Can Pitch Be Measured On A Ratio Scale?

Florian Kattner and Wolfgang Ellermeier
Technische Universität Darmstadt
kattner@psychologie.tu-darmstadt.de

Abstract

Stevens' direct scaling methods have been widely used in basic and applied psychophysical research. They rest on the assumption that observers are capable of processing ratios on a subjective intensity scale. Axioms fundamental to this assumption have only recently been formulated, and empirically tested for a limited number of sensory continua, most notably loudness. The present study attempts to extend this line of research to the perception of pitch. N=13 participants were asked to adjust pitch intervals defined by pure tones in a frequency range between 264 and 699 Hz to specific ratios (e.g., 1/2 or 2/3) of a standard interval. Whereas most participants' adjustments were in accordance with the axiom of monotonicity, there were a few significant violations of commutativity and multiplicativity, particularly if the standard interval exceeded an octave. In contrast to what has been found for loudness, however, multiplicativity still held for most observers. Systematic effects of stimulus range and musical training appear to distinguish pitch scaling from other quantifiable sensations.

In a typical application of direct psychophysical scaling, the observer is asked to assign numbers to the intensities of sensations (Stevens, 1956, 1975). That is, the participant is either responding with a numeral to the intensity of stimulus that is being presented (magnitude estimation), or the numerical values are presented and the participant is asked to adjust the stimulus intensity accordingly (magnitude production). In order to obtain valid estimates of sensations by these direct scaling methods, the estimated or adjusted magnitudes must be meaningful on a ratio scale. Fundamental conditions (axioms) to allow the interpretation of an observer's scaling behavior have been formulated by Narens (1996). Some of these axioms - monotonicity, commutativity, and multiplicativity - are empirically testable. An extension of the axiomatization formulated by Narens (1996) is inherent to Luce's theory of psychophysical scaling, and it allows to fractionate magnitudes of sensations.

According to Narens (1996) and Luce (2002), magnitude estimates or productions are valid only if the commutativity property (or threshold proportion commutativity Luce, 2002) can be proven empirically. Commutativity holds if the outcome of two successive adjustments (e.g., 2× as loud and 3× as loud) is independent of the order in which these adjustments are made. Furthermore, if the adjustments are consistent also with the multiplicative property, then the numbers assigned by the participant may be interpreted as the mathematical numbers they stand for. Multiplicativity (or the probability reduction property Luce, 2002) holds, if two successive adjustments (e.g., 2× as loud and 3× as loud) result in the same stimulus intensity as a single adjustment based on the mathematical product of the two numerical values (e.g., 6× as loud).

These axioms have been tested recently, mostly for the perception of loudness. In a magnitude production experiment, Ellermeier and Faulhammer (2000) showed that whereas commutativity holds (with the exception of one participant), the multiplicativity assumption was violated. That is, doubling the loudness of a 1000-Hz sine tone and then tripling the outcome resulted in the same sound pressure level as first tripling and then doubling the tone. However, in all participants, the resulting sound pressure level exceeded the level which results from a single adjustment to produce a tone that is six times as loud. Similar results have been obtained for the fractionation of loudness (Zimmer, 2005), and in other sensory domains (e.g., in visual size perception; Augustin & Maier, 2008).

The present study is an attempt to test the axioms of commutativity and multiplicativity for the perception of pitch. Perhaps, pitch scaling has been neglected so far because it has been suggested that pitch is not a 'prothetic' sensory continuum characterized by a power function of stimulus intensity (e.g., Stevens & Galanter, 1957). On the other hand, it has been shown that cross-modal matching of pitch with a 'prothetic' continuum (duration) is possible (e.g., Painton, Cullinan, & Mencke, 1977). One problem with the scaling of pitch is that it involves two sensory phenomena, (a) the order of sounds from low to high (tone height) and (b) the similarity between tones that are separated by an octave (tone chroma). To avoid confounds between tone chroma and tone height, participants were asked to fractionate pitch intervals rather than to produce absolute pitch heights in the present study. Additionally, two different starting intervals were compared, one that corresponds to an octave interval, and another one that exceeds an octave interval.

A generalized ratio production procedure (Steingrimsson & Luce, 2005a) was used to obtain fractions of pitch intervals. In this procedure, a lower tone y and a higher tone x is presented as the standard interval, and the participant is asked to match a comparison interval from y to z to a given fraction p of the standard interval. Thus, z can be defined as a tone that makes the subjective interval from y to z stand in the ratio p to the standard interval from y to x (equation 1).

$$x \circ_p y := z \tag{1}$$

With regard to this procedure, the commutativity and the multiplicativity axioms can be formalized as follows:

Commutativity Axiom:
If $x \circ_p y := a$, $a \circ_q y := b$, $x \circ_q y := c$, and $c \circ_p y := d$, then $b = d$

Multiplicativity Axiom:
If $x \circ_p y := a$, $a \circ_q y := b$, and $r = pq$, then $x \circ_r y := b$

Method

Participants

$N = 13$ participants (7 female) were recruited for individual testing in a single-walled sound-attenuated listening room (IAC). Ages ranged between 20 and 55 years ($M = 29.9; SD = 9.6$). All participants had normal hearing with thresholds not exceeding 20 dB HL at any of the audiometric frequencies between 125 and 8000 Hz. Three musically trained participants were included. They reported more than 7 years of musical instruction and continued musical activity.

Stimuli and Apparatus

445 pure sine tones were generated digitally with a sampling rate of 44.1-kHz for all integer frequencies between 259 and 703 Hz. Each tone had a duration of 250 ms including 20-ms cosine-shaped rise and decay ramps. The signals were passed through a Behringer HA 8000 Powerplay PRO-8 headphone amplifier and played back with Beyerdynamics DT 990 PRO (250 Ohm) headphones. The sounds were attenuated to comfortable sound pressure levels of about 70 dB (500 Hz).

Procedure

The experiment consisted of 150 pitch-production trials divided into 3 test sessions. In each session, the participants completed 50 trials, and they were allowed to take a break after 30 trials. In the first session, there was an additional training block consisting of 10 trials which was not included in the analysis. In each trial, two pitch intervals (a lower tone followed by a higher tone) were presented successively to the participant via headphones. The first pair of tones was the standard pitch interval, and the second pair was the comparison pitch interval. The first tones of both intervals were always identical and lower in frequency than the second tones. In accordance with the procedure described by Steingrimsson and Luce (2005b, p. 312), the two tones of an interval were separated by a silent gap of 450 ms, and the two intervals were separated by a 750-ms gap.

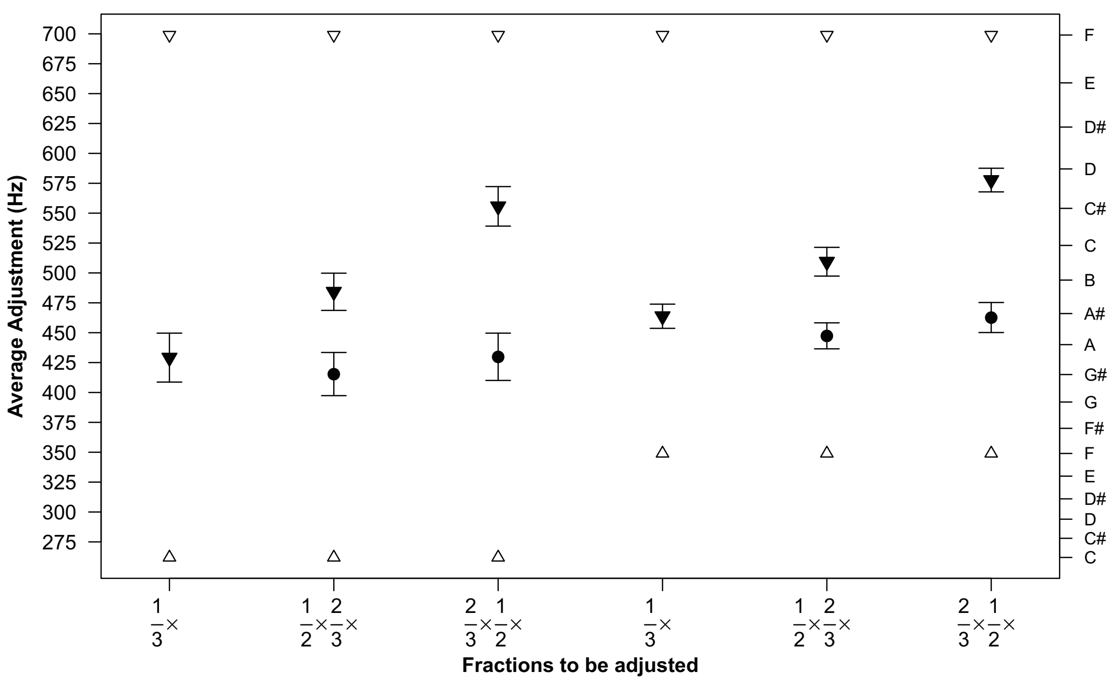
The participants' task was to adjust the size of the comparison interval to a given fraction of the standard pitch interval ($\frac{1}{3}\times$, $\frac{1}{2}\times$ or $\frac{2}{3}\times$) by changing the frequency of the higher tone in the comparison interval. At the beginning of each trial, the to-be-adjusted tone had a randomly chosen frequency lying between the lower tone and 699 Hz. A short instruction together with the respective fraction was displayed on the screen during each trial. The frequency of the higher tone of the comparison interval could be decreased/increased by pressing the left or right cursor key, respectively. The frequency increment/decrement Δf_i was a logarithmic function of the frequency f_i of the tone in that trial: $\Delta f_i = (10 \cdot (log(f_i) - log(262)) + 1)Hz$. By pressing the "Shift" key together with the cursor key, the adjustments could be accelerated (in that case, Δf_i was multiplied by 10). After each adjustment, both pitch intervals were played again (with the last tone having changed in frequency). The participants were encouraged to repeat the adjustments until they were content with the size of the comparison interval. There was no time limit to the task. The adjustments were confirmed by pressing the "Enter" key. The next trial started after a 1-s delay.

Blocks of 10 different types of adjustments were repeated 15 times each. The order of the adjustments within each block was randomized. For two different frequency ranges (starting intervals: 262-699 Hz / C4-F5 and 329-699 Hz / F4-F5), the participants had to produce comparison intervals that were $\frac{1}{3}\times$, $\frac{1}{2}\times$ and $\frac{2}{3}\times$ as large. Additionally, there were two successive adjustments in which either $\frac{2}{3}\times$ or $\frac{1}{2}\times$ of a previously fractionated interval ($\frac{1}{2}\times$ or $\frac{2}{3}\times$ of the starting interval, respectively) had to be adjusted.

Results

The participants made $M = 11.8$ key presses per adjustment on average (ranging between $M_{CL} = 5.4$ and $M_{XS} = 25.3$ individual key presses). There were $M = 5.8 (SD = 3.5)$ accelerated pitch changes, and $M = 5.3 (SD = 3.1)$ small pitch changes per adjustment. In $M = 9.6\%$ (individually ranging between 0% and $M_{JK} = 40.7\%$) of the trials, the participants did not use small-step adjustments. The mean standard deviation of the 15 adjustments per type of trial was 46.2 Hz (ranging between $SD_{MH} = 4.0$ Hz and $SD_{CL} = 77.5$ Hz).

Figure 1. Arithmetic means of individual median adjustments (with standard errors) as a function of the fractionation condition. The open triangles indicate the frequency range; simple fractionations of the starting interval are marked by filled triangles; the results of successive adjustments are depicted by filled circles.



For the total sample, the average median adjustments in the ten fractionation conditions are illustrated in Figure 1. First of all, it can be seen that the instructions to produce one $\frac{1}{3}\times$, $\frac{1}{2}\times$, or $\frac{2}{3}\times$ the starting pitch interval resulted in distinguishable and monotonically increasing adjustments (filled triangles in Fig. 1). Friedman tests revealed that the median $\frac{1}{3}\times$, $\frac{1}{2}\times$, and $\frac{2}{3}\times$ adjustments differed significantly for the large (262-699 Hz), $\chi^2(2) = 24.15; p < .001$, and for the small starting interval (349-699 Hz), $\chi^2(2) = 24.15; p < .001$.

Furthermore, similar comparison intervals were adjusted when the participants had to successively produce $\frac{2}{3}\times$ and $\frac{1}{2}\times$ of the starting interval irrespective of the order (filled circles in Fig. 1). Wilcoxon signed rank tests (two-sided) on the listeners' median adjustments revealed that $\frac{2}{3} \times \frac{1}{2}\times$ and $\frac{1}{2} \times \frac{2}{3}\times$ adjustments did not differ significantly with both starting intervals, $V = 27; p = .21$, and $V = 22; p = .11$, respectively. That is, for the total sample commutativity was not violated.

Additional Wilcoxon signed-rank tests (two-sided) revealed that the participants' median successive adjustments ($\frac{2}{3} \times \frac{1}{2}\times$ and $\frac{1}{2} \times \frac{2}{3}\times$) and their median $\frac{1}{3}\times$ adjustments did not differ significantly both for the large, $V = 44.5; p = .97$, and the small starting interval, $V = 68; p = .12$. This indicates that multiplicativity holds for the total sample, as well.

Violations of the axioms of commutativity and multiplicativity were further tested statistically for the adjustments of each participant. The median individual adjustments together with the significance of axiom violations (two-sided Wilcoxon rank-sum tests) can be seen in Table 1.

Table 1: Median adjustments of the higher tone of the comparison interval (Hz) and evaluation of violations of the axioms of commutativity (p_c) and multiplicativity (p_m) for each listener (monotonicity was significant in all participants). Statistically significant ($p < .1$) axiom violations are printed in boldface. Musically trained participants are indicated by an asterisk.

	Starting interval									
	262-699 Hz					349-699 Hz				
Participant	$\frac{1}{2} \times \frac{2}{3}\times$	$\frac{2}{3} \times \frac{1}{2}\times$	$\frac{1}{3}\times$	p_c	p_m	$\frac{1}{2} \times \frac{2}{3}\times$	$\frac{2}{3} \times \frac{1}{2}\times$	$\frac{1}{3}\times$	p_c	p_m
AK	587	579	578	.31	1	569	572	546	.79	**<.001**
AS*	378	371	372	**.06**	.11	440	448	452	**.01**	**.06**
CL	436	493	549	.46	**.01**	437	513	538	.48	**.04**
CW	426	389	456	.36	**.09**	437	440	450	.16	.11
DD	405	411	379	.63	.87	453	450	477	.69	.10
FK	400	404	376	1	.25	433	430	440	.97	.28
JK	388	408	385	.11	.69	413	426	440	.69	.28
JS	368	398	471	.36	.11	418	473	463	**.08**	1
MH*	360	356	364	**.04**	**.07**	437	441	441	.20	.34
MS	433	422	430	.52	.79	425	456	453	.10	.84
SD	411	435	441	**.01**	.31	447	433	439	.72	.84
WE	483	560	453	**.01**	**.02**	468	518	452	.46	.46
XS*	325	362	325	.21	.32	439	415	438	.11	.80

Discussion

The present investigation shows that instructing listeners to generate different fractions of pitch intervals produces outcomes consistent with the axiom of monotonicity.

There were, however, violations of commutativity, particularly for the large frequency range (4 of 13 participants), though not consistent in magnitude and direction (s. Table 1). This finding is in contrast to what has been reported for other sensory continua (e.g., loudness, Ellermeier & Faulhammer, 2000), and it indicates that the magnitude production of pitch intervals may not be valid on a ratio scale. However, when the frequency range was restricted to an octave, commutativity was violated by only two participants. This suggests that a ratio scale of subjective pitch intervals may exist within but not beyond an octave interval.

Moreover, there were violations of the multiplicativity axiom in some participants. Such violations have also been reported for other sensations (e.g., Augustin & Maier, 2008; Ellermeier & Faulhammer, 2000; Zimmer, 2005), and they indicate that the participants did not use the fractions like scientific numbers. Again, there were more axiom violations for the large starting interval than for the octave starting interval, indicating that tone chroma might interfere with the consistent adjustment of pitch ratios.

References

Augustin, T., & Maier, K. (2008). Empirical evaluation of the axioms of multiplicativity, commutativity, and monotonicity in ratio production of area. *Acta Psychologica*, *129*, 208-216.

Ellermeier, W., & Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevens's ratio-scaling approach: I. Loudness production. *Perception and Psychophysics*, *62*, 1505-1511.

Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, *109*, 520-532.

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*, 109-129.

Painton, S. W., Cullinan, W. L., & Mencke, E. O. (1977). Individual pitch functions and pitch-duration cross-dimensional matching. *Perception and Psychophysics*, *21*, 469-476.

Steingrimsson, R., & Luce, R. D. (2005a). Evaluating a model of global psychophysical judgments - I: Behavioral properties of summations and productions. *Journal of Mathematical Psychology*, *49*, 290-307.

Steingrimsson, R., & Luce, R. D. (2005b). Evaluating a model of global psychophysical judgments - II: Behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, *49*, 308-319.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes - loudness. *American Journal of Psychology*, *69*, 1-25.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: Wiley.

Stevens, S. S., & Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, *54*, 377-411.

Zimmer, K. (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception and Psychophysics*, *67*, 569-579.