Synesthesia-like Mappings of Lightness, Pitch, and Melodic Interval

Timothy L. Hubbard

The American Journal of Psychology, Vol. 109, No. 2. (Summer, 1996), pp. 219-238.

Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28199622%29109%3A2%3C219%3ASMOLPA%3E2.0.CO%3B2-0

The American Journal of Psychology is currently published by University of Illinois Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. 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/journals/illinois.html.

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.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Wed Oct 31 10:10:38 2007

Synesthesia-like mappings of lightness, pitch, and melodic interval

TIMOTHY L. HUBBARD
University of Arizona

Synesthesia-like mappings between visual lightness and auditory pitch and between visual lightness and melodic interval were examined. When subjects rated how visual lightnesses and auditory pitches "fit together," lighter stimuli fit better with higher pitches, and darker stimuli fit better with lower pitches. These patterns were stronger against black than against white visual backgrounds; however, effects of visual background were eliminated when subjects had a large set of lightness levels from which to choose the visual lightness level that best fit a given auditory pitch or melodic interval. When subjects chose which visual lightness best fit or matched a melodic interval, lighter stimuli were chosen for ascending melodic intervals, and darker stimuli were chosen for descending melodic intervals. Larger melodic intervals produced more extreme (lighter or darker) choices. Auditory pitch exhibits meaningful synesthesia-like mappings with visual lightness when unidimensionally varied in frequency and when multidimensionally varied in interval size and direction.

When some persons are presented with a stimulus in one modality, they often experience sensations or images that are more typically associated with a second modality. Such experiences are called *synesthesia* (for reviews, see Cytowic, 1989; Marks, 1975, 1978). Synesthetic mappings between vision and audition have been studied using a variety of visual and auditory stimuli. One of the more commonly studied synesthetic mappings is between visual lightness and auditory pitch. Auditory stimuli that are lower in frequency typically evoke visual sensations of stimuli that are darker, and auditory stimuli that are higher in frequency typically evoke visual sensations of stimuli that are lighter (e.g., Karwoski, Odbert, & Osgood, 1942; Marks, 1975, 1978; Ortmann, 1933).

Although true synesthesia is relatively rare, many nonsynesthetes are reliably and consistently able to match qualities or quantities of stimuli in one modality with qualities or quantities of stimuli in a second modality (e.g., Marks, 1974; Wicker, 1968). In true synesthesia the cross-modality experience occurs involuntarily, but in nonsynesthetes these matches or judged comparisons are made deliberately. Even so, the nature of the choices made by nonsynesthetes may reveal something

AMERICAN JOURNAL OF PSYCHOLOGY
Summer 1996, Vol. 109, No. 2, pp. 219–238

about the representation underlying perceptual experience in general and synesthesia in particular and also about synesthesia-like experiences such as metaphor (e.g., Marks, 1978, 1982) and cross-modality similarity (e.g., Marks, 1987). To the extent that patterns produced by nonsynesthetes mimic the patterns reported by synesthetes, an explanation of synesthesia at a more abstract or nonsensory level is supported; to the extent that the patterns produced by nonsynesthetes differ from those produced by synesthetes, an explanation of synesthesia at a more concrete or sensory level is supported.

When nonsynesthetic subjects are asked to match visual lightness with either auditory pitch or loudness, they often produce matches very similar to the correspondences reported by synesthetic subjects. Marks (1974, 1975) has interpreted the similarity in auditory-visual mapping between synesthetes and nonsynesthetes as suggesting that stimuli in the different modalities may tap a common connotative meaning mediated by higher cognitive (possibly linguistic) processes; hence, synesthesia may therefore represent meaning in its most sensory form. On this view, the similarity between synesthetes and nonsynesthetes occurs because the language of the nonsynesthetes preserves the same types of relationships found in the more sensory coding of synesthetes (e.g., perhaps both loudness and brightness are coded by the underlying rate of neural activity). Claims of an underlying linguistic mediation of synesthesia-like mapping are bolstered by findings of cross-cultural agreement as to the connotative meaning of various colors (e.g., Osgood, 1960). Evidence from classical psychophysical cross-modality scaling studies is also consistent with the notion of a common mediator between stimuli in different modalities; for example, Stevens (1961, see also Stevens, 1975) has shown that subjects can equate the sensation magnitudes of stimuli drawn from two different modalities, thus implying that different modalities are comparable (although whether either directly to each other or to some third intervening more abstract quality is not yet clear from the psychophysical data).

Some aspects of nonsynesthetic performance, however, are not similar to synesthetic performance. For example, Cytowic and Wood (1982) reported that synesthetic subjects tended to use a more restricted and more asymmetrical response range than nonsynesthetic subjects and that some nonsynesthetic controls were unclear as to how to map one dimension onto another and reported picking seemingly arbitrary strategies. Such differences are consistent with either differences in mapping between the different modalities for synesthetes and nonsynesthetes or with a lack of systematic mapping in nonsynesthetes. Failure to find a consistent one-to-one mapping led Cytowic (1989) to suggest that synesthesia is not a mediated phenomenon, because a higher order cognitive or linguistic mediator would be assumed to function similarly in both synesthetes and nonsynesthetes.

Although Marks suggested that synesthesia-like mappings are mediated by higher order cognitive or linguistic processes and that semantic processes may play a role in or even control synesthesia-like mappings, Cytowic (1989) argued that nonsynesthetes actually engage in a great deal of synesthesia, but that those synesthetic processes normally do not reach consciousness. Marks's view might be characterized as a top-down strategy that adds synesthesia-like mappings relatively late in cognitive processing, whereas Cytowic's view involves a bottom-up strategy that places synesthesia-like mappings very early in perceptual processing and then edits those mappings out. Part of the apparent conflict between the views of Marks and Cytowic may arise from differences in the linguistic codability of the stimulus dimensions each investigator focused on in his studies. Specifically, the gustatory stimuli in Cytowic and Wood's (1982) studies might not be as familiar or as easily labeled verbally as the visual and auditory stimuli in Marks's (1974) studies; therefore, the image of the gustatory stimuli may have been kept in a more sensory coding format longer than a visual or auditory stimulus might have been.[1] It would then not be surprising if qualities of the stimuli that are less easily verbalized are less likely to result in higher order linguistic mediation.

Many of the experiments on synesthesia-like mappings between auditory pitch and visual lightness in nonsynesthetes presented subjects with an experimenter-specified visual stimulus (e.g., a patch of gray) and required them to match a movable auditory stimulus (e.g., adjust the pitch or loudness of a tone generator) to the visual stimulus. Although this method has produced fairly reliable data, it is not known to what extent the previously reported patterns are dependent upon this specific methodology. One purpose of the current experiments is to attempt to replicate the previously obtained patterns through different methodologies: for example, by presenting visual and auditory stimuli simultaneously and having subjects rate how well the stimuli "fit together," and by presenting an experimenter-specified auditory stimulus and having subjects choose which visual stimulus from a set of visual stimuli "fits best" with the auditory stimulus.

Many of the experiments in synesthesia-like matching between auditory pitch and visual lightness in nonsynesthetes examined only the mapping between the absolute lightness of the visual stimuli and either the pitch or loudness of the auditory stimuli; however, any effects of the contrast between the visual target and the visual background on judgments of the synesthesia-like mapping between the visual and auditory stimuli were not explicitly examined. Unless the visual background is varied, contrast between the visual target and the visual background will be confounded with luminance of the visual target. When the visual background has been varied, these changes in the background have typically occurred between experimental sessions and not during a single session. Therefore, it is possible that any effects of background may have been obscured by differences in the perceptual salience of the background or in the perceptual set of the subjects across sessions. A second purpose of the current experiments is to compare explicitly the effects of absolute lightness with any effect of contrast between the visual stimuli and visual background on the judged fit or match between visual and auditory stimuli within a single experimental session.

Although auditory pitch has been used in numerous investigations of synesthesia and synesthesia-like mappings, the perceived pitch has often been treated as a unidimensional stimulus varying only along the frequency domain. However, the literature on music cognition has revealed that pitch is not a unidimensional stimulus corresponding solely to frequency (for a brief overview, see Hubbard & Stoeckig, 1992), but consists of at least two dimensions (Shepard, 1982) corresponding to pitch height (absolute frequency) and tone chroma (relative location of the pitch within a scale collapsed across octaves). Given that tone chroma may be an important aspect of the representation of pitch, it is possible that synesthesia-like connections or influences could be affected by not just the absolute frequency of the pitch (pitch height) but also by the relationship between pitches (differences in interval size and tone chroma). A third purpose of the current experiments, then, is to examine the more multidimensional aspects of pitch (e.g., size and direction of musical intervals) and whether specific interval sizes or interval directions are judged as matching or fitting better with specific visual lightness levels.

EXPERIMENT 1

Previous investigators required subjects to match the pitch or loudness of an adjustable auditory stimulus to the lightness of a fixed visual target; whether the visual target was presented against a white or a black visual background did not significantly affect the matching (e.g., Marks, 1974). It is possible, though, that any effects of visual background may have been obscured by considering only the lightness of the target and not the contrast between the target and background. In this experiment subjects were presented with visual lightness target stimuli varying in luminance, with white or black visual backgrounds, and with auditory pure tone stimuli varying in frequency. Subjects judged how well the visual and auditory stimuli fit together, and the data were analyzed by comparing both the luminance of the target and the contrast of the visual target and background with the ratings of how well each visual stimulus fit with each auditory stimulus.

METHOD

Subjects

The subjects in all experiments were undergraduates from Eastern Oregon State College who participated for partial credit in an introductory psychology course. Subjects were not screened for synesthetic abilities. There were 13 participants in Experiment 1.

Apparatus

All stimuli were generated by an Apple Macintosh IIcx microcomputer equipped with an Apple RGB color monitor. The RGB color monitor was approximately 60 cm from the subjects, but they could adjust this distance slightly to achieve maximum comfort and confidence in their responses.

Stimuli

The visual stimuli consisted of squares differing in luminance. The squares measured 200 pixels (approximately 8.33°) along each side and were centered on the RGB monitor screen. The squares were created by varying the firing intensity of the color guns in the RGB monitor. On each trial the three guns fired at the same intensity level, and that level was either 10, 30, 50, 70, or 90% of their total capacity, resulting in stimuli ranging in perceived lightness from a very dark gray to a very light gray. The background color on the monitor surrounding the square was either white or black on each trial. The brightness control for the monitor was adjusted to its maximum level throughout the experiment. The auditory stimuli consisted of sine wave tones of eight different frequencies: 200, 300, 450, 675, 1012.5, 1518.75, 2278.125, and 3417.188 Hz. These tones were spaced 700 cents apart so that the perceived interval sizes between successive pitches (i.e., a fifth) were equal. On each trial a single visual luminance and auditory frequency were simultaneously presented for a duration of 4 s. Each subject received 400 trials (2 backgrounds × 5 luminances × 8 frequencies × 5 replications) in a different random order.

Procedure

The subjects were first given a practice session consisting of 12 trials randomly chosen from the experimental trials. Subjects pressed a designated key to begin each trial. A single visual stimulus and a single auditory stimulus were presented simultaneously. After the stimuli vanished, subjects were asked to rate how well the lightness and the pitch fit together.Ratings ranged from 1 (*lightness and pitch did not fit together at all*) to 9 (*lightness and pitch fit together very well*). Subjects were instructed to enter intermediate numbers for intermediate levels of fit. No specific type of ordering, definition, or example of fit was provided.

RESULTS AND DISCUSSION

The fitness ratings were analyzed in a 2 (Background) × 5 (Luminance) × 8 (Frequency) repeated measures ANOVA. White backgrounds ($M = 4.91$) led to significantly lower overall fitness ratings than did black backgrounds ($M = 5.27$), $F(1, 12) = 7.33$, $MSE = 4.81$, $p < .02$. As shown in Figure 1, the Background × Luminance × Frequency interaction was marginally significant, $F(28, 336) = 1.43$, $MSE = 0.90$, $p = .08$, such that a trend for a wider response range was found with a black background. Frequency significantly influenced fitness ratings, $F(7, 84) = 4.59$, $MSE = 13.59$, $p < .01$; in general, higher frequencies received lower fitness ratings than lower frequencies. The Luminance × Frequency interaction was highly significant, $F(28, 336) = 7.53$, $MSE = 2.39$, $p < .01$. As shown in Figure 1, more luminous (i.e., lighter) visual stimuli received higher fitness ratings when presented with higher auditory frequencies and less luminous (i.e., darker) visual stimuli received higher fitness ratings when presented with lower auditory frequencies. This pattern was especially pronounced when the background was black. No other factors were significant, all *F*s < 1.1, *p*s > .39.

The preceding analysis considered the absolute level of luminance, but we may also examine how the degree of contrast between the background and the square influenced fitness ratings. If the data are reanalyzed in a 2 (Background) × 5 (Contrast) × 8 (Frequency) repeated measures ANOVA, we again see that white backgrounds produced lower fitness ratings than black backgrounds, $F(1, 12) = 7.33$, $MSE = 4.81$, $p < .02$, and higher frequencies produced higher fitness ratings than lower frequencies, $F(7, 84) = 62.33$, $MSE = 13.59$, $p < .01$. The Background × Contrast × Frequency interaction was also highly significant, $F(28, 336) = 7.51$, $MSE = 2.41$, $p < .01$. As seen in Figure 2, when the contrast was low, relatively greater effects of frequency were seen against a black background than were seen against a white background. No other factors were significant, all *F*s < 1.78, *p*s > .10.

Figure 1 shows the decline in rated fitness of less luminous visual stimuli and the increase in rated fitness of more luminous visual stimuli with increasing auditory frequency, but the same pattern might not be quite as evident in Figure 2. Closer examination of the top panel of Figure 2 shows that high contrast (i.e., dark) visual stimuli against a white background decreased in rated fitness with increasing auditory frequency; closer examination of the bottom panel of Figure 2 shows that low contrast (i.e., dark) visual stimuli against a black background also decreased

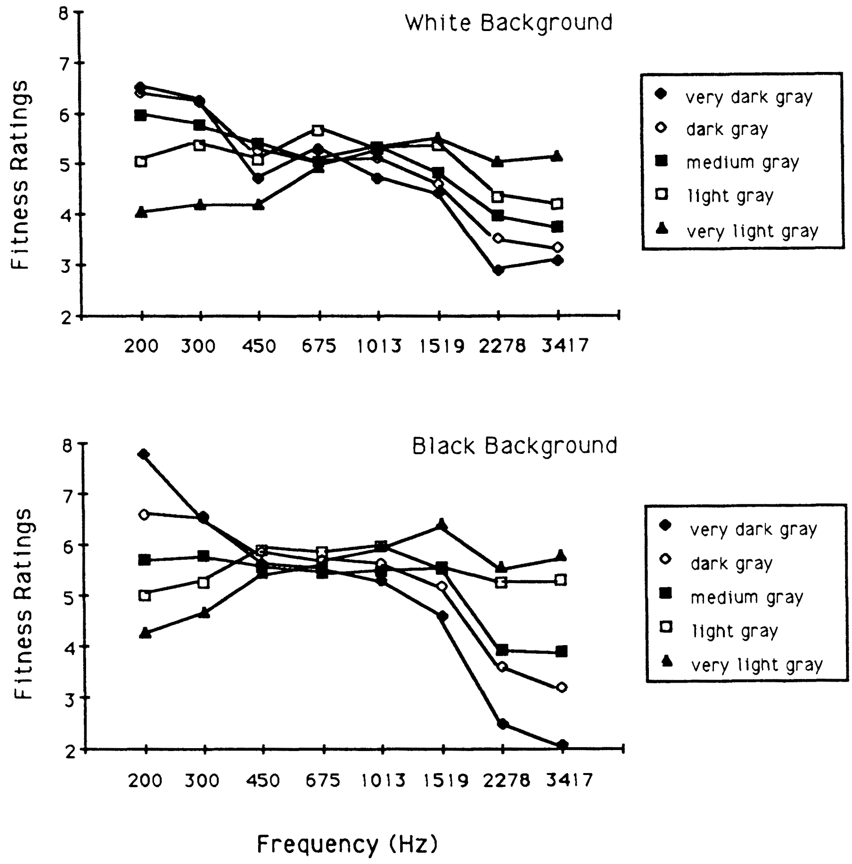

Figure 1. Fitness ratings of visual luminance stimuli as a function of auditory frequency in Experiment 1

in rated fitness with increasing auditory frequency. Thus, both analyses and all four panels of Figures 1 and 2 show a consistent result: Lighter visual stimuli received higher fitness ratings when paired with higher auditory pitches and received lower fitness ratings when paired with lower auditory pitches, whereas darker visual stimuli received lower fitness ratings when paired with higher auditory pitches and received higher fitness ratings when paired with lower auditory pitches. The higher fitness ratings for light square/high pitch pairs and for dark square/low pitch pairs obtained in Experiment 1 are consistent with the previous literature on both synesthesia and synesthesia-like metaphorical mapping, and so the connections between increases in lightness and

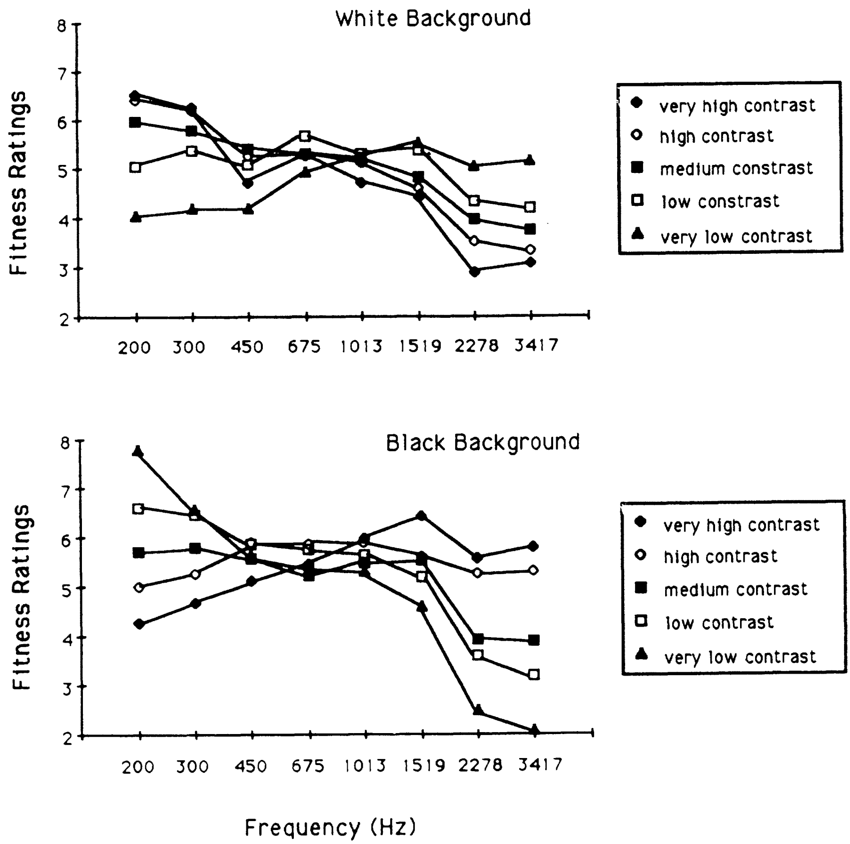

Figure 2. Fitness ratings of visual contrast stimuli as a function of auditory frequency in Experiment 1

increases in pitch do not appear to be an artifact of the previous methodologies in which subjects adjusted the auditory stimulus to match the visual stimulus.

In Experiment 1, visual background influenced subjects' judgments but this influence of background is at odds with the failure to find such an effect of background in Marks (1974). Why might the current experiments find an effect of background, whereas Marks (1974) did not find such an effect? Two possible explanations involve methodological differences between the study by Marks and Experiment 1. One possibility is that Marks's failure to find an effect of background was based on a comparison between different experimental sessions, whereas the effect of background in Experiment 1 was based on a comparison within a single experimental session. A background that remains constant across trials within an experimental session may not be as perceptually salient as a background that varies across trials within an experimental session; therefore, the contrast between the background and the visual stimuli in Experiment 1 may have been more perceptually salient than the contrast between the background and the visual stimuli in Marks's experiments. A more perceptually salient background might receive greater attentional or other processing, and so might then be able to exert a greater overall influence on perception than a less perceptually salient background.

A second possible explanation for the difference between Marks's experiments and Experiment 1 involves the degree to which subjects could adjust the stimuli. Marks's subjects were presented with an experimenter-specified visual luminance and then produced the best matching auditory frequency, whereas in Experiment 1 subjects were presented with a single auditory frequency and a single visual luminance that were both experimenter-specified and then judged how well the two stimuli fit together. In Marks's experiments subjects could adjust (via requests to the experimenter) the frequency of the auditory stimuli, whereas in Experiment 1 subjects could not adjust either the visual or the auditory stimuli. Perhaps if subjects are presented with a range of values or given a chance to pick one value from a range, context may become less salient as more attention is required by the processing of the wider range of alternative possible stimulus values in the adjustable modality.

EXPERIMENT 2

In this experiment, subjects were presented with the same auditory stimuli used in Experiment 1 and chose the luminance that best fit each individual auditory stimulus. Specifically, on each trial a visual gray scale figure consisting of several adjacent columns of differing luminances was presented on the RGB monitor, and subjects indicated which of the columns seemed to fit best with the auditory frequency presented on that trial. This methodology paralleled that of Marks in that subjects determined which value of one modality best matched an experimenter-specified value in a second modality, but this methodology was the opposite of Marks's in that subjects were presented with an auditory frequency and chose a visual luminance. If having a range of stimuli to choose from minimizes the effect of the background, then we should expect a diminished (or no) effect of background.

METHOD

Subjects

The subjects were 14 undergraduates drawn from the same pool used in Experiment 1. None had participated in the earlier experiment.

Apparatus

The apparatus was the same as in Experiment 1.

Stimuli

The visual stimulus consisted of a gray scale figure. The entire figure was 550 pixels wide and 390 pixels high (approximately 22.92° × 16.25°) and was centered on the monitor screen. The gray scale figure was divided into 11 equal-sized adjacent columns, with each column measuring 50 pixels by 390 pixels (approximately 2.08° × 16.25°). Each column was a different luminance, and luminance was manipulated by changing the overall firing rate of the three color guns in the RGB monitor. The leftmost column had all guns firing at 100% (thus producing white), and each subsequent adjacent column decreased the firing rate by 10% (of the total firing capacity) until the rightmost column had all guns firing at 0% (thus producing black). The numbers 1–11 were printed above the columns, with the number 1 printed above the leftmost column, the number 2 above the adjacent column immediately to the right, and so on, to the number 11 printed above the rightmost column. The gray scale figure was the same on all trials and was presented against either a white or a black background on each trial. If the background was white, a thin black line (1 pixel wide) was drawn around the perimeter of the entire set of columns and the numbers above the columns were printed in black; if the background was black, a thin white line (1 pixel wide) was drawn around the perimeter of the entire set of columns and the numbers above the columns were printed in white. The auditory stimuli were the same as in Experiment 1. Each subject received 160 trials (2 backgrounds × 8 pitches × 10 replications) in a different random order.

Procedure

The subjects were first given a practice session consisting of 12 trials randomly chosen from the experimental trials. Subjects pressed a designated key to begin each trial. The visual gray scale figure and the auditory stimulus were presented simultaneously. Subjects were requested to choose which of the luminance columns seemed to fit best with the auditory pitch, but no specific type of ordering, definition, or example of fit was provided. The visual gray scale figure remained visible until the subject responded.

RESULTS AND DISCUSSION

The lightness preferences were analyzed in a 2 (Background) × 8 (Frequency) repeated measures ANOVA. No factors were significant, all Fs < 1.51, ps $> .17$. Closer examination of the data revealed two distinct patterns, however: One group of subjects ($n = 8$, henceforth referred to as the descending group) consistently chose more luminous (i.e., lighter) columns as best fitting the higher frequencies, and a second group of subjects ($n = 6$, henceforth referred to as the ascending group) consistently chose less luminous (i.e., darker) columns as best fitting the higher frequencies. Separate ANOVAS for these two groups revealed large effects of frequency (ps $< .01$), but no effects of background or a Background × Frequency interaction. The data for both descending and ascending groups are plotted in Figure 3. There was no effect of background on subjects' lightness preferences.

Given that Experiment 2 was similar to that of Marks (1974) in that one stimulus was fixed and subjects matched a second stimulus to the first, the lack of a background effect in Experiment 2 is consistent with the hypothesis that providing subjects with a range of stimuli in one modality from which to choose the best match to a fixed stimulus in a second modality decreases the effects of the background. Such a pattern might result from the limited capacities of attentional resources; specifically, when subjects have to consider a variety of possible stimulus values for one modality in making a match with a second modality, it is assumed that more attention would be required than when subjects have to consider only one value from each modality. In the former case a large portion (if not all) of available attention would be required by the stimuli, whereas in the latter case some attention could also be given to the context.

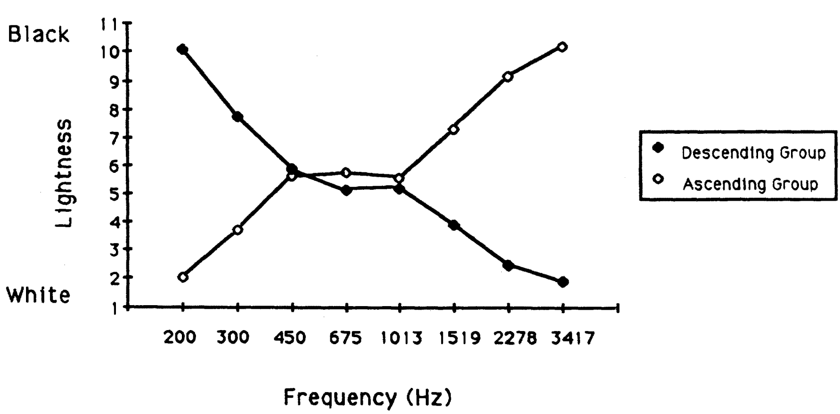

Figure 3. Lightness preferences of visual luminance stimuli as a function of auditory frequency in Experiment 2

The lack of a background effect in Experiment 2 may alternatively be due simply to the larger overall proportion of the CRT monitor occupied by the gray scale figure. In Experiment 1 the visual stimulus was a square measuring 200 pixels per side (40,000 square pixels) and in Experiment 2 the visual stimulus was a gray scale figure measuring 550 × 390 pixels (214,500 square pixels); therefore, the gray scale figure used in Experiment 2 clearly occupied a larger percentage of the screen than the square used in Experiment 1. It seems reasonable to conclude that the larger percentage of the screen given over to the background in Experiment 1 may contribute (at least in part) to the larger effects of background in Experiment 1. This account of background is weakened, however, by the realization that even though the size of the overall gray scale figure was larger than the visual square in Experiment 1, the areas of the individual columns within the gray scale figure were each smaller than the area of the visual square used in Experiment 1.

The preferred amount of visual lightness was strongly related to auditory pitch in Experiment 2, but the nature of this relationship varied across subjects as meaningful mappings existed within subjects but not across subjects. Marks (1974) reports that all of his subjects produced higher auditory frequencies for greater visual luminances. The differences between Experiment 2 and the study by Marks may have resulted from methodological differences as to which modality was fixed and which modality was adjustable or presented choice options. Regardless of the explanation for the empirical differences between Experiment 2 and the study by Marks, however, it should be noted that the data in Experiment 2 are more consistent with Marks's top-down view than with Cytowic's bottom-up view, because it is possible that different semantic associations could be learned (or produced by demand characteristics) but less plausible that an early modular perceptual processing could produce two completely opposite patterns.

EXPERIMENT 3

Although most studies of synesthesia or synesthesia-like mappings using auditory materials have focused on either perceived pitch or loudness and have presented single tones in isolation, two notable exceptions are studies by Odbert, Karwoski, and Eckerson (1942) and Lehman (1972). Odbert et al. (1942) presented nonsynesthete subjects with short segments of orchestral recordings. Subjects chose which adjective (from a list) best fit that selection and then judged whether the music evoked any sense of color. Subjects who agreed on mood tended to also agree on color, and subjects who disagreed on mood also tended to disagree on color. Lehman (1972) presented subjects with brief excerpts of classical music and had subjects rate how well a series of adjectives and colors described each excerpt. Models of the underlying cognitive spaces were computed, and the data were compatible with the suggestion that one of the axes of the cognitive space corresponded to a modal axis. The models suggested that synesthetes collapsed across the modal axis and that nonsynesthetes did not (at least involuntarily) collapse across the modal axis.

It is not yet clear which aspects of the orchestral music (e.g., pitch, loudness, key, contour, interval, timbre, etc.) were responsible for the seeming synesthesia-like responses in the data of Odbert et al. (1942) and Lehman (1972). Marks (1974) has documented synesthesia-like mappings between visual lightness and auditory pitch, but even if pitch was primarily responsible for the patterns in the Odbert et al. and Lehman data, the direction and size of the changes in pitch within or across their musical excerpts were uncontrolled. It is possible that any synesthesia-like responses were influenced by not just the absolute frequency of the pitch(es), but also by the direction and size of pitch change. Accordingly, in Experiment 3 a melodic interval (i.e., two sequential pitches) was presented to determine whether any systematic relationship existed between the direction or size of pitch change and the lightness of a visual stimulus that was judged to best fit that interval.

METHOD

Subjects

The subjects were 13 undergraduates drawn from the same pool used in previous experiments. None had participated in either of the previous experiments.

Apparatus

The apparatus was the same as in Experiment 1.

Stimuli

The visual gray scale figure was the same as in Experiment 2, and given the lack of background effects in Experiment 2, the gray scale figure was presented against a black background on each trial. The auditory stimulus on each trial consisted of two notes played sequentially. Two frequency ranges were used; for the low range the starting frequency was 500 Hz, and for the high range the starting frequency was 1000 Hz. The first note lasted 2 s; the second note immediately followed the cessation of the first note and also lasted 2 s. The pitch interval between the two notes corresponded to either 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 semitones along an equal-tempered scale. Each subject received 288 trials (2 directions × 2 frequency ranges × 12 interval sizes × 6 replications) in a different random order.

Procedure

The subjects were first given a practice session consisting of 12 trials randomly chosen from the experimental trials. Subjects pressed a designated key to begin each trial. The visual gray scale figure and the melodic interval were presented simultaneously. Subjects were asked to choose which of the luminance columns seemed to fit best with the melodic interval, but no specific type of ordering, definition, or example of fit was provided. The visual gray scale figure remained visible until the subject responded.

RESULTS AND DISCUSSION

The lightness preferences were analyzed in a 2 (Direction) × 2 (Frequency Range) × 12 (Interval Size) repeated measures ANOVA. Ascending intervals (M = 4.71) led to lightness preferences for significantly more luminous visual stimuli than did descending intervals (M = 6.16), $F(1, 12) = 5.56$, $MSE = 58.51$, $p = .04$, and direction also interacted with interval size, $F(11, 132) = 5.23$, $MSE = 1.19$, $p < .01$. As shown in Figure 4, increases in interval size for ascending intervals led to lightness preferences for more luminous (i.e., lighter) visual stimuli, but increases in interval size for descending intervals led to lightness preferences for less luminous (i.e., darker) visual stimuli. There was a marginally significant trend for intervals beginning on 500 Hz (M = 6.10) to produce lightness preferences for less luminous visual stimuli than intervals beginning on 1000 Hz (M = 4.77), $F(1,12) = 3.42$, $MSE = 81.41$, $p = .08$, a pattern corresponding to the previously discussed finding that lower pitches are mapped as darker. No other factors were significant, all Fs < 1.56, ps $> .11$.

A clear relationship was seen between the direction and size of an auditory melodic interval and the visual luminosity judged as fitting best with that interval. Lighter visual stimuli were judged to fit best with ascending intervals and darker visual stimuli were judged to fit best with descending intervals. Additionally, the size of the melodic interval influenced the lightness preferences. Larger melodic intervals (within the octave range used) led to preferences for more extreme levels of lightness or darkness; specifically, visually lighter stimuli were preferred for larger ascending intervals than for smaller ascending intervals, and visually darker stimuli were preferred for larger descending intervals than for smaller descending intervals.

These data complement the Odbert et al. (1942) and Lehman (1972) data in several ways. First, the data form a link between the previous research using single tones in isolation and the research using musical excerpts. Second, the data demonstrate a lawful correspondence be-

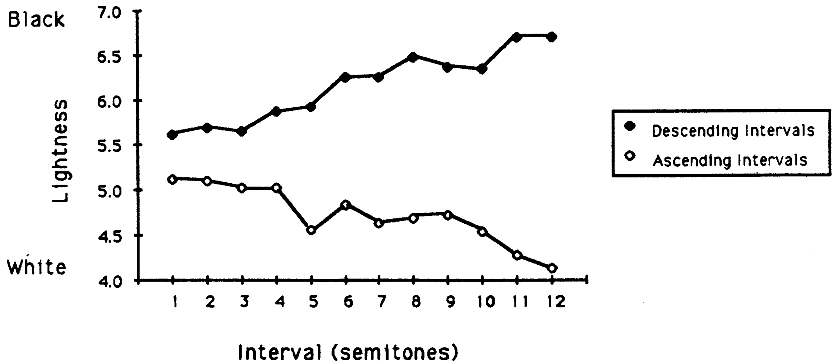

Figure 4. Lightness preferences of visual luminance stimuli as a function of interval size in Experiment 3

tween the size and direction of change in the musical stimulus and judgments of visual stimuli and provide controls that were noticeably absent in the previous literature. Such controls should assist in developing a more complete specification of the synesthesia-like relationship between visual lightness and auditory pitch. Third, although Odbert et al. reported that the color judgments approximated a color circle and Lehman reported that the cluster analysis resembled a color solid, it may simplify the underlying mapping to consider only luminance or perceived lightness rather than hue. To the extent that luminance is systematically mapped onto interval size, the task of explaining the remainder of the synesthesia-like responses evoked by music is reduced.

Given that the starting frequencies in Experiment 3 were limited (either 500 or 1000 Hz), larger ascending intervals terminated on frequencies that were necessarily higher than the terminating frequencies of the smaller ascending intervals, and the larger descending intervals terminated on frequencies that were necessarily lower than the terminating frequencies of the smaller descending intervals. The more extreme ending frequencies of the larger intervals might account for the more extreme lightness and darkness judgments given to the larger intervals, because the lightness preferences might then reflect the extremity of the final pitch of the interval rather than the size of the interval per se. Were this effect of frequency range solely responsible for the observed pattern, however, we would have expected intervals beginning with a 1000-Hz tone to have evoked lightness preferences for lighter visual stimuli than intervals beginning with a 500-Hz tone. Given that the effects of frequency range approached (but did not attain) significance in Experiment 3, effects of frequency range should not yet be ruled out.

A second possible explanation for the more extreme judgments with larger intervals builds on the notion of *representational momentum,* that is, the common finding that memory for the position of a moving target is generally shifted in the direction of anticipated future motion of the target (for review, see Hubbard, 1995). If subjects perceived the melodic interval as reflecting a single pitch source that traveled from the first pitch to the second pitch of the interval, and given that all of the intervals sounded for an identical duration, a single pitch source would have had to travel at a faster velocity to reach the terminal pitch of the larger intervals. Faster target velocities have been shown to lead to larger forward shifts in remembered position (e.g., Freyd & Finke, 1985; Freyd, Kelly, & DeKay, 1990; Hubbard & Bharucha, 1988), and so any shift in memory for the second tone would be greater for larger intervals than for smaller intervals. A larger forward shift would correspond to choosing a more extreme lightness or darkness. If, however, subjects perceived the melodic interval as reflecting two different pitch sources that sounded in succession, an explanation based on representational momentum is less plausible.

GENERAL DISCUSSION

Lighter visual stimuli were judged to fit better with higher auditory pitches and darker visual stimuli were judged to fit better with lower auditory pitches. Although this pattern was found regardless of the visual background upon which the visual stimuli were presented, the pattern was stronger against black backgrounds than against white backgrounds. This effect of visual background was eliminated when subjects could choose from a set of visual lightness levels the lightness that best matched a given auditory pitch or melodic interval. When subjects judged the fit of various visual lightness stimuli to various melodic intervals of different sizes and directions, ascending melodic intervals were judged to fit better with relatively lighter visual stimuli and descending melodic intervals were judged to fit better with relatively darker visual stimuli. Additionally, larger intervals fit better with more extreme lightness or darkness stimuli; more specifically, larger ascending intervals were judged to fit better with visually lighter stimuli than were smaller ascending intervals and larger descending intervals were judged to fit better with visually darker stimuli than were smaller descending intervals.

With relatively simple stimuli such as single auditory frequencies and visual luminances, the historical data may be conceived of as supporting a simple mediation based on a common intensity value (e.g., visual and auditory brightness, Marks, 1974). However, with somewhat more complex stimuli such as (the probably) multidimensional gustatory stimuli, the data are less supportive of a consistent mediation (e.g., Cytowic, 1989). To test the mediation idea more completely, it would seem necessary to examine synesthesia-like responses with more complex (yet better controlled) stimuli similar to the classical recordings used by both Odbert et al. (1942) and Lehman (1972). Presumedly subjects would have greater semantic knowledge about a complex stimulus such as music, even if only from tacit knowledge built up solely from exposure and not from explicit instruction (see Bharucha 1987, 1991; Krumhansl, 1990), than about a simple stimulus such as a patch of gray. The processing of a stimulus might be more easily penetrated to produce metaphoric or synesthesia-like judgments, therefore, if subjects have more semantic (or otherwise meaningful) information about that stimulus.

Semantic penetration might be accomplished by spreading activation from an adjacent node (e.g., from the mood-circle to the color-circle in Odbert et al., 1942) or by collapsing across the dimension specifying source (e.g., the modal axis in Lehman, 1972). In nonsynesthetes the penetrating information and the sensory experience are separable and the observers are aware of the separation, but in synesthetes both the penetrating information and the current percept are active with no clear separation between the two representations. Nonsynesthetes could exploit the separation between the sensory experience and the penetrating information by using coordinates along the modal axis as guides in the selection of appropriate metaphors or cross-modal comparison values. Both Odbert et al. and Lehman suggest that the cognitive space for adjectives (moods) may correspond to the cognitive space for color. Although such a correspondence may seem to be a de facto synesthesia, unless the corresponding quality in the nonpresented modality is necessarily and involuntarily evoked upon stimulation in the presented modality, synesthesia is not necessarily produced. Such an equivalence of cognitive spaces, however, may help explain commonalities in metaphor such as those detailed by Osgood (1960) and support Marks's notion of a top-down linguistic mediation.

It remains possible, however, that neither strictly top-down nor bottom-up models of synesthesia are appropriate; for example, perhaps synesthesia is indeed a normal and early part of perceptual processing engaged in by all people (as suggested by Cytowic), but that experience (and the knowledge gained by experience) may then function to tune the content of the synesthesia. In the majority of people such tuning may result in suppression or inhibition of the synesthesia, but in synesthetes the tuning produces specific and idiosyncratic responses and may also supply the linguistic relevant connotations (i.e., "red hot"). Such learned suppression or inhibition of synesthetic responding may also account for why synesthesia is more commonly reported in children than in adults (for a brief review, see Marks, 1975), because children might not have completed the fine tuning of their perceptual experience. In the current data such tuning may have contributed to the pattern seen in Experiment 3, because the intervals used were typical of the Western harmonic tradition. Such tuning would perhaps be even more relevant in studies such as Odbert et al.'s (1942) and Lehman's (1972) which used even more complex (i.e., meaningful) musical excerpts. To the extent that synesthesia-like mapping or judgments are mediated or linguistic, changes in meaning should lead to changes in mapping, but to the extent that true synesthesia is sensory rather than linguistic, changes in meaning should not lead to changes in mapping.

In summary, systematic mappings can be found between lightness, pitch, and melodic interval. When a single visual luminance and a single auditory frequency are presented, more luminous stimuli fit better with higher frequencies than with lower frequencies, and less luminous stimuli fit better with lower frequencies than with higher frequencies. These patterns are stronger when the visual stimulus is presented against a black than a white visual background, but this effect of background is eliminated when subjects are given a range of visual stimuli varying in luminance and allowed to choose the best-fitting luminance for a specific auditory frequency. When subjects choose the best fitting luminance for an ascending or descending melodic interval, more luminous stimuli are chosen for ascending intervals and less luminous stimuli are chosen for descending intervals, and more extreme luminance values (i.e., lighter or darker) are chosen for larger interval sizes. Not only does pitch contribute to meaningful synesthesia-like mappings with visual lightness when pitch is considered as a unidimensional stimulus varying only in frequency,but pitch also contributes to meaningful mappings with visual lightness when it is treated as a more multidimensional stimulus varying in melodic interval size and direction.

Notes

Portions of these data were presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC (November 1993). The author thanks Lisa Frantzen for assistance in data collection, and Lawrence Marks and W. Jay Dowling for helpful comments.

Correspondence concerning this article should be addressed to Timothy Hubbard, who is now at the Department of Psychology, Texas Christian University, Forth Worth, TX 76129. Received for publication March 24, 1994; revision received September 27, 1994.

1. It is possible that the more easily a value along some stimulus dimension may be linguistically labeled, the more likely subjects may be to discard the sensory image in favor of the linguistic label, but that the less easily a value along some stimulus dimension may be linguistically labeled, the less likely subjects may be to discard the sensory image (see discussions in Hubbard, 1994; Hubbard & Stoeckig, 1992). In any case, sensory coding may preserve distinctions or discriminations that are not preserved by linguistic coding (Hubbard, in press).

References

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological perspective. In P. Todd & G. Loy (Eds.), *Music and connectionism* (pp.84–99). Cambridge, MA: MIT Press.

Cytowic, R. E. (1989). *Synesthesia: A union of the senses.* New York: Springer-Verlag.

Cytowic, R. E., & Wood, F. B. (1982). Synesthesia II: Psychophysical relationships in the synesthesia of geometrically shaped taste and colored hearing. *Brain and Cognition, 1,* 36–49.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society, 23,* 443–446.

Freyd, J. J., Kelly, M. H., & DeKay, M. L. (1990). Representational momentum in memory for pitch. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1107–1117.

Hubbard, T. L. (1994). Memory psychophysics. *Psychological Research/Psychologische Forschung, 56,* 237–250.

Hubbard, T. L. (1995). Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonomic Bulletin & Review, 2,* 322–338.

Hubbard, T. L. (in press). The importance of a consideration of qualia to imagery and cognition. *Consciousness and Cognition.*

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics, 44,* 211–221.

Hubbard, T. L, & Stoeckig, K. (1992). The representation of pitch in musical imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 199–235). Hillsdale, NJ: Erlbaum.

Karwoski, T. F., Odbert, H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking: II. The role of form in visual responses to music. *Journal of General Psychology, 26,* 199–222.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Lehman, R. S. (1972). A multivariate model of synesthesia. *Multivariate Behavioral Research, 7,* 403–439.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology, 87,* 173–188.

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin, 82,* 303–331.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities.* New York: Academic Press.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 177–193.

Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 384–394.

Odbert, H. S., Karwoski, T. F., & Eckerson, A. B. (1942). Studies in synesthetic thinking: I. Musical and verbal associations of color and mood. *Journal of General Psychology, 2,* 153–173.

Ortmann, O. (1933). Theories of synesthesia in the light of a case of colored hearing. *Human Biology, 5,* 155–211.

Osgood, C. E. (1960). The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Science, 5,* 146–169.

Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 343–390). New York: Academic Press.

Stevens, S. S. (1961). The psychophysics of sensory function. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 1–33). Cambridge, MA: MIT Press.

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

Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space. *American Journal of Psychology, 81,* 178–188.

http://www.jstor.org

LINKED CITATIONS

You have printed the following article:

Synesthesia-like Mappings of Lightness, Pitch, and Melodic Interval
Timothy L. Hubbard
The American Journal of Psychology, Vol. 109, No. 2. (Summer, 1996), pp. 219-238.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28199622%29109%3A2%3C219%3ASMOLPA%3E2.0.CO%3B2-0

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

On Associations of Light and Sound: The Mediation of Brightness, Pitch, and Loudness
Lawrence E. Marks
The American Journal of Psychology, Vol. 87, No. 1/2. (Mar. - Jun., 1974), pp. 173-188.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28197403%2F06%2987%3A1%2F2%3C173%3AOAOLAS%3E2.0.CO%3B2-4

Mapping the Intersensory Regions of Perceptual Space
Frank W. Wicker
The American Journal of Psychology, Vol. 81, No. 2. (Jun., 1968), pp. 178-188.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28196806%2981%3A2%3C178%3AMTIROP%3E2.0.CO%3B2-B