INVESTIGATING COLOUR-SOUND MAPPING IN CHILDREN AND ADULTS: A PILOT STUDY

Costanza Cenerini

Università Campus Bio-Medico di Roma
Via Alvaro del Portillo 21
Rome, Italy

Luca Vollero

Università Campus Bio-Medico di Roma
Via Alvaro del Portillo 21
Rome, Italy

Giorgio Pennazza

Università Campus Bio-Medico di Roma
Via Alvaro del Portillo 21
Rome, Italy

Marco Santonico

Università Campus Bio-Medico di Roma
Via Alvaro del Portillo 21
Rome, Italy

Nicola Di Stefano

National Research Council
Institute for COgnitive Sciences and Technologies
Rome, Italy

Flavio Keller

Università Campus Bio-Medico di Roma
Via Alvaro del Portillo 21
Rome, Italy

ABSTRACT

This paper presents an experimental study that investigates the mapping between color and sound in two groups of subjects, consisting of 8 children and 8 adults, respectively. The color was decomposed using the Hue-Saturation-Brightness representation, and the musical stimuli consisted of chords determined by three parameters: the Root Tone, the Mode, and the Octave. The results showed that both groups preferred the mapping Hue-Octave, Saturation-Mode, Brightness-Root Tone. The ANOVA test was applied to the results, and it showed a trend towards significance in children's preferences, whereas adults' results were not significant. Once the best mapping is defined, it will be implemented in a portable, embedded system for sensory substitution of sight with sound. This paper provides a promising starting point for further research in the field of affective sonification.

1. INTRODUCTION

In recent years, the concept of sonification of color has gained much attention as a means of helping individuals with visual impairments to obtain information about the world around them. The process of converting color into audio has been explored through various devices and theories, with the aim of providing visually impaired individuals a way to perceive the visual world. However, despite the technological advancements in this field, the lack of emotional engagement in these devices has been noted as a major drawback, leading to low acceptance among users.

The literature on the topic of sonification of color is rich and diverse, covering a wide range of subjects from the design of auditory color spaces [1] to the development of virtual reality tools for training and evaluation [2]. The field has seen the creation of various devices such as the ColorPhone [3], the Sensory Substitution Device [1], and the See ColOr [4], as well as studies on the design of sonification methods [5].

Moreover, research has also explored the integration of multiple senses in sonification, with a focus on the design of audiovisual sonifications [6]. In the realm of art and multimedia installations, the Voice of Sisyphus [7] and the Traces of Modal Synergy [8] are notable examples of the use of sonification in creating engaging and emotional experiences.

The evaluation of auditory displays and sonification of textured images has been an area of interest, with studies by Martins [9, 10] and Evreinov [11] exploring the use of tonal volume and spot-mapping in imaging sonification. The sonification of color and depth in mobility aids for the blind has also been explored [12], and the perception of compensatory devices has been studied in the context of sensory substitution and sensorimotor extension [13].

By building upon the existing body of literature and exploring new avenues, we hope to contribute to the advancement of the field of color sonification and create more effective and engaging devices for visually impaired individuals.

Specifically, the objective of this paper is to propose a new kind of approach of mapping color into sound and to test the preferences of two groups of children and adults on its difference combinations, in order to determine whether the two groups have similar preferences.

This work is licensed under Creative Commons Attribution – Non Commercial 4.0 International License. The full terms of the License are available at http://creativecommons.org/licenses/by-nc/4.0/

2. ASSOCIATION CRITERION DESIGN

The first step in designing the association was to select the color model and the type of musical stimulus. To do so, we considered the literature on synesthetes who experience colored hearing, where the sight of color automatically leads to the involuntary experience of sound, and studies on crossmodal correspondences between visual and acoustic dimensions.

2.1. Color model

After reviewing different studies on the color-sound association, we chose to adopt the Hue, Saturation, Brightness (HSB) color model: Hue refers to the basic color tone, Saturation defines the intensity or purity of a color, with higher saturation indicating a more vivid color, and Brightness refers to the perceived brightness of a color, with higher brightness indicating a lighter color. According to a recent review on colored hearing [14], these three components are very commonly associated with the correspondence between color and sound:

1. Hue is often associated with timbre [15, 16] and historically linked to pitch [17, 18].
2. Saturation has been reported to be associated with loudness and pitch [19].
3. Brightness is often seen in relation to pitch and loudness [20, 21].

2.2. Musical model

Since the objective of this work is to design an association criterion that allows users to feel the emotion they would upon seeing a color while listening to music, it was crucial to choose a musical model that could elicit a sentiment. For this reason, we decided to use an auditory stimulus that reproduces sounds of common instruments with notes and harmony in accordance with the rules of western music.

The key characteristics were two: it had to be representable on a scale with sufficiently high resolution so that each color could be linked to a single sound, and it had to remain unchanged over time regardless of the surroundings, as sensory substitution devices are to be taken everywhere. From the various mappings that were reported above, the obvious choice for the music model would be to use a representation including pitch, loudness, and timbre. Out of these, pitch is the only one that satisfies the aforementioned conditions: timbre can be difficult to discern and is not easily representable in a scale, and loudness is heavily influenced by outside noises.

For all these reasons, the selected musical stimulus was a chord, which was also one of the forms of Scriabin's synesthesia [22]. Chords were generated according to occidental classical music harmony. The following parameters of the chord were selected:

1. The root tone (RT), which is one out of the 12 notes in the chromatic scale (i.e. C, C#/D, E, ...).
2. The octave on the piano (OC), which is one out of 7 (i.e. C1, C2, ...).
3. The mode of the chord (MO), which is one out of up to 12 (i.e. minor, major, ...).

The chord is played with all the notes at the same time and can be heard with a frequency of one chord per second.

2.3. Correspondence between stimuli

Once the models for color and music were chosen, one out of the 6 possible mappings of the musical-color parameter associations (see Table 1) had to be chosen. Since literature on the topic does not provide a clear indication, the best course of action was to gather data from subjects in a trial. In this way, the outcome of the study would be an association logic that elicits the same emotions for most of the population so that there is the greatest chance that devices developed using such logic will not be rejected for lack of emotional stimulus.

When citing the possible mappings later in this paper, they will be called with the three musical parameters respectively associated with Hue, Saturation, and Brightness (for example, the mapping Hue-OC, Saturation-RT, Brightness-MO will be called OC-RT-MO).

Hue	Saturation	Brightness
OC	MO	RT
OC	RT	MO
MO	OC	RT
RT	OC	MO
MO	RT	OC
RT	MO	OC

Table 1: Possible mappings.

3. TAILORING OF THE LOGIC

In order to select which musical parameter had to be associated with each color parameter, we designed an experimental trial. This study received the approval of the university's ethical committee on February 16, 2022, with the clinical studies register number 2021.236.

3.1. Protocol

The test was administered remotely via a web app. The graphical interface was developed using `p5.js`[1], the audio was generated with `soundfont-player`[2], and the backend was developed using `node.js`[3].

Subjects were shown two color spectra (see Figure 1) and could move the cursor on them to hear a chord played by a piano repeatedly. The two scales had different color-music mappings, so each color played differently in the two scales. They were asked to set the audio and screen settings to their preferences and to choose which of the two scales sounded more pleasant. They repeated the selection three times for each of the 6 possible mappings, with two of the values of the HSB model set and only the last one changed in the scale. For example, saturation and brightness were set to an intermediate value and they were shown a scale of hue. The total number of selections to be completed was 45, and they were presented to all subjects in the same order. the test lasted an average of 40 minutes, with a minimum time of 30 minutes and a maximum of 60 minutes.

Each scale selection was saved in the results, which could be downloaded as a JSON file at the end of the test and sent to the

[1] `https://p5js.org/`
[2] `https://www.npmjs.com/package/soundfont-player`
[3] `https://nodejs.org/`

author. If at any point they felt tired, they could click on the "I'm done!" button and save their results. They could then resume the test at a later time.

3.2. Participants

Two groups of participants took part in the study: a group of 8 children aged between 7 and 11 years old (with a mean age of 9.5 years, 37.5% were female) and a group of adults aged between 25 and 36 years old (with a mean age of 28 years, 50% were female). A summary of participant demographics is provided in Table2

(a) Children Group

Subject	Sex	Age
1	M	11
2	M	9
3	F	8
4	M	9
5	F	11
6	F	11
7	M	7
8	M	8

(b) Adults Group

Subject	Sex	Age
1	F	29
2	F	28
3	F	25
4	M	26
5	M	27
6	M	25
7	M	36
8	F	28

Table 2: Description of the participants

Before starting the test, all participants signed the privacy policy and informed consent form.

4. DATA ANALYSIS

4.1. Data collection

For each participant, we extracted two pieces of information:

- The number of times each mapping was chosen: each pairing was presented to the subjects 5 times for comparison with the other possible mapping. For each couple of mappings, they had to choose which of the two mappings they preferred three times, one for hue, saturation, and brightness. For each comparison between mappings, we considered the mapping selected if it was chosen at least 2 out of 3 times for the color parameters comparison.
- The number of times each color-musical parameter pairing was chosen: each pairing could be selected a maximum of 5 times.

Once we had extracted these values from each subject, we computed the mean and standard error of the results for each group and for all subjects to determine the most voted mappings. We also applied the ANOVA test to the mappings for both groups and all subjects to verify the significance of at least one preferred mapping. Before performing the test, the Lilliefors test was applied to the mapping results to verify their Gaussian distribution.

4.2. Results

The Lilliefors non ha dato risultati significativi for the children's, adults', and combined groups, with p-values of 0.0026, 0.0106, and 0.001, respectively. Results from the ANOVA are shown in Table 3 The ANOVA results for children suggest a trend towards a difference in the ratings of the mappings, with a value close to 0.05. In contrast, the differences in adults' choice of mappings

Population	df	F	p-value
Children	5	2.31	0.06
Adults	5	1.76	0.14
Children and adults	5	1.68	0.15

Table 3: Results of ANOVA test on mappings: a value minor than 0.05 would indicate that the test rejects the null hypothesis of equal means at the 5% significance level

are not significant. For this reason, when combining the results of children with those of adults, the significance of the children's group is lost.

The results of the data collection are shown in Figure **??**.

The differences between the two groups can be seen in Table 4

Children's mappings' results can be described as follows:

- A strong preference for the OC-MO-RT mapping (mean value = 3.25) with a low variance (0.45), indicating compactness in the group.
- A slightly less preference for the RT-OC-MO mapping (mean value = 3.00) with a low standard error (0.41).
- A weak preference of the MO-RT-OC mapping (mean value = 1.5) with a low variance (0.42).
- An equally weak preference of the RT-MO-OC mapping (mean value = 1.8) with a very low standard error (0.31).

Their results on pairings are:

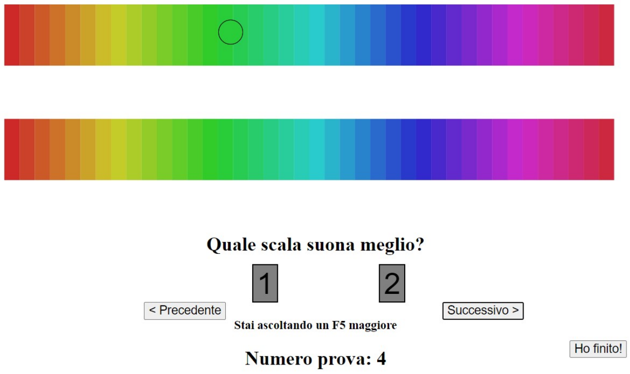
- The best sonification of the Hue is through the Root Tone (mean value = 5.8), followed by the Octave (5.5). Both these values, however, have a slightly high value of standard error (i.e., 0.68 and 0.79 respectively), suggesting lower agreement in the group. On the other hand, the pairing Hue-Mode is the least voted with high agreement.
- The best pairing for saturation is the Octave, which collected a high number of preferences with a low standard error (mean value = 5.5, standard error = 0.35), followed by the Root Tone (mean value = 5.0, standard error = 0.47).
- Lastly, the most voted pairings for brightness were the Root Tone (mean value = 6.0), whose standard error is a little high (0.61). The Mode also received a high score (mean value = 5.0) with high agreement.

As for the adults, their preferences on mappings are described below:

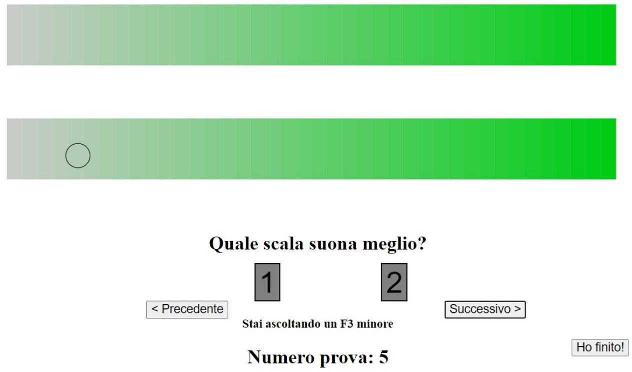
- The preferred mapping is OC-MO-RT (mean value = 3.3, standard error = 0.56).
- RT-MO-OC received a similar score (3.25), but with a lower standard error (0.38), indicating a higher agreement in the group.
- They show a weaker preference for the mapping RT-OC-MO (mean value = 1.75) with a low standard error (0.40).
- Similarly, they voted the mapping OC-RT-MO the least with a high agreement (mean value = 2.1, standard error = 0.40).

Their choices on the pairings are listed below:

- For the sonification of Hue, they like Octave and Mode equally (5.1), with the first one having a higher standard error,



(a) Shades of hue



(b) Shades of saturation

(c) Shades of brightness

Figure 1: Graphical Interface of the test: The two scales play different chords for the same color. Participants have to choose which one sounds better to them, they can go back to the previous selection and move on to the next one. They are shown the name of the musical chord they are hearing and the number of the selection they are currently at and they can click the "Finish!" button if they are tired. As all the participants were from Italy, the test was administered in Italian.

indicating the more agreement in the second choice.

- All the values of the saturation have fairly high standard error (ranging from 0.68 to 0.65), suggesting a general low agreement in this parameter. Despite this, a strong preference for the mode (mean value = 6.3) and a weak preference of the Octave (mean value = 3.6) can be seen.
- The higher pairing for brightness, which is the Octave (mean value = 5.8) have a very high standard error (1.08), indicating a high disagreement in the group.

When analyzing the two groups as one, the following results emerged:

- There is a preference for the mapping OC-MO-RT (mean value = 3.3) with a relatively low standard error (i.e. 0.47) indicating agreement in the whole group.
- There is a weak preference of the mapping MO-RT-OC (mean value = 1.9).
- The pairings value belong in a very small range (from 4.6 to 5.4) and not much can be said about preferences.

5. DISCUSSION

Results from the two groups show some differences, which can be interpreted as an indication of a different perception of color and music. In particular, there is a high discrepancy in their preferences in two mappings:

- RT-MO-OC, which was the second most voted by adults but the least voted by children.
- RT-OC-MO, which was the second most voted by children but the least voted by adults.

These differences can be explained by the pairing with saturation: children prefer the pairing with the Octave the most, while it is the worst for adults, and adults prefer the pairing with the Mode the most, while it is the worst for children.

Despite this, both groups indicate with high agreement that the mapping OC-MO-RT is the preferred one. This can be explained by the strong preference for the pairing Hue-OC in both groups, for Brightness-RT in children, and for Saturation-MO in adults.

This result suggests that, despite their differences, children and adults agree on this type of mapping, which could be the best one for color sonification.

It is important to note that, despite the agreement, the ANOVA test revealed that the results were close to be significant for the

Table 4: Results of data collection

(a) Mapping ratings in children

Hue	Saturation	Brightness	Mean
OC	MO	RT	3,3 ± 0,45
OC	RT	MO	2,9 ± 0,61
MO	OC	RT	2,6 ± 0,50
RT	OC	MO	3,0 ± 0,46
MO	RT	OC	1,5 ± 0,42
RT	MO	OC	1,8 ± 0,31

(b) Mapping ratings in adults

Hue	Saturation	Brightness	Mean
OC	MO	RT	3,3 ± 0,56
OC	RT	MO	2,1 ± 0,40
MO	OC	RT	2,4 ± 0,45
RT	OC	MO	1,75 ± 0,40
MO	RT	OC	2,25 ± 0,62
RT	MO	OC	3,25 ± 0,38

(c) Pairings ratings in children

	OC	MO	RT
Hue	5,5 ± 0,79	3,8 ± 0,46	5,8 ± 0,68
Saturation	5,5 ± 0,35	4,5 ± 0,59	5,0 ± 0,47
Brightness	4,0 ± 0,73	5,0 ± 0,40	6,0 ± 0,61

(d) Pairings ratings in adults

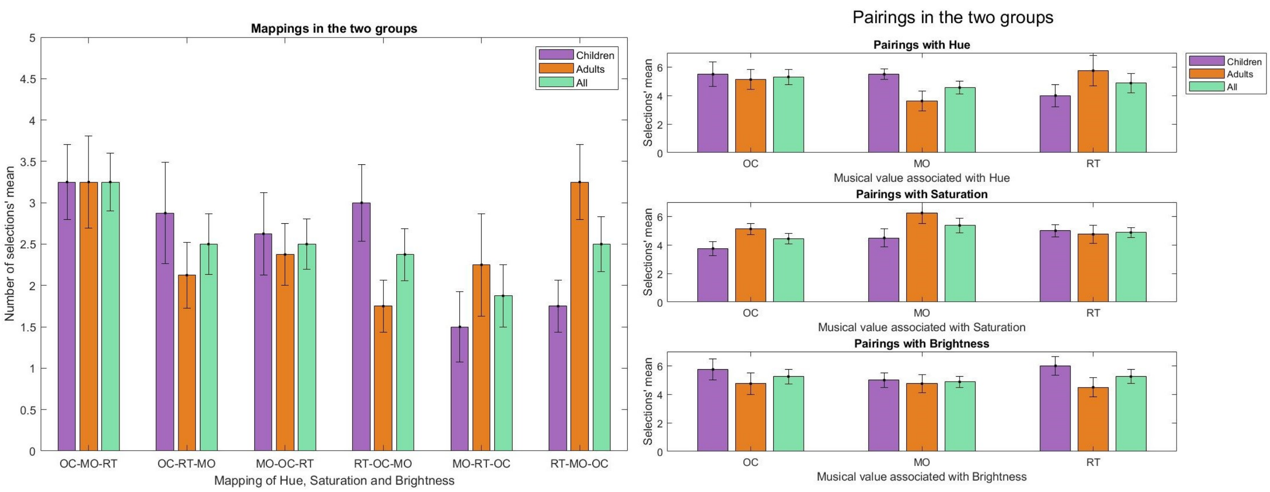
	OC	MO	RT
Hue	5,1 ± 0,69	5,1 ± 0,40	4,8 ± 0,75
Saturation	3,6 ± 0,68	6,3 ± 0,73	4,8 ± 0,65
Brightness	5,8 ± 1,08	4,8 ± 0,62	4,5 ± 0,68

(e) Mappings ratings in subjects

Hue	Saturation	Brightness	Mean
OC	MO	RT	3,3 ± 0,47
OC	RT	MO	2,5 ± 0,53
MO	OC	RT	2,5 ± 0,70
RT	OC	MO	2,4 ± 0,44
MO	RT	OC	1,9 ± 0,64
RT	MO	OC	2,5 ± 0,52

(e) Pairings ratings in subjects

	OC	MO	RT
Hue	5,3 ± 0,79	4,4 ± 0,48	5,3 ± 0,76
Saturation	4,6 ± 0,60	5,4 ± 0,70	4,9 ± 0,58
Brightness	4,8 ± 0,93	4,9 ± 0,48	5,2 ± 0,73



(a) Differences in mappings

(b) Differences in pairings

Figure 2: Differences between children and adults

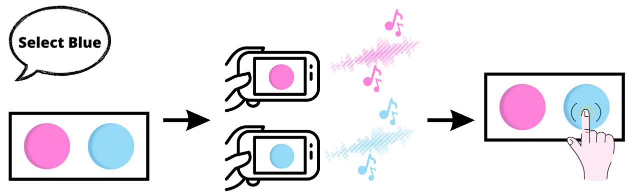


Figure 3: An example of a possible application in the field of sensory substitution is using sonification to help subjects with visual impairments distinguish between two colors and select the desired one

children's group, whereas the results of the adults and the entire group were not significant. Therefore, the preference for mapping can only be expected in this particular group.

The limited number of subjects considered must be taken into account when drawing conclusions, which can only be preliminary at this point and should be confirmed by a higher number of data.

6. CONCLUSION AND FUTURE DEVELOPMENTS

In this paper, the authors presented the design of an association criterion between color and music with the objective of finding the most natural to all subjects.

Three musical parameters and the HSB color representation were chosen for this task, and an experimental protocol was developed to find the best mapping. The protocol was applied to two groups of subjects: the first consisted of 8 children, and the second one of 8 adults. Their preferences on mappings and pairings were registered and confronted.

Results show that children and adults present several differences in their preferences, but both like the mapping Hue-Octave, Saturation-Mode, Brightness-Root Tone the best. This preference is a result of the strong liking of the pairing Hue-Octave by both groups, of the Saturation-Mode by the adults, and of the Brightness-Root Tone by the children.

Their agreement on the same mapping is promising for the identification of a new mapping that will enable a more natural sonification, but it is yet a preliminary result that will have to be confirmed by a higher number of data collected from more subjects.

The aforementioned mapping was derived from the subjects' preferences, and therefore, it may be worthwhile to investigate whether this approach could serve as a viable means for conveying the emotional responses evoked by color through sound. By addressing the issue of the limited adoption of sensory substitution devices resulting from the failure to transmit emotional information between sensory modalities, this mapping technique holds potential as a promising solution.

6.1. Application

Once the best mapping is defined, it will be implemented in a non-invasive, portable, embedded system. In particular, this mapping will be combined with sound alteration techniques for audio localization of the acoustic source in the space to provide the user with information on the background without looking at it (Figure 3). The obvious application for this type of device is sensory substitution of sight with sound: this new affective approach aims to solve the scarce use of these devices that is caused by the inability that they have to convey emotion from one sense to another

7. REFERENCES

[1] D. Osinski and D. R. Hjelme, "A sensory substitution device inspired by the human visual system," in *2018 11th International Conference on Human System Interaction (HSI)*. IEEE, 2018.

[2] P. Bizoń *et al.*, "Virco: A virtual reality tool for long-term training and evaluation of the cognitive skills development in an interactive sensory substitution environment," in *2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2019.

[3] P. Bizoń-Angov *et al.*, "Visual echolocation concept for the colorophone sensory substitution device using virtual reality," *Sensors*, vol. 21, no. 1, p. 237, 2021.

[4] J. D. Gomez, G. Bologna, and T. Pun, "See color: an extended sensory substitution device for the visually impaired," *Journal of Assistive Technologies*, vol. 8, no. 2, pp. 77–94, 2014.

[5] W. S. Yeo and J. Berger, "A framework for designing image sonification methods," in *Proceedings of international conference on auditory display*, 2005.

[6] V. Caiola, S. Lenzi, and D. Riccò, "Audiovisual sonifications: A design map for multisensory integration in data representation," 2022.

[7] R. M. McGee, J. Dickinson, and G. Legrady, "Voice of sisyphus: An image sonification multimedia installation," 2012.

[8] N. Rönnberg and J. Lowgren, "Traces of modal synergy: studying interactive musical sonification of images in general-audience use," 2019.

[9] A. C. G. Martins *et al.*, "Auditory display and sonification of textured image," 1996.

[10] A. C. G. Martins and R. M. Rangayyan, "Experimental evaluation of auditory display and sonification of textured images," 1997.

[11] G. Evreinov, "Spotty: imaging sonification based on spot-mapping and tonal volume," 2001.

[12] T. Pun, B. Deville, and G. Bologna, "Sonification of colour and depth in a mobility aid for blind people," 2010.

[13] M. Auvray and E. Myin, "Perception with compensatory devices: From sensory substitution to sensorimotor extension," *Cognitive Science*, vol. 33, no. 6, pp. 1036–1058, 2009.

[14] C. Spence and N. Di Stefano, "Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound," *i-Perception*, vol. 13, no. 3, p. 20416695221092802, 2022.

[15] A. Lavignac, *Music and musicians*. Henry Holt and Company, 1899.

[16] Z. Wallmark, "Semantic crosstalk in timbre perception," *Music & Science*, vol. 2, p. 2059204319846617, 2019.

[17] J. W. v. Goethe, *Theory of colours*. John Murray, 1840, original work published 1810.

[18] H. v. Helmholtz, *Handbuch der Physiologischen Optik [Handbook of Physiological Optics].* Voss, 1867.

[19] A. Anikin and N. Johansson, "Implicit associations between individual properties of color and sound," *Attention, Perception, & Psychophysics*, vol. 81, pp. 764–777, 2019.

[20] L. E. Marks, "On cross-modal similarity: Auditory–visual interactions in speeded discrimination." *Journal of experimental psychology: Human Perception and Performance*, vol. 13, no. 3, p. 384, 1987.

[21] ——, "On cross-modal similarity: the perceptual structure of pitch, loudness, and brightness." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 15, no. 3, p. 586, 1989.

[22] B. M. Galeyev and I. L. Vanechkina, "Was scriabin a synesthete?" *Leonardo*, vol. 34, no. 4, pp. 357–361, 2001.