a

Journal of High School Science

Exploring perceptual associations between color and music

Bai E[1], Belfi A.M

Abstract

Previous work has established associations between senses, such as associations between pitch and spatial height or timbre (the specific sound of a certain instrument) and shape. Specifically, Palmer, Langlois, and Schloss (1) have found associations between the colors yellow and blue and major and minor chords, respectively, that are mediated by emotions. We conducted two experiments to test whether these associations exist on a perceptual level (i.e. associations that can be made without participants consciously thinking about the stimuli). In two experiments, participants were presented with audio-visual pairs of congruent (blue-minor or yellow -major) and incongruent (blue-major or yellow -minor) stimuli. The stimuli were presented one after another and participants were asked to decide which stimulus occurred first. We hypothesized that participants would have lower accuracy for congruent pairs, as perceptual binding would make it more difficult to distinguish the time between each. Experiment 1 tested the associations between color and musical modality with blue and yellow squares for visual stimuli and C major and C minor chords for auditory stimuli. Experiment 2 tested possible associations between emotional visual stimuli (happy and sad faces) and chord modality using the same procedure. We identified the major-happy and minor-sad pairs as congruent and major- sad and minor-happy pairs as incongruent. We were unable to find an effect of congruency in either case. These results suggest that the association between color and music is made based on higher order cognitive processes and does not have a perceptual basis.

Keywords: perception, music, modality, cross modal associations, color

---

[1]Corresponding author: Elena Bai, Rolla High School, 900 Bulldog Run, Rolla, MO 65401.
Current affiliation: McGill University, 845 Rue Sherbrooke O, Montréal, QC H3A 0G4.
belena60665@gmail.com, (573) 202 8246
This work was performed under the supervision of Dr. Ami Belfi, Department of Psychological Science, Missouri University of Science and Technology, Rolla, MO.

Introduction

While we often think about our senses as individual and unrelated, they are far from that. We blend the information that we receive from each of our senses to create the world that we experience: Information from one sense amplifies or adds context to that collected from another sense. Suspenseful scenes in films are accompanied by the appropriate tense music and light-hearted scenes are often accompanied by cheerful music to intensify the emotions in the audience.

However, one cannot simply pair any music with any situation. Imagine a scene in which a person is hiding from a murderer accompanied by cheerful, upbeat music- it would be strange (this has been used by directors to create the clash between conflicting emotions). So, certain music is better associated with certain emotions, a concept that can be applied to other types of cross modal experiences. Similar correlations have been found between different senses and with simpler stimuli, such as the correlation between high tones and smaller visual stimuli (2). Further correlations have been found between taste and music style, such as the association of sourness with higher notes and bitterness with low notes (3), and various scents with colors such as the associations between lavender and green or cinnamon and red (4).

Possibly the most explored sensory associations are those between auditory and visual stimuli. Past work has studied the correlations between pitch and timbre in auditory stimuli and color and shape in visual stimuli. Participants were asked to match auditory stimuli that varied in pitch and instrument timbre (the unique sound of a particular instrument) with one of two sets of visual stimuli, the first of varying shape and color and the second of the same shapes with varying grayscale. Participants showed strong agreement on which timbres closely matched which shapes. It was found that the cello, guitar, piano, and marimba were more closely associated with the rounder shapes. These "rounder" timbres were more closely associated with blue-green tones and lighter grayscales and higher pitches were associated with lighter grayscales (5). Further experiments have identified the association of visual luminance with auditory pitch and intervals (6). It was found that participants thought higher pitched audio and ascending intervals of notes (the second one higher in pitch) fit better with more luminous visual stimuli. Stevens and Marks (7) found correlations between louder auditory stimuli and brighter visual stimuli.

Many associations have been found between music and color based on musical tempo, mode (major or minor), and timbre with color saturation, brightness, and hue. Another study examined color-music associations between 8 colors and 24 melodies from Bach's Well-Tempered Clavier. They found that blue was most associated with minor pieces with slow tempos, while yellow was most associated with major pieces with fast tempos (8). Associations between saturation and tempo, lightness and tempo, lightness and mode, blueness (as opposed to yellowness) and tempo, and warmth/ coolness and tempo (9). Further studies have shown associations between timbre and brightness (10) and saturation (11).

However, there still exists the question of what creates these correlations. That is, why do we associate certain aspects of music and color and how are these associations created? It has been found that participants associated yellow hues and major pieces, which were both later shown to be associated with happiness, and blue hues with minor pieces, both associated with sadness (1). This suggests that these music-color associations were mediated by emotion, rather than being directly associated.

Another important question regarding these color-music correlations is whether they are made on a perceptual level. Would these associations exist when tested through unconscious, perceptual binding tests? Or does it require conscious thought to establish them? Since past tasks have typically required participants to choose the stimuli that match the best, there has been no indication as to the level on which these associations occur.

Thus, to identify any associations that can be made on an unconscious, perceptual level, we conducted a Temporal Order Judgement (TOJ) task. In a TOJ task, participants are presented with a pair of stimuli occurring at a certain stimulus onset asynchrony (SOA) and asked to judge which stimulus occurred first. The TOJ task is based on the unity assumption, which causes sets of multisensory stimuli to be perceived as coming from the same source if they have correlating physical properties, such as pitch and spatial height or size and pitch (2). When a pair of stimuli that are heavily associated or congruent, such as a large visual stimulus and a low-pitched auditory stimulus, they are more likely to be perceived as coming from the same source. Stimuli coming from the same source are more likely to be perceived as occurring at the same time. Due to this, participants will have a harder time deciding which stimulus came first in congruent pairs, such as high-pitched tones and small visual stimuli, than in incongruent pairs, such as high-pitched tones and large visual stimuli. For example, a TOJ experiment was conducted in which auditory stimuli consisted of a 9 ms burst of white noise played from either the left or right side of a fixation point, and visual stimuli consisted of a red light lighting up for 9 ms from either the left or right side of the fixation point. It was harder for participants to determine which stimulus came first in trials in which the auditory and visual stimuli were presented from the same side than for those in which they were presented from different sides (12). Thus, participants were more likely to perceive the stimuli as occurring at the same time when they had corresponding physical properties. In this way, we can test whether the previously found associations between modality and color exist on a perceptual level.

We conducted two TOJ experiments to test the associations between musical modality (major or minor) and colors and the mediating role of emotion in these associations. The first experiment used blue and yellow squares as visual stimuli and major and minor chords as auditory stimuli. In accordance with previous work (1), the congruent pairs were yellow - major and blue-minor, with yellow -minor and blue-major being the incongruent pairs. We predicted that participants would be less accurate for congruent than incongruent pairs, which would suggest a perceptual association. This is because we predict the congruent pairs to be more heavily associated than the incongruent pairs, leading to greater perceptual binding. The second experiment was designed to test whether the associations between emotion and modality were made on the perceptual level. We used sad and happy faces for visual stimuli and the same major and minor chords for auditory stimuli. Much like the first experiment, we predicted that participants would be less accurate for happy-major and sad-minor pairs (congruent) than happy-minor and sad-major pairs (incongruent). This would indicate that music-emotion associations can be made on a basic, unconscious level.

Experiment 1

Methods

Participants

Participants were 17 undergraduates from the Missouri University of Science and Technology (16 M, 1 F). Participants were aged between 18 and 22 years (M=19, SD=1.3). The experiment took approximately 45 minutes and participants received 1 credit hour upon completion. We excluded 1 participant that missed more than 2 of the 10 catch trials. The total number of participants in the final analysis was 16 (15 M, 1 F).

Procedure

This experiment was programmed using PsychoPy and conducted online using Pavlovia (13).

Procedures listed below have been approved by the Institutional Review Board at Missouri University of Science and Technology. Participants were instructed before the experiment to wear headphones turned up to a comfortable volume. At the beginning of the experiment, participants were required to read informed consent information and give consent to participate.

In each trial, an audio-visual pair of stimuli was presented at varying stimulus onset asynchrony (SOA) of 0, ±76, ±133, ±200, ±267, ±333, ± 467ms, with positive SOA values indicating the visual occurring before the audio and negative SOA values indicating the audio occurring before the visual. SOA values were based off previous work (14). Visual stimuli consisted of dark blue (#173F9C) or yellow (#FFFF69) squares of .125 height units (based on the height of the experiment window, which is equal to the size of the participant's computer monitor) with duration 100ms. Auditory stimuli consisted of C major and C minor chords, created using Matlab scripts from previous work, each with a duration of 100ms. Each stimulus was presented against a background that remained grey throughout the experiment.

The congruent pairs were yellow-major and blue-minor, while the incongruent pairs were yellow-minor and blue-major, for a total of 4 possible pairs. There were 14 SOA values (0 counted twice) and 10 trials for each option, or 14 (SOAs) * 4 (pairs) * 10 (trials). This, in addition to 10 'catch' trials with an SOA of 1500ms used to identify inattentive participants (an issue with online research, see 15), equals 570 total trials for each participant. To prevent participant fatigue, participants were instructed to take a break every 114 trials.

Participants were presented with the pair of stimuli and asked to judge which one came first, pressing 'a' on their keyboards to indicate the auditory stimulus and 'v' to indicate visual stimulus. After the participant answered either ‘a’ or ‘v’, there was a 1 second fixation ("+") and the next pair of stimuli was presented in 300 ms.

Results

We examined the average accuracy at each SOAs to see if there was an effect of congruency on accuracy. To do this, we conducted a repeated- measures ANOVA, with SOA and congruency (congruent or incongruent) as the between-subjects variables. The ANOVA found a significant effect of SOA [$F=26.977$, $p<.001$, partial $\eta^2=.643$]. Accuracy increased as SOA increased (see Figure 1). However, there was no significant effect of congruency [$F= 0.568$, $p=0.463$, partial $\eta^2=.037$] or interaction between SOA and congruency [$F=0.599$, $p=0.828$, partial $\eta^2=.038$].

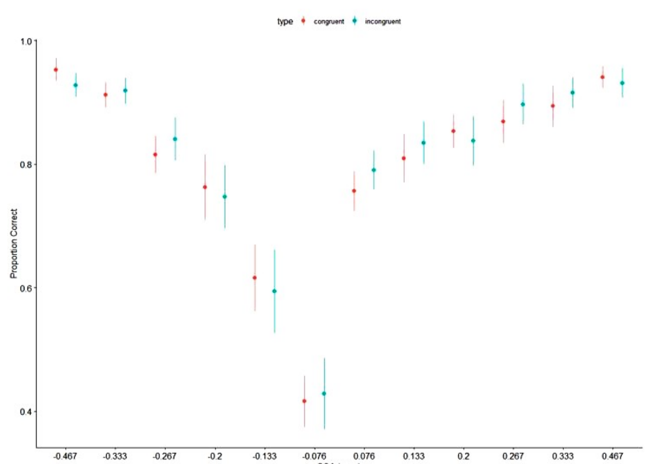


Figure 1. Accuracy at each SOA in Experiment 1. The y-axis displays the proportion correct, while the x-axis displays the SOA. Points indicate the averages across participants, and the error bars depict the standard error of the mean.

We then calculated the psychometric curve for the individual participants across congruent and incongruent trials (see Appendix A). The psychometric curve displays the proportion of trials in which the participant answered that the visual stimulus was presented first and is used to determine the JND. For the psychometric curve for the entire group, see Figure 2. The JND is the difference between the SOA value at which participants answered correctly 75% of the time and that at which they answered correctly 50% of the time. It shows the gap in SOA where the participant is unsure of which stimulus occurred first and should be larger if two stimuli are more closely associated in this task.

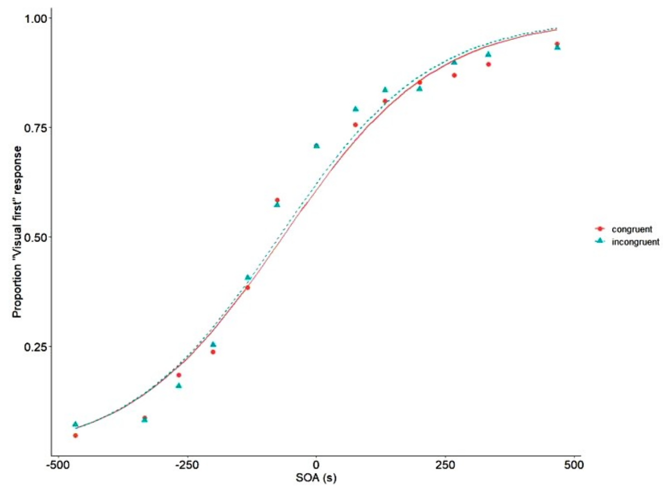


Figure 2. Psychometric curve for all participants in Experiment 1. This displays the average psychometric curve across all participants.

Using these JND values, we conducted a paired-samples t-test to find differences in JND between congruent and incongruent trials (see Figure 3). The average JND value in congruent trials was 168 ms and that of the incongruent trials was 160 ms, with standard deviations of 102 and 94 ms, respectively. However, we found no significant difference between congruent and incongruent JND values ($t=0.863$, $p=0.40$, 95% CI: [-.01, .03]).

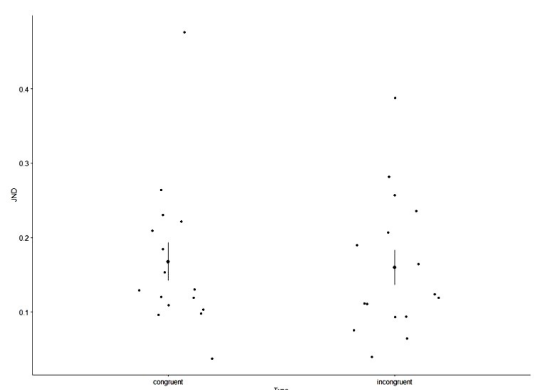


Figure 3. Comparing JND in congruent and incongruent trials (Experiment 1). Large circles represent the group averages, and error bars indicate the standard error of the mean. Individual points represent individual subjects. Jittering is to prevent the points from overlapping and does not indicate any difference in the data on the x-axis.

Discussion

The results from Experiment 1 suggest that there is no effect of congruency on participants' JND and accuracy. The resulting JND values are slightly higher than those of past studies

(14), which could be because the stimuli used in this experiment were more complex and it would be necessary to use more attention to distinguish between major and minor chords (16). Furthermore, as can be seen in Figure 1, the accuracy is higher for trials with higher absolute SOA and lower for SOA values closer to 0. This is a typical pattern, as larger absolute values in SOA indicate a larger gap between the onset of the first and second stimulus, making the task less difficult as the gap increases. Further, there was no significant effect of stimulus order on the participants' accuracy. In other words, participants' accuracy is roughly the same between audio first and visual first trials. Thus, while the data is regular and does not point to any large experimental errors, it indicates that there was no perceptual binding between the pairs of stimuli.

The lack of a clear effect of congruency on JND, as opposed to previous work, suggests that these associations are more complex than those tested in previous work or that they do not have a perceptual basis. It is likely that participants' associations of blue and yellow with minor and major modalities, respectively, was influenced by a cultural or semantic bias rather than existing on a perceptual level.

Experiment 2

In Experiment 1, we were unable to find any significant effect of congruency on JND, suggesting that the associations between chords and colors are not perceptual in nature. However, previous work has found that these associations are mediated by emotion

(1). Thus, we decided to conduct another TOJ experiment to explore the potential for perceptual associations between music and emotion. In Experiment 2, we used pictures of a man displaying happiness and sadness to establish emotion and conducted a similar TOJ experiment to Experiment 1.

Methods

Participants

Participants consisted of 16 undergraduate students at the Missouri University of Science and Technology (M=7, F=9). They were between 18-22 years of age (M=19, SD=1.02). The experiment took approximately 45 minutes and each participant received 1 credit hour upon completion.

Procedure

Experiment 2 used the same methods as Experiment 1 with stimuli chosen to test possible connections between emotion and chord modality. For our emotional stimuli, we chose the face of the same man displaying happiness and sadness from the Karolinska Directed Emotional Faces set (17; See Appendix C). Faces were presented at a duration of 100 ms and a size of .5*.6 height units. The same C major and C minor chords from Experiment 1 were used as auditory stimuli at a duration of 100ms. All other procedures were identical to Experiment 1.

Results

As in Experiment 1, we first conducted a repeated-measures ANOVA to investigate possible effects of SOA and congruency on participants' accuracy. We found a significant effect of SOA [$F=17.517$, $p<0.001$, partial $\eta^2=0.593$], but no effect of congruency [$F=0.019$, $p=0.89$, partial $\eta^2=.002$] or interaction between SOA and congruency [$F=1.112$, $p=0.36$, partial $\eta^2=.085$]. Next, the psychometric curves were calculated and the JNDs found. The average JND across participants for the congruent pairings was 133 ms and that of the incongruent pairings was 129 ms, with standard deviations of 54 and 69 ms, respectively.

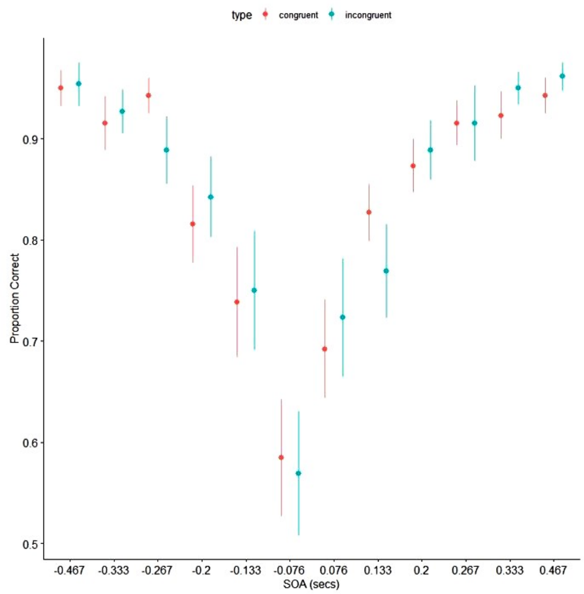


Figure 4. Accuracy at each SOA in Experiment 2. The y-axis displays the proportion correct, while the x-axis displays the SOA. Points indicate the averages across participants, and the error bars depict the standard error of the mean.

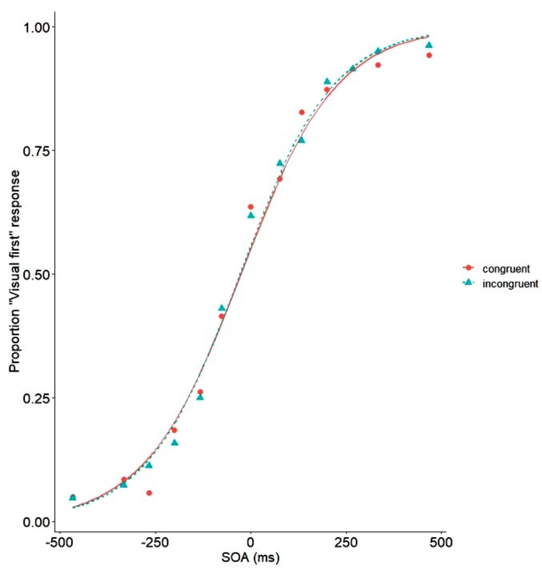


Figure 5. Psychometric curve for all participants in Experiment 2.

Then, we conducted a paired-samples t-test to compare the JND in congruent and incongruent trials. We found no significant effect of congruency on JND (t=-.3332, p=.7447, 95% CI: [-.018, 0.25]).

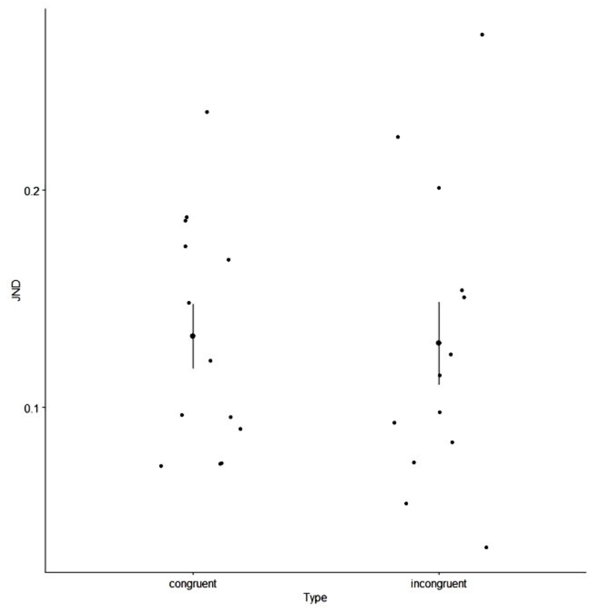


Figure 6. Comparing JND in congruent and incongruent trials (Experiment 2). Large circles represent the group averages, and error bars indicate the standard error of the mean. Individual points represent individual subjects. Jittering is to prevent the points from overlapping and does not indicate any difference in the data on the x-axis.

Discussion

We were again unable to find any significant effect of congruency on participants' accuracy or JND. However, Figure 4 displays an expected graph shape, which decreased when SOA values approached 0 from both sides and with no effect of stimulus order. This, like the results of Experiment 1, demonstrates that although the data displays normal behavior, it does not point to any effect of congruency.

Our goal in this experiment was to directly test the associations between music and emotion. However, showing visual stimuli of happy and sad faces may not have carried the same emotional weight perceptual level. Previous studies have included pictures of erotic couples for positive stimuli and threatening scenes for negative stimuli (18). Specific emotions, such as happiness and sadness may be more difficult to evoke with pictures.

Thus, despite modality- emotion correspondences found in previous work, our results show no effect of congruency on SOA. This suggests that the music emotion association found in previous studies was not perceptual in nature.

General Discussion

Neither Experiment 1 nor Experiment 2 demonstrated any significant difference in JND between groups. This does not support our hypothesis that yellow visual stimuli and major chords and blue visual stimuli and minor chords would be perceptually bound together, as perceptual binding should result in larger JND values for congruent pairs. Although we found no associations between either color and music modality or emotion and music modality, all JND values were within an expected range as determined by similar experiments. Our data also followed the expected trend, with accuracy increasing as the space between the stimuli increased and the proportion of "visual first" responses at each SOA forming a psychometric curve.

The lack of an effect of congruency on JND and accuracy in either experiment suggest that there is no perceptual binding in music-color or music-emotion pairs. If there were implicit correlations between music modality and color or emotion, participants would have been more accurate in incongruent trials than congruent trials. This suggests that the associations between music, color, and emotion are made only through conscious effort. However, one study observed these associations throughout the United States and Mexico (19). In order to generalize to other populations, it may be helpful in future studies to explore these associations, both perceptual and decision-based, with larger sample sizes that include participants from different cultural backgrounds. Future research could also broaden the participant demographics in other ways, for example, by investigating potential sex differences in these effects. It has been found that men typically have lower Temporal Order Threshold (TOT) values, which are similar to JND values in this study (16). Since the proportion of men and women were not equal in our present studies, and we did not intend to assess sex effects when planning our sample size, we are not able to directly test for sex effects in the current dataset. Finally, while our study did not examine the effect of participants' musical training or experience on their performance in the present task, a previous study that examined this found no significant effect of musical experience on performance in a similar task (21).

Alternatively, because of the many variables that make a piece of music, it may be that a combination of variables would be better associated with a certain emotion or color. For example, a slow song in a minor key may be more likely to be interpreted as sad. However, it is more difficult to test for perceptual associations in stimuli with more variables, such as tempo, dynamics, and note length. Thus, either because the emotional mediation in color-music associations does not exist on a perceptual level or because of the complexity of music, there were no implicit associations found between either modality and color or modality and emotion.

In previous papers examining the associations between music and color, experimenters have used lines of melody rather than isolated chords (1). Thus, this experiment may suffer from a lack of musical context or other aspects of music, such as the tempo. Although there are other audio-visual TOJ experiments examining aspects of music such as the pitch and volume, the associations require no apparent emotional mediation. Because of this, they may not require the same amount of context as the current experiments. Furthermore, these associations may simply be unable to be made on a perceptual level. Past studies have either asked participants to choose what color best matches a melody or rate the degree to which they match, establishing no associations that do not require effortful decision making.

However, there are limitations to our study. We only included 16 participants in the final analysis of both Experiments 1 and 2. While this is typical of studies involving TOJ tasks, as there are many trials for each participant, a larger sample size should be a consideration in the future (14; 20). As can been seen in Appendix A, individual performance on the task varied greatly. Previous studies have found that participants' performance on TOJ tasks can be affected a number of factors, including gender, age, intellectual resources (assessed with Mosaik test), alertness, and vigilance (16). Additionally, since the study was conducted online, we were not able to ensure that the participants were paying attention when completing the task. We attempted to account for this using the catch trials and excluding participants based on accuracy at the highest SOA, but it is possible that participants' data was affected by lack of attention.

To conclude, through this study, we examined potential perceptual associations between color and music modality as well as emotion and music. Because our results point to no association between either pair of variables, it is likely that the demonstrated associations are made with effort and that there is no inherent factor that associates them. However, there is evidence that these blue/minor and yellow/major associations exist across cultures, and it would be interesting to examine the possibility of a common root for these associations.

References

1. Palmer, S. E., Langlois, T. A., & Schloss, K. B. (2016). Music-to-color associations of single-line piano melodies in non-synesthetes. *Multisensory Research*, *29*(1-3), 157-193.

2. Chuen, L., & Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude envelope cues. *Attention, Perception, & Psychophysics*, *78*(5), 1512-1528.

3. Crisinel, A. S., & Spence, C. (2009). Implicit association between basic tastes and pitch. *Neuroscience letters*, *464*(1), 39-42.

4. Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *The American journal of psychology*, 335-351.

5. Adeli, M., Rouat, J., & Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Frontiers in human neuroscience*, *8*, 352.

6. Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American journal of psychology*, 219-238.

7. Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences of the United States of America*, *54*(2), 407.

8. Isbilen, E. S., & Krumhansl, C. L. (2016). The color of music: Emotion-mediated associations to Bach's Well-tempered Clavier. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 149.

9. Palmer, S. E., Langlois, T., Tsang, T., Schloss, K. B., & Levitin, D. J. (2011). Color, music, and emotion. *Journal of Vision*, *11*(11), 391-391.

10. Caivano, J. L. (1994). Color and sound: physical and psychophysical relations. *Color Research & Application*, *19*(2), 126-133.

11. Griscom, W. S., & Palmer, S. E. (2012). The color of musical sounds: Color associates of harmony and timbre in non-synesthetes. *Journal of Vision*, *12*(9), 74-74.

12. Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research*, *152*(2), 198-210.

13. Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8-13.

14. Parise, C. V., & Spence, C. (2009). 'When birds of a feather flock together': Synesthetic correspondences modulate audiovisual integration in non- synesthetes. *PloS one*, *4*(5), e5664.

15. Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods*, *17*(3), 437.

16. Szymaszek, A., Sereda, M., Pöppel, E., & Szelag, E. (2009). Individual differences in the perception of temporal order: the effect of age and cognition. *Cognitive neuropsychology*, *26*(2), 135-147.

17. Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*, *91*(630), 2-2.

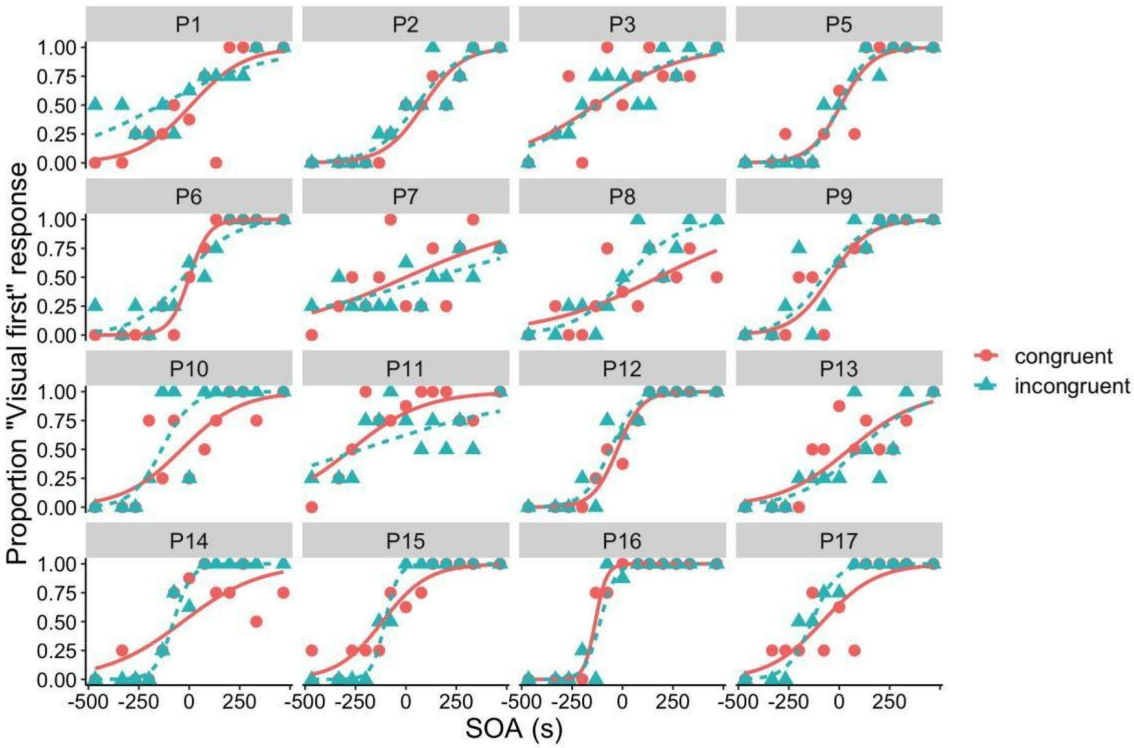
18. Shi, Z., Jia, L., & Mueller, H. J. (2012). Modulation of tactile duration judgments by emotional pictures. *Frontiers in integrative neuroscience*, *6*, 24.

19. Prado-León, L. R., Schloss, K. B., & Palmer, S. E. (2014). Color, music, and emotion in Mexican and US populations. *New Directions in Colour Studies. Amsterdam: John Benjamins*.

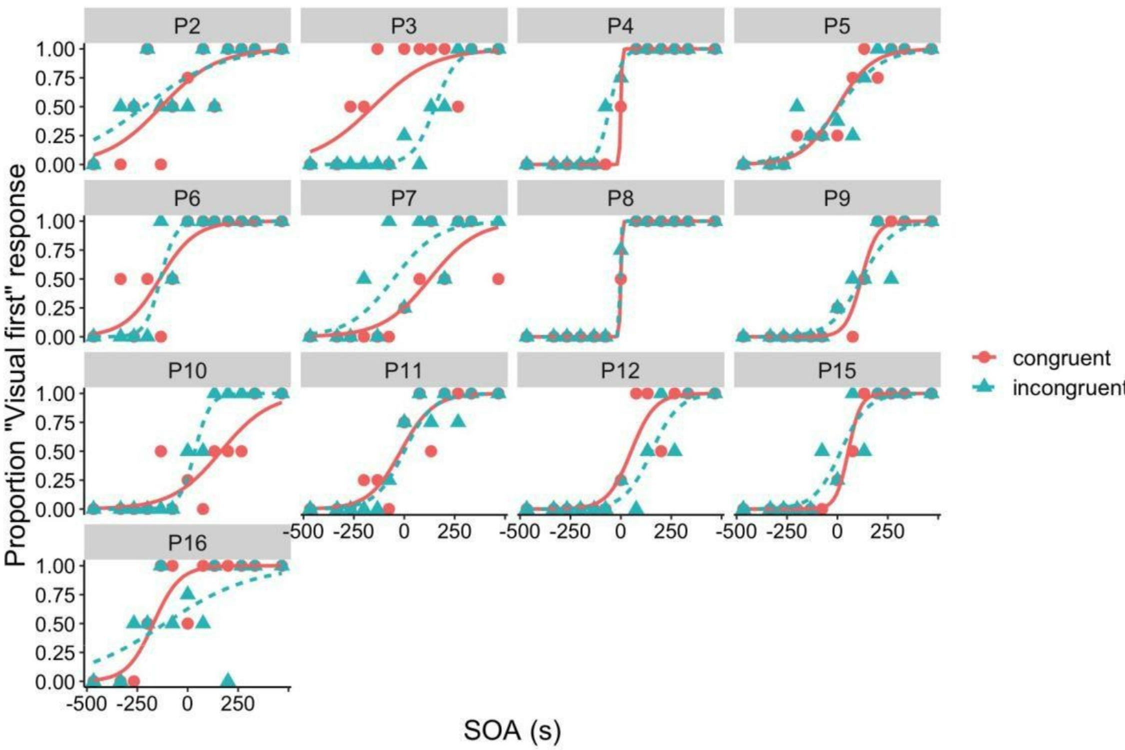
20. Habets, B., Bruns, P., & Röder, B. (2017). Experience with crossmodal statistics reduces the sensitivity for audio-visual temporal asynchrony. *Scientific reports*, *7*(1), 1-7.

21. Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain research*, *1111*(1), 134-142.

Appendix A: Individual psychometric curves from Experiment 1



Appendix B: Individual psychometric curves from Experiment 2



Appendix C

Happy face from Karolinska Direct Faces set (Experiment 2)

Sad face from Karolinska Direct Faces set (Experiment 2)