added to the response sheets. Except at transitional regions between these plateaus, the responses for each interval were quite consistent. The transi-

2. The analysis is indifferent to the direction of the difference. However, only a single case was found with a significantly higher mean estimate assigned to the smaller of a pair of intervals (120 versus 160 cents, for subject S3).

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

A

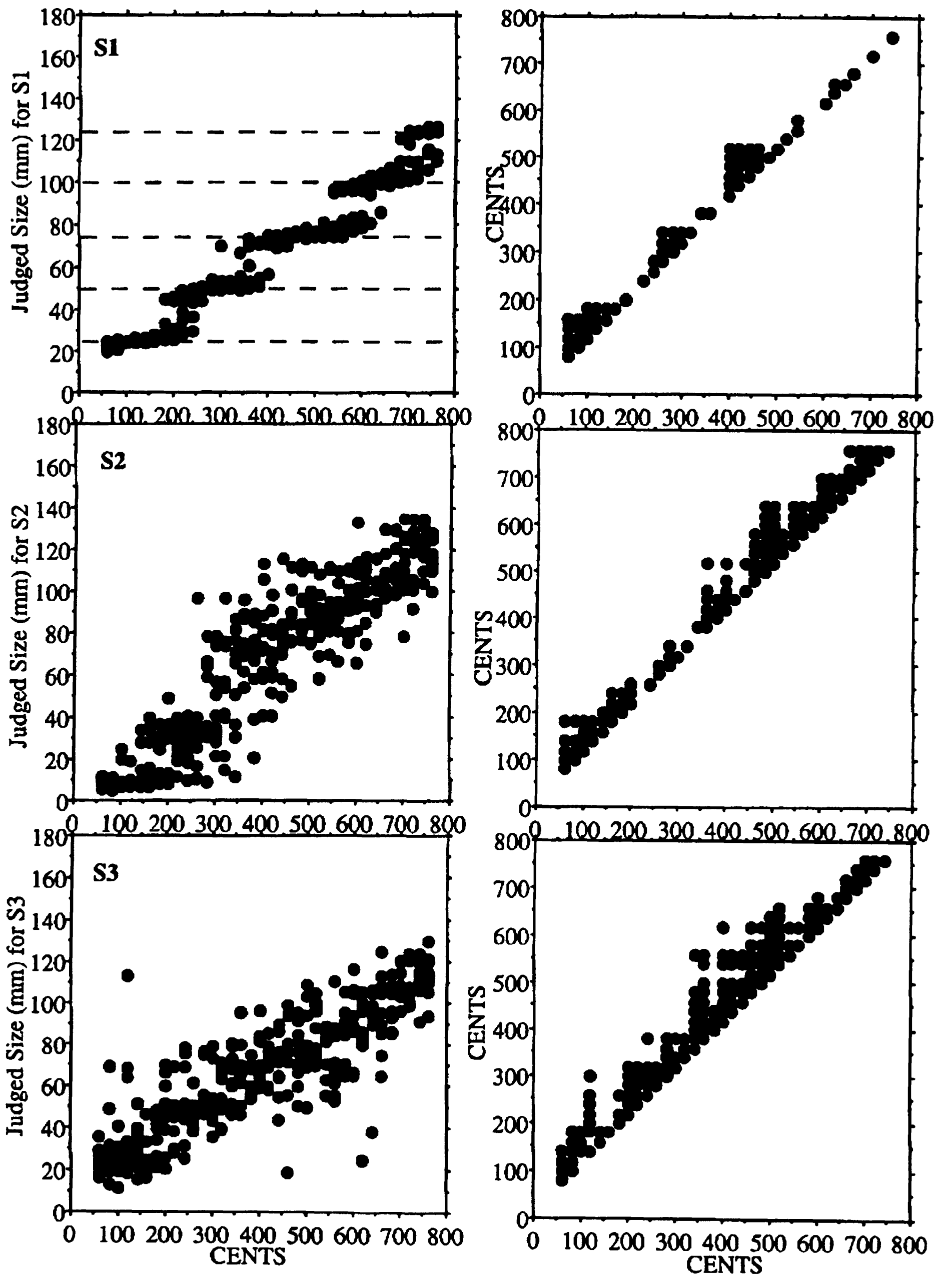

Fig. 2. A and B, Magnitude estimates (left graph) and interval confusions (right graph) for the Javanese participants (S1–S6). The interval confusion plots show, for each interval, all larger intervals that are not given significantly different responses. Horizontal dashed lines indicate the location of the guidelines or tick marks added to the response sheets by S1, S4, and S6.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

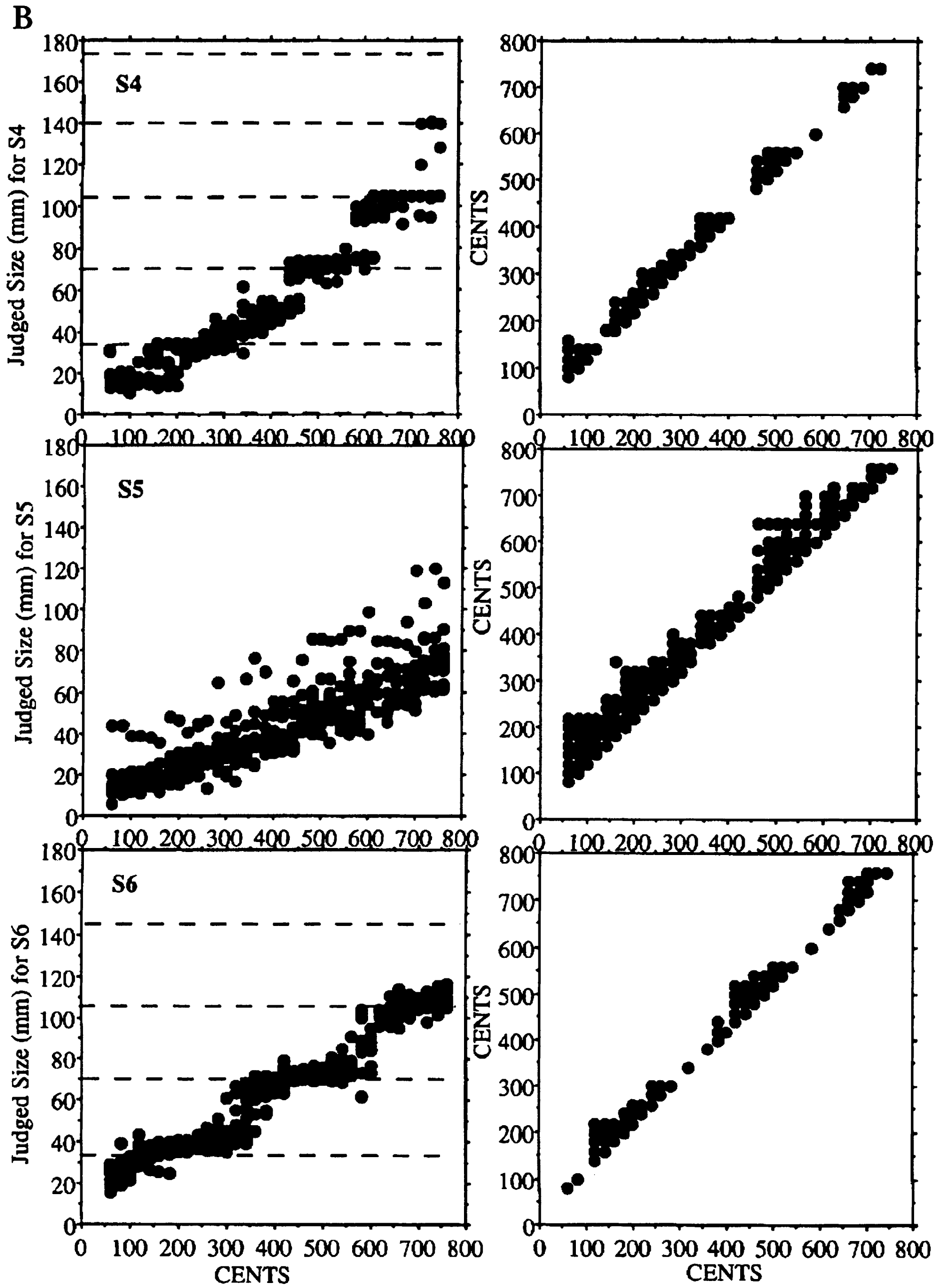

Fig. 2. (Continued)

tional regions between plateaus covered narrow ranges of intervals and tended to be marked by bimodal distributions of responses. These listeners had the smallest average standard deviations in their judgments (see Table 2). This meant that relatively few intervals were confused with one an-

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

A

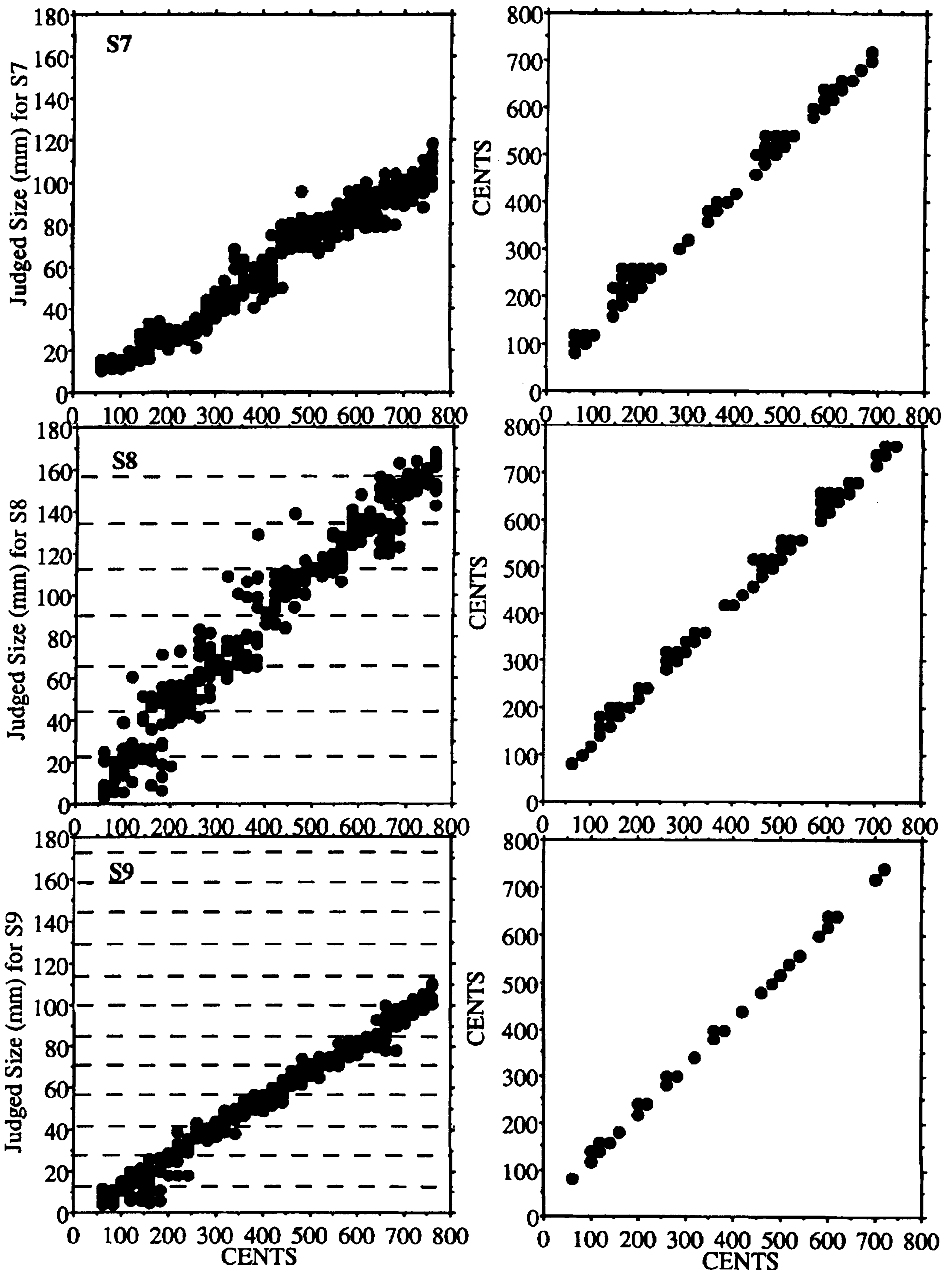

Fig. 3. A and B, Magnitude estimates (left graph) and interval confusions (right graph) for the Western participants (S7–S12). Horizontal dashed lines indicate the location of the tick marks added to the response sheets by S8, S9, and S12.

other. However, the interval confusions that did appear formed patterns that corresponded with intervals in gamelan tuning systems (Table 3).

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

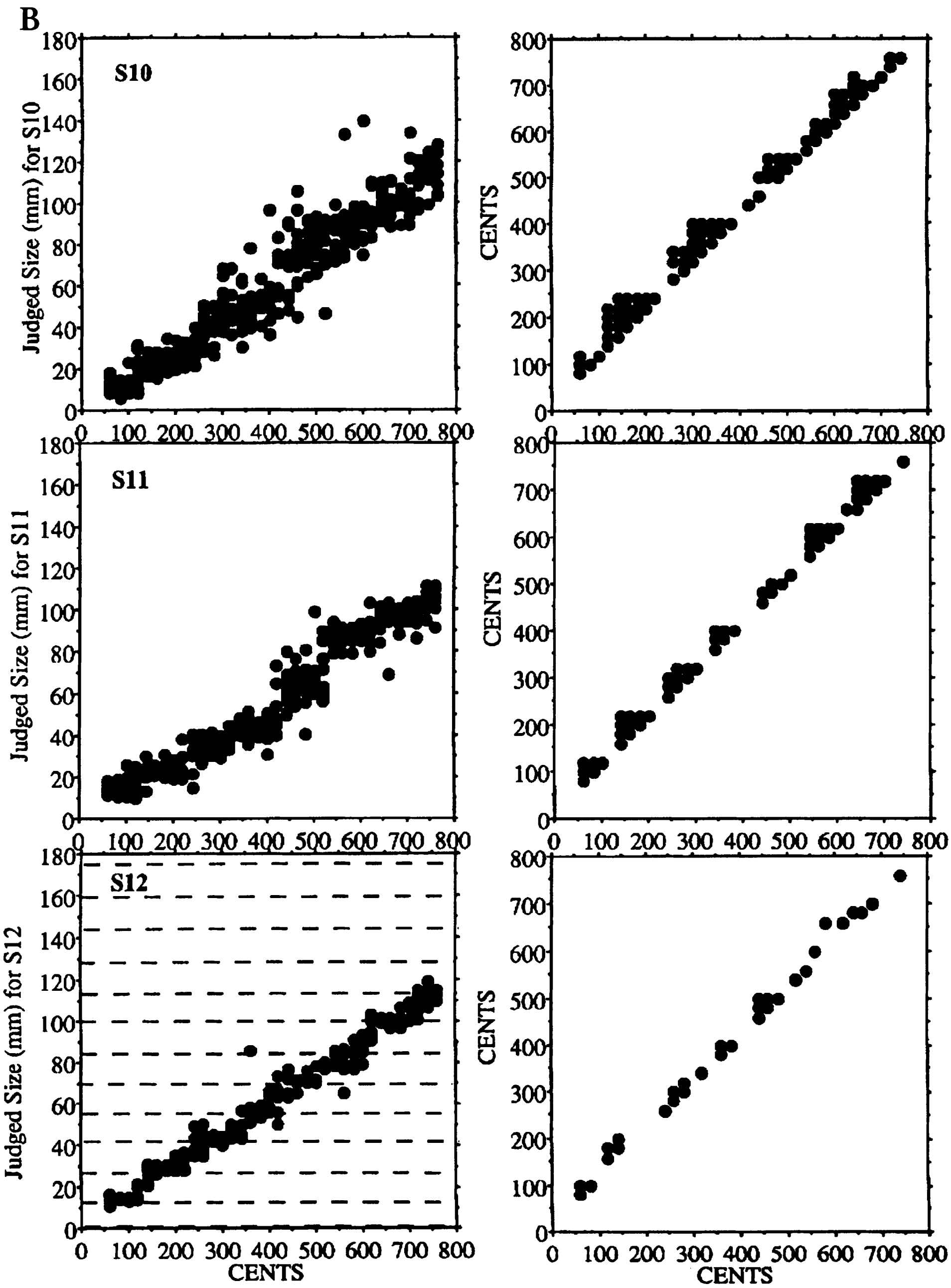

Fig. 3. (Continued)

S1 labeled the guidelines added to the response sheet (in 25-mm intervals) by tones of the *pélog* tuning system: the base tone was labeled *lima* (5), and the vertical guidelines were labeled *nem* (6), *barang* (7), *panunggul*

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

TABLE 3
Regions of Confusion for Three Javanese Musicians Compared with Pitch Measurements from Gamelan Instruments

	Region of Confusion			Relevant Gamelan Interval		
	Range	Spread	Midpoint	Range	Mean	Tuning
S1	60–160	100	110	95–145	114	*pélog* 5–6
	220–340	120	280	245–305	281	*pélog* 5–7
	400–520	120	460	? (491–577	534)	*pélog* 5–1
S4	460–540	80	500	470–515	495	*sléndro* 6–2
	640–720	80	680	? (715–755	735)	*sléndro* 6–3
				or ? (655–705	680)	*pélog* 6–3
S6	120–300	180	210	235–265	250	*sléndro* 6–1
	380–540	160	460	470–515	495	*sléndro* 6–2
	640–740	100	690	715–755	735	*sléndro* 6–3

(1), and *gulu* (2). He prepared himself for each trial by humming short ascending sequences of these tones. Consistent with this, a number of intervals in the range of 60–160 cents were confused with one another; this range includes values typical of the 5–6 interval in the *pélog* scale. Another set of intervals in the range of 220–340 cents were confused with one another, including values typical of the 5–7 interval. A third region was around 400–520 cents, somewhat smaller than the 5–1 interval. Beyond this range, the intervals were discriminated quite reliably.

S4 drew vertical lines (at approximately 35-mm intervals). He reported hearing the base tone of the intervals as *nem* (6), so his lines correspond to the *sléndro* pitches 1, 2, 3, and 5. However, S4's comments indicated that he was also listening in terms of *pélog* intervals, which he tried to relate to the *sléndro* intervals. This became apparent when he remarked that he had been working under the erroneous assumption that the tone *gulu* (2) in *pélog* is lower than *gulu* (2) in *sléndro*, and he adjusted his responses accordingly. Two distinct regions of confusion appear in his responses: one in the range from 460 to 540 cents, including the *sléndro* 6–2 interval; and the other in the range from 640 to 720 cents, somewhat smaller than the *sléndro* interval 6–3, and approximately equal to the corresponding *pélog* interval. Responses for smaller intervals generally increased with interval size, although the region of confusion that appeared for the smallest intervals may suggest that a *miring* pitch was being used in that range.

Finally, S6 drew tick marks at the top of each page (at 36-mm intervals), which he said represented the *sléndro* tuning system, with the base tone of the intervals corresponding to *nem* (6). Consistent with this, three distinct regions of confusion appear in his responses: one in the range from 120 to 300 cents, which would include the interval 6–1; one in the range from 380 to 540 cents, enclosing the 6–2 interval; and one in the range of 640 to 740

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

cents, somewhat smaller than the 6–3 interval. In all cases, the midpoint of the region was somewhat lower than the mean of the relevant interval measurements.

Figure 3 presents the results for the six Western musicians, S7–S12. Three (S8, S9, and S12) added tick marks to the response sheet, corresponding to tones of the diatonic scale. These were marked at intervals of 22.5, 14.5, and 14.5 mm, for the three listeners, respectively. The other three listeners did not alter the response sheet. Two of the listeners who added the tick marks, S9 and S12, produced extremely linear functions with small standard deviations. This meant that essentially all intervals (except those that differed by 20 cents) were reliably distinguished from one another, and no regions of confusion appeared. At the other extreme was S11 with a clear pattern of interval confusions. The regions were approximately 40–80 cents in width and correspond to the steps of the chromatic scale: m2, M2, m3, M3, P4, TT, and P5. In almost every case, the theoretical value of the tempered interval appears as the upper bound of the region, not as its center.

The three remaining listeners (S7, S8, and S10) produced responses between these extremes. S7 showed some confusion of intervals within the range from 140 to 240 cents, which includes the M2, and in the range from 440 to 520, which includes the P4. Some confusion also occurred around the m2 (60–100 cents), the M3 (340–400 cents), and the TT (560–640 cents). S8 also produced some regions of confusion, in this case around the m3 (260–340 cents), the P4 (460–520 cents), and the TT (560–640 cents). The responses of S10 showed two relatively large regions of confusion in the ranges of 140–220 cents, including the M2, and 260–360 cents, including the m3.

Discussion

The results of the experiment suggest that neither the artificiality of the experimental task nor the timbres of the tones used prevented the Javanese and Western musicians from applying their skills of musical interval perception. For each individual, the average magnitude estimates generally increased with interval size. The correlations between the estimates and the actual interval size were high and were essentially equal for Javanese and Western listeners. Thus, the stimulus timbre apparently gave rise to a clear perception of pitch for both Javanese and Western musicians. This is consistent with the findings of Carterette, Kendall, and DeVale (1993) and suggests that Schneider and Beurmann's (1993) claim that inharmonic partials impair the perception of pitch is limited to only certain gamelan instruments.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

However, the musicians who participated in the experiment showed considerable variability in the strategies they adopted in response to the task. These variations did not appear to be general differences between Javanese and Western listeners. In both groups, half of the listeners spontaneously altered the response form by adding guidelines or tick marks. The three Javanese musicians who did this showed plateaus in their response functions and corresponding regions of confusion. In contrast, two of the three Western musicians who added tick marks produced extremely reliable responses and essentially all intervals were distinguished from one another. It would seem, then, that this response strategy in and of itself did not determine whether or not regions of confusion appeared in the results.

Rather, it appeared that the musicians in both groups differ in the extent to which they related the intervals to tones in musical scales. Three of the Javanese musicians produced responses suggesting that they were assimilating the intervals to those found in Javanese music: one musician to tones of the *pélog* tuning system, one to tones of the *sléndro* tuning system, and one perhaps to a mixture of the two tuning systems. The other Javanese listeners showed no evidence that their judgments were influenced by tones in the musical scales. Two Western listeners reliably distinguished essentially all intervals from one another, with essentially no confusions. Three listeners showed regions of confusion around some but not all tones of the chromatic scale. The last listener showed regions that clearly corresponded to each tone of the chromatic scale. Thus, the two groups of listeners exhibited approximately equal variability in the degree to which the magnitude estimation judgments were made with reference to tones in the musical scales.

Because our central question concerns the relationship between the Javanese scale systems and the internal interval standards of the Javanese musicians, we focus on their results in more detail. As noted above, Javanese scales differ from Western scales in terms of the number and size of scale steps and the degree of tolerance of tuning variation in fixed-pitch instruments. Corresponding differences in the regions of confusion, if they exist, may appear in both the *location* of the regions and their *width*. Three regions of confusion were found for S1. The first two of these coincide quite clearly with intervals in the *pélog* scale; the third appears shifted downward relative to the expected corresponding interval. Two distinct regions appeared in the responses of S4, who reported trying to use *sléndro* and *pélog* systems simultaneously. The first coincides quite well with tuning measurements of a *sléndro* scale interval. The second could be accounted for in terms of either *sléndro* or *pélog* intervals. Three distinct regions appeared in the responses of S6. These regions corresponded to *sléndro* intervals, although displaced somewhat lower than the corresponding intervals.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

The regions of confusion for the Javanese musicians were somewhat wider than those for the Western musicians. This could be due to either greater variations in the tunings of intervals in gamelan music or larger spacing between scale tones. Because of the greater range of variation found in Javanese gamelan intervals, Javanese musicians would need to be able to recognize the equivalence of these variants across different gamelan ensembles. Thus, they may use their internal interval standards more flexibly than do Western listeners. Consistent with this, the narrowest regions of confusion were found for the Western listeners. The regions of confusion were somewhat larger for the musician (S1) using a *pélog* reference, which has somewhat larger intervals. The regions of confusion were largest for the musician (S6) using a *sléndro* reference, which has the largest intervals. These results extend Cooke's hypothesis (1992) that links the size of scale steps, the breadth of interval categories, and variability of tuning; these factors also appear to be related to the width of internal interval standards.

Perhaps the most striking difference was that the Javanese musicians seem to have access to two distinct sets of standards, whereas the Western musicians have access to only one. S1 selected *pélog*, whereas S6 selected *sléndro*. The question naturally arises as to whether musicians can use both sets of standards simultaneously. In this interval perception task, S4 claimed to be consulting both standards, but apparently had difficulty relating intervals between the two tuning systems. During the course of the experiment, he realized that he had miscalculated the size of a certain *pélog* interval relative to the *sléndro* scale. This suggests that he had been relying on a relatively abstract representation of the relationship between the two tuning systems, not a perceptual one. Given that the tones of the two systems are almost never mixed in practice, it is plausible that he had no perceptual image of the relations between them to draw upon.

In summary, *sléndro* and *pélog* seem to provide Javanese musicians with two distinct perceptual sets. In this respect, the Javanese situation seems somewhat comparable to the Vietnamese situation, as studied by Keefe et al. (1991), in which different modal standards could be evoked by the musician. In addition, the present results suggest that the two perceptual sets can be combined only with difficulty. To use a linguistic analogy, Javanese musicians may be "bilingual" in *sléndro* and *pélog* (cf. Elman, Diehl, & Buchwald, 1977) but exhibit costs in "code switching" from one to the other.

A number of subsidiary results are interesting in light of earlier studies of interval perception in Western musicians. Two of the Western listeners (S9, S12) exhibited impressively accurate responses in the magnitude estimation task, although they clearly approached the task in terms of equal-tempered intervals. These results contrast with those of Siegel and Siegel (1977b), a contrast that is striking given the similarity of their method to

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

our own. The differences found may be attributable to the fact that our study used a fixed base tone for all the intervals, whereas theirs varied the base tone. Alternatively, or in addition, the difference may also be a consequence of the particular musicians' expertise. In this experiment, S9 is a harpsichord player practiced in historical temperaments. S12, a string player, possesses absolute pitch. Both have extensive experience with fine tuning variations within interval categories.

The accuracy of these two listeners, when expressed in terms of estimated d' (MacMillan, 1987; Braida & Durlach, 1972), is generally comparable to those in the study of Burns and Campbell (1994). In their study, the average separation for a d' of 1 was 31 cents; for the best participant (RP3), it was 25 cents. What is distinctive in the performance of S9 and S12, then, is not their average sensitivity, but the shape of the sensitivity functions. They exhibited fairly uniform sensitivity along the interval continuum, whereas the sensitivity functions for Burns and Campbell's listeners (1994, p. 2710) showed a peak-trough form.

However, consistent with the results presented in Burns and Ward (1978), Siegel and Siegel (1977b), and Burns and Campbell (1994), we found intersubject variability in both the precise location and width of regions of confusion. This was true of both Western and Javanese musicians. We also found that the regions of confusion were not always centered on the theoretical values. For both the Javanese musician S6 and the Western musician S11, the regions were shifted downward relative to the theoretical values. In addition, for some musicians, certain ranges of intervals contained regions of confusion, whereas other ranges displayed fairly linear patterns. Usually, the results for larger intervals showed more linearity than the results for smaller intervals. These findings suggest that stable internal interval standards exist in only some parts of the interval range.

The results of the present experiment have a number of implications for cross-cultural studies of interval perception that may be done in the future. We found an unexpected degree of variation in the response patterns across individuals. These differences could not be attributed to distinctive tendencies of the Javanese and Western musicians. Within-group variation was as large as between-group variation and necessitated presenting the results on an individual basis. This meant that only a few cases were available for some of the comparisons of interest, for example, between the width of internal interval categories for the two Javanese tuning systems. Moreover, we observed that individuals approached the task by adopting different strategies. Reporting these in some detail was motivated by an interest in understanding how expert musicians bring their skills to bear on a controlled experimental task such as that used here.

Both these observations suggest the importance of increasing the number of participants beyond the number used here and in other similar stud-

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

ies (e.g., six subjects in Siegel & Siegel, 1977b; five in Burns & Ward, 1978; four in Burns & Campbell, 1994). It also suggests that it is important to consider the kinds of musical skills that are acquired in a particular musical culture and their relationship to the experimental situation. In the West, for example, ear-training exercises—a form of interval identification task—are required of music students, who thus are in a sense pretrained for similar tasks in psychological experiments (Burns & Campbell, 1994, p. 2710). Lastly, consideration needs to be given to the selection of participants, even from among trained musicians. Western musicians with comparable performance experience vary somewhat in tests of relative pitch (Howard, Rosen, & Broad, 1992), and only a small subset of conservatory students met Siegel and Siegel's (1977b, p. 400) requirements of relative pitch. In a music culture such as Java's, where the extent of formal music education is relatively restricted, and where ear-training does not form part of the curriculum, musical experience may be even less reliable as a basis for selection. These considerations suggest that cross-cultural studies call for more detailed description of the individual's musical skills, their approach to the experimental task, and their individual results than is typical in psychological research.[3]

References

Acker, B., Pastore, R., & Hall, M. (1995). Within-category discrimination of musical chords: Perceptual magnet or anchor? *Perception & Psychophysics, 57,* 863–874.

Berliner, P. (1978). *The soul of mbira.* Berkeley: University of California Press.

Braida, L. D., & Durlach, N. I. (1972). Intensity perception. II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America, 51,* 483–502.

Burns, E. M. (1977). *The perception of musical intervals (frequency ratios).* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Burns, E. M., & Campbell, S. (1994). Frequency and frequency-ratio resolution by possessors of absolute and relative pitch: Examples of categorical perception? *Journal of the Acoustical Society of America, 96,* 2704-2719.

Burns, E. M., & Ward, W. (1978). Categorical perception—phenomenon or epiphenomenon. *Journal of the Acoustical Society of America, 63,* 456–468.

3. We thank the musicians who participated in this study. This research was begun while the first author held a Fulbright-Hayes doctoral dissertation fellowship from the U.S. Department of Education. The research in Indonesia was conducted under the auspices of the Indonesian Institute of Sciences, with local sponsorship in Surakarta from the Indonesian College of the Arts. The first author is deeply grateful to the personnel of these institutions for their gracious support. He is also grateful to R. Weisling for invaluable logistical support and advice in the field and to Matthew Arciniega for computer assistance. Fellowship support from the Society for the Humanities made possible the first author's presence at Cornell University during 1994–1995. Initial planning of this project was carried out while the second author was a fellow at the Center for Advanced Study in the Behavioral Sciences, with support from the National Science Foundation and the James McKeen Cattell Fund. We are also grateful to James Kippen, an anonymous reviewer, and especially Edward M. Burns for their careful readings of an earlier version of this paper and for their many detailed and helpful suggestions.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

Burns, E. M., & Ward, W. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press.

Carterette, E., Kendall, R., & DeVale, S. (1993). Comparative acoustical and psychoacoustical analyses of gamelan instrument tones. *Journal of the Acoustical Society of Japan, 14,* 383–396.

Cooke, P. (1992). Report on pitch perception experiments carried out in Buganda and Busoga (Uganda). *African Music, 7,* 119–125.

Dowling, W. J., Lung, K., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41,* 642–656.

Ellis, A. (1885). On the musical scales of various nations. *Journal of the Royal Society of Arts, 33,* 485–527.

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America, 62,* 971–974.

Francès, R. (1988). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum.

Hood, M. (1966). Slendro and pelog redefined. *Selected Reports of the Institute for Ethnomusicology, 1,* 28–48.

Howard, D., Rosen, S., & Broad, V. (1992). Major/minor triad identification and discrimination by musically trained and untrained listeners. *Music Perception, 10,* 205–220.

Jairazbhoy, N. A., & Stone, A. W. (1963). Intonation in present-day North Indian classical music. *Bulletin of the School of Oriental and African Studies, 26,* 119–132.

Keefe, D., Burns, E. M., & Nguyen, P. (1991). Vietnamese modal scales of the *dan tranh*. *Music Perception, 8,* 449–468.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance, 5,* 579–594.

Levy, M. (1982). *Intonation in North Indian music: A select comparison of theories with contemporary practice.* New Delhi: Biblia Impex.

MacMillan, N. A. (1987). Beyond the categorical/continuous distinction: A psychophysical approach to processing modes. In Harnad, S. (Ed.), *Categorical perception.* Cambridge: Cambridge University Press.

Miyazaki, K. (1995). Perception of relative pitch with different references: Some absolute-pitch listeners can't tell musical interval names. *Perception & Psychophysics, 57,* 962–970.

Morton, D. (1974). Vocal tones in traditional Thai music. *Selected Reports of the Institute for Ethnomusicology, 2,* 166–175.

Morton, D. (1976). *The traditional music of Thailand.* Berkeley: University of California Press.

Powers, H. (1980). India. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians*. London: MacMillan.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music, 18,* 60–72.

Rakowski, A. (1991). Context-dependent intonation variants of melodic intervals. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech and brain.* Stockholm: Department of Speech Communication and Music Acoustics, Royal Institute of Technology.

Schneider, A., & Beurmann, A. (1993). Notes on the acoustics and tuning of gamelan instruments. In B. Arps (Ed.), *Performance in Java and Bali.* London: School of Oriental and African Studies, University of London.

Shepard, R. & Jordan, D. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science, 226,* 1333–1334.

Siegel, J. & Siegel, W. (1977a). Absolute identification of notes and intervals by musicians. *Perception & Psychophysics, 21,* 143–152.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions

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

Surjodiningrat, W., Sudarjana, P. J., & Susanto, A. (1972). *Tone measurements of outstanding Javanese gamelans in Jogjakarta and Surakarta.* 2nd rev. ed. Yogyakarta, Indonesia: Gadjah Mada University Press.

Tracey, H. (1970). *Chopi musicians.* London: Oxford University Press.

Wachsmann, K. P. (1950). An equal-stepped tuning in a Ganda harp. *Nature, 165,* 40–41.

Wachsmann, K. P. (1967). Pen-equidistance and accurate pitch: A problem from the source of the Nile. In L. Finscher & C.-H. Mahling (Eds.), *Festschrift für Walter Wiora.* Basel: Bärenreiter Kassel.

Wapnick, J., Bourassa, G. & Sampson, J. (1982). The perception of tonal intervals in isolation and in melodic context. *Psychomusicology, 2,* 21–37.

Zatorre, R. (1983). Category-boundary effects and speeded sorting with a harmonic musical-interval continuum: Evidence for dual processing. *Journal of Experimental Psychology: Human Perception & Performance, 9,* 739–752.

Zatorre, R., & Halpern, A. (1979). Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception & Psychophysics, 26,* 384–395.

This content downloaded from 128.84.127.139 on Thu, 4 Apr 2013 15:41:28 PM
All use subject to JSTOR Terms and Conditions