Music-to-Color Associations of Single-Line Piano Melodies in Non-synesthetes

Stephen E. Palmer [1,*], **Thomas A. Langlois** [1] **and Karen B. Schloss** [2]

[1] Department of Psychology, University of California, Berkeley, CA 94720-1650, USA

[2] Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI 02912, USA

Abstract

Prior research has shown that non-synesthetes' color associations to classical orchestral music are strongly mediated by emotion. The present study examines similar cross-modal music-to-color associations for much better controlled musical stimuli: 64 single-line piano melodies that were generated from four basic melodies by Mozart, whose global musical parameters were manipulated in tempo (slow/fast), note-density (sparse/dense), mode (major/minor) and pitch-height (low/high). Participants first chose the three colors (from 37) that they judged to be most consistent with (and, later, the three that were most inconsistent with) the music they were hearing. They later rated each melody and each color for the strength of its association along four emotional dimensions: *happy/sad*, *agitated/calm*, *angry/not-angry* and *strong/weak*. The cross-modal choices showed that faster music in the major mode was associated with lighter, more saturated, yellower (warmer) colors than slower music in the minor mode. These results replicate and extend those of Palmer *et al.* (2013, *Proc. Natl Acad. Sci.* 110, 8836–8841) with more precisely controlled musical stimuli. Further results replicated strong evidence for emotional mediation of these cross-modal associations, in that the emotional ratings of the melodies were very highly correlated with the emotional associations of the colors chosen as going best/worst with the melodies ($r = 0.92, 0.85, 0.82$ and 0.70 for *happy/sad*, *strong/weak*, *angry/not-angry* and *agitated/calm*, respectively). The results are discussed in terms of common emotional associations forming a cross-modal bridge between highly disparate sensory inputs.

---

* To whom correspondence should be addressed. E-mail: sepalmer@gmail.com

1. Introduction

Achieving a unified, coherent experience of the external world necessarily involves integrating information across the five, qualitatively distinct, sensory modalities: seeing, hearing, touching, tasting and smelling. How might this remarkable feat be accomplished? A key requirement is solving the multi-sensory binding problem: identifying which information in one sensory modality arises from the same external source as information in another sensory modality (e.g., Spence and Driver, 2004). Such cross-modal correspondences are driven to a large extent by spatio-temporal correspondences: i.e., sensory events whose sources co-occur in space and time are likely to arise from the same environmental object or event (Calvert *et al.*, 2004; Spence and Driver, 2004). However, the human brain appears to use a much wider variety of associations to guide this process (e.g., Driver and Spence, 1998; Spence, 2011). To take one example of an auditory–visual correspondence, people tend to hear lower-pitched sounds as coming from larger-sized objects (e.g., McMahon and Bonner, 1983). Extensive empirical research has identified a number of other auditory–visual correspondences between low-level stimuli, such as single sounds and uniform patches of color (see Spence, 2011, for a review). Louder sounds, for example, tend to be associated with brighter colors (Bond and Stevens, 1969; Stevens and Marks, 1965) and larger-sized objects (Lewkowicz and Turkewitz, 1980), whereas higher-pitched sounds tend to be associated with lighter colors (Hubbard, 1996; Marks, 1974; Wicker, 1968), smaller objects (Marks *et al.*, 1987) and higher positions (Evans and Treisman, 2010; Lidji *et al.*, 2007; Rusconi *et al.*, 2006).

The primary evidence for such cross-modal correspondences is phenomenological: e.g., when objects of different sizes are visually present, people tend to report that higher frequency sounds appear to emanate from smaller objects and that lower frequency sounds tend to emanate from larger objects. There is also objective evidence from the existence of cross-modal Stroop effects, in which the speed and/or accuracy of a perceptual discrimination task in one modality is increased by the consistency (or decreased by the inconsistency) of a distractor stimulus in an irrelevant modality (e.g., Bernstein and Edelstein, 1971; Evans and Treisman, 2010; Rusconi *et al.*, 2006). For example, when participants must rapidly discriminate whether an auditory tone is high or low in pitch, a simultaneously presented visual object that is large tends to improve performance when the tone is low and diminish performance when the tone is high and *vice versa* (see Spence, 2011, for a review).

Why might these auditory–visual correspondences exist? Three hypotheses have been distinguished based on structural, statistical, and semantic relations (see Spence, 2011, for a review), each of which appears to have empirical support for certain dimensions and modalities. Structural associations are a

factor when pairs of dimensions have analogous neural coding systems in the brain. For example, louder tones may be coded as higher firing rates in the centers of the auditory system that represent the volume of sounds, and brighter lights as higher firing rates in the centers of the visual system that represent the intensity of lights. In such a case, there would be a natural, structural correspondence between the perception of louder tones and brighter lights, even though there is no logically necessary relation between these poles of the two bipolar dimensions. Statistical associations arise from pairs of dimensions that are correlated in the external world, such as the previously mentioned example of lower-frequency tones and larger-sized objects. The size of the resonating body of stringed instruments (bass, cello, viola, and violin) is one obvious example. Semantic associations arise from pairs of dimensions that have similar meanings. They often include a linguistic component, such as 'brighter' sounds being associated with 'brighter' colors, but any common semantic association would be sufficient, even in the absence of linguistically nameable features.

It is important to note that these three hypotheses are not mutually exclusive. The fact that neural codings of dimensions X and Y are correlated in the brain (the structural hypothesis) does not preclude either the possibility that they might also covary in the environment (the statistical hypothesis) or the possibility that they might be semantically related (the semantic hypothesis). Similarly, the fact that dimensions X and Y might covary in the environment (the statistical hypothesis) does not preclude the possibility that they might be semantically related (the semantic hypothesis). Each hypothesis is properly considered a logically independent possibility, such that none, one, two, or all three might be true simultaneously. The belief that evidence supporting one hypothesis implies anything about the truth or falsity of the other hypotheses is therefore unjustified.

A possible fourth hypothesis, about which there is far less theoretical discussion or empirical evidence, is that cross-modal correspondences might be mediated by emotion. There have been some reports and speculations of emotional mediation in cross-modal associations (e.g., Collier, 1996), and particularly in matching odors to non-olfactory stimuli such as colors (Schifferstein and Tanudjaja, 2004) and shapes (Seo *et al.*, 2010). This evidence tends to focus on the dimensions of 'valence' (*pleasant/unpleasant*, *good/bad* or *like/dislike*) and 'activity' (*active/passive* or *strong/weak*), both of which seem relevant to almost anything (Osgood *et al.*, 1957). Emotional mediation, at least as we will consider it here, is a much broader concept, however, including specific emotional dimensions, such as *happy/sad*, *agitated/calm*, and *angry/not-angry*. This larger range of specific emotional associations (see Note 1) has been even more neglected than valence and activity as the basis of possible emotional mediation effects. Somewhat surprisingly, Spence

(2011) mentions emotional mediation only once in his comprehensive tutorial on cross-modal associations and does not comment on its relation to the three types of correspondences mentioned above. Indeed, emotional correspondences might be considered a special case of semantic correspondences, for example, if one takes the semantic content of emotional responses to be just one aspect of a sensory event's meaning. However, it might equally well be considered a separate category of cross-modal correspondences, since emotions are often considered as categorically different from semantics and cognition, especially when they involve aspects of the phenomenological experiences that emotions typically generate (e.g., Zajonc, 1980).

The idea that specific emotional content might mediate cross-modal associations goes back at least to Arnheim (1986), who suggested that mappings between any pair of sensory dimensions should be possible if they have common underlying emotional associations. He mentioned associations between music and color as a specific example, and some early, relatively informal studies reported limited support for this idea. For example, Karwoski *et al.* (1942) reported that participants produced similar drawings in response to twelve different pieces of music played on a clarinet, similarities that were claimed to be well-explained by categorizing the pieces and drawings on emotional dimensions, such as *strong*, *happy*, and *exciting*. Further investigations using more sophisticated methods have produced evidence for more specific connections between music and color (e.g., Barbiere *et al.*, 2007; Bresin, 2005; Sebba, 1991), including suggestions that emotions may be involved in the associations.

Palmer *et al.* (2013a) recently reported a more thorough and rigorous investigation of music-to-color associations that included explicit tests for emotional mediation. They asked participants to listen to 18 brief selections of classical orchestral music — six each by Bach, Mozart, and Brahms — that varied in tempo (slow/medium/fast) and mode (major/minor). While listening, the participants chose the five colors that were most consistent with (and later the five colors that were least consistent with) each musical selection. The colors were selected from among 37 colors that were systematically sampled in visual appearance, differing in saturation, lightness, redness/greenness, and yellowness/blueness (see Fig. 1). The results showed that both US and Mexican participants chose lighter, yellower (warmer), and more saturated colors as going better with faster music in the major mode, but darker, bluer (cooler), and less saturated (grayer) colors as going better with slower music in the minor mode.

More relevant to the emotional mediation hypothesis, Palmer and colleagues also reported strong evidence that these music-to-color associations were related to emotion. They computed the correlations between explicit emotional ratings of the 18 musical selections and explicit emotional ratings of
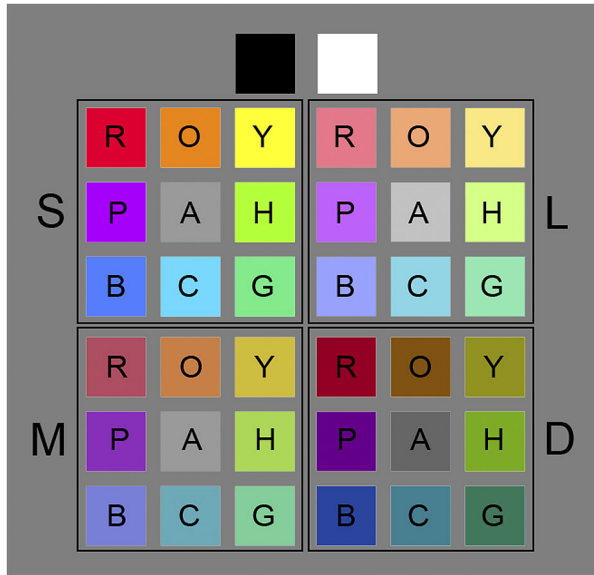
Figure 1. The 37 colors used in this experiment. The spatial array had black and white at the top, with the other colors arranged in quadrants that constituted the colors for each 'cut' (saturated, light, muted, and dark, going clockwise from top-left). Within these quadrants, the hues were ordered clockwise from the upper-left corner, starting with long-wavelengths: red (R), orange (O), yellow (Y), chartreuse (H), green (G), cyan (C), blue (B), and purple (P). (NB: The letters and the lines around the S, M, L, and D colors were not present in the display presented during the experiment).

the colors people chose as going best (and worst) with those musical selections for four dimensions: *happy/sad*, *angry/calm*, *strong/weak*, and *lively/dreary*. Remarkably, 94% of variance ($r = 0.97$) in the *happy/sad* ratings of the colors chosen as going best/worst with a given musical selection could be predicted from the *happy/sad* ratings of the musical selection. Similarly high correlations were evident for the other three emotional dimensions they studied ($r = 0.89$ for *angry/calm*, 0.96 for *strong/weak*, and 0.99 for *lively/dreary*).

These results support what Palmer *et al.* (2013a) called the 'emotional mediation hypothesis': the proposition that cross-modal mappings between music and color are mediated, at least in part, by accessing shared emotional associations of the two sets of stimuli. Although this evidence is correlational and therefore provides no support for inferring that these emotional commonalities actually *caused* people to choose the colors in the way they did, two additional experiments reported by Palmer *et al.* (2013a) provided further evidence supporting this conclusion. One experiment showed that similarly strong emotional correlations were present when people made cross-modal matches from the same selections of classical orchestral music to pictures of emotional faces

(e.g., happy-looking faces were chosen as going best with happy-sounding music). The other experiment showed that similarly strong emotional correlations were present when people made matches from pictures of the same emotional expressions on faces to patches of color (e.g., happy-looking colors were chosen as going best with happy-looking faces). Because the pictures of emotional facial expressions were of the same person's face and differed only in their perceived emotional expression, it seems almost certain that common emotional content was responsible for the matches people made in all three tasks.

There is an important ambiguity in the results reported by Palmer *et al.* (2013a), however, concerning the specific effects of tempo and major/minor mode in influencing music-to-color associations. The musical stimuli these researchers used were commercial recordings that varied widely in many musical and auditory features unrelated to tempo and mode that may have co-varied with them (e.g., differences in volume, pitch, timbre, rhythm, melodic structure, and harmonic structure). As a result, one cannot conclude that faster tempos and major mode actually produced the emotional associations and/or color associations that Palmer and colleagues reported. To disentangle these potentially confounding effects, we performed the experiment described below, in which we created musical stimuli that varied systematically and independently on four global musical factors of interest: tempo, mode, pitch-height, and note-density.

We chose to study four single-line melodies composed by Mozart because we could very precisely control their musical features (see Fig. S1 in the online Supplementary Materials). All four melodies were digitally rendered as sequences of synthesized piano notes with appropriately and precisely controlled durations, pitches, and volumes with no variation in these properties. These four basic melodies were then transformed to create 16 different versions of each melody for a total of 64 melodies: 4 basic melodies (1/2/3/4) × 2 modes (major/minor) × 2 pitch-heights (low/high) × 2 tempi (slow/fast) × 2 note-densities (sparse/dense), as specified in Table 1. The changes from major to minor mode were performed by standard key-change transpositions into the keys of D-major and D-minor. Each melody was played both with D_4 as the tonic (low pitch-height) and with D_5 as the tonic (high pitch-height). The fast tempo was approximately twice as fast as the slow tempo (see Note 2), and note-density was varied by composing variations on the melodic themes with approximately twice as many notes/beat. Our variations were based on Mozart's own variations of these themes, but ours were significantly simpler and more regular than Mozart's compositions. The notes that were added to the basic melodies generally produced twice the note-rate of the original by interpolating a single note between each pair of notes in the original melody.

Table 1.

The stimulus design for the 16 versions of each of the four Mozart melodies as presented to the participants (see text for details)

NOTE DENSITY	TEMPO	MODE Major — PITCH HEIGHT		MODE Minor — PITCH HEIGHT		NOTE RATE
		Low: D_4	High: D_5	Low: D_4	High: D_5	
Theme 1 NPB[1]	Slow 140 BPM[2]					Slow 140 NPM[3]
Theme 1 NPB[1]	Fast 240 BPM[2]					Medium 240 NPM[3]
Variation 2 NPB[1]	Slow 120 BPM[2]					Medium 240 NPM[3]
Variation 2 NPB[1]	Fast 220 BPM[2]					Fast 440 NPM[3]

[1]Notes/beat   [2]Beats/minute   [3]Notes/minute

2. Methods

2.1. Participants

Twenty-one undergraduate students (14 women, seven men) were tested at U.C. Berkeley through the Psychology Department's Research Participation Pool (RPP) for course credit. One subject's data was not complete for all three experimental tasks and therefore could not be included in the analyses, leaving a total of 20 participants. All reported normal or corrected-to-normal spatial vision and normal hearing. None were color deficient when screened with the Dvorine Pseudo-Isochromatic Plates. None reported having synesthesia when asked during the debriefing following the experimental tasks. No biographical information about musical or artistic training was obtained. All gave informed consent, and the Committee for the Protection of Human Subjects at the University of California, Berkeley, approved the experimental protocol.

2.2. Design

The design of the experiment was primarily defined by the orthogonal combinations of five stimulus factors that were used to generate the 64 melodies used as auditory stimuli: basic melodies (1/2/3/4), mode (major/minor), pitch-height (low/high), tempo (slow/fast), and note-density (sparse/dense, where the themes were note-sparse and the variations were note-dense; see below) (see Table 1). Each of these 64 melodies was used in two of the three different

tasks as defined below: the music-to-color association task and the emotional rating task for the 64 melodies. The third task was the corresponding emotional rating task for the 37 colors.

2.3. Stimuli

2.3.1. Auditory Stimuli

The 64 melodies were all based on four themes adapted from compositions by Mozart: (1) the theme from the Nine Variations in C Major on 'Lison Dormait' from Dezede's 'Julie' (K. 264) for solo piano, (2) the theme from the Duport variations in D Major (K. 573) for solo piano, (3) the theme from the six variations in G minor on 'Helas j'ai perdu mon amant' (K. 360) for piano and violin, and (4) a theme taken from the fourth movement of the Serenade in C minor (K. 388). Themes 1 and 2 were originally written in the major mode and were transposed into the minor mode; themes 3 and 4 were originally written in the minor mode and were transposed into the major mode. Both the major and minor versions of each theme were then embellished by the second author as variations in the manner of Mozart using Mozart's own variations on those same melodies as a guide. The present variations contained the same underlying theme as the basic melody, but with about twice the note-density of the basic melody. Whereas the notes in the theme for each melody were mostly quarter-notes, the variations were written almost entirely in eighth-notes that contained the theme, but embedded it within 'interpolated' eighth-notes nearby in the scale. The original variations on these themes composed by Mozart were used as a guide and incorporated where that was consistent with our goal of having the note-density of the variations be essentially double that of the basic themes. All four theme and variation pairs were transposed to the key of D. The resulting 16 melodies were rendered in a low octave (whose tonic was D_4, the D above middle-C) and a high octave (whose tonic was D_5). The themes were rendered at a slow and a fast tempo (140 BPM and 240 BPM respectively), and the variations were also rendered at a slow and a fast tempo (120 BPM and 220 BPM, respectively). The same tempi were used for all four themes and variations in both the major and minor modes. These tempo manipulations yielded three distinct modal note-rates for all of the melodies, since the modal note-rate for the fast themes (rendered at 240 BPM) was identical to the modal note-rate for the slow variations (rendered at 120 BPM), as shown in Table 1 (see Note 3). The ranges for each of the high register themes and variations were A4–D6 for both the theme and variation of Melody 1, A4–D6 for the theme and D4–D6 for the variation of Melody 2, D4–D6 for the theme and D4–E6 for the variation of Melody 3, and A4–D6 for the theme and A4–C6 for the variation of Melody 4.

The variations were primarily written by the second author as improvisations that were intended to resemble actual variations on those themes written

by Mozart, with feedback from a classically trained musician in the Music Department at Berkeley (see Note 4). The second author generated the final audio files in the Linux audio and MIDI sequencer *Rosegarden*. These were then transposed to the key of D using enhanced piano sounds to improve the sounds produced by *Rosegarden* (see Note 5). (For more details on the temporal statistics of the melodies, as well as the actual sheet music for each of the themes and variations, see Fig. S1 and Table S1 in the online Supplementary Materials.) There were no variations in the duration or loudness of notes across the different versions. The recordings used in the experiment can be found at http://dx.doi.org/10.6084/m9.figshare.1355942.

2.3.2. *Visual Stimuli*

The 37 colors used in the experiment are the same as those used by Palmer *et al.* (2013a) (see Fig. 1). They were chosen to sample a color appearance space roughly analogous to the Natural Color System (NCS) space by systematically varying the hue, saturation and lightness of the colors. The eight hues included the four unique hues: red (R), green (G), blue (B) and yellow (Y). To those we added four intermediate hues with approximately equal amounts of the adjacent unique hues: orange (O), chartreuse (H), cyan (C) and purple (P). These hues were sampled at four 'cuts' (saturation/lightness levels) through color space, with the saturated (S) colors being the maximally saturated colors that could be produced on the monitor used by Palmer *et al.* (2013a), muted (M) colors being approximately halfway between each S-color and neutral-gray, light (L) colors being approximately halfway between each S-color and white, and dark (D) colors being approximately halfway between each S-color and black in Munsell space (see Table S1). We also included white, black, and the three grays whose lightnesses were approximately the average lightnesses of the L-, M-, and D-colors (Fig. 1 contains 38 colored squares because the grays for the S and M cuts are the same). The colors were initially chosen from the Munsell Book of Colors, Glossy Series, and were translated into CIE *xyY* coordinates to generate them on our computer using the Munsell Renotation Table (see Table S2 in the online Supplementary Materials). The computer displays were presented on a 21.5″ iMac computer monitor with a resolution of 1680 × 1050 using the software Presentation (www.neurobs.com) that was calibrated using a Minolta CS100 Chroma Meter. The background was always neutral gray [International Commission on Illumination (CIE) $x = 0.312$, $y = 0.318$, $Y = 19.26$].

Participants viewed a computer screen from about 70 cm inside a darkened booth. In the first experimental task (the music-to-color association task), the 37 colors were displayed as squares 60 × 60 pixels wide, in the spatial arrangement shown in Fig. 1, located at the center of the screen (see Note 6). The melodies were played continuously over Sennheiser HD 270 earphones

until the participant had made his/her required choices of the three best-fitting and three worst-fitting colors. In the second experimental task of rating the emotional associations of the colors, single colors from the 37 BCP colors were displayed as 100×100 pixel squares, above a thin 500 pixel wide slider scale located below the single colored square and horizontally-centered on the screen. Each end of the slider scale contained the ends of one of four possible emotional dimensions (*happy/sad*, *agitated/calm*, *strong/weak*, *angry/not-angry*). For instance, the *happy/sad* emotional scale had the word *sad* on the left end and the word *happy* on the right end. A short, thin black vertical tick-mark could be moved along the horizontal response scale using the computer mouse. When clicked, the mouse registered at the current position of the vertical tick-mark along the scale (an x-coordinate between -200 and $+200$ pixels), that was scaled to the range of -100 to $+100$. In the third task of rating the emotional associations of the music, each version of each melody was played continuously until the participant responded. They made slider ratings of each melody along the same four emotional dimensions that they rated for the colors and in the same manner as they did for colors while listening to the relevant musical selection, making a single global judgment for each piece.

2.4. Procedure

There were three separate tasks in the experiment. In the music-to-color association task, participants listened to one of the 64 melodies and selected the three colors that they judged as going best with the melody and then the three colors that they judged as going worst with the melody. After hearing the instructions but before beginning the experimental trials, they were played a representative sample of the melodies so that they had an understanding of the range of the musical stimuli they would hear in the experiment. On each trial, they were instructed to select the three best colors, in order (best, second best and third best), and the three worst colors, in order (worst, second worst, and third worst), by clicking on the colors in the array at the center of the screen, while the music played. Each time they clicked on a colored square, that square disappeared from the screen, leaving the remaining squares as possible further choices. All the colors would return to the screen after the three 'best' colors had been picked, at which time participants were prompted to select the three 'worst' colors for the same musical selection.

In the emotional color rating task, participants rated each of the colors along four bipolar emotion dimensions (*happy/sad*, *agitated/calm*, *strong/weak*, *angry/not-angry*) by moving a slider along a response scale for each emotional dimension and clicking somewhere along the scale to register their response. The line-mark scale was located below the colored square, and participants rated the same color on four separate trials, one for each emotional dimension. These trials were blocked by emotion dimension (i.e., participants rated

all 37 colors on one dimension before going on to the next) but randomized over colors within blocks. The order of the blocks was also randomized across participants.

In the emotional music rating task, participants rated each of the 64 melodies along the same four emotional dimensions in the same way as for the colors while the music played continuously from the beginning of the selection until the participant clicked on their slider rating position to indicate their rating. There was an initial 2 s period during which the music played without participant being able to respond to insure that they heard a representative sample of the melody to be rated. Otherwise, the melodies were repeated indefinitely in a loop until a response was made. At the start of the music-to-color association task, and the emotional music rating task, a representative sample of the melodies was presented along with the instructions to insure that participants had a good feel for the range of the music they would be hearing.

3. Results and Discussion

The results of this experiment are numerous and complex. We have therefore chosen to structure our report and interpretation of them in a combined 'Results and Discussion' section, in which a set of closely related results and their statistical analyses are immediately followed by our interpretation. The intent is to avoid readers having to jump back and forth between a Results section and a Discussion section to connect the relevant data with their interpretation. This Results and Discussion section is divided into five parts with subheadings. First we present and discuss the results for the music-to-color association task. Next, we present and discuss the emotional correlations that assess the emotional mediation hypothesis. These correlations are followed by the rating data on which they are based: the emotional associations of the music and the emotional associations of the colors, including multidimensional scaling (MDS) results to discover whether similar emotional dimensions underlie both sets of data. Finally, we discuss more broadly the implications of the findings of the study in a General Discussion section.

3.1. Music-to-Color Associations

The music-to-color association data were analyzed using a method analogous to that reported by Palmer *et al.* (2013a), but modified to reflect the fact that only three colors were chosen as most consistent and only three colors as least consistent (instead of the five most and five least consistent colors — see Note 7). For each melody, m, and each color dimension, d (where d is saturation, lightness, red–green, or yellow–blue), the music–color association ($MCA_{d,m}$) index was computed from a weighted average of the values of the

three most consistent ($C_{d,m}$) minus a weighted average of the three most in-consistent ($I_{d,m}$) color choices for melody *m* as follows:

$$MCA_{d,m} = C_{d,m} - I_{d,m}, \tag{1}$$

$$C_{d,m} = (3c_{d,m,1} + 2c_{d,m,2} + 1c_{d,m,3})/6, \tag{2}$$

$$I_{d,m} = (3i_{d,m,1} + 2i_{d,m,2} + 1i_{d,m,3})/6, \tag{3}$$

where $c_{d,m,1}$ is the value of the first (best) color association for melody *m* on dimension *d*, $c_{d,m,2}$ is the value of the second-best color association for melody *m* on dimension *d*, and $c_{d,m,3}$ is the value of the third-best color association for melody *m* on dimension *d*, with $i_{d,m,1}$, $i_{d,m,2}$, and $i_{d,m,3}$ representing the corresponding values of the first-, second-, and third-worst color associations for melody *m* on dimension *d*. The dimensional values of the colors in these computations were the average values of explicit, bipolar, line-mark ratings made by 48 additional participants on each of these four color appearance dimensions: *saturated/desaturated*, *light/dark*, *yellow–blue*, and *red–green*. These ratings were made by the 48 participants described in Palmer *et al.* (2013a), and the values are the same as those used in that publication.

The results for each color appearance dimension are plotted in Fig. 2 in terms of average *MCA* scores as a function of note-rate, tempo, note-density, mode, and pitch-height. The variable on the *x*-axis is labeled as 'note-rate', because it is the combination of tempo (*slow/fast*, as defined by beats-per-minute) and note-density (the *sparse/dense* versions of the basic melodies, as defined by the number of notes-per-minute for notes of the modal note-value) (see Note 8). Note-rate has three levels (see Table 1), as defined on a loga-rithmic scale, of the rate at which modal notes (quarter-notes for themes and eighth-notes for variations) were sounded, which we will call *slow*, *medium*, and *fast*. This combined variable of note-rate was considered preferable to plotting the separate variables of tempo and note-density because the fast versions of the note-sparse basic melodies (at 240 notes/minute) gave sim-ilar results to the slow version of the note-dense variations (also at 240 notes/minute), and because the resulting three-level factor generally produced monotonic and often nearly linear variations in the *MCA* values of the cho-sen colors (see Fig. 2). Indeed, multiple linear regression analysis (see below) showed that after the variance due to note-rate was removed, tempo and note-density did not account for a significant portion of the remaining variance.

The data were analyzed in two ways: repeated measures ANOVAs and forward stepping, multiple linear regression analyses. One ANOVA was per-formed for each of the four color appearance dimensions (*saturated/desat-urated*, *light/dark*, *red/green*, and *yellow/blue*) (see Note 9), each of which included the four manipulated musical variables — tempo (*slow/fast*), note-density (*sparse/dense*), mode (*major/minor*), and pitch-height (*low/high*) —
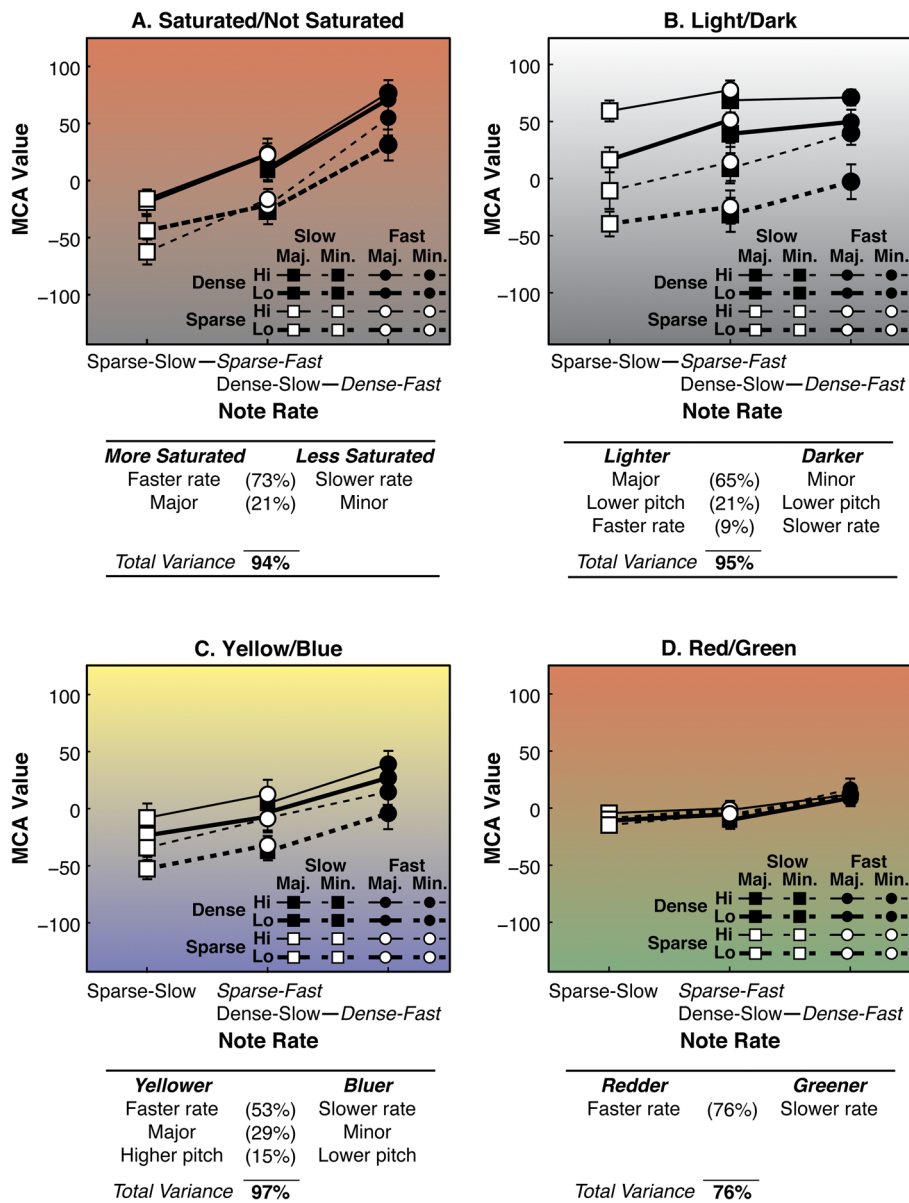
Figure 2. Results of the music-to-color association task. Average color appearance ratings (saturation, lightness, yellow–blue and red–green) for the colors chosen as going best with the melodies are plotted as a function of the note-rate, mode, and pitch-height of the melodies. Error bars represent standard errors of the means.

as well as the repeated-measures variable of participants. Note-rate was not included in the ANOVAs because we wanted to analyze the independent con-

tributions of tempo and note-density. These ANOVAs test the degree to which the average effects on each color dimension due to the manipulated variables are significantly different from chance when compared to the variability over participants (i.e., the error term for each main effect or interaction is the interaction of the corresponding factors with the participants factor). The multiple linear regressions are based on the average results (over participants) for the conditions in order to determine how much of the variance in the overall pattern of results can be explained by each musical variable in the experiment. For these regression analyses, we included the combined note-rate variable as well as the tempo and note-density variables to determine which of these variable(s) provided the best (most effective) account of the variance in the pattern of results. In every case, the combined note-rate variable superseded the tempo and note-density variables.

In all of the ANOVAs and regression analyses reported below, the *MCA* data for each participant and for each color dimension were first averaged over the four basic melodies, because we viewed the different melodies as a small sample of all possible 'Mozart-like' melodies within which we could examine the more global musical variables of interest (tempo, note density, mode, and pitch height). Because all four of these global musical variables were matched as closely as possible for the four different basic melodies, we expected little additional variance to be explained by the internal structure of the melodies (see Note 10).

For the saturation of the chosen colors (Fig. 2A), both tempo and note density had statistically reliable effects [$F(1, 19) = 52.27, 45.30$, $p < 0.0001$, 0.0001, $\eta_p^2 = 0.73, 0.70$, respectively], with faster tempos and higher note densities being paired with more saturated colors. There was also a smaller, but reliable interaction between these two variables [$F(1, 19) = 11.04$, $p = 0.004$, $\eta_p^2 = 0.37$], due to the slight increase in slope for the note-dense variations over the note-sparse themes as the tempo increased. Further, the saturation of the colors chosen as going best with melodies were affected by mode, with major melodies being paired with more saturated colors than minor melodies [$F(1, 19) = 16.36$, $p < 0.0001$, $\eta_p^2 = 0.46$]. When the mean ratings were subjected to a forward stepping multiple linear regression, note-rate alone accounted for 73% of the variance, and major/minor mode for an additional 21% of the variance, bringing the total of explained variance in the saturation of the chosen colors to 94% (multiple-$R = 0.97$, $p = 0.0001$). Tempo and note-density did not account for reliably more variance than note-rate alone for saturation or for any of the other three color appearance dimensions.

The results for the lightness of the colors chosen as going best with the melodies (Fig. 2B) shows a pattern that is interestingly different from that for saturation. Mode was the single most potent variable [$F(1, 19) = 51.11$, $p < 0.0001$, $\eta_p^2 = 0.73$], with major melodies being associated with lighter

colors than minor melodies. Quite unlike the saturation results, however, there was also a substantial effect of pitch-height [$F(1, 19) = 14.28$, $p < 0.001$, $\eta_p^2 = 0.43$], with melodies played in the higher octave being paired with lighter colors than the same melodies played in the lower octave. This result is consistent with previous reports that individual notes having higher pitches are associated with lighter colors (e.g., Hubbard, 1996; Marks, 1974; Wicker, 1968) and provides a valuable generalization of the previous findings over a broader musical context. Still, it is noteworthy that the larger-scale musical variable of major/minor mode produced a larger effect on the lightness of the associated colors than pitch-height did. Tempo also had a significant effect on lightness [$F(1, 19) = 20.84$, $p < 0.0001$, $\eta_p^2 = 0.52$], as it did for saturation, with faster tempos being paired with lighter colors, but, unlike saturation, note-density did not reliably increase the lightness of associated colors [$F(1, 19) = 6.25$, $p = 0.022$, $\eta_p^2 = 0.25$]. The forward stepping multiple linear regression showed that *major/minor* mode explained 65% of the variance, pitch-height explained an additional 21%, and note-rate explained a further 9%, for a total of 95% of the variance in lightness being explained by these three predictors (multiple-$R = 0.98$, $p = 0.0001$).

The pattern of effects on the yellowness/blueness of the colors people associated with the melodies (Fig. 2C) is similar to the pattern for saturation effects (Fig. 2A). The primary musical variables that affected the yellowness/blueness of the colors chosen as going best with the melodies were tempo and note-density, both of which had statistically reliable effects [$F(1, 19) = 37.41, 21.67$, $p < 0.0001, 0.0001$, $\eta_p^2 = 0.66, 0.53$, respectively], with faster tempos and higher note densities being paired with yellower colors. Further, the yellowness of the colors chosen as going best with melodies were affected by mode [$F(1, 19) = 14.00$, $p < 0.001$, $\eta_p^2 = 0.42$], with major melodies being paired with yellower colors and minor melodies with bluer ones. Finally, there was a small effect of pitch-height, with melodies in the lower octave being associated with bluer colors than those in the upper octave [$F(1, 19) = 8.59$, $p = 0.009$, $\eta_p^2 = 0.31$]. It is important to note that the yellower colors in our 37-color sample were systematically lighter than the bluer colors. This confound is unavoidable if one wants to include highly saturated colors (because darker yellows are necessarily much less saturated than yellows of the highest saturation), so any of these effects on yellowness/blueness may well include a lightness component. This seems especially likely for the pitch-height effect, which is strongly associated with lightness differences, both in the present data and in previously reported effects (Hubbard, 1996; Marks, 1974; Wicker, 1968). In the regression analysis, note-rate explained 53% of the variance, with *major/minor* mode explaining an additional 29%, and pitch-height for a final 15%, bringing the total of explained variance

in the yellowness/blueness of the chosen colors to 97% (multiple-$R = 0.99$, $p = 0.0001$).

Unlike the other three color dimensions, there is very little variance in the redness/greenness of the colors chosen as going best with the melodies (Fig. 2D). Only tempo and note-density had reliable effects on this color dimension [$F(1, 19) = 12.56, 7.21, \ p = 0.002, 0.015, \ \eta_p^2 = 0.40, 0.28$], with faster tempos and higher note-densities producing slightly redder color associates than slower tempos and lower note-densities. There was also a reliable interaction between these two variables [$F(1, 19) = 8.87, \ p = 0.008, \ \eta_p^2 = 0.32$] due to the fact that the slopes of the tempo functions for the denser variations were somewhat steeper than those for the sparser theme melodies. The compound predictor of note-rate that combines these two variables was the only variable entered into the regression equation, accounting for 76% of the variance in the redness/greenness of the color choices ($r = 0.87$, $p = 0.001$).

Overall, there is a surprising, but gratifying, degree of specificity and articulation in the relations among the various versions of these single-line piano melodies and the colors with which they are associated. All four of the manipulated variables (tempo, note-density, mode, and pitch-height) reliably influenced at least two of the four color dimensions, but only tempo influenced all four. Moreover, the tempo and note-density variables combined nicely into a note-rate effect that seemed to capture most of the systematic variance of the two explicitly manipulated variables. In some cases there were small tempo × note-density interactions due to the slope of the functions being higher for the note-dense variations than for the note-sparse themes. These results taken together provide interesting and complex patterns of associations from music to color. We now consider how they might arise and how we might be able to understand them better.

3.2. The Emotional Mediation Hypothesis

One promising possibility is that the results might be simpler and more understandable if they were recast in terms of music-to-color associations being determined by some mediating factor. Given the prior results of Palmer *et al.* (2013a), the most obvious mediating factor to consider is emotion. Emotional mediation implies that people performing a music-to-color association task have emotional responses to the music they hear and then choose the colors whose emotional associations most closely match the musically driven emotions. To assess this possibility, we employed the same analysis as that reported by Palmer and colleagues: namely, computing correlations between the explicit ratings of the emotional qualities of each piece of music and the corresponding explicit ratings of the emotional qualities of the colors they chose as going best (and worst) with it. The latter quantities were calculated

from the three most consistent and three most inconsistent colors chosen for each of the 64 melodies using a computation analogous to the *MCA* values in equations (1)–(3), but substituting the emotional dimensions (*happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*) for the color dimensions (saturation, lightness, yellow–blue, and red–green).

We then computed the correlations between the average emotional ratings for each of the 64 melodies and the emotional values of the three-best and three-worst color associations for that melody (see Note 11). The scatter plots obtained are shown in Fig. 3 for each of the four emotional dimensions, with mode represented in terms of color (major as red; minor as blue), note-rate in terms of lightness (slow as dark; medium as medium; fast as light), and pitch-height in terms of symbol shape (low as circles; high as triangles). The correlations between the emotional ratings of the melodies and those of the colors associated with them are all positive and highly significant: *happy/sad*
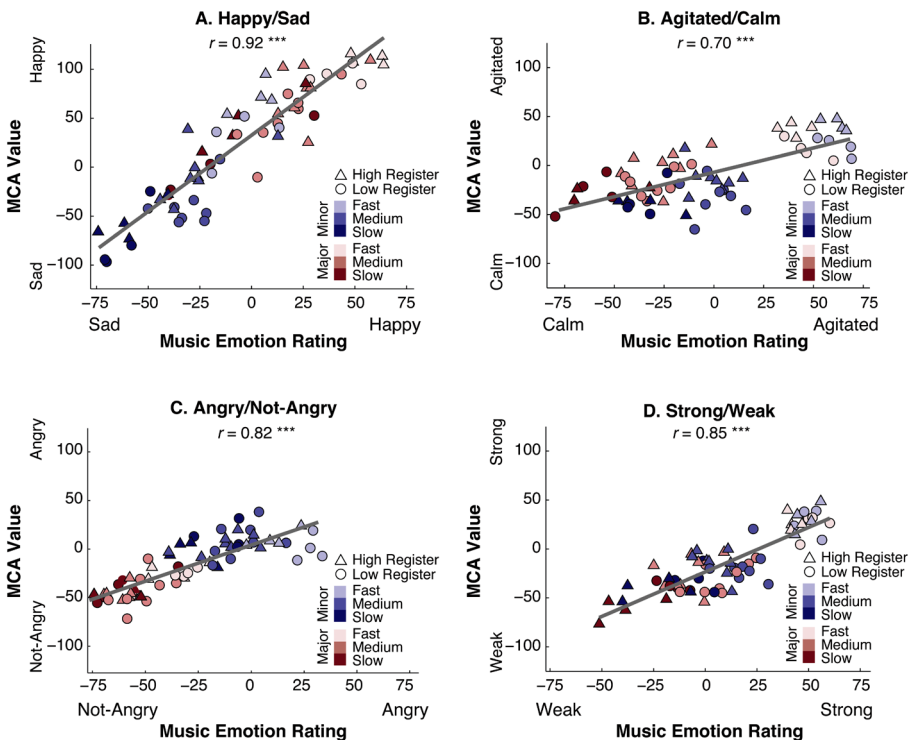

Figure 3. Scatter plots of the relation between the emotional ratings of the music and the emotional ratings of the colors chosen as going best with the melodies as a function of mode (red = major; blue = minor), note-rate (slow = dark; medium = medium; fast = light) and pitch-height (low = circles; high = triangles) for the four emotional dimensions rated (happy/sad, agitated/calm, angry/not-angry, and strong/weak).

($r = 0.92$, $t(62) = 8.78$, $p < 0.0001$, accounting for 85% of the variance), *strong/weak* ($r = 0.85$, $t(62) = 6.04$, $p < 0.0001$, accounting for 72% of the variance), *angry/not-angry* ($r = 0.82$, $t(62) = 5.36$, $p < 0.0001$, accounting for 67% of the variance), and *agitated/calm* ($r = 0.70$, $t(62) = 3.67$, $p = 0.0005$, accounting for 58% of the variance). Although these correlations are slightly lower than those reported by Palmer *et al.* (2013a), they robustly replicate the corresponding correlations previously reported for 18 selections of classical orchestral music (0.97 for *happy/sad*, 0.89 for *angry/calm*, and 0.96 for *strong/weak*). It is unclear why the correlations are lower for the present melodies data, but several factors are likely to be relevant. One obvious possibility is that the magnitude of the correlations may be influenced by the differences in the emotional nature of the music. Because the classical orchestral pieces are much more emotionally varied and expressive than the present single-line piano melodies, employing variations in loudness, orchestral timbre, and harmonic structure that were not available in the present piano melodies, the restricted emotional ranges in the present music may have caused the lower correlations. Another possible factor is sample size: the present data came from only 20 participants *versus* 48 participants in Palmer *et al.*'s (2013a) study. The smaller current sample presumably provided less stable and accurate estimates of the emotional ratings, thus potentially lowering the correlations based on them. In any case, the main point is that, consistent with the emotional mediation hypothesis, very high positive correlations were evident in both studies between the emotional ratings of the music and the emotional ratings of the color chosen as going best/worst with the music.

3.3. Music and Emotion

Emotional mediation implies that there must be musical features or dimensions that correspond to different dimensions of emotional responses to the music, and, likewise, that there must be color appearance features or dimensions that correspond to the same or similar emotional associations to the colors. We now evaluate these possibilities further using the data from the explicit emotional ratings obtained in the second and third tasks described in the Methods section.

Figure 4 shows the average emotional ratings (rather than the average color ratings, as plotted in Fig. 2) of the colors that were chosen for each of the 16 types of variations of the four basic piano melodies, plotted separately for each of the four measured emotional dimensions. We analyzed them in the same way as the color appearance ratings, except that the rating dimensions (see equations (1)–(3)) were identified with the emotional ratings of the colors (*happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*) rather than the color appearance ratings (saturation, lightness, red–green, and yellow–blue).
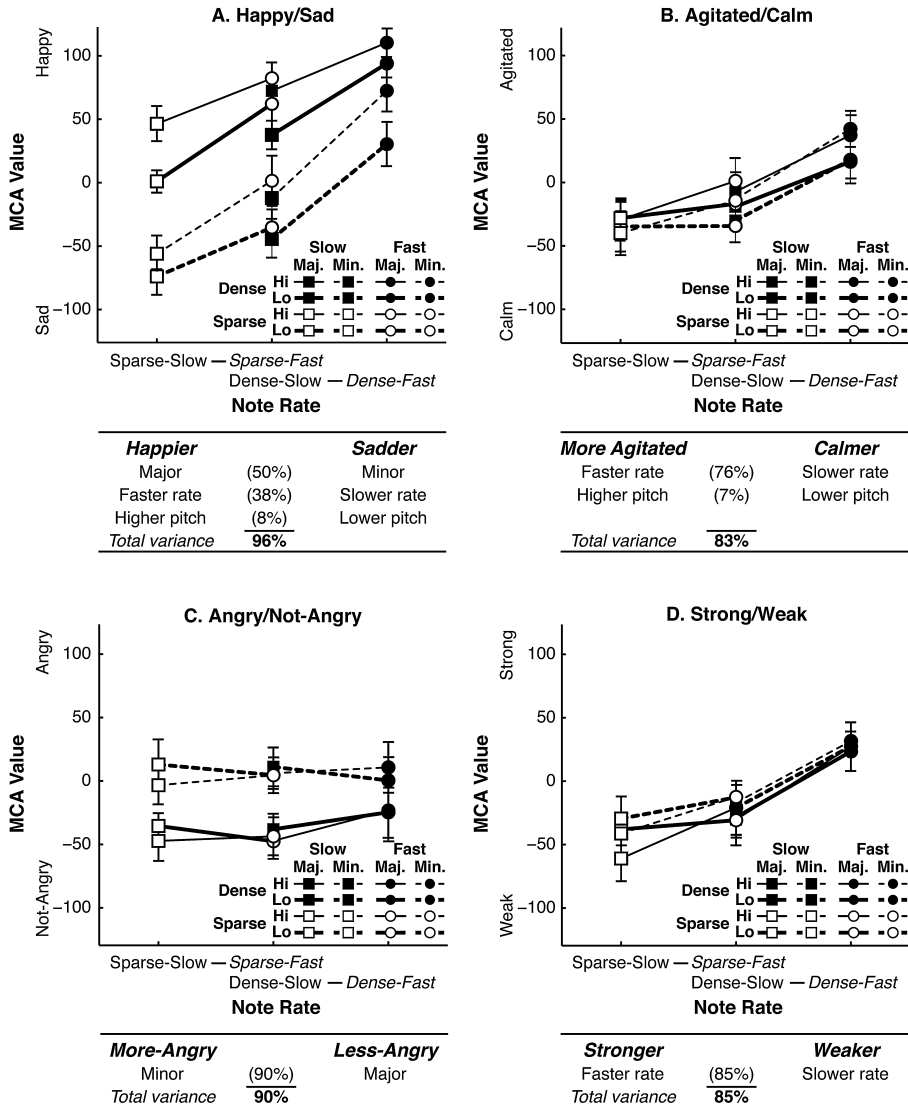
Figure 4. Average emotional ratings (*happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*) of the colors chosen as going best with the 16 types of melodies as a function of the note-rate, mode, and pitch-height of the melodies. Error bars represent standard errors of the means. The tables below each graph show the results of forward-stepping multiple linear regression analyses to predict the Music-to-Color Association (*MCA*) values of the colors chosen as going best/worst with the melodies from the musical attributes of the melodies.

A repeated measures ANOVA of the *happy/sad* ratings of the colors chosen as going best with the 16 types of melodies (Fig. 4A) were affected reliably by all four of the manipulated musical variables [$F(1, 19) = 43.13, 29.53, 66.33,$

17.56, $p < 0.0001, 0.0001, 0.0001, 0.0005$, $\eta_p^2 = 0.69, 0.61, 0.78, 0.48$, for mode, tempo, note-density and pitch-height, respectively], with no reliable interactions among them. A forward stepping multiple linear regression analysis showed *major/minor* mode to be the first variable entered, explaining 50% of the variance, with note-rate explaining an additional 38% and pitch-height explaining a further 8%, totaling 96% of the variance explained by these three variables (multiple-$R = 0.98$, $p < 0.0001$).

The *agitated/calm* ratings (Fig. 4B) were primarily driven by tempo, note-density and their interaction ($F(1, 19) = 8.24, 13.87, 13.03$, $p = 0.01, 0.001, 0.002$, $\eta_p^2 = 0.30, 0.42, 0.41$, respectively), with faster, note-dense versions of the melodies being perceived as more agitated than slower, note-sparse ones. The interaction reflects the fact that the tempo-induced increments for the note-dense variations were greater than those for the note-sparse basic melodies. No other main effects or interactions influenced the *agitated/calm* ratings significantly. The forward stepping multiple linear regression analysis showed that the combined variable of note-rate explained 76% of the variance in these ratings and pitch-height explained an additional 7%, with higher-pitched melodies being rated as more agitated than lower-pitched melodies. These two variables together accounted for 83% of the variance (multiple-$R = 0.91$, $p < 0.0001$).

The only musical variable that reliably influenced *angry/not-angry* ratings (Fig. 4C) in the ANOVA was major/minor mode [$F(1, 19) = 10.14$, $p = 0.005$, $\eta_p^2 = 0.35$], with minor melodies sounding angrier than major ones. Likewise, the only factor entered into the multiple linear regression equation was mode, which accounted for 90% of the variance in the *angry/not-angry* ratings of the colors chosen as going best with the melodies ($r = 0.95$, $p < 0.0001$).

A repeated measures ANOVA showed that average *strong/weak* ratings (Fig. 4D) of the colors chosen as going best with the melodies, were, like *agitated/calm* ratings, influenced by tempo, note-density and their interaction [$F(1, 19) = 18.63, 16.48, 12.73$, $p = 0.0004, 0.0007, 0.002$, $\eta_p^2 = 0.50, 0.46, 0.40$, respectively], with faster speeds being rated as stronger. The interaction is again due to higher tempo-induced slopes for the note-dense variations than for the note-sparse themes. Consistent with these ANOVA results, the multiple linear regression analysis included only the composite note-rate factor, which accounted for 85% of the variance ($r = 0.92$, $p < 0.001$).

We can also evaluate how the musical variables affect the emotional content of these melodies by examining how the direct ratings of emotion are influenced by the 16 different types of melodies. These data differ from those just described in that they do not involve the associated colors that were chosen as going best with the melodies, but simply the average direct ratings of

the melodies on the emotional dimensions. The results are plotted in Fig. 5 for each of the four emotional dimensions.

First, note the similarity of the results plotted in Figs 4 and 5. If emotional mediation is occurring, as we believe it is, then both sets of data should be measuring similar things, since Fig. 4 plots indirect ratings of the melodies' emotionality as reflected in the emotional ratings of the colors participants chose as going best (and worst) with the music, and Fig. 5 plots direct, explicit ratings of the emotional associations of the music. The overall correlation between the two data sets for the 64 melodies is 0.70 [$t(62) = 7.72$, $p < 0.0001$], with the effects of the musical variables being somewhat greater in the direct ratings than in the ones measured indirectly through colors.

The repeated-measures ANOVAs for the direct ratings of the four emotional dimensions (*happy/sad*, *agitated/calm*, *angry-not-angry*, and *weak/strong*) on which each melody was evaluated showed similar effects to the indirect ratings of the colors chosen as going with the melodies. For *happy/sad* ratings, all four variables produced main effects with no interactions [$F(1, 19) = 42.41, 23.53, 65.52, 13.08$, $p < 0.0001, 0.0001, 0.0001, 0.002$, $\eta_p^2 = 0.69$, $0.55, 0.78, 0.41$, for tempo, note-density, major/minor mode and pitch-height, respectively], with faster, denser, major and higher melodies producing happier ratings. The multiple linear regression analysis showed that *major/minor* mode was the most effective variable, accounting for 57% of the variance, with note-rate adding an additional 40% and pitch-height another 2%, bringing the total to 99% of the variance explained by these three variables (multiple-$R = 0.995$, $p < 0.0001$).

For *agitated/calm* ratings, there were main effects of tempo, note-density and major/minor mode [$F(1, 19) = 120.16, 110.23, 20.46$, $p < 0.0001$, $0.0001, 0.0002$, $\eta_p^2 = 0.86, 0.85$ and 0.52, respectively], with faster, denser, minor melodies producing ratings that indicate greater agitation. There was also a reliable interaction between tempo and note-density [$F(1, 19) = 35.72$, $p < 0.0001$, $\eta_p^2 = 0.65$] due to the greater slope for the note-dense variations with increased tempo. The multiple linear regression likewise showed a potent effect of note-rate, which accounted for 84% of the variance, plus an additional 11% for *major/minor* mode, for a total of 95% of the variance explained by these two factors.

For *angry/not-angry* ratings, there were main effects of *major/minor* mode, note-density and pitch-height [$F(1, 19) = 40.74, 12.87, 8.90$, $p = 0.0001, 0.002, 0.008$, $\eta_p^2 = 0.68, 0.40, 0.32$, respectively], with faster, denser, lower-pitched minor melodies producing ratings that indicate greater anger. Again, there was a two-way interaction between tempo and note-density [$F(1, 19) = 8.32$, $p = 0.009$, $\eta_p^2 = 0.30$] due to the higher slope for the note-dense variations. The multiple linear regression showed that 74% of the variance was explained by *major/minor* mode, an additional 19% by note-rate
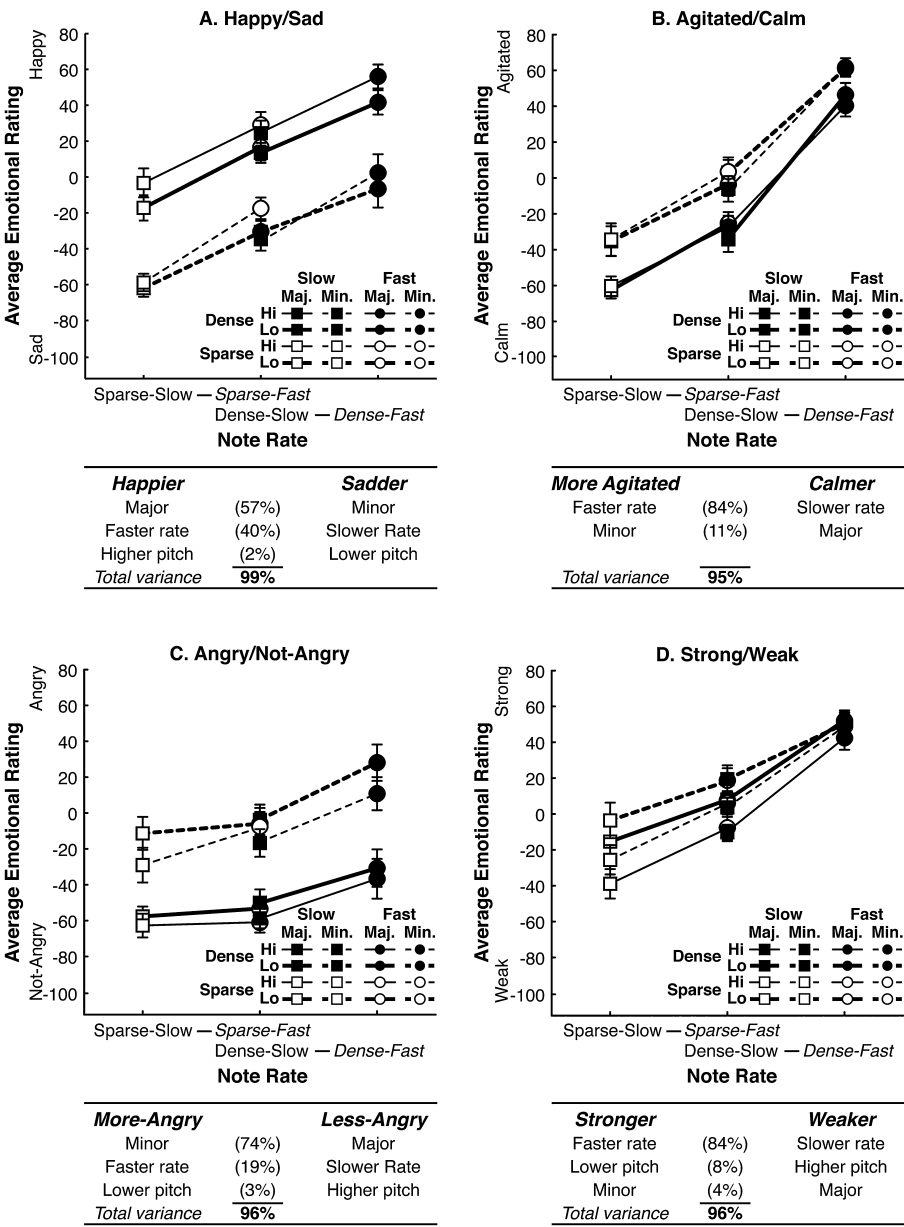
Figure 5. Average emotional ratings of the melodies as a function of note-rate, mode, and pitch-height, including the results of multiple linear regression analyses for each emotional dimension. Error bars represent standard errors of the mean. The tables below each graph show the results of forward-stepping multiple linear regression analyses to predict the emotional ratings from the musical features of the melodies.

and a final 3% by pitch-height, for a total of 96% of the variance explained by these three factors (multiple-$R = 0.98$, $p < 0.0001$), with faster, lower-pitched minor melodies receiving angrier ratings.

Finally, for *weak/strong* ratings, there were main effects of tempo, note-density, and pitch-height [$F(1, 19) = 45.11, 30.80, 10.73$, $p < 0.0001$, $0.0001, 0.004$, $\eta_p^2 = 0.70, 0.62, 0.36$, respectively], plus an interaction between tempo and note-density [$F(1, 19) = 13.54$, $p = 0.002$, $\eta_p^2 = 0.42$] due to the higher slope for the note-dense variations. The multiple linear regression showed that 84% of the variance was explained by note-rate, an additional 8% by pitch-height, and a final 4% by major/minor mode, for a total of 96% of the variance explained by these three factors (multiple-$R = 0.98$, $p < 0.0001$), with faster, lower-pitched, minor melodies being rated as stronger than slower, higher-pitched, major melodies.

The final analyses we performed of the direct emotional ratings of the melodies and colors were emotional multidimensional scalings (MDSs). The input to the mdscal function in MATLAB was the symmetric 16×16 matrix whose entries were the pairwise correlations between the four average emotional ratings for each of the 16 versions of the four basic melodies, with high correlations corresponding to small inter-point distances. The 2-D solution shown in Fig. 6A was very good (stress = 0.01). Dimension 1 along the *x*-axis clearly differentiates between major and minor modes, whereas Dimension 2 along the *y*-axis appears to reflect a combination of note-rate and pitch-height that might be termed 'energy', with faster note rates and higher pitches being more energetic. To interpret the emotional dimensions of this solution, we computed, for each emotional dimension, the correlation between
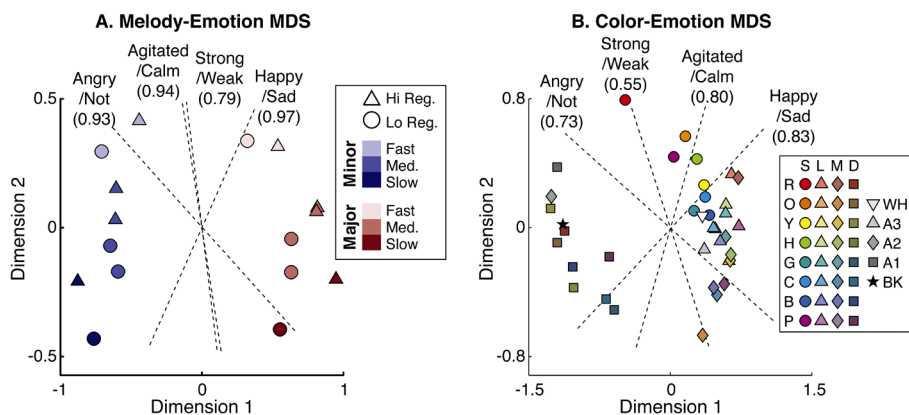

Figure 6. Emotional multidimensional spaces (MDSs) for the 16 melody types and the 37 colors, showing the best-fitting central axis orientations for the rated emotional dimensions (*happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*).

the average emotional ratings of the 16 melody types and the projections of the corresponding 16 points along an oriented line through the center at each of 360 orientations one-degree apart. The maximum of these correlation functions are shown in Fig. 6A as lines through the center for each emotional dimension. It is evident that the best-fitting, most nearly orthogonal axes are *happy/sad* and *angry/not-angry*.

This interpretation differs somewhat from the emotional MDS reported by Palmer *et al.* (2013a) for classical orchestral music, in which the emotional dimensions appeared to be *happy/sad* and *strong/weak*. One plausible explanation of the difference is that *strong/weak* may not be a particularly powerful or salient emotional dimension for these single-line piano melodies that were rendered with loudness-equated synthesized piano tones. Indeed, the present stimuli's lack of many of the most powerful musical variables that typically differentiate between *strong* and *weak* emotional responses to music, such as differences in loudness (pianissimo *vs.* fortissimo), orchestration (trumpets and tympani *vs.* flutes and harps), rubato (speeding up *vs.* slowing down), and harmonic structure (harmonious *vs.* dissonant chords). *Angry/not-angry* appears to be a more relevant feature for this simpler, smaller-scale music, even though few of the melodies were rated as particularly angry (see Fig. 5C). Another possible explanation is that the particular emotional dimensions used in the MDSs of the two studies were different. In particular, Palmer *et al.* (2013a) used *angry/calm* as a single dimension, whereas the present study broke this into two dimensions: *agitated/calm* and *angry/not-angry*. The present study also dropped the *lively/dreary* dimension used by Palmer and colleagues because it was so redundant with the *happy/sad* dimension. It is unclear exactly why these changes might have affected the orientations of the best-fitting emotional axes within the MDS solution as they did.

3.4. Color and Emotion

We also analyzed the 37 colors in terms of their ratings for each of the four emotional dimensions we studied: *happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*. The average ratings are plotted in Fig. 7 as a function of the eight hues along the *x*-axis and four 'cuts' (saturated, light, muted, and dark) as separate curves, plus the five achromatic colors (white, back, and the intermediate grays that were matched to the average lightness of the light, muted, and dark cuts) at a separate position along the *x*-axis. We analyzed the relation between the emotional and color dimensions using forward stepping multiple linear regressions, one for each emotion, because the colors were not orthogonal and therefore are not easily adapted to ANOVA, primarily because the achromatic series of five grays do not fit into a hue category. Instead, we used the ratings of each color on the four different dimensions of color (saturation, lightness, red–green, and yellow–blue) from the ratings previously reported by
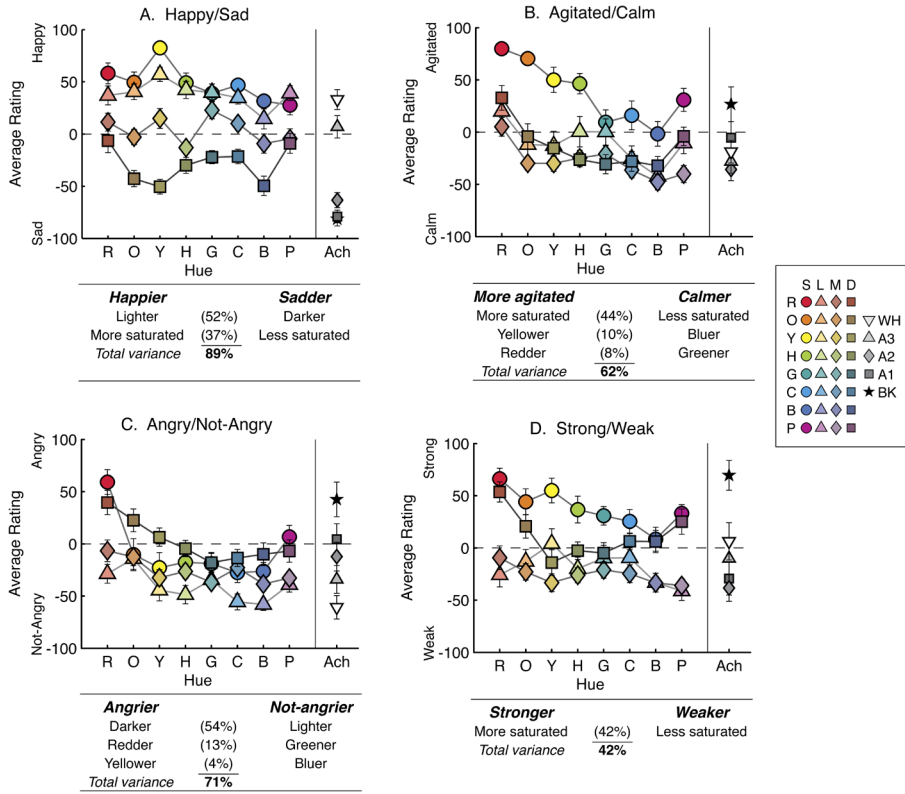
Figure 7. Ratings of the 37 colors for each of the four emotional dimensions: *happy/sad*, *agitated/calm*, *angry/not-angry*, and *strong/weak*. The tables below each graph show the results of forward-stepping multiple linear regression analyses to predict the emotional ratings from the color appearance ratings of the colors.

Palmer *et al.* (2013a) to predict the average emotional ratings from the present study.

For the *happy/sad* ratings (Fig. 7A), 52% of the variance was explained by the *light/dark* ratings, with lighter colors being rated as happier. An additional 27% of the variance can be explained by *saturated/desaturated* ratings, with more saturated colors being rated as happier, bringing the total variance explained to 89% (multiple-$R = 0.94$, $p < 0.0001$). Somewhat surprisingly, given that saturated yellow is usually rated as the single happiest color (see Fig. 7A), neither of the hue dimensions (red–green nor yellow–blue) reliably affected the *happy/sad* ratings. The reason is presumably that saturated yellow is the color with the highest combination of saturation and lightness, since all of the other most highly saturated colors are noticeably darker than saturated yellow.

Agitated/calm ratings (Fig. 7B) were best predicted by the saturation of colors, accounting for 44% of the variance, with more saturated colors being rated as more agitated. An additional 10% of the variance was due to yellow–blue ratings, with yellower colors being rated as more agitated. A final 8% of the variance was predicted by the red–green ratings, with redder colors being associated more strongly with agitation. Together these three variables accounted for 62% of the variance (multiple-$R = 0.79$, $p = 0.0003$).

For *angry/not-angry* (Fig. 7C), the most powerful predictor was, again, lightness, accounting for 54% of the variance, with darker colors being rated as angrier. Red–green explained another 13% of the variance, with redder colors being rated as angrier, and yellow–blue for a final 4%, with yellower colors being rated as angrier. These three variables together explain 71% of the variance in the emotional ratings of angriness (multiple-$R = 0.85$, $p < 0.0001$). *Strong/weak* ratings (Fig. 7D) were least predictable from the color ratings, with only 42% of the variance explained solely by color saturation ($r = 0.65$, $p = 0.006$).

We also performed an emotional MDS of the 37 colors (Fig. 6B) that was analogous to the emotional MDS of the melodies (Fig. 6A). We first computed a 37×37 matrix of the emotional similarities among the colors by correlating the emotional ratings of for each pair of colors, with high correlations corresponding to high similarities and short interpoint distances. We then used this matrix as the input to the ALSCAL algorithm, which found a very good fit to the 2-D solution shown in Fig. 6B (stress = 0.08). We also used the same procedure as described above for finding the best-fitting orientation of possible central axes for each of the rated dimensions of emotion. As was the case for the emotional MDS of the melodies (Fig. 6A), the pair of best-fitting, most orthogonal axes was *happy/sad* and *angry/not-angry*. The optimal *strong/weak* axis had a somewhat higher correlation with the data than the optimal *angry/not-angry* axis, but the latter was much more orthogonal to the optimal *happy/sad* axis than was the former. This space differs from the emotional MDS for the same 37 colors reported by Palmer *et al.* (2013a) primarily in the second dimension being identified as *angry/not-angry* (rather than *weak/strong*). It is unclear why this difference emerged, but, again, one possibility is that the particular emotional dimensions used in the MDSs of the two studies were different. As noted above with respect to corresponding issues with the emotional MDS of melodies, Palmer *et al.* (2013a) used *angry/calm* as a single dimension, whereas the present study broke this bipolar dimensions into two dimensions: *agitated/calm* and *angry/not-angry*. Another possibility is that the emotional ratings of the colors might have been influenced somewhat by the fact that the same participants made their emotional ratings of colors (and music) in the same experimental session, right after

performing the music-to-color association task. It is unclear why this might matter, however.

When one compares the 2-D emotional MDS solutions for the melodies and the colors, the main difference is merely that the angles between the best-fitting dimensional axes are compressed in the musical space relative to the color space, where they are more evenly spread out. The high degree of similarity of the corresponding dimensions in the two spaces implies that one can easily envision a process of making the best (and worst) music-to-color associations by mapping the musical features of the melodies into the emotional melody space and then 'reading out' the best fitting (and worst fitting) colors as the points closest to (and farthest from) from the corresponding referent positions in the color emotional space.

3.5. General Discussion

The present results confirm and extend certain findings previously reported by Palmer *et al.* (2013a) about the nature of cross-modal music-to-color associations in non-synesthetes, despite using a very different musical sample: synthesized single-line piano melodies by Mozart played at precisely controlled loudnesses, tempi, registers, and note-densities. First, despite substantial variation between individuals in the particular colors they choose as going best (and worst) with different samples of music, there are robust regularities in such mappings that generalize from the classical orchestral music used in the previous study to the simple piano melodies used in the present one. In particular, the present data replicate the findings that music at faster tempos and in the major mode tends to be associated with more saturated, lighter, yellower (warmer) colors, and music at slower tempos and in the minor mode tends to be associated with less saturated (grayer), darker, bluer (cooler) colors. These effects cannot be attributed to confounding variables such as loudness, instrumentation, rubato, and/or harmonic structure because none of those variables were present in the synthesized piano melodies.

The present data have also extended these results by showing that greater note-density (in the variations) has effects similar to those of faster tempos. Moreover, tempo and note-density, taken together, generally have roughly additive effects, although there are cases in which their effects are selectively modulated (e.g., the lightness of the color chosen are more strongly affected by tempo than note-rate). Another extension of the results described by Palmer *et al.* (2013a) is the selective influence of pitch-height on color lightness. As previously reported for individual tones (e.g., Hubbard, 1996; Marks, 1974; Wicker, 1968), higher pitches also tend to map to lighter colors in the present context of extended melodies. We found a further effect of pitch-height on the yellowness/blueness of the colors chosen as going with the melodies, but this

effect may well be a side-effect of the fact that the yellows in our color sample were considerably lighter than the blues.

The present results also replicated the evidence reported by Palmer *et al.* (2013a) supporting the emotional mediation hypothesis. In particular, the correlations between people's ratings of the emotional associations of the musical selections and their ratings of the emotional associations of the colors they chose as going best (and worst) with them were remarkably high: *happy/sad* ($r = 0.92$), *strong/weak* ($r = 0.85$), *angry/not-angry* ($r = 0.82$), and *agitated/calm* ($r = 0.70$). Although these correlations were slightly less extreme than those reported by Palmer *et al.* (2013a), the reduction is explicable in terms of the more restricted range of emotional expression available in these piano melodies than in the full classical orchestral recordings of music, which employ variations in loudness, orchestral timbre, harmonic structure, and compositional style (Bach *vs.* Mozart *vs.* Brahms) that were not present in the Mozartian piano melodies used here. Indeed, the reduced emotional range of the present piano melodies is so pronounced that it would be somewhat surprising if these simple melodies produced correlations as high as those reported by Palmer *et al.* (2013a) for full orchestral music.

In several ways, these robust emotional effects go considerably beyond previous reports of emotional mediation in cross-modal associations to odors (e.g., Schifferstein and Tanudjaja, 2004; Seo *et al.*, 2010). First, the present emotional effects are highly specific to particular 'real' emotions (e.g., happiness, sadness, anger, agitation, and calmness) rather than to the generalized affective dimensions that have been postulated to underlie all emotions: valence (*good/bad* or *pleasant/unpleasant*) and potency (*strong/weak* or *active/passive*). Presumably, almost everything can be mapped onto valence and potency (Mehrabian and Russell, 1974; Osgood *et al.*, 1957), but not everything maps onto specific emotions. In the domain of odors, for example, the most important dimension of people's odor-to-color associations is valence or pleasantness (Schifferstein and Tanudjaja, 2004), presumably because everyone has relatively definite feelings of the degree to which they *like/dislike* different smells and different colors. But the importance of valence in odor-to-color associations has been further interpreted as indicating emotional mediation of odor-to-color associations, even though it is uncertain whether this implies mediation by more clearly emotional dimensions, such as *happiness*, *sadness*, *anger*, *agitation* and *calmness*. Rather, it seems likely that *pleasantness* ratings are merely tapping into some dimension of preference (i.e., how much people *like/dislike* the odors). In the present case of music-to-color associations, we have unpublished evidence that aesthetic preference is relatively unimportant in choosing the colors that go best (and worst) with music. Using a more extended set of musical stimuli (34 different genres, from heavy metal to jazz, to country western, to Hindustani sitar music) participants were asked

to rate how much they *liked/disliked* each musical selection and how much they *liked/disliked* each color (as well as rating nine other emotion-related dimensions). But in this case, there was only a modest correlation between people's preferences for the music and their preferences for the colors they chose as going best (and worst) with the music. In other words, people did not choose colors they liked/disliked as going well with music they liked/disliked (Whiteford *et al.*, 2013). In any case, the current results provide strong additional support for the emotional mediation hypothesis with specific emotional effects beyond those of valence.

A further question of interest is how we are to interpret the present results in terms of the three-way classification of cross-modal associations as being due to structural, statistical, or semantic factors. As mentioned in the Introduction, these categories of association types should not be viewed as mutually exclusive, because any one being true does not rule out the possibility that any other might also be true. The present evidence supporting emotional mediation effects in music-to-color associations therefore does not imply the lack of structural, statistical, or semantic components to these cross-modal associations. Indeed, it seems quite possible that at least some of the other types of associations will also turn out to be present. For example, if music-to-color associations are largely based on emotional mediation, then it is likely to be true (a) that there will be correlations between the neural representations of the music and the colors (e.g., the neural coding of the emotional associations to colors will be similar to the neural coding of the emotional associations to music), (b) that there will be statistical correlations between the occurrences of colors and music in the environment (e.g., in movies, there may well be systematic differences in the emotional-tone of the colors that are seen, depending on the emotional-tone of the music that is being played), and (c) that common, semantically related terms will be used in both domains (e.g., people talk about music sounding 'sad' and 'happy' as well as colors looking 'sad' and 'happy'). The fact that the present results support emotional mediation therefore has no implications about these additional associative possibilities. Indeed, it is quite possible that these other forms of association actually arise from and depend on the emotional mediation effects we find in the cross-modal mappings we have reported here and in Palmer *et al.* (2013a).

What about the further issue of whether emotional mediation should be viewed as a fourth distinct type of association or as a specific subtype of semantic relation? There is clearly an aspect of emotion that fits within the general framework of semantics. Happy and sad are concepts people understand much as they understand *large* and *small*, *red* and *green*, and a myriad of other bipolar semantic dimensions that characterize different objects or even how the same object changes over time. In this sense, at least some emotional aspects of music-to-color associations would fall within a semantic

framework: namely those that are mediated by emotional concepts. Nevertheless, some other aspects of emotional mediation may be different from semantics in an important way, particularly if they involve the congruence of people's *emotional experiences* to different modalities of input. If people actually *feel* happy (or happy-ish) when they hear fast major music and when they see happy-looking colors, but *feel* sad (or sad-ish) when they hear sad-sounding music and when they see sad-looking colors, then the emotional consistency of their feeling states could be what drives them to choose the colors they do. Research shows that music affects certain physiological indicators of emotion (e.g., heart rate, blood pressure, skin conductance, and respiration rate) and that these measures correlate with participants' real-time ratings of their emotional reactions to the music they are hearing (e.g., Juslin and Sloboda, 2001; Krumhansl, 1997). These findings suggest that people actually do have emotion-like experiences while listening to music, and if they have similar kinds of emotion-like responses to colors while performing the music-to-color association task, the *felt* emotional component of the matching task would differ importantly from purely semantic mediation, which does not require people to *feel* the mediating conceptual quality (e.g., people may associate loud sounds with large objects because loud sounds can be described as 'large', but this does not necessarily imply that they *feel* or *see* visual largeness when they hear the sound).

The high correlations between the emotional ratings of the music and the emotional ratings of the colors that people chose as going best (or worst) with the music indicate a very strong, common emotional basis for the music-to-color associations made by non-synesthetes. But where do these associations to emotion come from? Why does fast music played in a higher register tend to sound agitated, and why does slow music in a lower register tend to sound sad (see Fig. 5)? The present results do not actually constrain possible answers, but many of the effects reported here seem consistent with how emotional responses influence various aspects of human behavior, as Huron (2012) has argued. As in so many domains, however, early expressions of this idea can be found in the writings of Aristotle about the mimetic aspect of music. Aristotle claimed, for example, that 'rhythm and melody supply imitations of anger and gentleness...' (*Politics*, book VIII, chapter 5, 1340a18–20), where Aristotle classifies anger and gentleness as emotions (*pathê*) (see Note 12).

This conjecture that emotion in music arises from 'imitations' of human emotions amounts to what might be termed 'musical anthropomorphism'. Perhaps the behavior of music is perceived as expressing a given emotion to the degree that it mirrors analogous behavior in humans who are expressing that emotion. For example, faster music at higher registers might be perceived as agitated because when people become agitated, their motor response rate tends to increase — pacing the floor, fidgeting with their hands, eyes darting vigi-

lantly around the environment, etc. — and their voice pitch tends to rise. If musical emotions are behaviorally anthropomorphic, it seems reasonable that faster note-rates and higher pitches would be plausible musical cues to agitation in music. Likewise, when people become calmer, their motor output tends to diminish in both speed and amplitude and their voice pitch tends to lower, making slower note-rates and lower pitches seem reasonable musical cues to calmness in music.

An analogous case can be made about the musical variables that drive *happy/sad* ratings of the melodies. When people are happy, they tend to be more animated, with somewhat increased vocal pitch, whereas when they are sad, they tend to be more lethargic, with corresponding tendencies toward lower vocal pitch. However, *major/minor* mode is an especially important musical cue to *happiness/sadness*, with major music sounding happier and minor music sounding sadder, despite the fact that it is unclear why this might be the case. The powerful effect of musical mode on how *happy/sad* music sounds has been noted for a long time (e.g., Heinlein, 1928; Hevner, 1935), but there seems to be no consensus about the cause other than culturally learned associations (Huron, 2008). Anger is also associated with faster note rates, but as a negative emotion, it tends to be associated with the minor mode for reasons that are not yet clear.

A better-defined and more easily researchable version of the general idea of what we are calling musical anthropomorphism is that music and emotion might be related through movement and dance. Sievers *et al.* (2013) have shown that music and movement have a common underlying spatio-temporal structure, both of which are related to emotion. They showed people in the USA and in an isolated region of Cambodia either auditory or visual displays in which the participants learned to manipulate variables such as the tempo (BPM), temporal jitter, and interval step-sizes of either the auditorily presented tone sequences or the visually presented motions using slider controls. Two different groups in each culture were instructed to adjust the parameters to make the perceived auditory or visual event best express five emotions: *angry*, *happy*, *peaceful*, *sad*, and *scared*. (Notice that all except *scared* are emotions studied here and by Palmer *et al.* (2013a), provided *peaceful* is roughly equivalent to *calm*.) They found that in both cultures, participants set the sliders to very similar values for the musical and movement displays for corresponding emotions, thus suggesting cross-modal correspondences between music and movements that are defined by emotions. Based on their results, the authors conclude that "music and movement can be understood in terms of a single dynamic model that shares features common to both modalities . . . made possible not only by the existence of prototypical emotion-specific dynamic contours, but also by isomorphic structural relationships between music and movement".

How might the relations between colors and emotions be understood? A similar situation exists here in the sense that some correspondences appear to be understandable in biological terms and others less so. Why, for example, might saturated red be judged as the angriest and most agitated color? Surely, biological considerations are relevant, especially the fact that human blood is vivid red in color. When someone becomes angry or agitated, his or her face tends to become redder and darker as it flushes with blood. This correlation of redness/darkness-of-face with anger and agitation seems a likely bridge between emotions in human interactions and colors. Another connection between anger and red is the fact that anger tends to lead to violence, violence to injury, and injury to fresh, red blood. Such connections between anger, agitation, and red through the color of blood and its various associations make red the prototypical color associated with these emotions. Most of the rest of the color ratings for anger appear to be largely based on the similarities of the other colors to this prototypical 'angry' red.

It is not so obvious why saturated yellow should be judged to be the happiest color, however, since people's faces do not become yellower when they are happier. One possibility is that yellow is considered a happy color because (a) the sun appears bright and yellowish, once the short wavelength blues have been filtered out of its white light (due to Raleigh diffraction), and (b) many people associate bright sun with happiness because they like good weather and the enjoyable things people do most pleasurably while the sun shines brightly, such as hiking, swimming, biking, sunbathing, and having picnics. Another possibility, however, is that the happiness of colors may be more directly tied to the color dimensions that support it: lightness (accounting for 52% of the variance) and saturation (and additional 37%) together accounting for 89% of the variance in *happy/sad* ratings of colors (see Fig. 7A). Saturated yellow then is the natural candidate for the happiest color simply because it is by far the lightest of the maximally saturated colors and the most saturated of the light colors. But why, then, are light and saturated colors judged to be happy? Perhaps lightness is regarded as related to happiness partly because humans are diurnal creatures who feel safer and happier during sunny, well-lighted days and more fearful and sad during dark and stormy nights. And perhaps saturation is regarded as related to happiness because people generally prefer saturated colors to desaturated (grayish) colors, based on their tendency to like objects whose colors are highly saturated more than objects whose colors are less saturated (Palmer and Schloss, 2010). It is difficult to pin down the reasons for such associations, however, given the many possibilities and the dearth of actual evidence.

Finally, the present results have potentially interesting implications for controversies about the nature of music-to-color synesthesia: a neurological condition in which people spontaneously experience visual colors while

they are listening to auditory music without any concurrent visual stimulation (e.g., Cytowic and Eagleman, 2009). There are debates in the synesthesia literature about several issues related to the present findings: e.g., whether synesthetic experiences in synesthetes are continuous with corresponding cross-modal associations in non-synesthetes (Martino and Marks, 2001; Ward, 2006), whether synesthetic experiences are due to semantic mediation (Hubbard, 1996), whether they are due to direct cross-modal connections in the brain (e.g., Ramachandran and Hubbard, 2001), or whether they are mediated through interactions with other parts of the brain (e.g., Grossenbacher and Lovelace, 2001). We have demonstrated (Palmer *et al.*, 2013a) and replicated (here) emotional mediation in the music-to-color association task among non-synesthetes. We can now test synesthetes with the same musical stimuli and a similar task — i.e., choosing the three colors that are most similar (and least similar) to the colors they actually experience while listening to the music — to examine the possibility of emotional effects in synesthetic experiences. If synesthetes also show corresponding emotional effects in the colors they report experiencing, the possibility of continuity with non-synesthetic associations and the possibility that synesthetic color experiences include indirect, mediating connections between from auditory and visual cortex are more plausible. Preliminary evidence suggests that synesthetes do indeed show emotional effects, albeit somewhat less strongly than non-synesthetes (Palmer *et al.*, 2013b; Whiteford *et al.*, 2013).

The present results have settled several important questions about the original results reported in Palmer *et al.* (2013a) by precisely controlling and isolating musical features of tempo, note-density, major/minor mode, and pitch height. Nevertheless, there are many fascinating questions that still need to be addressed. Both of the studies thus far reported concern people's cross-modal associations to classical music: classical orchestral music in Palmer *et al.* (2013a) and classical single-line piano melodies in the present study. Will similar evidence of emotional mediation be found if the set of musical samples is expanded to include various popular and international musical genres (e.g., heavy metal, jazz, country-and-western, Hindustani sitar, salsa, and Balkan folk music)? Will lower level musical stimuli, such as chords, two-note intervals, and even single-note instrumental timbres, also show emotional effects? How important are aesthetic factors in people's cross-modal music-to-color associations? To what degree are people's cross-modal music-to-color associations the same as those of people from different cultures with different musical traditions? And how do the colors non-synesthetes associate with different pieces of music compare with the colors that synesthetes experience when hearing the same music? The present paradigm provides an important research tool for answering such questions and for tracing out the possible emotional connections among them.

Notes

1. We intend the term 'emotional associations' to music and/or colors to be a relatively neutral and inclusive way to describe how people register emotionality in such stimuli, regardless of whether it is conscious or unconscious. We do not mean to take a position about whether they are merely recognizing emotion in music *versus* actually having emotional responses when listening to music or viewing colors. This is very much an open topic for future research.

2. Due to a programming error, the fast tempo (220 beats/minute) was not quite twice as fast as the slow tempo (120 beats/minute), but the difference is relatively slight.

3. We define the temporal parameters as follows. Tempi are specified as the number of beats per minute (BPM). The modal note-value for each version is defined as the most frequent number of notes per beat, which was a quarter note for all four themes and an eighth note for all four variations (see Table S1 in the Supplementary Materials for note frequencies). Modal note-density is defined as the number of modal notes per beat (1 for all four themes and 2 for all four variations) and note-rate is defined as the number of modal notes per minute.

4. We thank Tawny Tsang, who was an undergraduate Music major working in our lab at the time, for her assistance to Thomas Langlois in composing some of the themes and variations.

5. We thank Karle-Philip Zamor from the laboratory of Daniel J. Levitin at McGill University for performing this post-processing of the melodies to final form.

6. The colors were arranged systematically by hue and cut because previous pilot work showed that when the colors were randomly ordered, participants reported difficulty in finding the colors they wanted. The arrangement was the same for every participant on every trial in the present study because, in previous research, we also used an array that was rotated by 180°, such that the up/down and left/right ordering of the colors were both reversed and this change made no reliable difference in the results (see Palmer *et al.*, 2013a).

7. Re-analyses of the original results by Palmer *et al.* (2013a) showed that the results did not differ substantially if only the first three color choices were used rather than the first five color choices.

8. We define the modal note-value for a given melody as the note-value of the highest frequency category of note (i.e., quarter-note, eighth-note, etc.). The modal note value for all basic themes was a quarter-note and that for the all variations was an eighth-note (see Table S1 in the Supplementary Materials for the supporting data).

9. We adjusted the critical alpha according to the Bonferroni correction, because we measured four different color dimensions for each color (adjusted alpha = 0.0125).

10. A further set of linear regressions were conducted on the complete set of 64 melodies that included the different basic melodies as four nominal variables (*melody-1/not-melody-1*, etc.). The results showed that all of the global musical variables were entered into the regression equation before any of the melody variables were added. Moreover, the increases in the percentage of variance explained by the melody variables were relatively small: an additional 1% (of 88%) for saturation, an additional 10% (of 88%) for lightness, an additional 6% (of 80%) for yellow–blue and no additional variance (of 45%) for red–green.

11. Here we use all 64 melodies (rather than averaging over the four basic melodies to give 16 melody types) because the color association effects are not restricted by the four musical variables we manipulated (tempo, note-density, mode, and pitch-height), but may include other factors (e.g., chromatic intervals and melodic structure) that differ across the basic melodies.

12. We thank Klaus Corcelius and Timothy Clarke of the Philosophy Department at U.C. Berkeley and an anonymous reviewer for their help in making the connection to Aristotle. We also acknowledge that there is a long and complex history of ideas about music and emotion in philosophy, including the modern of writings of Langer (1942) and Kivy (1980), but the issues they raise are tangential to the present discussion.

References

Arnheim, R. (1986). *New Essays on the Psychology of Art*. University of California Press, Berkeley, CA, USA.

Barbiere, J. M., Vidal, A. and Zellner, D. A. (2007). The color of music: correspondence through emotion, *Empir. Stud. Arts* **25**, 193–208.

Bernstein, I. H. and Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time, *J. Exp. Psychol.* **87**, 241–247.

Bond, B. and Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds, *Percept. Psychophys.* **6**, 337–339.

Bresin, R. (2005). What is the color of that music performance?, in: *Proceedings of the International Computer Music Conference — ICMC 2005*, San Francisco, CA, USA, pp. 367–370.

Calvert, G., Spence, C. and Stein, B. E. (Eds) (2004). *The Handbook of Multisensory Processes*. MIT Press, Cambridge, MA, USA.

Collier, G. L. (1996). Affective synesthesia: extracting emotion space from simple perceptual stimuli, *Motiv. Emot.* **20**, 1–32.

Cytowic, R. E. and Eagleman, D. (2009). *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*. MIT Press, Cambridge, MA, USA.

Driver, J. and Spence, C. (1998). Cross-modal links in spatial attention, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**(1373), 1319–1331.

Evans, K. K. and Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features, *J. Vis.* **10**, 6, 1–12.

Grossenbacher, P. G. and Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints, *Trends Cogn. Sci.* **5**, 36–41.

Heinlein, C. P. (1928). The affective characteristics of the major and minor modes in music, *J. Comp. Psychol.* **8**, 101–142.

Hevner, K. (1935). The affective character of the major and minor modes in music, *Am. J. Psychol.* **47**, 103–118.

Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch and melodic interval, *Am. J. Psychol.* **109**, 219–238.

Huron, D. (2008). A comparison of average pitch height and interval size in major- and minor-key themes: evidence consistent with affect-related pitch prosody, *Empir. Musicol. Rev.* **3**, 59–63.

Huron, D. (2012). Understanding music-related emotion: lessons from ethology, in: *Proceedings of the 12th International Conference on Music Perception and Cognition*, Thessaloniki, Greece, pp. 473–482.

Juslin, P. N. and Sloboda, J. A. (2001). *Music and Emotion: Theory and Research*. Oxford University Press, Oxford, UK.

Karwoski, T. F., Odbert, H. S. and Osgood, C. E. (1942). Studies in synesthetic thinking: II. The rôle of form in visual responses to music, *J. Gen. Psychol.* **26**, 199–222.

Kivy, P. (1980). *The Corded Shell: Reflections on Musical Expression*. Princeton University Press, Princeton, NJ, USA.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology, *Can. J. Exp. Psychol.* **51**, 336.

Langer, S. K. (1942). *Philosophy in a New Key: a Study in the Symbolism of Reason, Rite, and Art*. Harvard University Press, Cambridge, MA, USA, p. 214.

Lewkowicz, D. J. and Turkewitz, G. (1980). Cross-modal equivalence in early infancy: auditory–visual intensity matching, *Dev. Psychol.* **16**, 597–607.

Lidji, P., Kolinsky, R., Lochy, A. and Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head? *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 1189–1207.

Marks, L. E. (1974). On associations of light and sound: the mediation of brightness, pitch and loudness, *Am. J. Psychol.* **87**, 173–188.

Marks, L. E., Hammeal, R. J. and Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor, *Monogr. Soc. Res. Child Dev.* **52**, 1–102.

Martino, G. and Marks, L. E. (2001). Synesthesia: strong and weak, *Curr. Dir. Psychol. Sci.* **10**, 61–65.

McMahon, T. A. and Bonner, J. T. (1983). *On Size and Life*. Scientific American Books/W. H. Freeman & Co., New York, NY, USA.

Mehrabian, A. and Russell, J. A. (1974). *An Approach to Environmental Psychology*. MIT Press, Cambridge, MA, USA.

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press, Urbana, IL, USA.

Palmer, S. E. and Schloss, K. B. (2010). An ecological valence theory of human color preference, *Proc. Natl Acad. Sci.* **107**, 8877–8882.

Palmer, S. E., Schloss, K. B., Xu, Z. and Prado-Leon, L. (2013a). Music–color associations are mediated by emotion, *Proc. Natl Acad. Sci.* **110**, 8836–8841.

Palmer, S. E., Schloss, K. B. and Whiteford, K. L. (2013b). Music–color associations to classical music in synesthetes and non-synesthetes: the surprising role of emotion, in: *13th Annual Meeting of the Vision Science Society*, Naples, FL, USA.

Ramachandran, V. S. and Hubbard, E. M. (2001). Synaesthesia — a window into perception, thought and language, *J. Cons. Stud.* **8**(12), 3–34.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C. and Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect, *Cognition* **99**, 113–129.

Schifferstein, H. N. J. and Tanudjaja, I. (2004). Visualising fragrances through colours: the mediating role of emotions, *Perception* **33**, 1249–1266.

Sebba, R. (1991). Structural correspondence between music and color, *Color Res. Appl.* **16**, 81–88.

Seo, H. S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G. and Hummel, T. (2010). Cross-modal integration between odors and abstract symbols, *Neurosci. Lett.* **478**, 175–178.

Sievers, B., Polansky, L., Casey, M. and Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion, *Proc. Natl Acad. Sci.* **110**, 70–75.

Spence, C. (2011). Crossmodal correspondences: a tutorial review, *Atten. Percept. Psychophys.* **73**, 971–995.

Spence, C. and Driver, J. (2004). *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford, UK, USA.

Stevens, J. C. and Marks, L. E. (1965). Cross-modality matching of brightness and loudness, *Proc. Natl Acad. Sci.* **54**, 407–411.

Ward, J., Huckstep, B. and Tsakanikos, E. (2006). Sound–colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all? *Cortex* **42**, 264–280.

Whiteford, K. L., Schloss, K. B. and Palmer, S. E. (2013). Music–color associations from Bach to the blues: emotional mediation in synesthetes and non-synesthetes, in: *13th Annual Meeting of the Vision Science Society*, Naples, FL, Italy.

Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space, *Am. J. Psychol.* **81**, 178–188.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences, *Am. Psychol.* **35**, 151–175.