Using Crossmodal Correspondence Between Colors and Music to Enhance Online Art Exhibition Visitors' Experience

Qian Guo[1] and Tingting Jiang[1,2](✉)

[1] School of Information Management, Wuhan University, Wuhan, Hubei, China
`tij@whu.edu.cn`
[2] Center for Studies of Information Resources, Wuhan University, Wuhan, Hubei, China

Abstract. Background music has been widely used in online art exhibitions to enhance visitors' art appreciation experience. At present, what music to use is highly dependent on exhibition designers' personal understanding, experience, or even intuition. In fact, it is possible to systematically identify matching music for given artwork due to the crossmodal correspondence between visual and auditory inputs. This paper presents an exploratory study of the crossmodal correspondence between color hues and music tempos and its effects on online art appreciation experience based on two experiments. According to Experiment 1, warm colors were congruent with fast music, whereas cool colors were congruent with slow music. In Experiment 2, congruent and incongruent background music was used in an online exhibition of oil paintings, and several findings were engendered. For warm-colored paintings, congruent (fast) music helped visitors remembered significant more paintings. For cool-colored paintings, visitors collected or shared more paintings when hearing congruent (slow) music but viewed the paintings for longer durations when hearing incongruent (fast) music. The findings not only enrich the understanding of the roles of audio-visual crossmodal correspondence, but also inform online art exhibition designers of how to make background music work in harmony with artwork to better engage visitors.

Keywords: Crossmodal correspondence · Colors · Music tempos · Online art exhibition

1 Introduction

Traditional exhibitions have appealed primarily to just one sense as vision. To further improve public interest in the artifacts displayed, now exhibitions have been integrating and stimulating multiple human senses rather than just a feast for eyes [1]. Many museums and art galleries also provide online access to the contents of their collections for the public to help the collections widespread. Different from physical exhibitions which can integrate the design of visual, auditory, olfactory, and other senses cues to trigger multisensory experiences [2], the sensory design for online exhibitions is restricted due to technical limitations. Visual and auditory stimuli are the most common cues that can

be seen in online exhibitions. Besides, it has been affirmed that visual and auditory cues are the most powerful multisensory cues combination that positively affects user experience and behavior on a digital exhibition website [3]. Generally, collections, e.g., paintings, are the typical visual cues and background music is the typical auditory cue in online exhibitions. Existing studies explored the associations between art collections and background music in terms of their semantic or content, e.g., affective meaning [4, 5]. Besides, formal features, which can be defined independently of content [6], have also found can influence human behavior. Color has the effectiveness in attracting attention and influencing consumers' purchase intentions memory. For example, blue is more effective than red in promotion recall of information [7]. Background music tempo also influences user's behavior, for example, fast music can enhance optimal exercising [8].

When humans interact with multisensory environments, their sensory modalities are related to one another and can influence each other, which is called *crossmodal interaction* psychologically. As a particular class of crossmodal interaction, *crossmodal correspondence* is a common phenomenon in humans that refers to the non-arbitrary perceptual associations between different stimulus features [9, 10]. For example, high lightness is associated with high pitch while low lightness is associated with low pitch. The majority of previous studies about crossmodal correspondence are devoted to revealing the associations between stimuli from different sensory modalities [11, 12]. Only a few studies have explored the influence of crossmodal correspondence on attention, perception, and behavior in a specific context [13–15]. Particularly, the congruence between stimulus features is beneficial to information perception, processing, and memory, and such influence has been explored in the context of products package and retailing [14, 16, 17]. In art, olfactory-color crossmodal has attracted researchers' attention [18], incongruent scents and artworks increase arousal relative to congruent scents and artworks [19]. Unfortunately, to the best of our knowledge, despite background music being widely used in exhibitions, there is little work exploring the role of visual-auditory crossmodal correspondence play on user behavior and memory.

Exploring the relationship between crossmodal correspondence and user behavior and memory in digital art exhibitions could help the design of the websites and art communication with visitors, this study aims to verify the association between color and background music and explores the influence of audiovisual crossmodal correspondence on users' user behavior and memory when they view art paintings. The research questions are:

RQ1. How do visual color and music tempo match?
RQ2. Does visual color-music tempo crossmodal correspondence play a role in user behavior and memory?

To answer these questions, two experiments were conducted. First, it is necessary to investigate whether the correspondence between color and music tempo could be identified. Thus, study 1 investigated whether people reliably matched music tempo with color (hue/saturation/lightness) in China. As the perception of color hue is more intuitive than lightness and saturation in a painting, study 2 further explored whether such association between color hue and background music tempo could affect user behavior and memory in an online art exhibition. A2 (background music tempo: fast/slow) *2

(color hue: warm/cool) mixed-subjects design was used in study 2, while the background music was treated as a between-subjects factor and color hue treated as a within-subject factor. User behavior was measured via two physical interactions, dwell time on painting and user actions with the painting (collecting or sharing), while memory a recognition test.

2 Related Work

2.1 The Effects of Background Music

The influences of background music have been explored by researchers in multiple disciplines. It has been found that music can positively affect consumer purchase behavior [20] and promote second language acquisition [21], but it may also distract students during the learning process [22]. In the exhibitions, music can be a catalyst for the exhibit to convey messages in it and influence user experience and behavior. Music affected the emotion and pace of the museum visit, e.g., relaxing music may induce a slower pace of visit [4, 23].

The impact of background music on tasks depends on the semantic congruency and formal congruency between music and the task [24]. The semantic congruency emphasizes the similarity of content between auditory and visual, which helps to communicate the meaning of the content. In the supermarket, French music would guide consumers to buy French wine, while German music would guide consumers to buy German wine [25]. Background music can generate more visual attention, higher cognitive workload, and withdrawal reactions when it is incongruent with advertisement content than when they are congruent [5]. In addition, the formal congruency provides a united perceptual form to auditory and visual information, which can also influence the responses of users [24]. For example, when users are exposed to online advertise or a website with fast music and warm color background, they would feel more aroused and pleasure than those who experienced slow tempo and cool color [26, 27]. The formal congruency is focused on this study.

2.2 Crossmodal Correspondence Between Color and Music Tempo

One important line of previous studies of crossmodal correspondence has devoted to revealing the associations between stimuli from different sensory modalities. Most of them focused on simple properties of elementary stimuli, such as visual color hue and auditory pitch [12, 28]. Complex stimuli that are more common in everyday life have not received much attention. Self-report methods have been adopted to determine the associations. Specifically, the participants were asked to directly select the most/least consistent stimuli for a given stimulus. With respect to visual stimuli, color is the most widely investigated property. Hue, saturation, and lightness are three basic dimensions of color. Colors have all been found to be associated with the auditory stimuli in related studies, including pitch, loudness, timbre of a sound, and mode and tempo of music [12, 29, 30].

The music-to-color associations are mediated by emotional associations [31, 32]. Fast music is considered more arousing and happier than slow one. Warm color can

increase arousal while cool color tends to induce calmness and peace. Besides, Palmer et al. [31] found that U.S. and Mexican participants tended to associate red/yellow color (warm color) with fast tempo classic music while blue color (cool color) with slow tempo; High lightness/saturation with fast tempo while low lightness/saturation with slow tempo. Qi et al. [28] found that Chinese undergraduate and graduate students matched higher-pitched sounds from Chinese instruments with red rather than black or gray. However, to the best of our knowledge, how do color and tempo associate in Chinese remains unknown.

2.3   Crossmodal Correspondence and User Responses

The crossmodal correspondence between visual and auditory has been found to affect people's responses. As for visual attention, it is shown that congruency between auditory pitch and visual lightness can influence visual search performance by means of top-down facilitation [16]. High-pitch sounds can cause users to fixate on light objects faster and longer while low-pitch sounds cause users to fixate on dark objects faster and longer, which was observed in undergraduates [14]. Besides, high pitch can guide user attention to a higher location [33]. The congruency effect can also be found in infants: they looked at an object whose height or size is congruent with pitch longer than an object whose height or size is incongruent with pitch [34, 35].

Crossmodal correspondence can also affect users' behavior intention and memory. When exposed to high pitch music in a supermarket, consumers were more likely to purchase products from light decorative shelves, while low pitch music was opposite [14]. Metatla et al. [15] examined the role of crossmodal display in gameplay and found that users had higher engagement levels with the congruent display than incongruent. Their results also indicated that congruent display had a positive performance effect on the memory task. Besides, in the context of art, the association between scent and artworks has an impact on undergraduate visitors' perception: incongruence can enhance the level of attention and increase arousal [19].

3   Study 1: Crossmodal Correspondence Between Music Tempo and Color

Study 1 aims to replicate and extend previous findings of crossmodal correspondence between music tempo and color in Chinese participants. A within-subject design was conducted in a classroom at a Chinese university. All the participants were asked to choose five colors to match or mismatch music in order they were listening to.

3.1   Participants

The participants of this experiment were recruited in a Chinese university. It was required that the participants had no defective color vision and or acoustic impairment. A total of 50 participants (28 females and 22 males, Mage = 20.06 years) meeting both the requirements participant in the study 1. All of them are Chinese students but none of them were professional artists or professional composer/theorist. Among them, 27 participants

self-report unfamiliar with music theory, and 23 participants self-report kind familiar with music theory. Each participant signed a consent form before the experiment. This study was approved by the University Research Ethics Committee and strictly followed the general ethical guidelines.

3.2 Materials

Colors. The colors stimuli were chosen based on Palmer et al. [31], which included eight hues (red (R), orange (O), yellow(Y), chartreuse(H), green (G), cyan (C), blue (B), and purple (P)) sampled at four cuts (saturation/lightness levels): saturated (S), light (L), muted (M), and dark (D). Colors were initially sampled from Munsell space, with the highly saturated colors within each hue. Then less-saturated versions of those hues at varying lightness levels were chosen, the M colors being approximately halfway between S colors and neutral gray, the L-colors were approximately halfway between S-colors and white, and the D colors were approximately halfway between S-colors and black. White, black, and three grays whose lightness was approximately the average lightness were also included (see Fig. 1a).

Music. As this study mainly focused on music tempo, the auditory stimuli were two classic music with no Lyrics, chosen from Bach's Brandenburg Concerto, no. 2 and no. 6, both in major mode but different in the tempo: one is fast tempo, and another is slow tempo.
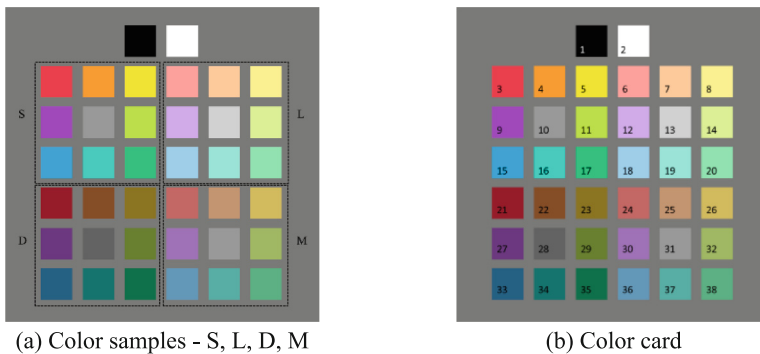

(a) Color samples - S, L, D, M   (b) Color card

Fig. 1. The display of 37 colors

3.3 Procedure

The experiment was conducted in a classroom. All participants were concentrated in the classroom and seated in their places. Color cards were presented in front of them with each color numbered (see Fig. 1b). Then they were told to choose the five colors in order that were most consistent with the music, and the five colors in order that were inconsistent with the music they listened. The two pieces of music (lasting 50 s) were both played twice so that the participants had enough time to make choice.

3.4   Results

All of the color were transformed into CIELAB color space and CIELCH color space. The three coordinates of CIELAB represent the lightness of the color, the position between red and green, and the position between yellow and blue, while one coordinate of CIELCH represent saturation of the color. As for the color chosen with music, the four dimensions (D_d represents the value of each dimension and d represents each dimension: Red/Green, Yellow/Blue, Lightness, and Saturation) were calculated by a linearly weighted average of the ratings chosen with the music: the ratings of the five colors chosen as most consistent with the music (C_d) minus an analogous weighted average of the ratings of the five colors chosen to be inconsistent with that music (I_d) [31].

$$C_d = \left(5c_{1,d} + 4\,c_{2,d} + 3c_{3,d} + 2c_{4,d} + 1c_{5,d}\right)/15 \tag{1}$$

$$I_d = \left(5i_{1,d} + 4i_{2,d} + 3\,i_{3,d} + 2\,i_{4,d} + 1i_{5,d}\right)/15 \tag{2}$$

$$D_d = C_d - I_d \tag{3}$$

where C_j represents the value participants picked as the most consistent with music, where j ranges from 1 to 5, and I_j represents the value participants picked as the most inconsistent with music, where j ranges from 1 to 5.

One-way ANOVAs was used to examine the relationship between the music tempo and the four-color dimensions. According to the results (see Fig. 2), faster tempo was associated with lighter ($F_{(1,98)} = 42.303$, $p = 0.000$), more saturated ($F_{(1,98)} = 100.636$, $p = 0.000$), yellower ($F_{(1,98)} = 21.707$, $p = 0.000$), and redder ($F_{(1,98)} = 151.059$, $p = 0.000$) colors. Slower tempo was associated with darker, less saturated, bluer, and greener colors.

## 4   Study 2: Relationships Between Crossmodal Correspondence and User Responses

The results in study 1 showed the existence of crossmodal correspondence between music tempo and color (hue, lightness, and saturation), which is consistent with existing research [31, 32]. Specifically, fast tempo was more strongly associated with warmer colors (yellower/redder colors) and slow tempo was more associated with cooler colors (bluer/greener colors). As the perception of color hue is more intuitive than lightness and saturation in a painting, study 2 would further test the role of crossmodal correspondence between color hue and music tempo in human responses. Study 2 used a 2 (music tempo: fast/slow) *2 (color: warm/cool) mixed-subjects design, with music tempo treated as a between-subjects factor and color treated as a within-subject factor.
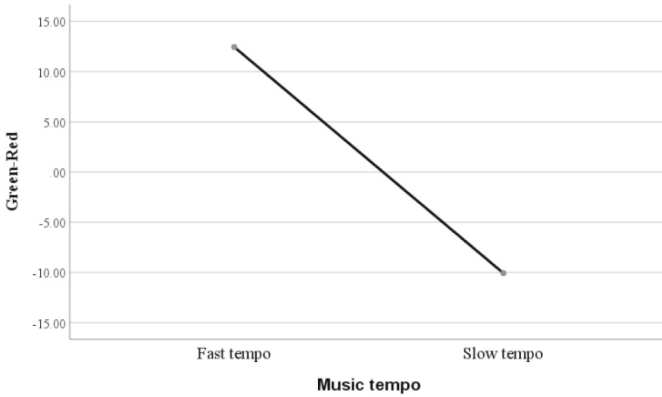
4.1   Participants

A recruitment advertisement that specified the purpose and requirements of the experiment was posted on social media. It was required that the potential participants had
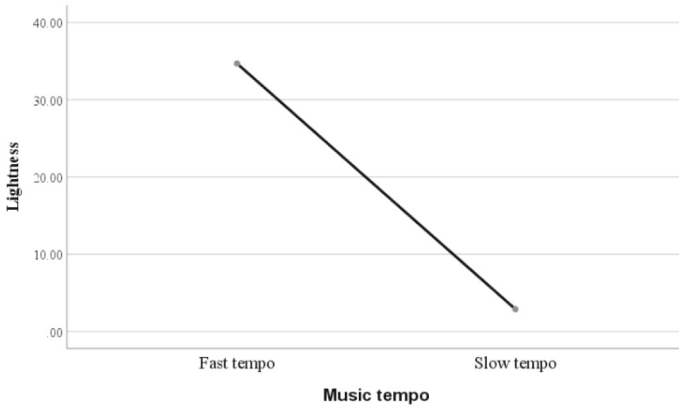
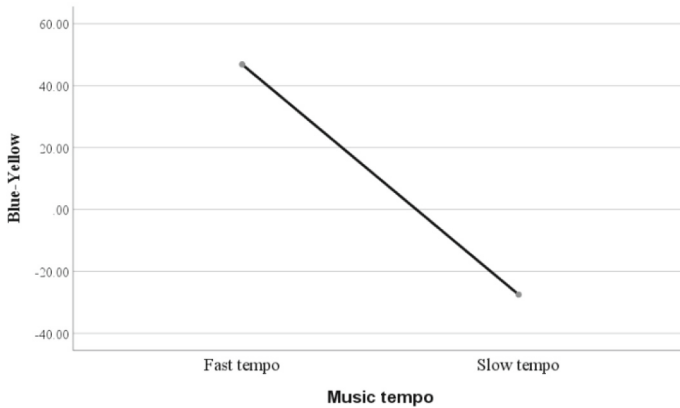(a)   Color lightness and music tempo
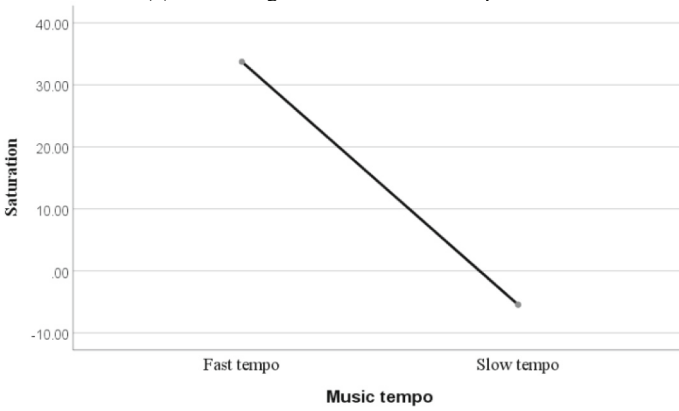
(b)   Color saturation and music tempo

(c)   Red/green and music tempo

Fig. 2. Mean value of color and music tempo

(d)   Yellow/blue and music tempo

Fig. 2. (*continued*)

a normal sense of visual and hearing. In addition, music or art professionals were not asked as they may have different than normal experiences when viewing the paintings, which helped minimize possible individual differences. As a result, a total of 39 Chinese participants (29 females, 10 males, Mage = 22.1 years) met these requirements and then participated in the experiment, all aged between 18 and 30. Among them, 35 participants were occasional visitors of art museums/galleries online or offline and 4 never visit. The familiarity with music theory or oil painting was assessed with the question: "Rate how much you are familiar with music theory" and "Rate how much you are familiar with oil paintings" (1-very unfamiliar, 7-very familiar). According to the results, all participants did not have a broad knowledge of music theory (Mean rank = 2.41) or oil paintings (Mean rank = 2.54). All participants signed a consent form before the experiment and received 15 RMB as compensation after it. This study was approved by the University Research Ethics Committee and strictly followed the general ethical guidelines.

4.2   Materials

This study built a mockup online art exhibition using the prototyping tool Axure. The exhibition consisted of 20 landscape oil paintings, with 10 warm-colored ones (dominated by red or yellow colors) and 10 cool-colored ones (dominated by blue or green colors). These paintings were selected by three professional visual designers based on three criteria: (1) lack of distinction; (2) excluding humans and animals; and (3) roughly the same levels of color lightness or saturation. Each painting was displayed at the center of a webpage and accompanied by textual information indicating the title, author, year, and country of the painting. The length of textual information was controlled to the same to avoid potential influence. Besides, two buttons, *collect* and *share*, were placed below the textual information to enable visitors to interact with the painting. Additionally, the *previous* and *next* buttons under the painting helped visitors navigate in the exhibition.

The warm-colored paintings alternated with the cool-colored ones in the exhibition. Two schemes of sequence were created based on a 2 * 2 Latin Square, one starting with

a warm-colored painting and the other a cool-colored painting. Background music was embedded in the online art exhibition and auto played upon visitors' arrival. The two pieces of music examined in Experiment 1, i.e., Bach's Brandenburg Concerto No. 2 and No. 6, were used in the exhibition as fast and slow auditory materials, respectively.

4.3 Procedure

The experiment was conducted online, and the participants were randomly allocated to the fast (n = 20) or slow (n = 19) tempo background music group. The experiment consisted of three parts. First, all the participants were told that the whole process would be remotely monitored, and screen recorded by the researchers. This permission was obtained from the participants. Then they were instructed to see a landscape oil painting exhibition while hearing background music in a website prototype and they can view paintings, collect, and share the paintings just as the way they behave as normal. After viewing each painting, they needed to evaluate the color perception, ranging from 1 as cool color (blue/green) to 7 as warm color (yellow/red). After viewing all the paintings, they needed to evaluate the tempo of the background music ranging from 1 (slow) to 7 (fast), and how much they like the music ranging from 1 (dislike) to 7 (like). The second part was three minutes distractor task, during which they worked on math problems [36]. The third part was a recognition test. Each participant was presented with a total of 40 paintings and asked to recognize from them 20 target paintings that had appeared in the online art exhibition. The remaining 20 paintings were irrelevant to the exhibition and randomly arranged in the 40 paintings as distractors. For each painting, the participants needed to indicate whether they had viewed the painting during art appreciation. They were given three options, i.e., "yes", "no", and "not sure". The entire experiment lasted 20 min on average.

4.4 Results

This study first checked if the color and music tempo was successfully manipulated. The Mann-Whitney U test was adopted to compare subjective ratings between warm and cool colors, fast and slow music, and the liking degree of the music as those data had no specific distribution [37]. The results showed that participants felt warm-colored paintings are yellower/redder and cool paintings are bluer/greener ($Z = -22.896$, $p = 0.000$, Mean rank: warm $568.38 >$ cool 207.16). The perceived tempo in the fast music group was faster than the slow music group ($Z = -4.818$, $p = 0.000$, Mean rank: fast $28.38 >$ slow 11.18). The degree of liking of music did not differ between the fast and slow tempo groups ($Z = -1.085$, $p = 0.296$). Those results suggested that the manipulation of the experiment was successful.

Secondly, to assess the differences in dwell time, behavior, and retention between warm-colored and cool-colored paintings, paired t test was conducted while the data satisfied parametric assumptions, and Wilcoxon signed rank test was conducted while the data did not satisfy parametric assumptions [38]. One-way ANOVA was used for assessing the differences between fast and slow tempo groups.

Dwell Time. This study defined dwell time as the time spent viewing a painting, an indicator of the user's involvement in interacting with the painting. The mean dwell time on cool-colored paintings (M = 10.27 s) was longer than warm-colored paintings (M = 10.04 s). The fast tempo group had a longer mean dwell time (M = 10.32 s) than the slow group (M = 9.98 s). However, there existed no significant difference in mean dwell time between warm-colored and cool-colored paintings (t = −0.622, p = 0.537), as well as fast and slow tempo ($F(1,37) = 0.04$, P = 0.843). Results further revealed that mean dwell time on warm-colored paintings in fast tempo group did not differ from slow tempo group ($F(1,37) = 0.010$, p = 0.919), either on cool-colored paintings ($F(1,37) = 0.236$, p = 0.630).

Next, the congruency effect was analyzed in terms of two tempo groups. As seen in Fig. 3, participants viewed paintings with incongruent color paintings longer than congruent. The difference was significant in fast tempo group ($M_{congruent} = 9.95$ vs. $M_{incongruent} = 10.7$; t = −2.142, p = 0.045 < 0.05), whereas not in slow tempo group (t = −0.527, p = 0.605). These results suggested that the crossmodal correspondence between color and music tempo can play a role on users' dwell time. Specifically, when the background music was fast, the participants spent more time on cool-colored paintings than warm-colored paintings.
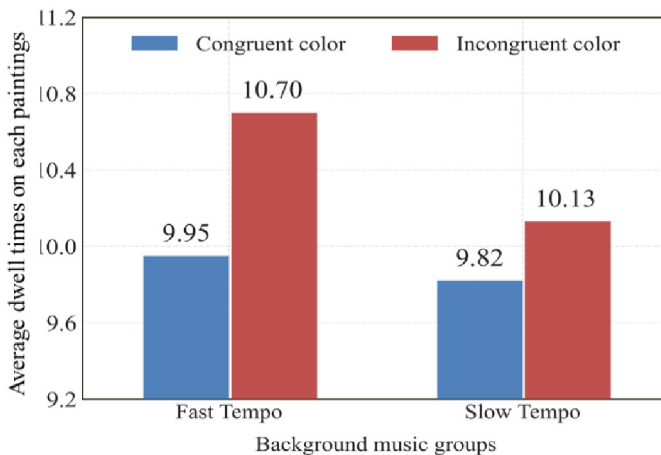

Fig. 3. Average dwell time on each painting

Collecting and Sharing. When viewing the paintings, users may collect or share, indicating that they approach the paintings more intensively. The number of collects or shares by each participant was extracted. No significant differences were found between fast tempo and slow tempo groups ($F (1,37) = 0.480$, P = 0.493). Cool-colored paintings attracted more intense approach clicks than warm-colored paintings ($M_{warm} = 2$ vs. $M_{cool} = 2.90$; Z = −2.310, p = 0.021 < 0.05). Results further revealed that clicks on warm-colored paintings in fast tempo group did not differ from the slow tempo group

$(F (1,37) = 0.733, p = 0.397)$, either on cool-colored paintings $(F (1,37) = 0.166, p = 0.686)$.

The congruency effect was analyzed in terms of two tempo groups. As seen in Fig. 4, participants collected or shared more paintings with incongruent color than congruent with fast tempo condition, while more paintings with congruent color than incongruent in slow tempo condition. The differences were significant in slow tempo group ($M_{congruent}$ = 2.74 vs. $M_{incongruent}$ = 1.8; t = 2.150, p = 0.045 < 0.05) but not in fast tempo group ($Z = -1.256$, p = 0.209). That is, when the background music was slow, the participants collected or shared more cool-colored paintings than warm-colored paintings.
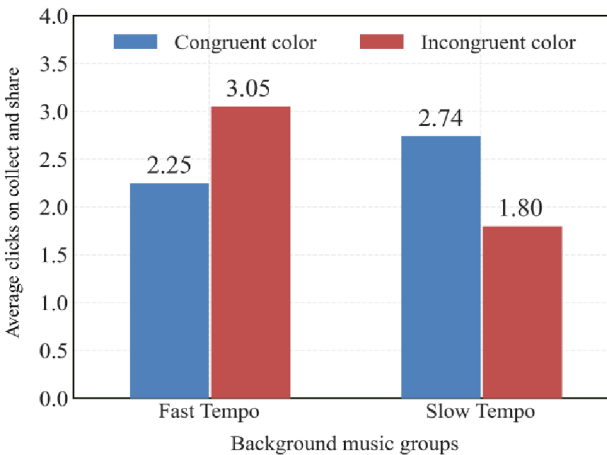

Fig. 4. Average clicks on collect and share in one visiting

Memory. Memory can be revealed by a recognition memory test [39]. Participants were considered to remember the painting if they correctly identified both the target painting and the distractor painting. The results showed no significant differences between fast tempo and slow tempo groups ($F(1,37) = 0.108$, P = 0.744). Warm-colored paintings were remembered more than cool-colored paintings (M_{warm} = 5.62 vs. M_{cool} = 4.67; $Z = -2.769$, p = 0.006 < 0.05). Results further revealed that the number of memories of warm-colored paintings in the fast tempo group did not differ from the slow tempo group ($F(1,37) = 0.344$, p = 0.561), either of cool-colored paintings ($F(1,37) = 0.003$, p = 0.959).

The congruency effect was analyzed in terms of two tempo groups. As seen in Fig. 5, the number of correct memories on paintings with congruent color was more than with incongruent. The difference was significant in fast tempo group ($M_{congruent}$ = 5.85 vs. $M_{incongruent}$ = 4.65; t = 2.812, p = 0.011) but not in slow tempo group (t = −1.556, p = 0.137). That is, when the background music was fast, the participants would remember more warm-colored paintings than cool-colored ones.
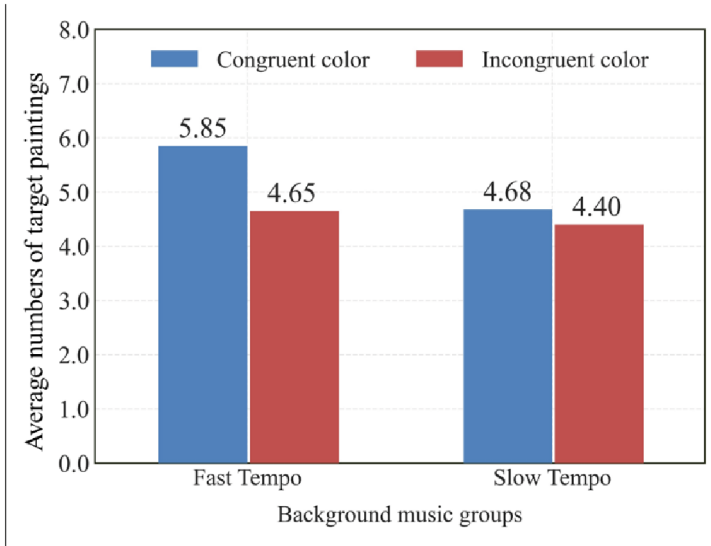
Fig. 5. Average numbers of target paintings recognized

5 Discussion

5.1 Crossmodal Correspondence in Online Art Exhibition

Background music is commonly used in art galleries, museums, and so on. However, how visual-auditory senses are interactively matched, and the consequences of such interaction remain unknown. This study was particularly interested in the crossmodal correspondence between color of paintings and background music and the relationship between them and user responses in an online art exhibition.

Different from the positive congruency effect on viewing times found in [34], this study found that the dwell time was longer on incongruent color paintings than congruent, but only in the fast music environment. When consumers are in online retail stores, warm color is considered to result in high arousal and pleasure than cool color, and fast music is associated with a higher level of arousal and pleasure. The match between warm color and fast music can enhance customers' emotions and create a desired environment [27]. However, the perception of art might be different from the evaluation of objects and it might be more complex than in a retailing setting: consumers' judgment may be influenced mainly by rational factor in retailing setting [40] while art exhibitions is an aesthetic experience [41]. A similar study that focused on visual-olfactory crossmodal correspondence found the negative influence of congruence on attention in an art exhibition because the arousal diminished and incongruence would enhance the level of attention [19].In addition, several studies have shown that faster music can speed up customers' activity [42], this study further found that if the fast music color does not consistent with the painting, users' activity may slow down. Compared to paintings with congruent music, paintings with incongruent music would force users from no conscious control (thinking fast) to conscious control (thinking slow) [43], but this incongruent

effect maybe only last a short time as the difference was small between congruent and incongruent color paintings, less than 1 s in this study.

The congruency effect on clicking collect and share was obvious in this study. Both behaviors indicated that users approach the painting more intensively. The results showed that users were more likely to collect and share cool-colored paintings than warm-colored paintings. This difference was significant in the slow music group but not in the fast music group, which indicated that fast music may weaken this relationship because of the incongruence between fast music and cool color.

The congruency effect also has an impact on users' memory, which is consistent with previous findings [44]. This study found that warm-colored paintings were remembered more than cool-colored paintings. This effect was still significant in fast music but not in slow music. That is, slow music may weaken this relationship due to the mismatch between slow music and warm-colored paintings and promote more memory on cool-colored paintings. Different from visual-olfactory crossmodal correspondence that scent destroyed the experience and memory in the art gallery [19], fast music would enhance the memory of warm-colored paintings.

5.2 Implications

This study presents two theoretical contributions to the literature on crossmodal correspondence and the interaction of visual and auditory cues. First, this study confirmed the robustness of crossmodal correspondence between color and music tempo: warm color and fast tempo were crossmodally matched; cool color and slow tempo were crossmodally matched. Second, this study adds understanding in the multisensory in online art exhibitions by examining the role of crossmodal correspondence between the color hue of paintings and music tempo play on users' behavior and memory formation.

The major findings of this study also have practical implications for multisensory design in online art exhibitions. Visual and auditory stimuli are common cues in digital art exhibitions. An effective multisensory design would help the artworks widely spread. The current findings indicate that the creation of a sensory between painting's color hue and background music tempo may enhance user behavior and memory when they visit digital art exhibitions. It is worth noting that the balance between congruence and incongruence needs to be taken into account. For example, if an online art exhibition designer wants users to attend to paintings, cool-colored paintings should be displayed in fast music environment. But if the designer wants users to collect or share more paintings, cool-colored paintings should be displayed in slow background music. In addition, audio-visual technology has been applied in the majority of current VR/AR [45]. Our results also provide implications for Virtual Reality (VR) and Augmented Reality (AR) design to improve user experience. Multisensory stimuli integrated in VR/AR have made it possible to develop an immersive environment [46]. In the future, the crossmodal correspondence between other sensory stimuli can be explored to help further understand users' experience in VR/AR.

6 Conclusions

This study investigated the association between visual color and music tempo and how such sensory correspondence contributes to user behavior and memory. The results in Study 1 showed that warm colors were reliably matched with fast music and cool colors with slow music. In addition, study 1 also found the association between high lightness/saturation and fast tempo, as well as low lightness/saturation and slow tempo. In terms of the relationship between crossmodal correspondence and user responses, several significant results were found in study 2: In the fast music environment, users spend more time on paintings with incongruent color than that with congruent color, but they tend to remember more paintings with congruent color than that with incongruent color. In the slow music environment, users would collect or share more paintings with congruent color than that incongruent color. This research revealed findings that the balance between congruence and incongruence should be considered. This calls for more research towards the potential mediators between crossmodal correspondence and user responses. In addition, our future work will investigate the effects of crossmodal correspondence between other sensory features, such as lightness/saturation and music tempo, on user responses. Besides, some potential mediators, such as subjective experiences, arousal, pleasure, and immersion can also be investigated in future works to further explore the influence mechanism.

References

1. Vi, C.T., Ablart, D., Gatti, E., et al.: Not just seeing, but also feeling art: mid-air haptic experiences integrated in a multisensory art exhibition. Int. J. Hum Comput Stud. **108**, 1–14 (2017)
2. Spence, C.: Scenting the anosmic cube: on the use of ambient scent in the context of the art gallery or museum. i-Perception **11**(6), 1–26 (2020)
3. Guo, K., Fan, A., Lehto, X., et al.: Immersive digital tourism: the role of multisensory cues in digital museum experiences. J. Hosp. Tour. Res., 1–23 (2021)
4. Chen, C.-L., Tsai, C.-G.: The influence of background music on the visitor museum experience: a case study of the Laiho memorial museum. Taiwan. Visit. Stud. **18**(2), 183–195 (2015)
5. Ausín, J.M., Bigne, E., Marín, J., et al.: The background music-content congruence of TV advertisements: a neurophysiological study. Eur. Res. Manag. Bus. Econ. **27**(2), 100154 (2021)
6. Kim, K.-H., Iwamiya, S.-I.: Formal congruency between telop patterns and sound effects. Music Percept. **25**(5), 429–448 (2008)
7. Yu, C.-E., Xie, S.Y., Wen, J.: Coloring the destination: The role of color psychology on Instagram. Tour. Manag. **80**, 104110 (2020)
8. Edworthy, J., Waring, H.: The effects of music tempo and loudness level on treadmill exercise. Ergonomics **49**(15), 1597–1610 (2006)

9. Spence, C.: Crossmodal correspondences: a tutorial review. Atten. Percept. Psychophys. **73**(4), 971–995 (2011)

10. Lin, A., Scheller, M., Feng, F., et al.: Feeling colours: crossmodal correspondences between tangible 3D objects, colours and emotions. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (2021)

11. Sun, X., Li, X., Ji, L., et al.: An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. PeerJ **6**, e4443 (2018)

12. Jonas, C., Spiller, M.J., Hibbard, P.: Summation of visual attributes in auditory–visual cross-modal correspondences. Psychon. Bull. Rev. **24**(4), 1104–1112 (2017). https://doi.org/10.3758/s13423-016-1215-2

13. Motoki, K., Velasco, C.: Taste-shape correspondences in context. Food Qual. Prefer. **88**, 104082 (2021)

14. Hagtvedt, H., Brasel, S.A.: Cross-modal communication: sound frequency influences consumer responses to color lightness. J. Mark. Res. **53**(4), 551–562 (2016)

15. Metatla, O., Correia, N.N., Martin, F., et al.: Tap the shapetones: Exploring the effects of cross-modal congruence in an audio-visual interface. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, New York, NY. ACM Press (2016)

16. Klapetek, A., Ngo, M.K., Spence, C.: Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? Atten. Percept. Psychophys. **74**(6), 1154–1167 (2012). https://doi.org/10.3758/s13414-012-0317-9

17. Velasco, C., Woods, A.T., Petit, O., et al.: Crossmodal correspondences between taste and shape, and their implications for product packaging: a review. Food Qual. Prefer. **52**, 17–26 (2016)

18. Spence, C.: Olfactory-colour crossmodal correspondences in art, science, and design. Cognit. Res. Princ. Implic. **5**(1), 1–21 (2020). https://doi.org/10.1186/s41235-020-00246-1

19. Cirrincione, A., Estes, Z., Carù, A.: The effect of ambient scent on the experience of art: not as good as it smells. Psychol. Mark. **31**(8), 615–627 (2014)

20. Andersson, P.K., Kristensson, P., Wästlund, E., et al.: Let the music play or not: the influence of background music on consumer behavior. J. Retail. Consum. Serv. **19**(6), 553–560 (2012)

21. Kang, H.J., Williamson, V.J.: Background music can aid second language learning. Psychol. Music **42**(5), 728–747 (2014)

22. Flowers, P.J., O'Neill, A.A.M.: Self-reported distractions of middle school students in listening to music and prose. J. Res. Music Educ. **53**(4), 308–321 (2005)

23. Webb, R.C.: Music, mood, and museums: a review of the consumer literature on background music. Visit. Stud. **8**(2), 15–29 (1996)

24. Tan, S.-L., Cohen, A.J., Lipscomb, S.D., et al.: The Psychology of Music in Multimedia. Oxford University Press, Oxford (2013)

25. North, A.C., Hargreaves, D.J., McKendrick, J.: The influence of in-store music on wine selections. J. Appl. Psychol. **84**(2), 271 (1999)

26. Wu, C.-S., Cheng, F.-F., Yen, D.C.: The atmospheric factors of online storefront environment design: an empirical experiment in Taiwan. Inf. Manag. **45**(7), 493–498 (2008)

27. Cheng, F.-F., Wu, C.-S., Yen, D.C.: The effect of online store atmosphere on consumer's emotional responses–an experimental study of music and colour. Behav. Inf. Technol. **28**(4), 323–334 (2009)

28. Qi, Y., Huang, F., Li, Z., et al.: Crossmodal correspondences in the sounds of Chinese instruments. Perception **49**(1), 81–97 (2020)

29. Anikin, A., Johansson, N.: Implicit associations between individual properties of color and sound. Atten. Percept. Psychophys. **81**(3), 764–777 (2018). https://doi.org/10.3758/s13414-018-01639-7

30. Adeli, M., Rouat, J., Molotchnikoff, S.: Audiovisual correspondence between musical timbre and visual shapes. Front. Hum. Neurosci. **8**, 352 (2014)

31. Palmer, S.E., Schloss, K.B., Xu, Z., et al.: Music–color associations are mediated by emotion. Proc. Natl. Acad. Sci. **110**(22), 8836–8841 (2013)
32. Whiteford, K.L., Schloss, K.B., Helwig, N.E., et al.: Color, music, and emotion: bach to the blues. i-Perception **9**(6), 1–27 (2018)
33. Chiou, R., Rich, A.N.: Cross-modality correspondence between pitch and spatial location modulates attentional orienting. Perception **41**(3), 339–353 (2012)
34. Walker, P., Bremner, J.G., Mason, U., et al.: Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. Psychol. Sci. **21**(1), 21–25 (2010)
35. Fernández-Prieto, I., Navarra, J., Pons, F.: How big is this sound? Crossmodal association between pitch and size in infants. Infant Behav. Dev. **38**, 77–81 (2015)
36. Bookbinder, S., Brainerd, C.: Emotionally negative pictures enhance gist memory. Emotion **17**(1), 102 (2017)
37. McKnight, P.E., Najab, J.: Mann-Whitney U test. The Corsini encyclopedia of psychology, 1 (2010)
38. McCrum-Gardner, E.: Which is the correct statistical test to use? Br. J. Oral Maxillofac. Surg. **46**(1), 38–41 (2008)
39. Wichmann, F.A., Sharpe, L.T., Gegenfurtner, K.R.: The contributions of color to recognition memory for natural scenes. J. Exp. Psychol. Learn. Mem. Cognit. **28**(3), 509 (2002)
40. Lee, K., Joshi, K.: Examining the use of status quo bias perspective in IS research: need for re-conceptualizing and incorporating biases. Inf. Syst. J. **27**(6), 733–752 (2017)
41. Bedford, L.: The Art of Museum Exhibitions: How Story and Imagination Create Aesthetic Experiences. Routledge (2016)
42. North, A.C., Hargreaves, D.J., Krause, A.E.: Music and Consumer Behaviour. Oxford University Press, Oxford (2009)
43. Daniel, K.: Thinking, Fast and Slow. Macmillan, New York (2011)
44. North, A.C., Sheridan, L.P., Areni, C.S.: Music congruity effects on product memory, perception, and choice. J. Retail. **92**(1), 83–95 (2016)
45. Ranasinghe, N., Jain, P., Thi Ngoc Tram, N., et al.: Season traveller: multisensory narration for enhancing the virtual reality experience. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (2018)
46. Halabi, O., Saleh, M.: Augmented reality flavor: cross-modal mapping across gustation, olfaction, and vision. Multimedia Tools Appl. **80**(30), 36423–36441 (2021). https://doi.org/10.1007/s11042-021-11321-0