frontiers | Research Topics

Crossmodal correspondence

Edited by
Na Chen, Thomas Alrik Sørensen
and Charles Spence

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-4701-4
DOI 10.3389/978-2-8325-4701-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Crossmodal correspondence

Topic editors
Na Chen – Bar-Ilan University, Israel
Thomas Alrik Sørensen – Aalborg University, Denmark
Charles Spence – University of Oxford, United Kingdom

Citation
Chen, N., Sørensen, T. A., Spence, C., eds. (2024). *Crossmodal correspondence*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4701-4

Table of contents

TYPE Editorial
PUBLISHED 21 March 2024
DOI 10.3389/fpsyg.2024.1385480

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Snehlata Jaswal,
Sikkim University, India

*CORRESPONDENCE
Na Chen
imminana7@gmail.com

RECEIVED 12 February 2024
ACCEPTED 11 March 2024
PUBLISHED 21 March 2024

CITATION
Chen N, Sørensen TA and Spence C (2024) Editorial: Crossmodal correspondence. *Front. Psychol.* 15:1385480. doi: 10.3389/fpsyg.2024.1385480

Editorial: Crossmodal correspondence

Na Chen [1]*, Thomas Alrik Sørensen [2] and Charles Spence [3]

[1]The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel, [2]Department of Communication and Psychology, Aalborg University, Aalborg, Denmark, [3]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

KEYWORDS

crossmodal correspondences, sound-shape, music-taste, development, color-taste

Editorial on the Research Topic
Crossmodal correspondence

Over the last decade or so, there has been an explosion of scientific research interest in the crossmodal correspondences, this the name given to the often surprising, yet consensual, associations that have been documented between a growing number of basic sensory features, attributes, and dimensions in different sensory modalities (see Spence, 2011, for a review). So, for example, people have been shown to match auditory pitch with visual elevation, size, lightness, etc. Over the years, several different mechanisms have been put forward to help explain the existence of such crossmodal correspondences, including the statistical account, the structural (or neurophysiological) account, a semantic/lexical account, and an account in terms of emotional-mediation. It is, however, important to note that these various explanations should not be treated as mutually exclusive. Indeed, several or perhaps all of them may help to explain a variable proportion of the various different correspondences that have been documented to date.

The 13 papers that were eventually accepted for this Research Topic effectively serve to highlight the global growth of research interest in this emerging phenomenon currently. What is also striking is how the articles collected together here range well-beyond the pitch-based audiovisual correspondences that attracted so much of the interest in correspondences research previously (see Spence and Sathian, 2020, for a review). Recently, researchers have increasingly started to investigate group differences in crossmodal correspondences as well as studying when in human development humans start to express a sensitivity to such crossmodal correspondences (see Spence, 2022, for a review). The latter approach is illustrated in this Research Topic by an exploratory study reported by Meng et al. in which crossmodal correspondences between visual features (such as shape/angularity and color) and tastes (e.g., bitter, sweet, sour, salty) were assessed in a group of pre-schoolers. Importantly, several important factors that have been suggested to affect/constrain the crossmodal correspondences include their context-dependence (Motoki and Velasco, 2021), their automaticity (Spence and Deroy, 2013b; Getz and Kubovy, 2018), and their bidirectionality (Motoki et al., 2023; Yang et al., 2023; Chen and Huang). These topics are addressed in several of the contributions here. Chen and Huang showed that the sound-shape correspondences are not completely automatic, but their modulation was bidirectionally symmetrical once it occurred. The bidirectionality of crossmodal correspondences means that effects on visual perception can be elicited by stimuli from other sensory modalities, such as gustatory and/or olfactory stimuli, as highlighted by a couple of the intriguing submissions in this Research Topic (e.g., Ward et al.; Yang et al., 2023). Ward et al. used an achromatic adjustment task to show that the presence of odors modulates color perception. Such crossmodal effects on vision are surprising inasmuch as vision is so often found to be the dominant sense in multisensory research (Hutmacher, 2019).

"Sonic seasoning", the generic name for the modification of the taste and flavor based on listening to music that corresponds crossmodally to the dominant tastes/flavors of food and drink that have been documented between sound and the chemical senses also continues to attract much research interest. The experimental paper from Xu et al. investigates the role of self-construal priming on the effectiveness of sonic seasoning. Meanwhile, the paper from Mesz et al. investigates emotional associations with music-based crossmodal correspondences.

Historically, the role of crossmodal correspondences in helping to solve the crossmodal binding problem has occupied the attention of many researchers (Chen and Spence, 2017). At the same time, the paper by Yang et al. (2023) demonstrates that the crossmodal correspondence between color and taste influences performance on the well-established Stroop task. However, what is striking about so many of the articles that are collected together here is how they attempt to apply our growing understanding of the crossmodal correspondences to a range of real-world applications: this includes everything from a consideration of the potential facilitatory role of designing human augmentation systems based on the crossmodal correspondences (see Pinardi et al.), through to the use of crossmodal correspondences to help describe, communicate about, and market wine (Crichton-Fock, Spence, Mora, et al.; Crichton-Fock, Spence, and Pettersson). At the same time, the crossmodal correspondences clearly also help to provide guidelines for the design of multisensory experiential events (see Velasco and Spence, 2022, for a review). Meanwhile, the article from Ogata et al. reports some intriguing results concerning the impact of the shape of chocolate on taste ratings, based on the literature on shape-taste crossmodal correspondences (see Spence, 2014).

One other area of continuing research interest concerns the relationship between crossmodal correspondences, synaesthesia, and mental imagery (Rader and Tellegen, 1987; Martino and Marks, 2001; Spence and Deroy, 2013a; Nanay, 2020). Relevant here, the paper by Hitsuwari and Nomura details research validating a Japanese version of the Plymouth Sensory Imagery Questionnaire which provides researchers with a means of assessing the strength of mental imagery in each of the senses.

Extending the scope of crossmodal correspondences research, this Research Topic also includes a couple of papers that might best be classified as sensory-conceptual/categorical correspondence (Chen et al.) and intramodal visual correspondences (Zelazny et al.). The first of these two papers demonstrates that the color red biases sex categorization of human bodies. One theme that emerges from the latter study, as well as from several other studies that have been published recently (e.g., Velasco et al., 2023) highlights the importance of providing participants with a wide enough range of options if one's goal is to identify the strongest correspondences, given that older studies with a narrow range of colors, say, may merely have picked up the best color amongst the range of options provided to the participant. Such methodological developments should help ensure that future theorizing about the correspondences is based on firm empirical foundations.

Taken together, the research papers that have been gathered together in this Research Topic clearly highlight the vibrant state of crossmodal correspondences research in both the theoretical and applied arenas. One exciting area in correspondences research that is not represented here relates to the emergence of studies assessing the sensitivity of various animals to crossmodal correspondences. So, for example, Loconsole et al. (2021, 2022) have recently published several studies demonstrating audiovisual crossmodal correspondences in both chicks and the tortoise (*Testudo hermanni*; Loconsole et al., 2023).

Author contributions

NC: Writing – review & editing. TS: Writing – review & editing. CS: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Chen, Y.-C., and Spence, C. (2017). Assessing the role of the 'unity assumption' on multisensory integration: a review. *Front. Psychol.* 8:445. doi: 10.3389/fpsyg.2017.00445

Getz, L. M., and Kubovy, M. (2018). Questioning the automaticity of audiovisual correspondences. *Cognition* 175, 101–108. doi: 10.1016/j.cognition.2018.02.015

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Front. Psychol.* 10:2246. doi: 10.3389/fpsyg.2019.02246

Loconsole, M., Gasparini, A., and Regolin, L. (2022). Pitch–luminance crossmodal correspondence in the baby chick: an investigation on predisposed and learned processes. *Vision* 6:24. doi: 10.3390/vision6020024

Loconsole, M., Pasculli, M. S., and Regolin, L. (2021). Space-luminance crossmodal correspondences in domestic chicks. *Vision Res.* 188, 26–31. doi: 10.1016/j.visres.2021.07.001

Loconsole, M., Stancher, G., and Versace, E. (2023). Crossmodal association between visual and acoustic cues in a tortoise (*Testudo hermanni*). *Biol. Lett.* 19:20230265. doi: 10.1098/rsbl.2023.0265

Martino, G., and Marks, L. E. (2001). Synesthesia: strong and weak. *Curr. Dir. Psychol. Sci.* 10, 61–65. doi: 10.1111/1467-8721.00116

Motoki, K., Marks, L. E., and Velasco, C. (2023). Reflections on cross-modal correspondences: current understanding and issues for future research. *Multisens. Res.* 37, 1–23. doi: 10.1163/22134808-bja10114

Motoki, K., and Velasco, C. (2021). Taste-shape correspondences in context. *Food Qual. Prefer.* 88:104082. doi: 10.1016/j.foodqual.2020.104082

Nanay, B. (2020). Synesthesia as (multimodal) mental imagery. *Multisens. Res.* 34, 281–296. doi: 10.1163/22134808-bja10027

Rader, C. M., and Tellegen, A. (1987). An investigation of synesthesia. *J. Pers. Soc. Psychol.* 52, 981–987. doi: 10.1037/0022-3514.52.5.981

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attent. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2014). Assessing the influence of shape and sound symbolism on the consumer's response to chocolate. *New Food* 17, 59–62.

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079

Spence, C., and Deroy, O. (2013a). "Crossmodal mental imagery," in *Multisensory Imagery: Theory and Applications*, eds S. Lacey, and R. Lawson (New York, NY: Springer), 157–183.

Spence, C., and Deroy, O. (2013b). How automatic are crossmodal correspondences? *Conscious. Cogn.* 22, 245–260. doi: 10.1016/j.concog.2012.12.006

Spence, C., and Sathian, K. (2020). "Audiovisual crossmodal correspondences: behavioural consequences and neural underpinnings," in *Multisensory Perception: From Laboratory to Clinic*, eds K. Sathian, and V. S. Ramachandran (San Diego, CA: Elsevier), 239–258.

Velasco, C., Barbosa Escobar, F., Spence, C., and Olier, J. S. (2023). The taste of colours revisited. *Food Qual. Prefer.* 112:105009. doi: 10.1016/j.foodqual.2023.105009

Velasco, C., and Spence, C. (2022). "Capitalizing on the crossmodal correspondences between audition and olfaction in the design of multisensory experiences," in *Experiential Marketing in an Era of Hyper-Connectivity*, eds N. Pomirleanu, B. J. Mariadoss, and J. Schibrowsky (Newcastle upon Tyne: Cambridge Scholars Publishing), 85–113.

Yang, Y., Chen, N., Kobayashi, M., and Watanabe, K. (2023). Color-taste correspondence tested by the Stroop task. *Front. Psychol.* 15:1250781. doi: 10.3389/fpsyg.2024.1250781

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 03 April 2023
DOI 10.3389/fpsyg.2023.1041202

Check for updates

OPEN ACCESS

EDITED BY
Na Chen,
Bar-Ilan University, Israel

REVIEWED BY
Steve Keller,
Studio Resonate-SXM Media, United States
Nicola Di Stefano,
National Research Council (CNR), Italy
Bruno Mesz,
National University of Tres de Febrero, Argentina

*CORRESPONDENCE
Jianping Huang
✉ jphuang@suda.edu.cn
Hui Xu
✉ xuhui900830@163.com

[†]These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology

RECEIVED 10 September 2022
ACCEPTED 02 March 2023
PUBLISHED 03 April 2023

CITATION
Xu J, Guo X, Liu M, Xu H and Huang J (2023) Self-construal priming modulates sonic seasoning. *Front. Psychol.* 14:1041202. doi: 10.3389/fpsyg.2023.1041202

Self-construal priming modulates sonic seasoning

Jingxian Xu[1†], Xiyu Guo[1†], Mengying Liu[1], Hui Xu[2*] and Jianping Huang[1*]

[1]Department of Psychology, Soochow University, Suzhou, China, [2]School of Public Affairs, Zhejiang University, Hangzhou, China

Introduction: "Sonic seasoning" is when music influences the real taste experiences of consumers. "Self-construal" is how individuals perceive, understand, and interpret themselves. Numerous studies have shown that independent and interdependent self-construal priming can affect a person's cognition and behavior; however, their moderating effect on the sonic seasoning effect remains unclear.

Methods: This experiment was a 2 (self-construal priming: independent self-construal or interdependent self-construal) × 2 (chocolate: milk chocolate or dark chocolate) × 2 (emotional music: positive emotional music or negative emotional music) mixed design, and explored the moderating role of self-construal priming and the effect of emotional music on taste by comparing participants' evaluations of chocolates while listening to positive or negative music after different levels of self-construal priming.

Results: After initiating independent self-construal, participants increased their ratings of milk chocolate sweetness when listening to music that elicited positive emotions, $t(32) = 3.11$, $p = 0.004$, Cohen's $d = 0.54$, 95% CI = [0.33, 1.61]. In contrast, interdependent self-construal priming led participants to perceive dark chocolate as sweeter when they heard positive music, $t(29) = 3.63$, $p = 0.001$, Cohen's $d = 0.66$, 95%CI = [0.44, 1.56].

Discussion: This study provides evidence for improving people's individual eating experience and enjoyment of food.

KEYWORDS

self-construal priming, emotional music, sensory marketing, sonic seasoning, tasting experience

Introduction

Researchers have been increasingly interested in sonic seasoning in recent decades (Knöferle and Spence, 2012; Spence, 2017; Spence et al., 2019; Spence and Di Stefano, 2022). This term refers to the deliberate matching of music with flavor to enhance the multisensory tasting experience (Sedacca, 2016). For instance, music has been composed for specific flavors such as "sweet," "bitter," or "salty" based on the flavors of the food (Crisinel and Spence, 2010). Wang et al. (2017) indicated that spiciness was associated with auditory attributes of high pitch, fast tempo, and high levels of distortion. Wang et al. (2021) further explored the acoustical/musical attributes that best match saltiness and found that auditory attributes based on emotional associations (negative valence, minor mode, and high arousal), long decay durations, regular rhythm, and a high degree of auditory roughness were associated most strongly with saltiness. These findings on sonic seasoning have been used in marketing-led activations (Spence et al., 2021) and further understanding of how music influences consumers' perceptions of food. Crisinel et al. (2012) demonstrated for the first time that individuals evaluate food differently based on the music they hear while tasting it. Participants were asked to rate the taste of several bittersweet toffees while listening to sweet or bitter music. It was found that when they listened to the sweet soundtrack, the sweetness of the toffee was evaluated higher than when listening to the bitter soundtrack. Wang et al. (2020) added that chocolate ratings could be affected only when participants listened to sweet or bitter music before or during a chocolate tasting.

Other previous literature also affirms that when music is playing in the background while people eat, the emotions that individuals experience from or associated with the music are transferred to the tasting experience itself, where these emotions are considered to act as an important mediator during sonic seasoning demonstrations (Kantono et al., 2016, 2019; Reinoso-Carvalho et al., 2019, 2020). For instance, North (2012) showed that emotional connotations associated with background music influenced consumers' taste perception of wine. In addition, Kantono et al. (2016) showed that music evoked positive emotions (satisfaction, happiness, and amusement) that influenced the perceived sweetness of gelati, while music evoked negative emotions (contempt, disappointment, and disgust) that influenced the perceived bitterness of gelati. Reinoso-Carvalho et al. (2019) discovered that compared to music associated with positive emotion when participants listened to music associated with negative emotion, they rated the same beer as more bitter, with higher alcohol content, and were willing to pay more for the beer. The music that evokes positive emotions is often performed with high intensity, a brighter timbre, staccato articulation, and a rapid tempo. Conversely, music that evokes negative emotions is often performed with low intensity, a duller timbre, legato articulation, and a slow tempo (Quinto and Thompson, 2013).

Research showed that auditory information like music could impact the subjective evaluation of the taste of food (Spence and Di Stefano, 2022). For example, Reinoso-Carvalho et al. (2020) discovered that emotional music (i.e., music that evokes positive or negative emotions), as opposed to crossmodal music (i.e., soft or hard music; the former means imaging the music is consistent with the smooth food texture, the latter refers to imaging the music is consistent with the rough food texture), had a more prominent effect on food flavor and purchase intention. Specifically, participants rated the chocolate as sweeter, and the purchase intention was higher when music judged to elicit positive valence was played compared to music that was judged to elicit negative valence. However, the same chocolate was judged to be more bitter when exposed to music that conveyed negative rather than positive emotions. The researchers proposed that the sweeter (bitter) taste of chocolate induced by positive (negative) music could be explained by a theory similar to the "attention-shifting or redirection effect" (Johnson and Proctor, 2004). That is music that transmits positive or negative emotions can allow participants to transfer their feelings to the flavor perception of chocolate, which in turn affects their purchase intention. An analogous conclusion was proposed in the study by Ziv (2018) that participants generally considered cookies to taste better when listening to pleasant music.

As described above, people's perception of food is affected by contextual factors. Moreover, a large number of studies have suggested that individuals from different cultural backgrounds are influenced by context to different degrees (Wan et al., 2014; Jeong and Lee, 2021), which may be attributed to differences in cognitive style (Henrich, 2014). As one of the important cultural characteristics, self-construal is often used to explain cultural differences in human behavior, cognition, and emotion. It is primarily about how individuals perceive, understand, and interpret themselves. Markus and Kitayama (1991) suggested that independent self-construal is more predominant in Western culture, which conceptualizes the self as an autonomous and bonded entity, emphasizing self-independence and uniqueness. By contrast, interdependent self-construal is more predominant in East Asian culture, which conceptualizes the self as interconnecting and overlapping with others, emphasizing the importance of living in harmony with other groups and individuals. In general, interdependent self-construal facilitates the spontaneous association of a focal object with context compared to independent self-construal (Goh et al., 2007). Moreover, interdependent self-construal consumers believe that a higher price means a higher quality than independent self-construal consumers because the former are more susceptible to perceiving the association between product elements (Lalwani and Shavitt, 2013).

While self-construal can be judged by scales (Singelis, 1994) or inferred based on an individual's ethnicity (Van Baaren et al., 2003), researchers prefer to manipulate self-construal through priming methods when studying it in a laboratory (Kühnen and Oyserman, 2002; Reinoso-Carvalho et al., 2020). In particular, participants were asked to search first-person pronouns such as "I" or "mine" in a story to represent independent self-construal, while "we" or "our" imply interdependent self-construal. Considering the container that holds the food as a background, Huang et al. (2021) suggested that self-construal priming could modulate the influence of receptacles on food perception. Compared to independent self-construal priming, interdependent self-construal priming elicited a greater influence of the size of the plate on participants' willingness to pay (WTP) for noodles and the pleasantness ratings of the noodles served on the red plate. Therefore, considering emotional music as a contextual factor, it is reasonable to expect that self-construal priming could moderate the effect of emotional music on food perception.

This study investigates whether self-construal can modulate the effects of emotional music on food perception. Owing to how interdependent self-construal can improve the connection between the target and the contextual background (Masuda and Nisbett, 2001; Goh et al., 2007) and influence object processing that is dependent on context (Kühnen and Oyserman, 2002), we hypothesized that the participants' perception of food flavor would not be influenced by the background music under independent self-construal priming. Conversely, different emotional music would affect participants' evaluation of food after initiating the interdependent self-construal.

Methods

Participants

To estimate the sample size, we used the Easypower package (McGarvey, 2015) in R.4.2.0 (R Core Team, 2022) to conduct an *a priori* power analysis. According to the mixed design of 2 (self-construal priming: independent self-construal or interdependent self-construal) × 2 (chocolate: milk chocolate or dark chocolate) × 2 (emotional music: positive emotional music or negative emotional music). The sample size should not be <48 with an effect size of the quadruple interaction of 0.25, a statistical power of 0.95, and an alpha of 0.05. Considering the balance between different experimental conditions, a total of 141 college students from Soochow University took part in the experiment. However, we had to exclude the data of 13 participants from analyses because their answers to all the questions were the same. Therefore, 128 valid data were received (M_{age} = 20.12 years, SD_{age} = 0.16 years, ranging from 18 to 26 years; 35 males). All participants were right-handed and had normal or corrected eyesight, without color blindness or color weakness. All participants were paid 10 Chinese yuan after completing the experiment.

Apparatus and materials

The experiment used a 27-inch monitor with a resolution of 2,560 × 1,440 pixels and a refresh rate of 60 Hz. The online questionnaire was based on the Qualtrics platform (https://www.qualtrics.com), including basic information (e.g., sex, age, and degree of hunger) and the priming of self-construal, which refers to the study by Sui and Han (2007). We measured the degree of hunger based on previous literature that focused on food and/or beverages (Biswas et al., 2021; Moss and McSweeney, 2021) because hunger could potentially influence taste perceptions (Hanci and Altun, 2016). Participants were asked to carefully read a travel story that was presented randomly and then click on all the personal pronouns in the story (Grossmann and Jowhari, 2018).

In line with Reinoso-Carvalho et al. (2020), our study used milk chocolate (Callebaut N. 823, containing milk and at least 33.6% cocoa solids) and dark chocolate (Callebaut N. 811, containing no milk and at least 54.5% cocoa solids), as well as positive and negative emotional music from https://tinyurl.com/music-emotions-xcultural. Specifically, the negative emotional music sample could evoke more negative valence, and the positive emotional music sample could evoke more positive valence. The chocolate rating task consisted of the following four questions (7-point Likert scale): "What do you think of the sweetness of the chocolate?" "What do you think of the bitterness of the chocolate?" "How much do you like the chocolate in the experiment?" and "How likely are you to purchase the chocolate?" In addition, they were required to answer how much they would be willing to pay for a bar of chocolate in RMB (Renminbi, Chinese Yuan). The Chinese revision of the Positive and Negative Affect Scale (PANAS; Watson et al., 1988) developed by Qiu et al. (2008) was applied to determine participants' emotions. Specifically, participants had to rate on a 5-point scale positive (e.g., gratitude, energetic, active, cheerful, excited, enthusiastic, proud, happy, and joyful) or negative emotional words (e.g., scared, afraid, guilty, ashamed, nervous, irritable, angry, jittery, and sad) that appeared randomly. Scores from 1 to 5 represent very slightly or not at all, a little, moderately, quite a bit, and extremely, respectively.

An online pretest was conducted to ensure that the emotional music samples were effective (the negative emotional music sample evoked more negative valence, and the positive emotional music sample evoked more positive valence). A total of 34 participants (eight men and 26 women) between the ages of 19 and 25 years (M = 21.26 years, SD = 1.58) were recruited. Participants were required to listen to the emotional music sample and complete the Chinese PANAS mentioned earlier. The number of participants who listened to the positive music emotional sample is equal to the number of participants who listened to the negative emotional music sample. The results showed that participants rated significantly higher on the Positive (M = 2.30, SD = 0.87) than Negative Affect Scale (M = 1.29, SD = 0.35) after listening to the positive emotional music sample, $t_{(16)} = 4.34$, $p = 0.001$, Cohen's $d = 1.05$, 95% CI = [0.51, 1.50]. As expected, participants rated significantly higher on the Negative (M = 2.66, SD = 0.47) than Positive Affect Scale (M = 1.63, SD = 0.47) after listening to the negative emotional music sample, $t_{(16)} = 6.52$, $p < 0.001$, Cohen's $d = 1.58$, 95% CI = [0.69, 1.36]. These results indicated that the positive emotional music sample we used in this experiment could evoke more positive valence while the negative emotional music sample could evoke significantly more negative valence.

Design

This experiment was a 2 (self-construal priming: independent self-construal or interdependent self-construal) × 2 (chocolate: milk chocolate or dark chocolate) × 2 (emotional music: positive emotional music or negative emotional music) mixed design. Self-construal priming and chocolate were between-subject variables, and emotional music was a within-subject variable. The type of chocolate was randomly matched with the two pieces of music. The dependent variables were the rating of the sweetness, bitterness, liking, WTP, and purchase intention of the chocolate that they ate. Each participant ate no more than two pieces of chocolate twice.

Procedure

Participants were first asked to fill in personal information and complete the self-construal priming task. Then they had to complete the PANAS as baseline emotion. After that, participants rinsed their mouths with water and put on headphones. It should be noted that although the duration of positive and negative music was inconsistent (69 vs. 60 s), the volume of the two kinds of music was controlled at 70 ± 6 dB. The experiment requested participants to taste a bar of chocolate on a plate we had prepared while listening to a random type of music played through headphones. Meanwhile, they were allowed to continue savoring a second of the same chocolate if they had finished tasting the first piece before the end of the music. Next, participants were required to evaluate the chocolate they ate in the five dimensions mentioned above and accomplish the PANAS again. After rinsing their mouths again with water, participants completed a similar task of matching another type of music with another type of chocolate. All the analysis data are publicly available from the Open Science Framework repository (OSF) at https://osf.io/qhdgz/.

Results

Before formal analysis, we first conducted a manipulation check on the emotional scores of the participants after listening to positive

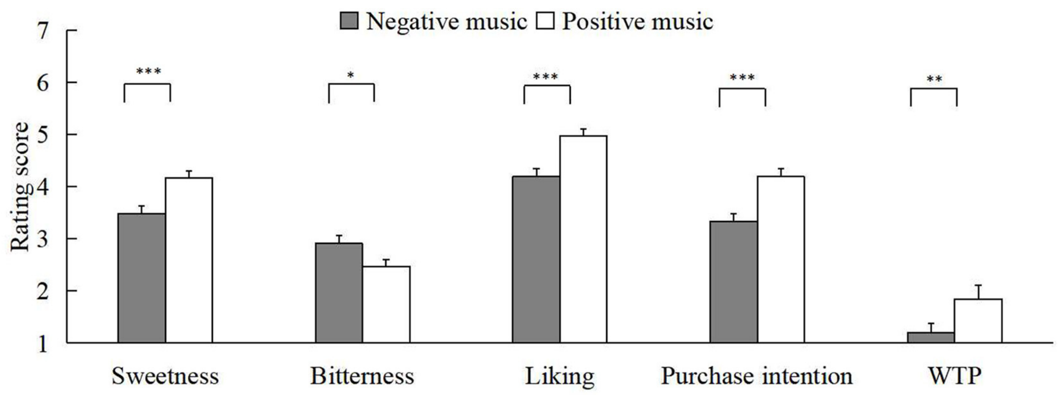

FIGURE 1
The main effect of emotional music on the five dimensions. Note that error bars show the standard errors of the means, *$p < 0.05$, **$p < 0.01$. and ***$p < 0.001$.

and negative emotional music. As expected, music that conveyed positive emotions induced significantly higher positive (M = 2.25, SD = 0.83) than negative feelings (M = 1.15, SD = 0.33), $t_{(127)} = 14.46$, $p < 0.001$, Cohen's $d = 1.28$, 95% CI = [0.95, 1.25]. Surprisingly, the participants also scored significantly higher on the Positive (M = 1.86, SD = 0.77) than Negative Affect Scale (M = 1.44, SD = 0.51) after listening to negative emotional music, $t_{(127)} = 4.74$, $p < 0.001$, Cohen's $d = 0.42$, 95% CI = [0.24, 0.59]. We further analyzed the emotional scores of the participants after listening to positive and negative emotional music. The results showed that positive emotional music (M = 2.25, SD = 0.83) evoked significantly more positive valence than negative emotional music (M = 1.86, SD = 0.77), $t_{(127)} = 6.02$, $p < 0.001$, Cohen's $d = 0.53$, 95% CI = [0.27, 0.53], while negative emotional music (M = 1.44, SD = 0.51) scored significantly higher negative valence than positive emotional music (M = 1.15, SD = 0.33), $t_{(127)} = 6.50$, $p < 0.001$, Cohen's $d = 0.57$, 95% CI = [0.20, 0.38].

We performed a 2 (self-construal priming: independent self-construal or interdependent self-construal) × 2 (chocolate: milk chocolate or dark chocolate) × 2 (emotional music: positive emotional music or negative emotional music) repeated-measures ANOVA on valid data. The results found that the main effect of emotional music was significant in all five aspects of chocolate evaluation [$F_{\text{sweetness}(1,124)} = 21.92$, $p < 0.001$, $\eta_p^2 = 0.15$; $F_{\text{bitterness}(1,124)} = 6.91$, $p = 0.010$, $\eta_p^2 = 0.05$; $F_{\text{liking}(1,124)} = 21.62$, $p < 0.001$, $\eta_p^2 = 0.15$; $F_{\text{purchaseintentions}(1,124)} = 24.97$, $p < 0.001$, $\eta_p^2 = 0.17$; $F_{\text{WTP}(1,124)} = 10.43$, $p = 0.002$, $\eta_p^2 = 0.08$]. As shown in Figure 1, compared to negative music, the participants rated the chocolate as having higher sweetness ($M_1 = 3.48$, $SD_1 = 1.68$; $M_2 = 4.16$, $SD_2 = 1.54$), liking ($M_1 = 4.19$, $SD_1 = 1.72$; $M_2 = 4.97$, $SD_2 = 1.43$), and lower bitterness ($M_1 = 2.91$, $SD_1 = 1.80$; $M_2 = 2.47$, $SD_2 = 1.56$) when the emotional music was positive. In addition, the purchase intentions ($M_1 = 4.19$, $SD_1 = 1.67$; $M_2 = 3.34$, $SD_2 = 1.63$) and WTP ($M_1 = 1.84$, $SD_1 = 2.97$; $M_2 = 1.20$, $SD_2 = 1.93$) for chocolate were further intensive when participants were exposed to positive rather than negative music. The experimental results also revealed that the type of chocolate had a significant main effect on the sweetness [$F_{(1,124)} = 33.18$, $p < 0.001$, $\eta_p^2 = 0.21$], bitterness [$F_{(1,124)} = 36.46$, $p < 0.001$, $\eta_p^2 = 0.23$], liking [$F_{(1,124)} = 17.45$, $p < 0.001$, $\eta_p^2 = 0.12$], and purchase intention [$F_{(1,124)} = 13.16$, $p < 0.001$, $\eta_p^2 = 0.10$] of chocolate (as shown in Figure 2). Specifically, participants considered milk chocolate to be sweeter ($M_{\text{dark}} = 3.16$, $SD_{\text{dark}} = 1.13$; $M_{\text{milk}} = 4.43$, $SD_{\text{milk}} = 1.32$) and less bitter ($M_{\text{dark}} = 3.37$, $SD_{\text{dark}} = 1.19$; $M_{\text{milk}} = 2.07$, $SD_{\text{milk}} = 1.24$) than dark chocolate. Then, they preferred milk chocolate ($M_{\text{dark}} = 4.11$, $SD_{\text{dark}} = 1.09$; $M_{\text{milk}} = 5.00$, $SD_{\text{milk}} = 1.27$) and more purchase intentions for it ($M_{\text{dark}} = 3.33$, $SD_{\text{dark}} = 1.15$; $M_{\text{milk}} = 4.16$, $SD_{\text{milk}} = 1.39$) compared to dark chocolate.

The results demonstrated that there was a significant triple interaction among emotional music, self-construal priming, and chocolate on the sweetness of chocolate, $F_{(1,124)} = 4.15$, $p = 0.04$, $\eta_p^2 = 0.03$. Based on this, we further analyzed the effects of self-construal priming and chocolate type on the evaluation of chocolate sweetness under two music conditions. As shown in Figure 3, positive music improved the sweetness of milk chocolate (M = 4.97, SD = 1.53) more than negative music (M = 4.00, SD = 1.75) when independent self-construal was primed, $t_{(32)} = 3.11$, $p = 0.004$, Cohen's $d = 0.54$, 95% CI = [0.33, 1.61]. However, music with positive emotions increased the sweetness of dark chocolate (M = 3.53, SD = 1.57) more than music with negative emotions (M = 2.53, SD = 1.25) under the priming of interdependent self-construal, $t_{(29)} = 3.63$, $p = 0.001$, Cohen's $d = 0.66$, 95% CI = [0.44, 1.56].

Discussion

This study aimed to explore the moderating role of different types of self-construal in the process of emotional music affecting food evaluation. The experimental results mainly suggested the following three points. First, positive music can improve participants' ratings of chocolate and purchase intention. Consistent with previous studies (Ziv, 2018; Kantono et al., 2019; Reinoso-Carvalho et al., 2019, 2020), our study elucidated that participants find chocolate to be sweeter after listening to positive music, which in turn increases their liking and purchase intention for it. Skaczkowski et al. (2016) argued that positive (negative) emotions evoked by positive (negative) music could be transferred

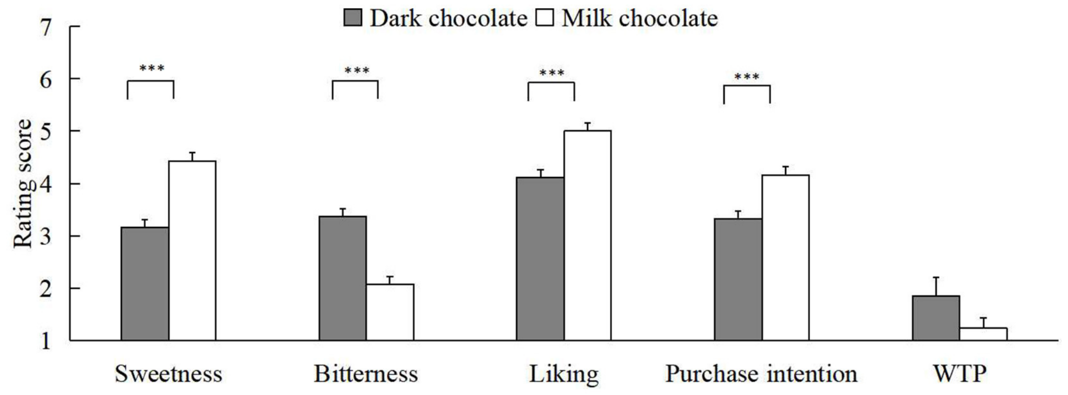

FIGURE 2
The main effect of chocolate is on the five dimensions. Note that error bars show the standard errors of the means, ***$p < 0.001$.

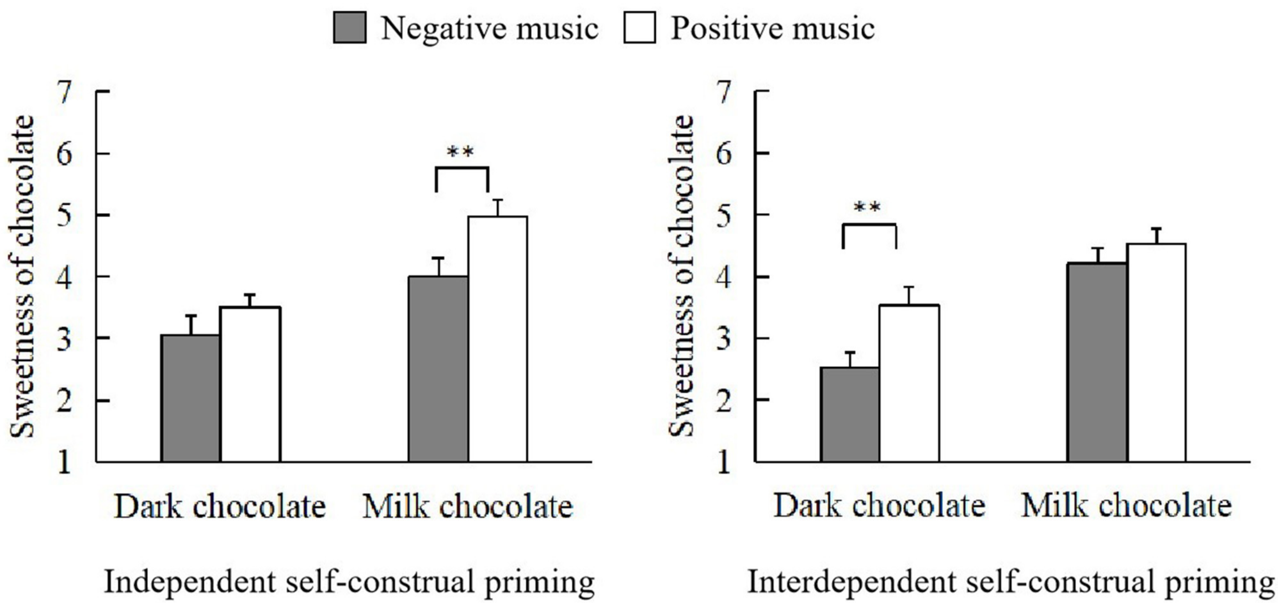

FIGURE 3
Triple interaction effect among emotional music, self-construal priming, and chocolate on the sweetness of chocolate. Note that error bars show the standard errors of the means, **$p < 0.01$.

to individuals' authentic tasting experience based on sensation transference effects.

Second, our results demonstrated that self-construal priming could modulate sonic seasoning. In particular, music with a positive emotion was more likely to enhance the participants' evaluation of the sweetness of milk chocolate under the priming of independent self-construal and dark chocolate under the priming of interdependent self-construal. However, independent self-construal priming did not influence participants' evaluation of dark chocolate sweetness. As researchers have suggested before, individuals with independent self-construal tend to adopt an analytic thinking style that emphasizes the independence of individual objects, whereas individuals with interdependent self-construal tend to adopt a holistic style of thinking emphasizing that the world is composed of interrelated elements (Nisbett et al., 2001; Monga and John, 2007, 2008). This may explain why different self-construal priming types have different influences on sonic seasoning. The self-construal of interdependence can guide the participants to adopt a holistic way of thinking to perceive the relationship between the background music and dark chocolate, which leads to the influence of positive music on the evaluation of dark chocolate's sweetness. However, an analysis of milk chocolate showed that the evaluation of the sweetness of milk chocolate was not affected by emotional music under the initiation of the self-construal of mutual dependence. The possible reason is that milk chocolate itself has been perceived as sweet, so there is no significant difference in the evaluation of sweetness between positive and negative music, even though the participants were under the priming of interdependent self-construal.

Previous studies have focused on the sonic seasoning of food by individuals from different countries (Reinoso-Carvalho et al., 2019, 2020) or on explaining the influence of self-construal priming on food perception (Hansen, 2019; Huang et al., 2021). The innovation of this study is to make a thorough inquiry into how self-construal priming impacts the relationship between music-evoked emotion and the perception of taste. Theoretically, by comparing independent and interdependent self-construal priming on food evaluations, the present study reveals the moderating role of self-construal priming types on the effect of emotional music on real food tasting, which could enrich the findings of sonic seasoning. For the marketing field, our study provides a certain theoretical reference for improving individual diet experience and boosting happiness in eating. In particular, marketers can control consumers' perception of flavors in food or beverage products by playing different styles of music according to their cultural background. In addition, it is a bright prospect for marketers to adopt words that represent different self-construals on the packaging of food to change consumers' cognition.

Like other studies, this study has unavoidable limitations. First, after listening to the negative music, participants did not report a higher negative valence than positive valence. There were several possible explanations. First, a previous study found that compared to an audio-only presentation, an audio-visual congruent presentation, which represented congruent audio and visual emotions (e.g., happy face and happy music), could lead to a more intense emotional response. This effect occurred in both positive and negative music, and the effect was larger for positive music (Pan et al., 2019). The baseline emotional score in our study showed that participants were significantly more positive before the experiment. Thus, it was possible that the negative emotional music was not capable of evoking a strong negative emotional response to make negative emotion dominant for the participants. Moreover, people listened to music while eating chocolate and then completed the PANAS, which might result in emotional scores that not only represent the valence evoked by the music but also the taste of chocolate. As a result, future experiments should measure participants' positive or negative valence to the music itself using the GEMS (Zentner et al., 2008) or GEMIAC questionnaire (Coutinho and Scherer, 2017) and evaluate the music with respect to different taste categories. In the current study, we focused on the effects of music that can induce positive or negative emotions on food perception. Although we can see the effects of the music on the five dimensions and between independent and interdependent self-construal priming we cannot gauge the extent to which these effects are significant because there was no control (i.e., a non-music condition) for comparison. As a result, future experiments could add a non-music condition.

Second, the participants in this study were university students aged ∼20 and were Chinese. We are not sure whether the findings of the study apply to people of other ages and cultures. Reinoso-Carvalho et al. (2020) found that the sonic seasoning effects were different on people from LATAM and Asia. Moreover, previous studies have found that sex differences in self-construal vary across different nations (Costa et al., 2001; Schmitt et al., 2008). Future experiments could expand the sample to different cultures and ages.

Third, this study showed that self-construal could modulate the effects of emotional music on food perception. Previous research discovered that crossmodal congruency between music and food/beverage had effects on the tasting experience (Wang and Spence, 2015; Reinoso-Carvalho et al., 2020; Motoki et al., 2022). For example, Wang and Spence (2015) demonstrated that the music chosen to be congruent with each wine was indeed rated a better match than other pieces of music, and the music significantly influenced the perceived acidity and fruitiness of the wine. Future experiments could investigate whether self-construal could modulate the effects of crossmodally congruent music on food perception.

Finally, Mawad et al. (2015) found that field-independent and field-dependent tendencies affected the visual attention processing of yogurt label information. Individuals with an independent tendency would pay more attention to the health label information of yogurt. It is well-known that individuals with field-independent tendencies are more detached from their surroundings and less susceptible to external cues than field-dependent individuals. Likewise, individuals who have experienced independent self-construal were also more focused on themselves. Therefore, we can further explore the influence of music on the process of an individual's attention to food and whether self-construal still modulates this process in future studies. Moreover, Peng-Li et al. (2020) discovered through eye tracking that sweet music improved gaze time to sweet food, while salty music increased their fixation time on salty food. Notably, Peng-Li et al. (2020) demonstrated the influence of different music on food with different flavors based on the cross-modality association of music and flavor. In light of this, follow-up studies should further examine the moderating role of self-construal in the effect of emotional music on an individual's attention to sweet or bitter foods.

Conclusion

The results of this study suggest that self-construal priming has a moderating role in how emotional music affects the authentic flavor experience. For example, under independent self-construal priming, there was no difference in the evaluation of the sweetness of dark chocolate between positive and negative music. However, positive music significantly improved participants' evaluation of the sweetness of milk chocolate. In contrast, interdependent self-construal priming obtained the converse result. Compared with negative music, positive music can enhance the sweetness of dark chocolate, but neither kind of music influences the sweetness of milk chocolate. This study provides evidence that could improve an individual's eating experience by promoting happiness when eating.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by Human Research Ethics Committee of the Department of Psychology of Soochow University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JX: methodology, investigation, data curation, and formal analysis. XG: formal analysis, methodology, writing, and original draft preparation. ML: writing, review, and editing. HX: supervision, writing, review, and editing. JH: conceptualization, supervision, writing, review, and editing. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China, awarded to JH (Grant No. 32100875).

Acknowledgments

The authors would like to thank Mengting Hu for their assistance in data collection. Comments concerning this article should be sent to JH at jphuang@suda.edu.cn.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1041202/full#supplementary-material

References

Biswas, D., Labrecque, L. I., and Lehmann, D. R. (2021). Effects of sequential sensory cues on food taste perception: Cross-modal interplay between visual and olfactory stimuli. *J. Consum. Psychol.* 31, 746–764. doi: 10.1002/jcpy.1231

Costa, Jr. P. T., Terracciano, A., and McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *J. Personal. Soc. Psychol.* 81, 322. doi: 10.1037/0022-3514.81.2.322

Coutinho, E., and Scherer, K. R. (2017). Introducing the GEneva Music-Induced Affect Checklist (GEMIAC): A brief instrument for the rapid assessment of musically induced emotions. *Music Percept.* 34, 371–386. doi: 10.1525/mp.2017.34.4.371

Crisinel, A., and Spence, C. (2010). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attent. Percept. Psychophys.* 72, 1994–2002. doi: 10.3758/APP.72.7.1994

Crisinel, A.-S., Cosser, S., King, S., Jones, R., Petrie, J., and Spence, C. (2012). A bittersweet symphony: Systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Qual. Pref.* 24, 201–204. doi: 10.1016/j.foodqual.2011.08.009

Goh, J. O., Chee, M. W., Tan, J. C., Venkatraman, V., Hebrank, A., Leshikar, E. D., et al. (2007). Age and culture modulate object processing and object—Scene binding in the ventral visual area. *Cogn. Affect. Behav. Neurosci.* 7, 44–52. doi: 10.3758/CABN.7.1.44

Grossmann, I., and Jowhari, N. (2018). Cognition and the self: Attempt of an independent close replication of the effects of self-construal priming on spatial memory recall. *J. Exp. Soc. Psychol.* 74, 65–73. doi: 10.1016/j.jesp.2017.08.005

Hanci, D., and Altun, H. (2016). Hunger state affects both olfactory abilities and gustatory sensitivity. *Eur. Archiv. Oto-Rhino-Laryngol.* 273, 1637–1641. doi: 10.1007/s00405-015-3589-6

Hansen, J. (2019). Construal level and cross-sensory influences: High-level construal increases the effect of color on drink perception. *J. Exp. Psychol.* 148, 890. doi: 10.1037/xge0000548

Henrich, J. (2014). Rice, psychology, and innovation. *Science* 344, 593–594. doi: 10.1126/science.1253815

Huang, J., Wang, C., and Wan, X. (2021). Self-construal priming modulates the influence of receptacles on food perception. *J. Sensory Stud.* 36, e12642. doi: 10.1111/joss.12642

Jeong, S., and Lee, J. (2021). Effects of cultural background on consumer perception and acceptability of foods and drinks: A review of latest cross-cultural studies. *Curr. Opin. Food Sci.* 42, 248–256. doi: 10.1016/j.cofs.2021.07.004

Johnson, A., and Proctor, R. W. (2004). *Attention: Theory and Practice.* Thousand Oaks, CA: Sage.

Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Skiredj, S., and Carr, B. T. (2019). Emotional and electrophysiological measures correlate to flavour perception in the presence of music. *Physiol. Behav.* 199, 154–164. doi: 10.1016/j.physbeh.2018.11.012

Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J., Carr, B. T., and Grazioli, G. (2016). The effect of background music on food pleasantness ratings. *Psychol. Music* 44, 1111–1125. doi: 10.1177/0305735615613149

Knöferle, K., and Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychon. Bullet. Rev.* 2012, 1–15. doi: 10.3758/s13423-012-0321-z

Kühnen, U., and Oyserman, D. (2002). Thinking about the self influences thinking in general: Cognitive consequences of salient self-concept. *J. Exp. Soc. Psychol.* 38, 492–499. doi: 10.1016/S0022-1031(02)00011-2

Lalwani, A. K., and Shavitt, S. (2013). You get what you pay for? Self-construal influences price-quality judgments. *J. Consumer Res.* 40, 255–267. doi: 10.1086/670034

Markus, H. R., and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

Masuda, T., and Nisbett, R. E. (2001). Attending holistically vs. analytically: Comparing the context sensitivity of Japanese and Americans. *J. Personal. Soc. Psychol.* 81, 922–934. doi: 10.1037/0022-3514.81.5.922

Mawad, F., Trías, M., Giménez, A., Maiche, A., and Ares, G. (2015). Influence of cognitive style on information processing and selection of yogurt labels: Insights from an eye-tracking study. *Food Res. Int.* 74, 1–9. doi: 10.1016/j.foodres.2015.04.023

McGarvey, A. (2015). *easypower: Sample Size Estimation for Experimental Designs.* R package version, 1.0.1.

Monga, A. B., and John, D. R. (2007). Cultural differences in brand extension evaluation: The influence of analytic versus holistic thinking. *J. Consumer Res.* 33, 529–536. doi: 10.1086/510227

Monga, A. B., and John, D. R. (2008). When does negative brand publicity hurt? The moderating influence of analytic vs. holistic thinking. *J. Consumer Psychol.* 18, 320–332. doi: 10.1016/j.jcps.2008.09.009

Moss, R., and McSweeney, M. B. (2021). Do consumers want seaweed in their food? A study evaluating emotional responses to foods containing seaweed. *Foods* 10, 2737. doi: 10.3390/foods10112737

Motoki, K., Takahashi, N., Velasco, C., and Spence, C. (2022). Is classical music sweeter than jazz? Crossmodal influences of background music and taste/flavour on healthy and indulgent food preferences. *Food Qual. Pref.* 96, 104380. doi: 10.1016/j.foodqual.2021.104380

Nisbett, R. E., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychol. Rev.* 108, 291. doi: 10.1037/0033-295X.108.2.291

North, A. C. (2012). The effect of background music on the taste of wine. *Br. J. Psychol.* 103, 293–301. doi: 10.1111/j.2044-8295.2011.02072.x

Pan, F., Zhang, L., Ou, Y., and Zhang, X. (2019). The audio-visual integration effect on music emotion: Behavioral and physiological evidence. *PLoS ONE* 14, e0217040. doi: 10.1371/journal.pone.0217040

Peng-Li, D., Byrne, D. V., Chan, R. C., and Wang, Q. J. (2020). The influence of taste-congruent soundtracks on visual attention and food choice: A cross-cultural eye-tracking study in Chinese and Danish consumers. *Food Qual. Pref.* 85, 103962. doi: 10.1016/j.foodqual.2020.103962

Qiu, L., Zheng, X., and Wang, Y. F. (2008). Revision of the positive affect and negative affect scale. *Chin. J. Appl. Psychol.* 14, 249–254.

Quinto, L., and Thompson, W. F. (2013). Composers and performers have different capacities to manipulate arousal and valence. *Psychomusicology* 23, 137–150. doi: 10.1037/a,0034775

R Core Team (2022). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available online at: https://www.r-project.org/

Reinoso-Carvalho, F., Dakduk, S., Wagemans, J., and Spence, C. (2019). Not just another pint! Measuring the influence of the emotion induced by music on the consumer's tasting experience. *Multisens. Res.* 32, 367–400. doi: 10.1163/22134808-20191374

Reinoso-Carvalho, F., Gunn, L., Molina, G., Narumi, T., Spence, C., Suzuki, Y., et al. (2020). A sprinkle of emotions vs. a pinch of crossmodality: Towards globally meaningful sonic seasoning strategies for enhanced multisensory tasting experiences. *J. Bus. Res.* 117, 389–399. doi: 10.1016/j.jbusres.2020.04.055

Schmitt, D. P., Realo, A., Voracek, M., and Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *J. Personal. Soc. Psychol.* 94, 168. doi: 10.1037/0022-3514.94.1.168

Sedacca, M. (2016). *Sonic Seasoning" is the Growing Scientific Field That Uses Sound To Make Food Taste Better.* Quartz. Available online at: https://qz.com/871605/sonic-season-changing-taste-with-sound/ (accessed September 5, 2021).

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personal. Soc. Psychol. Bullet.* 20, 580–591. doi: 10.1177/0146167294205014

Skaczkowski, G., Durkin, S., Kashima, Y., and Wakefield, M. (2016). The effect of packaging, branding and labeling on the experience of unhealthy food and drink: A review. *Appetite* 99, 219–234. doi: 10.1016/j.appet.2016.01.022

Spence, C. (2017). "Sonic seasoning," in *Audio Branding: Using Sound to Build Your Brand*, eds L. Minsky and C. Fahey (London: Kogan Page), 52–58.

Spence, C., and Di Stefano, N. (2022). Crossmodal harmony: Looking for the meaning of harmony beyond hearing. *i-Perception* 13, 20416695211073817. doi: 10.1177/20416695211073817

Spence, C., Reinoso-Carvalho, F., Velasco, C., and Wang, Q. J. (2019). Extrinsic auditory contributions to food perception and consumer behaviour: An interdisciplinary review. *Multisens. Res.* 32, 275–318. doi: 10.1163/22134808-20191403

Spence, C., Wang, Q. J., Reinoso-Carvalho, F., and Keller, S. (2021). Commercializing sonic seasoning in multisensory offline experiential events and online tasting experiences. *Front. Psychol.* 12, 740354. doi: 10.3389/fpsyg.2021.740354

Sui, J., and Han, S. (2007). Self-construal priming modulates neural substrates of self-awareness. *Psychol. Sci.* 18, 861–866. doi: 10.1111/j.1467-9280.2007.01992.x

Van Baaren, R. B., Maddux, W. W., Chartrand, T. L., De Bouter, C., and Van Knippenberg, A. (2003). It takes two to mimic: Behavioral consequences of self-construals. *J. Personal. Soc. Psychol.* 84, 1093. doi: 10.1037/0022-3514.84.5.1093

Wan, X., Velasco, C., Michel, C., Mu, B., Woods, A. T., and Spence, C. (2014). Does the type of receptacle influence the crossmodal association between colour and flavour? A cross-cultural comparison. *Flavour* 3, 1–7. doi: 10.1186/2044-7248-3-3

Wang, Q., and Spence, C. (2015). Assessing the effect of musical congruency on wine tasting in a live performance setting. *i-Perception* 6, 2041669515593027. doi: 10.1177/2041669515593027

Wang, Q. J., Keller, S., and Spence, C. (2017). Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack. *Food Qual. Pref.* 58, 1–9. doi: 10.1016/j.foodqual.2016.12.014

Wang, Q. J., Keller, S., and Spence, C. (2021). Metacognition and crossmodal correspondences between auditory attributes and saltiness in a large sample study. *Multisens. Res.* 34, 785–805. doi: 10.1163/22134808-bja10055

Wang, Q. J., Spence, C., and Knoeferle, K. (2020). Timing is everything: Onset timing moderates the crossmodal influence of background sound on taste perception. *J. Exp. Psychol.* 46, 1118–1126. doi: 10.1037/xhp0000820

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Personal. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion* 8, 494–521. doi: 10.1037/1528-3542.8.4.494

Ziv, N. (2018). Musical flavor: The effect of background music and presentation order on taste. *Eur. J. Market.* 52, 1485–1504. doi: 10.1108/EJM-07-2017-0427

TYPE Original Research
PUBLISHED 02 June 2023
DOI 10.3389/fpsyg.2023.1129903

OPEN ACCESS

EDITED BY
Kazuhiko Yokosawa,
Tsukuba Gakuin University, Japan

REVIEWED BY
Magdalena Szubielska,
The John Paul II Catholic University of Lublin, Poland
Romke Rouw,
University of Amsterdam, Netherlands

*CORRESPONDENCE
Aurore Zelazny
✉ aurore@ikp.aau.dk

RECEIVED 22 December 2022
ACCEPTED 17 April 2023
PUBLISHED 02 June 2023

CITATION
Zelazny A, Liu X and Sørensen TA (2023) Shape–color associations in an unrestricted color choice paradigm.
Front. Psychol. 14:1129903.
doi: 10.3389/fpsyg.2023.1129903

Shape–color associations in an unrestricted color choice paradigm

Aurore Zelazny[1,2]*, Xun Liu[3] and Thomas Alrik Sørensen[1,2]

[1]Centre for Cognitive Neuroscience, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark, [2]Sino-Danish College (SDC), University of Chinese Academy of Sciences, Beijing, China, [3]Department of Psychology, University of the Chinese Academy of Sciences, Beijing, China

Since Kandinsky's claim for fundamental shape–color associations, several studies have revealed that those tendencies were not generalizable to the entire population and that different associations were more prevalent. Past studies, however, lacked a methodology that allowed participants to freely report their shape–color preferences. Here, we report data from 7,517 Danish individuals, using a free choice full color wheel for five different geometrical shapes. We find significant shape–hue associations for circle-red/yellow, triangle-green/yellow, square-blue, and pentagon/hexagon-magenta. The significant shape–hue associations are also more saturated than non-significant ones for the circle, triangle, and square. At the conceptual level, basic shapes, which show stronger associations, are linked to primary colors, and non-basic shapes to secondary colors. Shape-color associations seem indeed to follow the Berlin-Kay stages of entry into languages. This pattern had previously been described for graphemes and weekday–color associations. The methodology employed in our study can be repeated in different cultural contexts in the future. We also provide another instance of color associations for ordinal concepts that follow the stages of entry into languages.

KEYWORDS

shape–color associations, cross-modal correspondence, concept learning, ordinal concepts, shape learning

Introduction

Color perception begins with light waves entering the eye. Interestingly, in some cases, color perception can occur in the absence of color information entering the retina. Notably, when presented with black letters, individuals with grapheme–color synesthesia do experience colors (Cytowič, 1993; Ásgeirsson et al., 2015). Although synesthesia is a fascinating phenomenon, the general population seems to also make associations to categories across sensory modalities. Cross-modal correspondence [or association (Spence, 2011; Malfatti, 2014)] refers to this tendency for a large portion of the non-synesthete population to join features from different modalities together (Spence, 2011). For example, high-pitched sounds are commonly associated with brighter colors (Marks, 1987). Although sharing some aspects with synesthesia (e.g., mapping from one modality to another modality), the two phenomena differ on several accounts. Cross-modal correspondences are universal, with little variation across individuals, and unconscious (Deroy and Spence, 2013), contrary to synesthesia, where associations are idiosyncratic and conscious. Research on cross-modal correspondence became more and more prolific over the past couple of decades, with much interest geared toward auditory-visual correspondence [see Spence (2011) for a review]. Nevertheless, a large variety of modality pairs have been investigated, including smell–color (Gilbert et al., 1996; Demattè et al., 2006a), shape–smell (Hanson-Vaux et al., 2013), sound–taste (Knöferle and Spence, 2012), shape–taste (Velasco et al., 2015), taste–shape (Velasco et al., 2016), vision–touch (Martino and Marks, 2000), or smell–touch (Demattè et al., 2006b), to cite a few.

Among studies on color-related correspondences are those examining shape–color associations. These originate in the arts when Kandinsky (Kandinsky, 1912[1946]) famously postulated a fundamental link between shapes and colors, linking the triangle to yellow, the circle to blue, and the square to red. These associations were empirically examined using a priming paradigm (Kharkhurin, 2012) as well as with the Implicit Association Task (IAT; Makin and Wuerger, 2013). Both studies demonstrated that Kandinsky's original proposal is not generalizable to the entire population. Nevertheless, although Kandinsky's suggested associations were not supported, it is well-established that some shape-color associations are broadly available in the general population. Table 1 provides a summary of experimental results for the three shapes proposed by Kandinsky. These studies point to a general trend, namely, red/yellow circle, yellow/red triangle, and blue/red square. However, because these studies employed various paradigms, a direct comparison between them is challenging. Indeed, while some studies used an IAT paradigm (Makin and Wuerger, 2013; Chen et al., 2015b), others preferred a direct report method, i.e., asking participants to choose which color best fits various shape stimuli (Jacobsen, 2002; Jacobsen and Wolsdorff, 2007; Albertazzi et al., 2013; Chen et al., 2015a, 2016, 2019; Dreksler and Spence, 2019; Hanada, 2019). Comparison between these direct report studies is also difficult, given that they differed in the number of color options they offered as answers to their participants. The most restricted ones (Jacobsen, 2002; Jacobsen and Wolsdorff, 2007) only offered three possible color answers (red, yellow, and blue), and each of them could only be used once. Other studies (Chen et al., 2016, 2019) instead offered four color chips (red, yellow, blue, and green), which could be reused for several shapes. Hanada (2019) proposed 15 color chips (light blue, blue, cyan, green, light green, mint green, light yellow, yellow, orange, pink, red, magenta, purple, violet, and brown) consisting of eight Berlin–Kay color categories (Berlin and Kay, 1969), plus seven more precise colors, to bring variety. Dreksler and Spence (2019) selected 36 color chips at maximum saturation, consisting of 10 color categories featuring eight Berlin–Kay color categories (Berlin and Kay, 1969), plus two more varieties of blue (red, orange, yellow, green, aquamarine, sky, blue, purple, pink, and brown) in three different levels of lightness, plus a 6-step grayscale. Finally, Albertazzi et al. (2013) and Chen et al. (2015a) used 40 preselected color chips, taken from the Natural Color System (NCS) at maximum saturation.

The use of pre-selected color chips, although facilitating data analysis, may in some cases have limited and potentially influenced the responses from the participants. In the case of three color options where participants were not allowed to reuse a color chip for two different shapes (Jacobsen, 2002; Jacobsen and Wolsdorff, 2007), it is possible that participants found themselves forced to report a shape–color association they did not experience due to the one-to-one pairing restriction.

TABLE 1 Summary of findings from previous shape–color association studies, compared to Kandinsky's proposal for the circle, triangle, and square.

References	Nationality	*N*	◯	△	□
Kandinsky (1912[1946])	Russian	1	B	Y	R
Jacobsen (2002)	German	200	Y	R	B
Jacobsen and Wolsdorff (2007)	German	74	R	Y	B
Kharkhurin (2012)	United Arab Emirates	284	NA	NA	B
Albertazzi et al. (2013)	Italian	60	R, Y	Y	R, B
Makin and Wuerger (2013)	British	36	NA	NA	NA
Chen et al. (2015a)	Japanese	138	R, Y, O	Y	B
Chen et al. (2015b)	Japanese	38	R	Y	B
Chen et al. (2016)	Japanese	91 + 95	R	Y	B, G
Dreksler and Spence (2019)	International	64	R, Y	R, Y	NA
Hanada (2019)	Japanese	50	R	Y	NA
Chen et al. (2019)	Chinese	99 + 117	R	Y	B

B: blue; G: green; O: orange; R: red; Y: yellow; NA: non-significant/non-tested association.

The color chips proposed by Albertazzi et al. (2013) and Chen et al. (2015a) were based on the 40 hues of the NCS at maximum saturation, which failed to provide color categories such as cyan, brown, or pink. On the other hand, Dreksler and Spence (2019) provided a restricted version of the color wheel, covering the eight Berlin–Kay hue categories, plus a grayscale. It contained, however, an over-representation of blue chips (nine chips), compared to other color categories (three chips each). Furthermore, the provided chips had not been controlled for perceptual differences, resulting in chips, within and across hues, that are close to perceptual similarity (Robertson, 1977).

Colors are the sum of three components: hue, lightness, and saturation. Hue is the aspect that we commonly call "color", such as red, blue, or green. Lightness refers to the amount of white or black in the color, while saturation refers to the degree of colorfulness of the hue, i.e., how much gray is contained in the color. In recent studies (Albertazzi et al., 2013; Chen et al., 2015a; Dreksler and Spence, 2019), the saturation aspect of colors was set to maximum, which did not allow the evaluation of the involvement of saturation in shape–color associations. Saturation and lightness seem though to play a role in cross-modal correspondences, as they were indeed found to highly influence the color-emotion associations (Schloss et al., 2020), saturation levels were related to haptic stimuli intensity (Slobodenyuk et al., 2015) and saturation differences reflected increases in loudness and frequency in sound-color associations (Hamilton-Fletcher et al., 2017).

Several explanations for the specific shape–color associations have been proposed. Shape–color associations have been hypothesized to be mediated by perceived temperatures, where shapes rated warmer were associated with warm colors, such as red and yellow, whereas shapes regarded as cold align more with the color blue (Albertazzi et al., 2013; Chen et al., 2015a). Malfatti et al. (2014), however, also found that emotions could as well mediate the shape–color associations, with angry shapes being associated with angry colors.

Prior research on color-related associations has found a tendency for ordinal concepts to be non-randomly associated with colors following Berlin and Kay (1969)'s six stages of entry into languages of color terms. Berlin and Kay (1969) established that color terms appear in languages following a precise order. Languages that only have two color terms will therefore feature words meaning "black" and "white", but languages that have three terms will have words that designate the colors "black", "white" and "red". Berlin and Kay (1969) managed thus to define six stages of entry into languages for color terms.

Ordinal concepts, such as numbers and days of the week (Shanon, 1982), as well as numbers and letters in both English and Arabic (van Leeuwen et al., 2016), were found to be associated with colors that follow the six stages defined by Berlin and Kay (1969). These studies thus demonstrate the involvement of higher-level concept mapping in color associations.

Previous shape–color studies did not allow for investigation of this question, as intrinsic order was not possible to be derived from the stimuli list. To test this hypothesis, we selected five shapes, with increasing complexity, as denoted by their number of sides (circle, triangle, square, hexagon, and pentagon). The order is even more clear for our tested Danish population as the shape names are directly derived from the number of sides they contain [trekant ("three-edge"), firekant ("four-edge"), femkant ("five-edge"), and sekskant ("six-edge")].

To investigate the shape–color associations in a direct report paradigm where we want to avoid the potential biases cited above, we performed a large-scale data collection in collaboration with the Trapholt Museum in Denmark (Zelazny and Sørensen, 2020), using a full color wheel inspired by work on synesthesia (Ásgeirsson et al., 2015) and work on color-related cross-modal associations involving various sensory modalities such as haptic (Lindborg, 2014; Slobodenyuk et al., 2015; Delazio et al., 2017) and sound (Hamilton-Fletcher et al., 2017). Museum visitors were invited to freely associate five pre-defined shapes (circle, triangle, square, hexagon, and pentagon) with the color they feel fits best. The full color spectrum, including the option to modify both lightness and saturation, was available for the participants. This method allows for a more thorough and genuine investigation of shape–color association tendencies without placing any restrictions on the participants' access to colors.

Methods

The study has been approved by Aalborg University and complies with the GDPR rules. The study falls under the Danish National Videnskabsetisk Komité Law (§ 14, part 2), by which behavioral studies not involving biological material are exempted from the regional ethics committee. The study was performed in accordance with the ethical standards of the Declaration of Helsinki (1964) and its subsequent amendments. Participants gave written informed consent to data collection.

Apparatus and stimuli

The experiment was part of a set-up containing four experiments, which were independent from each other. Only the results from one of the four experiments are used in the present study. The setup was designed to be self-explanatory and to be able to function without any supervision. The experiment was run on a Windows 10 Lenovo Miix 320 (10.1-inch screen) two-in-one computer, consisting of an FHD screen with a 1920 × 1200 resolution and a detachable keyboard, used in touchscreen mode. The set of experiments was run through a custom-made webpage, using Chrome in kiosk mode, coded using the Django web framework.

Three custom-made wooden frames were used to secure the computers in place and were attached to an adjacent wall, ∼ 90 cm from the floor. The wooden frames were aligned along the wall, 20 cm apart. The setup consisted of a total of six computers, each wooden frame hosting two computers. A printer was also used, which was fixated between the two computers from the rightmost wooden frame (Figure 1).

Viewing distance could not precisely be controlled for, therefore display dimensions will be described in millimeters rather than in visual angle degrees. The stimuli consisted of five two-dimensional shapes (circle, triangle, square, pentagon, and hexagon) against a gray background [RGB (219, 219, 219), 266 cd/m^2]. Each of the shapes was dimensioned as shown in Figure 2. The shapes were presented one after the other on the right-hand side of the screen and a full color wheel of 85 mm × 85 mm was displayed on the left-hand side. A lightness bar of 87 mm × 5 mm was positioned just below the color wheel. The color wheel and shape stimuli were separated by a 33-mm gap. Upon presentation, the outline of the shape stimulus was shown using a one-pixel black line and transparent filling. As soon as an area of the color wheel was touched, the shape was automatically filled in with the corresponding color and the black outline disappeared. Unlimited color changes were possible, and there was no time limit to respond. Once satisfied with the selected color, the participants had to press a 33 mm × 12 mm "Next" button displayed at the bottom right of the screen to see the next shape. A 20 mm × 10 mm counter was also inserted at the top right of the screen to maintain engagement. See Figure 3 for a screenshot of the experimental layout. Each shape was shown only once, and the presentation was randomized, resulting in five trials per participant. Only the CIELab values [i.e., the coordinates of the colors in the CIELab color space (International Organization for Standardization, 2018)] of the last selected color (i.e., the one displayed on the shape before pressing the "Next" button) were collected for each shape for each participant.

FIGURE 1
Experimental setup at the Trapholt Museum, Denmark. The experiment was part of a 4-experiment setup. The experiment described in this study was conducted at the station marked "1." Visitors received a unique ID number at the "Start" station and could review their responses at the "Result" station.

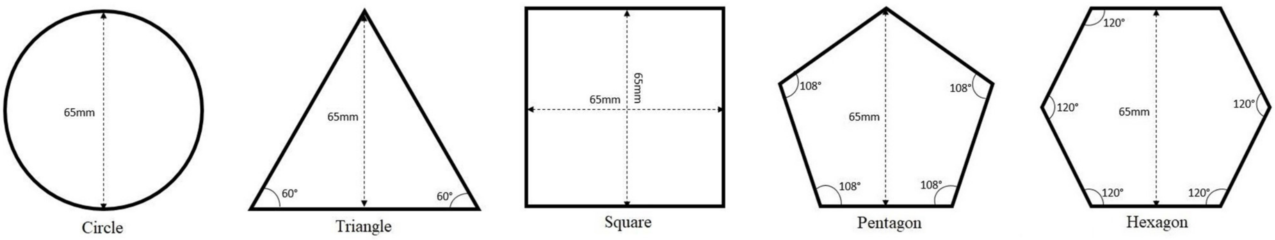

FIGURE 2
Five shapes used as stimuli.

Procedure

The data collection was performed at the Trapholt Museum (located in Kolding, Denmark) as part of the SENSEME exhibit, inviting visitors to explore their multisensory perceptions. The visitors could not be isolated to perform the experiment, nevertheless, particular attention was made to not place the setup close to any decoration or art elements that could remind the participants of any of the stimuli presented. As such, the setup was placed in a quiet area of the museum, facing a white wall to avoid distractions (see Figure 1). The setup was accessible without having to go through the various exhibition rooms and the experiment could also be performed as a first stop before entering the exhibition rooms. Furthermore, none of the art pieces presented at the SENSEME exhibit featured sensory responses to shapes and focused rather on multisensory responses to taste, touch, movement, and music.

The experiment reported here was part of a four-experiment set-up, the other three experiments asked participants to report their color experience to a set of nine graphemes (i.e., 1, 4, 8, A, D, N, 水, 竹, and 皿), to report their color experience to seven different pure tones (i.e., ranging from C4 to B4) and to report their spatial representation of a year. The experiments could be performed in any order; nevertheless, the shape–color association experiment being the one placed the closest to the "Start" station (Figure 1, see below a description of data collected at the Start station), 68% of the visitors completed this experiment before any of the three others. Each experiment could be performed only once by each participant.

The museum context for data collection comes with some limitations. The experimental paradigm had to be brief to ensure that the participants completed the task. For this reason, we tried to keep the testing time to <5 min in total, to lower the dropout rate. Similarly, we only tested five shapes to remain within the time constraints. The number of shapes was determined based on the maximum time estimated to provide a color answer for each stimulus using a color wheel.

The museum's floor plan did not allow us to isolate the visitors during testing. Paradigms based on fast stimuli presentation and requiring an isolated environment were avoided, such as experiments based on priming or implicit experience. Direct measures, where the participants were asked to report their experience, were therefore used here. For the reasons cited above,

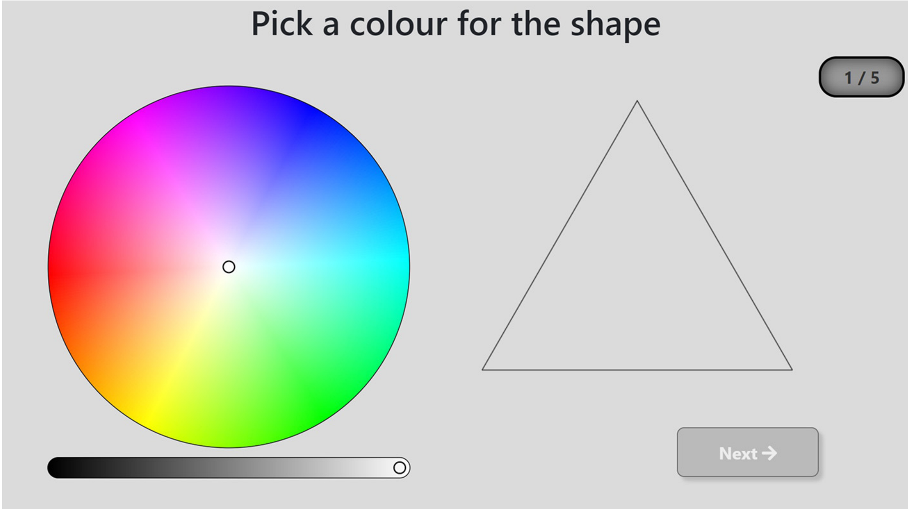

FIGURE 3
Screenshot of an experimental trial.

no time limit was introduced for participants to deliver their responses either.

The visitors were provided with a unique ID number after having completed a demographic questionnaire (Station marked as "Start" in Figure 1) and could also select their preferred language (English or Danish). Upon entering their ID number to start the "shape-color" experiment (Station marked as "1" in Figure 1), an instruction screen appeared. It explained the task ("*Pick the color that you feel fits each shape best, by pressing on the color wheel. When you are satisfied with the chosen color, press Next to see the next shape*") and featured a gif animation on how to perform it, using an illustrating example, a star, which was not used as a stimulus during the test. A "?" button was also there to provide a detailed step-by-step procedure. Upon pressing the "Begin" button, the visitors were brought to the experiment screen, starting the first of five trials.

After the test was completed, the visitors received some feedback, to try to maintain engagement in the experiments and lower the drop-off rate throughout the whole setup. The feedback consisted of a single screen featuring the five shapes each displayed in the color chosen by the visitor, along with reference results collected during the piloting session. The feedback was provided after task completion to avoid biasing the reports.

Participants

A total of 8,841 visitors completed the task. The data from 106 visitors presented issues (trials were not forwarded to the database) and were removed. An additional 196 participants were removed due to having reported colorblindness. Abnormal and not corrected-to-normal visual acuity were reported by 195 visitors, who were also removed from the analysis. Of the remaining 8,344 visitors, 7,550 reported to be of Danish nationality. Among those, 32 were removed due to reporting an age below 6 years, reflecting potentially a typo or parents performing the task for their child, and one was removed for reporting an aberrant age (i.e., 900 years). This resulted in a total of 7,517 visitors remaining for data analysis and comprised 5,186 women (mean age = 37.1; SD = 19.7; range: 6–89) and 2,331 men (mean age = 36; SD = 21.2; range: 6–100). Among the 7,517 visitors, a total of 993 reported having some type of synesthesia. These were nevertheless included in the analysis, as we could not formally validate their synesthesia and did not inquire about the type of synesthesia they had.

Results

The CIELab values allow measuring the three components of color: lightness, saturation, and hue. The CIELab color space is built around three axes: L, *a,* and *b*. Lightness is denoted by the value along the L axis. The *a* and *b* axes represent continua of colors ranging from green to red and blue to yellow, respectively. Saturation and hue are retrieved by converting the cartesian coordinates denoted by the *a* and *b* axes to polar coordinates, where the saturation is defined as the length S of the projection along the *a* and *b* axes, calculated using formula (1) and hue as the angle H relative to the positive pole of the *a* axis, calculated using formula (2).

$$S = \sqrt{a^2 + b^2} \tag{1}$$

$$H = \tan^{-1}\frac{b}{a} \tag{2}$$

Hue analysis

The participants' answers were not evenly distributed along the hue wheel (Figure 4), with some hues concentrating larger portions of answers than others. To determine what areas of the hue wheel were more often associated with each of the five shapes, the CIELab space was divided into 180 bins of 2° angles each. Responses with a saturation value below 1 were excluded, as they denote a grayscale response. These represented 326 responses (0.87% of the data). A chi-square test of independence was carried out on the frequency of each bin for each shape (5 × 180) to examine shape–hue association tendencies. The chi-square test revealed a significant association between hue bins and shapes [$\chi^2(716, N = 37{,}259) = 4937.7$, $p < 0.001$, $\phi_C = 0.18$, $CI_{.95}$ (0.163, 0.174)]. *Post-hoc* residual analysis, Bonferroni-corrected to the threshold $z > 4.03$ at $p < 0.05$ (Figure 5), showed that 50° to 63° hues [red hues; peaking at 52° ($z = 12.7$)], as well as 340° to 3°, plus 8° and 9° hues [yellow; peaking at 350° ($z = 7.8$)], were significantly more chosen for the circle. Hues corresponding to 310° to 317° [green; peaking at 314° ($z = 14.5$)] and 342° to 359° [yellow; peaking at 346° ($z = 7.6$)] were significantly more chosen for the triangle, and hues corresponding to 142° to 162°, plus 166°, 167°, 170° and 171° [blue; peaking at 144° ($z = 21.7$)] for the square. The pentagon was significantly more associated with hues corresponding to 128° and 129°, and 132° to 139° [magenta/purple; peaking at 128° ($z = 5.7$) and 136° ($z = 5.3$)]. The hexagon was significantly more associated with hues corresponding to 22° and 23° [orange ($z = 4.2$)], 114° and 115°, 118° to 121°, 124° and 125°, 128° to 133° [magenta/purple; peaking at 120° ($z = 5.2$) and 130° ($z = 6.9$)].

Saturation analysis

Maximum saturation is not denoted by the same *S* value for every hue angle in the CIELab space (e.g., full saturation at a 255° hue angle (cyan) corresponds to a *S* value of 40, while at a 143° hue angle (blue), it corresponds to a *S* value of 133.8). To make sure that the saturation analysis was not confounded by saturation differences across hues, participant responses' *S* values were normalized to a 0–100 scale, where 0 denotes a *S* value of 0 (grayscale) and 100 the maximum saturation value possible at a given hue angle.

A one-way repeated measures ANOVA was run on the normalized saturation, as a function of shape. A significant main effect of shape was found, $F(1, 3.85) = 21.58$, $p < 0.001$, $\eta_p^2 = 0.003$.

Post-hoc Bonferroni-corrected paired-sample *t*-tests were run to compare normalized saturation between shapes. The triangle was found to have a significantly higher saturation ($M_T = 85.2$; $SD_T = 16.8$) than all four other shapes ($M_C = 83.5$; $SD_C = 20.2$; $M_S = 83.7$; $SD_S = 20.5$; $M_P = 83.3$; $SD_P = 17.9$; $M_H = 83.1$; $SD_H = 18.4$), all *p*s < 0.001.

This tendency for the triangle to be associated with more saturated colors may be because the hues significantly associated with the triangle (310° to 317° and 342° to 359°) have a higher saturation. To test this, the data were split into four groups, whether the data point corresponds to a triangle trial or not (shape type: 2 levels) and whether the data point falls within the hues significantly associated with the triangle (310° to 317° and 342° to 359°) or not (hue bin: 2 levels). A two-by-two ANOVA was run on the normalized saturation. A main effect of shape type was found, $F(1, 37581) = 29.87$, $p < 0.001$, $\eta_p^2 \leq 0.001$, by which the triangle was more saturated (M = 85.2; SD = 16.8) than the other shapes (M = 83.4; SD = 19.3), as well as a main effect of hue bin was found, $F(1, 37581) = 206.57$, $p < 0.001$, $\eta_p^2 = 0.005$, showing that the bins 310°-317° and 342°-359° are significantly more saturated (M = 86.7; SD = 14.7) than all other bins (M = 83; SD = 19.7). A significant interaction of shape type and hue bin was also found, $F(1, 37581) = 4.27$, $p < 0.001$, $\eta_p^2 <= 0.001$. *Post-hoc* Bonferroni-corrected pairwise comparison was run on the two levels of the shape type condition and on the two levels of the hue bin condition to investigate the interaction. All comparisons were significant (significant hue bins for the triangle: M = 88.2; SD = 12.8; non-significant hue bins for the triangle: M = 83.9; SD = 18.2; significant hue bins for the other shapes: M = 86.1; SD = 15.4; non-significant hue bins for the other shapes: M = 82.9; SD = 19.9), with all *p*s < 0.001.

The hues significantly associated with the triangle (310°-317° and 342°-359°) tended to be more saturated overall, but the effect was even more present when those hues were picked for the triangle, compared to the other shapes.

Shape–hue associations and saturation interaction

As shown above, the yellow and green hues significantly associated with the triangle were more saturated when chosen as a response to the triangle than when chosen as a response to the other shapes. This raises the question of whether significant shape–hue associations are more saturated than non-significant ones for the four other shapes as well. To test this, each data point was marked as either being a choice that agrees with the significant shape–hue associations of the population or not. A two-by-two ANOVA was run on the normalized saturation as a function of shape (five levels) and significance (two levels). A main effect of significance was found, $F(1, 37575) = 482.64$, $p < 0.001$, $\eta_p^2 = 0.013$, indicating that significant shape–hue associations are more saturated than non-significant ones (significant associations: M = 87.6; SD = 14.9; non-significant associations: M = 82.6; SD = 19.7). A significant interaction of significance and shape was also found, $F(4, 37575) = 38.63$, $p < 0.001$, $\eta_p^2 = 0.004$. *Post-hoc* Bonferroni-corrected pairwise comparison was run on the two levels of the significance condition for each of the five shapes. Comparisons were significant for the circle, triangle, and square, showing that they were more saturated in their significant hues than in the non-significant ones, with all *p*s < 0.001 (Figure 6). No difference was found, however, for the pentagon and hexagon.

Lightness analysis

A one-way repeated measures ANOVA was run on the L value of the CIELab space, as a function of shape. A significant main effect of shape was found, $F(3.94, 29633.17) = 281.03$, $p < 0.001$, $\eta_p^2 = 0.036$. *Post-hoc* Bonferroni-corrected paired-sample *t*-tests were run to compare lightness values between shapes. The lightness values of all shapes were significantly different from each other, except for the pentagon and hexagon ($M_C = 69.2$; $SD_C = 20.1$; $M_T = 71.7$; $SD_T = 19.9$; $M_S = 61.2$; $SD_S = 22.2$; $M_P = 66.5$; $SD_P = 19.9$; $M_H = 66.4$; $SD_H = 19.8$), all *p*s < 0.001. The triangle appears therefore as the shape associated with the highest lightness value, followed by the circle. The pentagon and hexagon both share a lower lightness value than the circle. Finally, the square is the shape associated with the lowest lightness value.

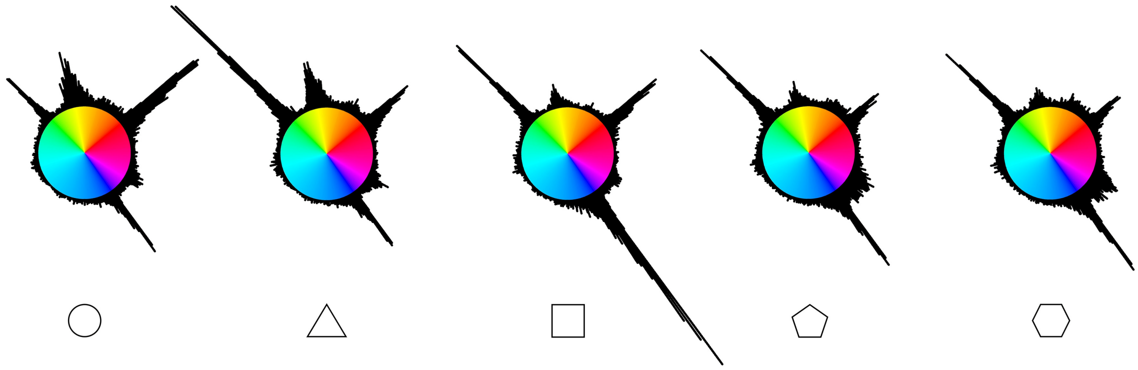

FIGURE 4
Cumulative responses for each shape along the a and b coordinates of the CIELab space. The color of each bin is represented at maximum saturation. $N = 7{,}517$.

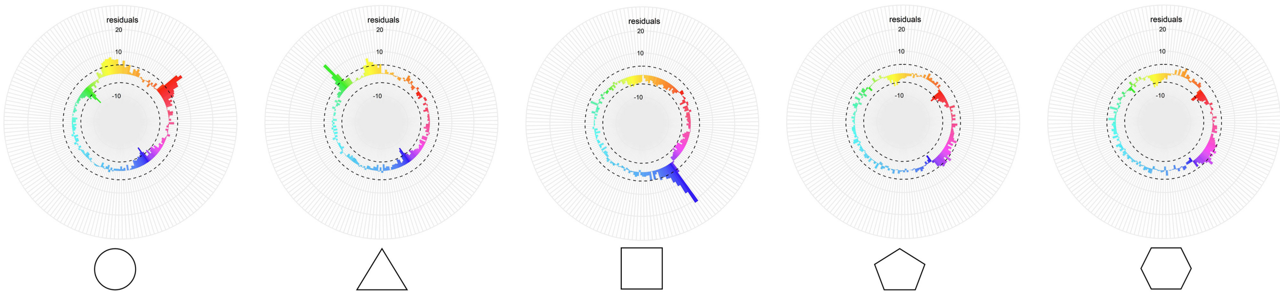

FIGURE 5
X^2-residuals for each hue bin and shape. The dashed line depicts the Bonferroni-corrected significance threshold. The colors displayed are the average color for each of the hue bins for the whole dataset.

Shape–hue associations and lightness interaction

Following our finding on saturation differing between significant and non-significant shape–hue associations, a similar analysis was run regarding lightness. The goal was to compare whether the lightness differs in significant shape–hue associations compared to non-significant ones, for each shape separately. A two-by-two ANOVA was run on the lightness value as a function of shape and significance. A main effect of significance was found, $F(1, 37575) = 103.98$, $p < 0.001$, $\eta_p^2 = 0.003$, indicating that non-significant shape–hue associations are brighter than significant ones (significant associations: M = 65.7; SD = 23.3; non-significant associations: M = 67.4; SD = 19.8). A significant interaction of significance and shape was found, $F(4, 37575) = 1710.38$, $p < 0.001$, $\eta_p^2 = 0.154$. *Post-hoc* Bonferroni-corrected pairwise comparison was run on the two levels of the significance condition for each of the five shapes. All comparisons were significant, with the circle and triangle being brighter in their significant hues than in the non-significant ones. Following a reversed pattern, the square, pentagon, and hexagon were less bright in their significant hues than in the non-significant ones, with all $ps < 0.001$ (Figure 7).

Color analysis

The three components of color (i.e., lightness, saturation, and hue) were this time considered together. The goal was to evaluate what areas of the color space were more often chosen for each shape. A density-based spatial cluster analysis was carried out using the DBSCAN algorithm (Ester et al., 1996). The technique searches for areas of high-density points, which come to form clusters, and leaves out outliers that occur in low-density areas. This method was favored over the k-mean algorithm, which relies on Euclidean distances for forming clusters, as a method to account for noise in the data. The two main parameters of the DBSCAN algorithm are the epsilon, which defines the minimum distance between two clusters, and minPts, which defines the minimum number of neighbors needed for cluster identification.

In the present study, six clustering analyses were performed: one for each shape, to identify whether the number and location of high-density areas vary between shapes, and one for the data as a whole, which then serves as a basis for further analysis (chi-square test of independence). For the five shapes analyses, the epsilon parameter was set at 6 and the minPts parameter at 50. For the whole data analysis, the epsilon parameter was set at 6 and the minPts parameter at 250. The mean CIElab values of the resulting clusters for the whole data analysis can be found in Table 2 and the

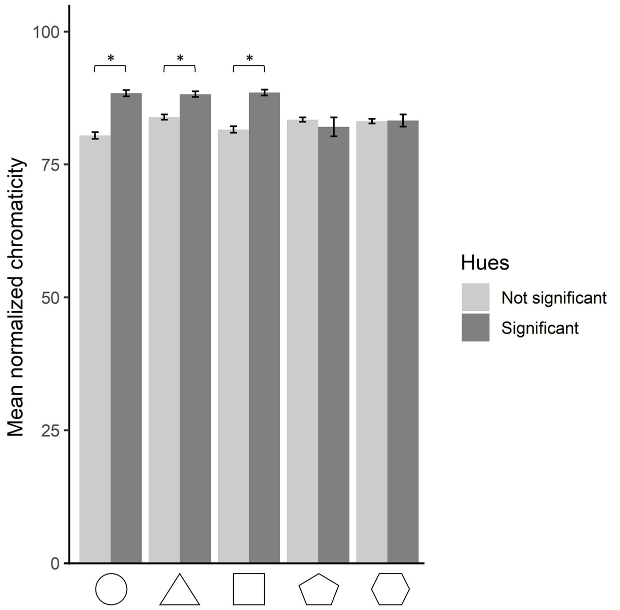

FIGURE 6
Mean normalized saturation for each shape, for hues resulting significant and non-significant from the hue analysis. The error bars depict 95% confidence intervals. Statistically significant comparisons at $p < 0.05$ are marked by an asterisk.

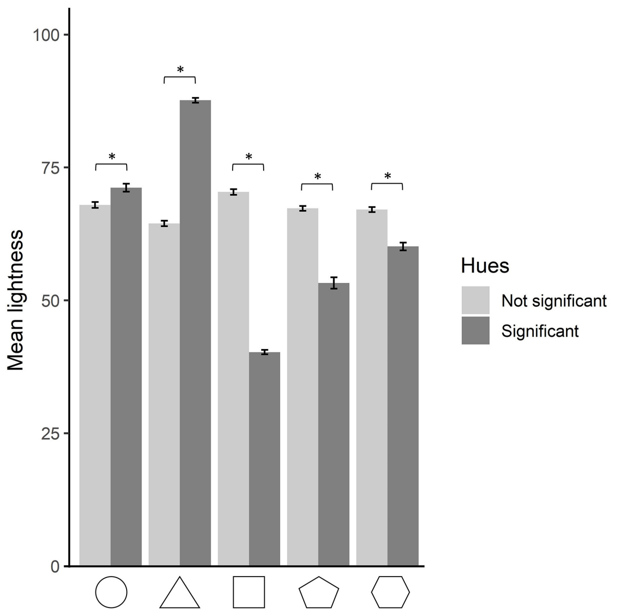

FIGURE 7
Mean lightness for each shape, for hues resulting significant and non-significant from the hue analysis. The error bars depict 95% confidence intervals. Statistically significant comparisons at $p < 0.05$ are marked by an asterisk.

disposition and colors of the data points are shown in Figure 8. The clusters resulting from the clustering analysis run on each of the five shapes independently are gathered in Supplementary Table S1 of the Supplementary material.

TABLE 2 Mean CIELab coordinate values for the eight clusters resulting from the whole-data analysis, along with their frequency count and color name.

Cluster	L	a	b	Color name	Frequency
1	42.09	57.96	−91.46	Blue	6,132
2	88.53	−77.97	72.75	Green	4,135
3	55	75.33	55.22	Red	4,788
4	89.65	−43.87	−11.45	Cyan	1,552
5	93.23	−15.9	85.19	Yellow	3,555
6	61.46	86.97	−46.92	Magenta	1,954
7	70.03	36.07	70.11	Orange	1,327
8	48.89	84.5	−79.89	Purple	312

Shape–cluster associations

A chi-square test of independence was carried out on the frequency of the clusters resulting from the whole-data analysis for each shape (5 × 8) to examine shape–cluster association tendencies. The chi-square test revealed a significant association between clusters and shapes ([(28, N = 23,755) = 3139.4, $p < 0.001$, $\phi_{C} = 0.18$, $CI_{.95}$ (0.175, 0.187)].

Post-hoc residual analysis, Bonferroni-corrected to the threshold, $z > 3.23$ at $p < 0.05$ (Figure 9, top panel), showed that the circle was significantly more chosen for cluster 3 (red; $z = 24.5$) and cluster 5 (yellow; $z = 20.2$). Cluster 2 (green; $z = 16.1$) and cluster 5 (yellow; $z = 14.4$) were significantly more chosen for the triangle, and cluster 1 (blue; $z = 31.3$) for the square. The pentagon was significantly more associated with cluster 4 (cyan; $z = 6.2$), cluster 6 (magenta; $z = 9$), cluster 7 (orange; $z = 4.7$), and cluster 8 (purple; $z = 9.2$), as well as the hexagon with cluster 4 (cyan; $z = 7.8$), cluster 6 (magenta; $z = 12.8$), cluster 7 (orange; $z = 5.2$), and cluster 8 (purple; $z = 5.2$).

Shape–cluster associations per age group

The age distribution of our group of participants follows a bimodal distribution, with a first group (younger group) ranging from 6 to 34 years of age ($N = 3{,}637$), and a second group (older group) ranging from 35 to 100 years of age ($N = 3{,}880$). Differences in shape–color associations found between studies conducted in Germany were imputed to sub-cultural differences, namely participants were on average 20 years younger (median age: 23 years; range: 17–32 years) in an early study (Jacobsen, 2002) compared to participants from a later study (mean age: 43 years, range: 23–63 years) (Jacobsen and Wolsdorff, 2007). In both studies, the square was associated with blue, but the younger participant group associated the circle with yellow and the triangle with red (Jacobsen, 2002), while the older participant group followed an opposite pattern, by associating the circle with red and the triangle with yellow (Jacobsen and Wolsdorff, 2007).

A log-linear analysis (Christensen, 1997) with a Poisson model was carried out on the frequency of the cluster for each shape for each age group. The three-way log-linear analysis of shape (5) ×

cluster (8) × group (2) revealed an interaction of shape, cluster, and age group [$\chi^2(28, N = 23{,}755) = 192.6, p < 0.001$].

A *post-hoc* chi-square test of independence on the frequency of each cluster (8) for each shape (5) was run for each age group (2) to examine significant shape–color associations for different age ranges. The results came out as significant for both age groups [Younger: $\chi^2(28, N = 10{,}652) = 1388.4, p < 0.001, \phi_C = 0.18$, $CI_{.95}$ (0.169, 0.188); Older: $\chi^2(28, N = 13{,}103) = 1952.7, p < 0.001$, $\phi_C = 0.193$, $CI_{.95}$ (0.183, 0.2)]. A Bonferroni-corrected threshold, $z > 3.23$ at $p < 0.05$, revealed that different associations between groups (Figure 9, middle and bottom panels) occurred for the circle regarding the orange association in the older group ($z = 4.78$), for the square with significant associations to red for the younger group ($z = 5.13$), for the pentagon with significant associations to orange ($z = 5.2$) for the younger group and cyan ($z = 6.09$) for the older group, and for the hexagon with significant associations to purple ($z = 4.24$) for the younger group. All other shape–color associations did not significantly differ between groups.

Color category analysis

Our goal was to test whether shape ordinality maps onto the six stages of entry into languages of color terms described by Berlin and Kay (1969). To do so, the CIELab values were converted into Berlin–Kay color categories (Berlin and Kay, 1969) using the online application Colournamer (Mylonas et al., 2010, 2013), a computation model trained by over 3,000 participants, which returns for color inputs the most likely color name in several languages and its Berlin–Kay color category in English. This method of converting 3D color space coordinates into Berlin–Kay color categories has been used in the recent studies assessing typical grapheme–color associations (Root et al., 2017, 2019; Rouw and Root, 2019). Upon visual inspection, the algorithm appears less precise for colors close to the grayscale axis (at a saturation below 1). Those color responses have therefore been corrected manually (629 data points, corresponding to 1.7% of the data).

The shape order has been numbered from 1 to 5, to denote their increasing number of sides, called complexity in the remainder of the analysis (circle:1; triangle: 2; square: 3; pentagon: 4; hexagon: 5).

Entry into languages order

A Spearman's Rho correlation analysis was run on the shape complexity as a function of color name entry into language order (Berlin and Kay, 1969). A correlation was found between entry into language order and shape complexity ($\rho = 0.17, p < 0.001$). A tendency for shape complexity order to follow color name entry into language order seems to appear.

Easy of generation order

Battig and Montague (1969) have shown that color name generation was not random, and that it followed a specific order. It could be possible therefore that less complex shapes are associated with easily generated color names, rather than names that enter first into the language. We, therefore, tested whether shape–color associations reflect shape complexity order being mapped from color name ease of generation order. A Spearman's Rho correlation analysis was run on the shape complexity order as a function of color name ease of generation order (Battig and Montague, 1969), following Simner et al. (2005)'s procedure. No correlation was found ($\rho = 0.006, p = 0.175$) between shape complexity and color category ease of generation order.

TABLE 3 The 11 color categories ranked by ease of generation, color name frequency, and entry into human languages orders, in descending order.

Ease of generation (Battig and Montague, 1969)	Color name frequency (DaTenTen20 corpus)	Entry into languages (Berlin and Kay, 1969)
Blue	Black	Black/white
Red	White	Red
Green	Green	Green/yellow
Yellow	Red	Blue
Orange	Blue	Brown
Black	Yellow	Orange/purple/gray/pink
Purple	Gray	
White	Brown	
Pink	Orange	
Brown	Pink	
Gray	Purple	

Color name frequency order

All color names do not have the same frequency in language. To investigate whether shape–color associations reflect shape complexity order being mapped from color name frequency order in language, color name frequencies in Danish were established based on the DaTenTen20 corpus (Danish web corpus) containing frequencies per 3.4 billion words. The Danish color names being inflected depending on the noun following them, the search was done on the root, gender-inflected, number-inflected, and determination-inflected variations of the eleven color categories from Berlin and Kay (1969). The resulting order can be found in Table 3. A Spearman's Rho correlation analysis was run on the shape complexity order, as a function of color name frequency order. A correlation was found between color name frequency order and shape complexity ($\rho = 0.057, p < 0.001$). These results suggested a slight tendency for shape complexity order to follow color name frequency.

Upon inspection, color name frequency order and entry into language order seem to resemble each other. A Spearman's Rho correlation analysis was run on the entry into language order, as a function of color name frequency order. A strong correlation was found between color name frequency order and entry into language order ($\rho = 0.753, p < 0.001$). These results suggested that the slight tendency for shape complexity order to follow color name frequency may be driven by the fact that it also follows the entry into language order tightly.

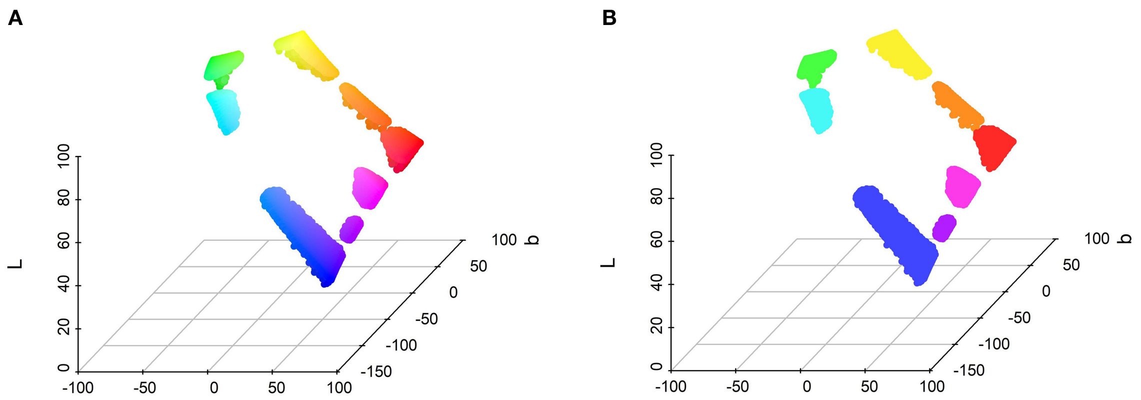

FIGURE 8
Positions in the CIELab space of the eight clusters from the analysis run on the whole dataset. **(A)** The color depicted corresponds to the corresponding color in the CIELab space. **(B)** The color depicted corresponds to the average color of each cluster.

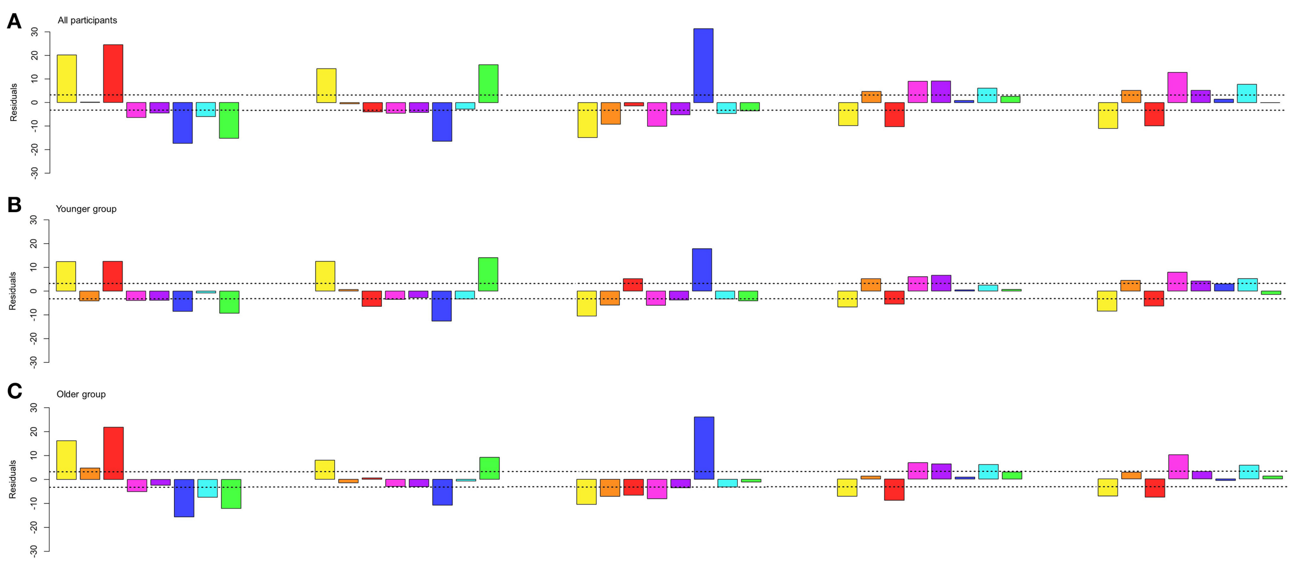

FIGURE 9
X^2-residuals for each cluster from the whole dataset and for each shape. The dash line depicts the Bonferroni-corrected significant threshold. The colors displayed are the average color for each cluster for the whole dataset. **(A)** Residuals for the analysis of all participants. **(B)** Residuals for the analysis for the 6- to 34-year-old group (younger group). **(C)** Residuals for the analysis for the 35- to 100-year-old group (older group).

Figure 10 shows the correlation between shape order and the ease of generation (left panel), color name frequency (center panel), and entry into language (right panel) orders.

Discussion

A large group of museum visitors have been asked to report their color intuitions regarding five shapes, using a full color wheel. The data were analyzed in terms of hue, saturation, and lightness, as well as color as a whole. Overall, Danish participants associated the circle more often with red and yellow, and preferred high lightness values. The triangle was associated with green and yellow, and featured high saturation and lightness values. The square was associated with the color blue and showed low lightness values. Finally, both the pentagon and hexagon were associated with magenta, as well as cyan, purple, and orange. These results indicate several points, first and foremost that the associations are partially in line with those presented in the past literature (Jacobsen, 2002; Jacobsen and Wolsdorff, 2007; Kharkhurin, 2012; Albertazzi et al., 2013; Makin and Wuerger, 2013; Chen et al., 2015a,b; Dreksler and Spence, 2019; see Table 1), where a tendency for red/yellow circle, yellow/red triangle, and blue/red square appeared. Second, despite those similarities to previous studies, the shape–color associations reported by Danish visitors do not fully match them. Notably, the association between green and triangle had never been attested before. It also goes against the general pattern of previous studies which saw the triangle being associated with yellow and red, red being the opponent color to green (Kay and McDaniel, 1978; Abramov and Gordon, 1994; Valberg, 2001; Witzel and Franklin, 2014). This variation could be due to cultural and/or methodological differences. Green was indeed not offered as a color

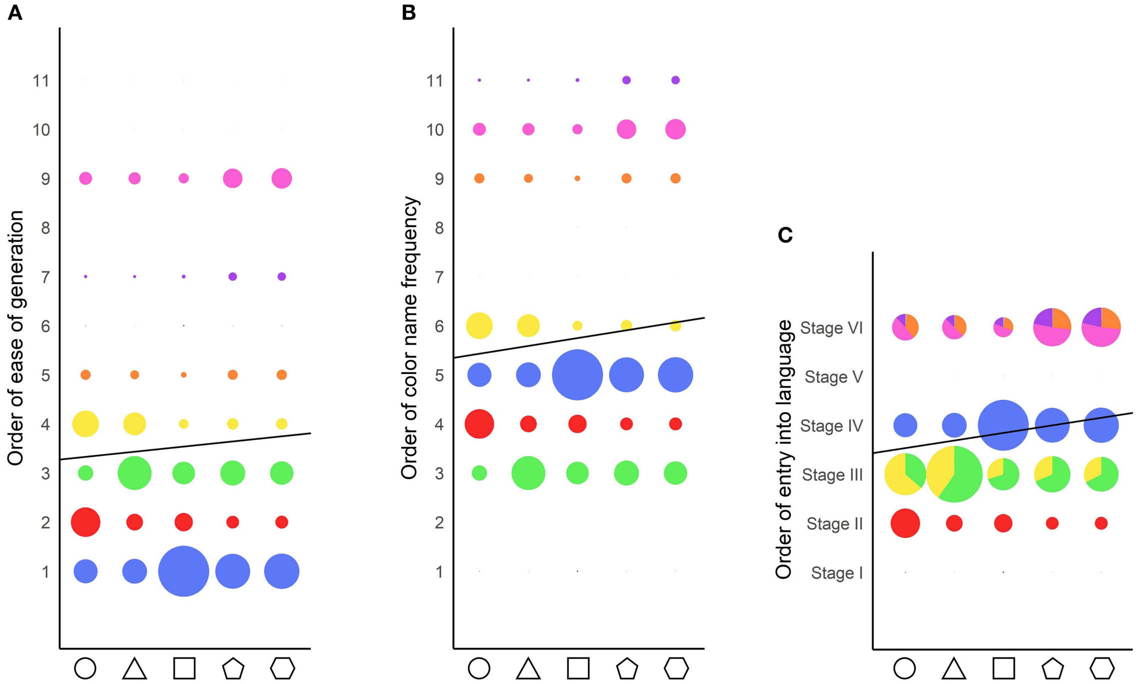

FIGURE 10
Correlation between shape order and the various color orders. The dot sizes reflect the frequency of response of each color category for each shape. The solid black line represents the correlation line. The colors depicted are the average color chosen by the participants for each color category. **(A)** Correlation of shape order with color name ease of generation order. **(B)** Correlation of shape order with color name frequency order. **(C)** Correlation of shape order with color name entry into language order. For stages containing more than one color category, the proportion of responses for each color category is denoted as the proportion of the color on the dots.

chip response by Jacobsen (2002), who found a preference for the triangle-red association. There were, however, an equal number of yellow, red, and green color chips in Dreksler and Spence (2019)'s paradigm, which nevertheless led to associations with red and yellow for the triangle. The triangle–green association may thus be a Danish particularity. Transposing the full color wheel paradigm in a different culture would help clarify this result.

Our color wheel paradigm allowed us to unveil a pattern regarding saturation never attested before due to constraints from previous methodologies. Indeed the saturation property of color choices was always controlled for, and color options were always presented at full saturation in previous studies (Albertazzi et al., 2013; Chen et al., 2015a; Dreksler and Spence, 2019). When saturation is allowed to be manipulated by participants, we observe that significant shape–hue associations are more saturated than non-significant ones for the circle, triangle, and square, while no difference occurs for the pentagon and hexagon.[1] Saturation can be defined as "how colorful a color is". The triangle overall produced more saturated colors than the four other shapes, but when looking at only significant shape–hue associations, the circle, triangle, and square showed similar mean saturations, despite all having different significant hue associations. Interestingly, highly saturated colors have been shown to be preferred to those of low and mid saturations (Palmer and Schloss, 2010) and to produce higher arousal levels (Wilms and Oberfeld, 2018). Saturation levels could therefore reflect the strength of the association at the group level. This coincides with the fact that the circle, triangle, and square have also stronger biases toward certain hues compared to the pentagon and hexagon, as shown by the residual values in Figures 5, 9. On and on, our data indicate that the shape–color associations are more clearly settled for the circle, triangle, and square than for the pentagon and hexagon.

The involvement of prototype categories

As described above, the circle, triangle, and square seem to be subjected to stronger shape–color associations than the pentagon and hexagon. Moreover, while the circle, triangle, and square each exhibit distinct color associations, the pentagon and hexagon are barely different from each other.

Our stimuli list, despite being small, had the particularity of containing only shapes that represent known geometrical concepts. More precisely, contrary to most recent studies on shape–color associations that featured a larger list of stimuli (Albertazzi et al., 2013; Chen et al., 2015a; Dreksler and Spence, 2019), our stimuli

1 It is important to mention that, on the color wheel, less saturated colors were closer to the center of the wheel, and more saturated ones, toward the outskirt. No saturation bar needed to be manipulated to affect saturation, making the highly saturated and less saturated colors equally reachable by one single press on the wheel. This was not the case with lightness, which could only be manipulated by using the lightness bar, thus involving more interactions with the screen.

had the property of all being concepts with a precise name. However, those five shapes are not concepts participants are equally familiar with. The circle, equilateral triangle, and square are, according to the Gestalt psychology principle of "good form," prototypes of shape categories (Rosch, 1973) and are therefore described as basic shapes. They differ from all other shapes (e.g., the pentagon, hexagon, or oval), which do not constitute prototype shape categories. In our study, we observe that the three prototypes of shape categories (circle, triangle, and square) are associated with the four primary color categories [red, blue, green, and yellow (Kay and McDaniel, 1978; Abramov and Gordon, 1994; Valberg, 2001; Witzel and Franklin, 2014)], whereas the non-prototypical shape categories (pentagon and hexagon) are associated with secondary color categories [namely, pink, purple, brown, and orange (Witzel and Franklin, 2014)]. This pattern seems very consistent across studies and cultures (Albertazzi et al., 2013; Chen et al., 2015a). The three prototypes of shape categories are, however, not randomly associated with the four primary color categories. They follow a pattern that is better explained by the stages of entry into languages (Berlin and Kay, 1969). A similar tendency has been observed in the general population for numbers and days of the week (Shanon, 1982), English and Arabic numbers and letters (van Leeuwen et al., 2016), and in synesthetes for letter frequency (Simner et al., 2005). Both intrinsic ordering and frequency have therefore been shown to map onto the Berlin and Kay (1969) stages of entry into languages. It is, however, not possible to determine whether it is intrinsic ordering or frequency that is at play regarding our shape stimuli, as they follow the same pattern, based on complexity order as measured by number of sides (Circle: 0; triangle: 3; square: 4) or frequency in language [Circle: 73,033; triangle: 36,773; square: 26,491/per 3.4 billion words (DaTenTen20)].

Past studies have hypothesized that grapheme–color associations derive from shape–color associations (Spector and Maurer, 2008, 2011) where toddlers tend to associate O and I with white but X and Z with black. Although synesthesia is highly idiosyncratic, synesthetes have been found to show a bias toward associating the letter O with the color white (Rich et al., 2005; Simner et al., 2005; Witthoft et al., 2015), and non-synesthetes have been found to associate the letter O to orange instead (Rich et al., 2005; Simner et al., 2005; Mankin and Simner, 2017). Non-synesthetes seem, therefore, to form letter–color associations based on phonology (i.e., O for orange) when possible. This has been witnessed for other initial letters of color words in adults (Rich et al., 2005; Simner et al., 2005; Spector and Maurer, 2011), and in children as early as from 7 years old (Spector and Maurer, 2008), but not in pre-literate children (Spector and Maurer, 2011) who based they letter–color associations on the visual properties of the letter O, and not on the sound of the letter neither. In our current study, participants were all above 6 years of age and significantly associated the circle with red and yellow. Furthermore, despite having 551 participants between 6 and 10 years of age, none of them associated the circle with white (i.e., the youngest participant who chose white for the circle was 13 years old). Interestingly, 19 of them associated the circle with orange. Our cluster analysis per age group showed that the older group (i.e., 35–100 years) significantly associated orange with the circle on top of red and yellow but not the younger (i.e., 6–34 years) group. Spector and Maurer (2008) found a significant association between the letter O and white in 7- to 9-year-old children and adults. However, it is important to note that the stimuli were only white and black and aimed, therefore, to probe preference for only those two colors. This means that for all age groups, the results only allowed to conclude that the letter O fits the color white more than it fits the color black. In other words, their paradigm did not allow to investigate other color associations to the letter O, such as red, yellow, orange, blue, etc. In line with our data indicating involvement of the conceptual identity of the stimuli, it is a possibility that interpreting the circle stimuli as a geometrical shape leads to color associations (i.e., red or yellow) that differ from the color associations that would be reported if the stimuli had been interpreted as a grapheme. It is also a possibility that the older group interpreted the circle stimuli as a letter more often than the younger group did, but our data do not allow us to evaluate this hypothesis. Furthermore, it would be interesting to replicate Spector and Maurer (2008)'s findings regarding associations with the letter O in pre-literate children in a paradigm offering more color options. Indeed, the circle in our data was significantly associated with colors at Stages II and III of entry into language, while the color white belongs to Stage I. Further research could elucidate whether in toddlers, who do not yet have a conceptual knowledge of shapes (Clements et al., 1999, 2018; Clements and Sarama, 2000, 2020), shape associations occur only for colors forming the Stage I of entry into language, and associations with colors at higher stages are developed as higher-level conceptual knowledge are formed.

Comparison with other mediating factors previously proposed

Shape–color associations have been proposed to be mediated by perceived temperatures (Albertazzi et al., 2013; Chen et al., 2015a) or emotions (Malfatti et al., 2014). A strong shape–temperature/emotion–color association does, however, not inform us about the directionality of the associations. Shape–color associations may emerge based on the shapes' perceived temperature/emotion, which in turn triggers a color matching this temperature/emotion. However, conversely, it could be the shapes' perceived colors that trigger a feeling of temperature/emotion.

Our study differs from the ones previously cited because it contained only shapes that could be named [as opposed to Dreksler and Spence (2019) and Malfatti (2014)'s stimuli lists]. Our findings nevertheless do not necessarily negate the previously proposed mediating factors. Indeed, comparing our study's results with Malfatti (2014) and Dreksler and Spence (2019) may be a case of comparing apples and oranges. In the same way as colors are low-level perceptual entities that can be decomposed into hue, lightness, and saturation, shapes can be decomposed into symmetry, angularity, and complexity. However, colors are also high-level mental categories, with names and prototypical members (Berlin and Kay, 1969; Rosch, 1973). Similarly, shapes also form high-level mental categories, with clear definitions and names (Rosch, 1973). As such, our study is more interested in shapes as high-level and learned categories, than as low-level perceptual properties. This distinction between shapes as perceptual information and conceptual information may be central to understanding the variety of results in the shape–color literature. As an example, Albertazzi et al. (2015) found when looking only at the angle aspect of shapes that 22.5° angles were associated with yellow, 45°, 90°, and 135° angles to green and yellow, and 157.5° angles to red and blue. However, they had previously found that the square (90° angles) was associated to red and blue (Albertazzi et al., 2013). This discrepancy could be due to looking at, on the one hand, low-level perceptual information (i.e., angles and symmetry), and on the other hand, at high-level concepts (shape categories).

Limitations

The current study featured a limited number of stimuli, due to the museum context requirements. It would be interesting in the future to offer a greater variety of shapes, including, for example, an inverted triangle, equilateral diamond, isosceles triangle, rectangle, and oval, which would allow for testing the weight of perceptual (square vs. equilateral diamond) and categorical (square vs. rectangle) differences on shape–color associations.

The museum setup, although offering the chance to test a high number of visitors, was not as controlled as a lab environment. Several factors could have at times influenced the color decisions of the visitors, such as background noises (Ward et al., 2006; Hamilton-Fletcher et al., 2017), subtle lightness changes, or surrounding odors (Gilbert et al., 1996; Schifferstein and Tanudjaja, 2004; Demattè et al., 2006a). Furthermore, visitors could have been influenced by discussions with other visitors. However, our high number of participants should allow to diminish the impact of such noise in our data. The experiment probing shape–color associations was run on only one computer, diminishing the possibility of simply copying other visitors' answers.

Our sample also featured twice more women than men. Color-related gender difference has been found at the perceptual (Bimler et al., 2004), categorization (Sørensen and Kyllingsbæk, 2012), and linguistics (Simpson and Tarrant, 1991; Greene and Gynther, 1995; Mylonas et al., 2014; Fider and Komarova, 2019) levels. It would, therefore, be interesting to investigate shape–color associations in a gender-balanced paradigm.

Finally, our study focuses on only one culture. The same paradigm could be transposed to a different cultural population.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not provided for this study on human participants because the study falls under the Danish National Videnskabsetisk Komité Law (§ 14, part 2), by which behavioral studies not involving biological material are exempted from regional Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

TAS and AZ contributed to the study design and data collection. AZ performed the testing and data analysis and drafted the manuscript. AZ did the interpretation under the supervision of TAS and XL. TAS and XL provided critical revisions. All authors approved the final version of the manuscript for submission.

Funding

TAS was partially funded by the Danish Independent Research Council grant for the DFF 2 project entitled Synaesthesia—the Roles of Association Learning and of Differential Brain Development (grant number 6107-00268B). XL was funded by the National Natural Science Foundation of China (NSFC, grant number NSFC 62061136001) and the German Research Foundation (DFG, grant number TRR-169).

Acknowledgments

We would like to thank the Trapholt Museum for its collaboration with the SenseMe exhibit 2019–2020. We would also like to thank Guillaume Dauster, Michael Damsbo Lyngs, Peter Møller Nielsen, Mathias Huus Olsen, and Tobias Elgaard Fleinert who helped develop the data collection platform. Finally, we would like to thank Daniel Barratt for his valuable input.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1129903/full#supplementary-material

References

Abramov, I., and Gordon, J. (1994). Color appearance: on seeing red-or yellow, or green, or blue. *Annu. Rev. Psychol.* 45, 451–485. doi: 10.1146/annurev.ps.45.020194.002315

Albertazzi, L., Da Pos, O., Canal, L., Micciolo, R., Malfatti, M., and Vescovi, M. (2013). The hue of shapes. *J. Exper. Psychol.: Hum. Percept. Perform.* 39, 37–47. doi: 10.1037/a0028816

Albertazzi, L., Malfatti, M., Canal, L., and Micciolo, R. (2015). The hue of angles—was kandinsky right? *Brill* 3, 81–92. doi: 10.1163/22134913-000 02025

Ásgeirsson, Á. G., Nordfang, M., and Sørensen, T. A. (2015). Components of attention in grapheme-color synesthesia: a modeling approach. *PLoS ONE* 10, e0134456. doi: 10.1371/journal.pone.0134456

Battig, W. F., and Montague, W. E. (1969). Category norms of verbal items in 56 categories A replication and extension of the Connecticut category norms. *J. Exp. Psychol.* 80, 1–46. doi: 10.1037/h0027577

Berlin, B., and Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution.* Berkeley, CA; Los Angeles, CA: University of California Press.

Bimler, D., Kirkland, J., and Jameson, K. A. (2004). Quantifying variations in personal color spaces: are there sex differences in color vision? *Color Res. Appl.* 29, 128–134. doi: 10.1002/col.10232

Chen, N., Jiang, X., and Watanabe, K. (2019). "Color-shape association in chinese people," in *11th International Conference on Knowledge and Smart Technology* (Phuket). doi: 10.1109/KST.2019.8687799

Chen, N., Tanaka, K., Matsuyoshi, D., Nagamori, Y., Namatame, M., and Watanabe, K. (2016). Color-shape association in deaf and hearing people. *Front. Psychol.* 7, 355. doi: 10.3389/fpsyg.2016.00355

Chen, N., Tanaka, K., Matsuyoshi, D., and Watanabe, K. (2015a). Associations between color and shape in Japanese observers. *Psychol. Aesthet. Creat. Arts* 9, 101–110. doi: 10.1037/a0038056

Chen, N., Tanaka, K., and Watanabe, K. (2015b). Color-shape associations revealed with implicit association tests. *PLoS ONE* 10, e0116954. doi: 10.1371/journal.pone.0116954

Christensen, R. (1997). *Log-Linear Models and Logistic Regression.* New York, NY: Springer-Verlag. doi: 10.2307/2669721

Clements, D., and Sarama, J. (2000). Young children's ideas about geometric shapes. *Teach. Children Math.* 6, 482–488. doi: 10.5951/TCM.6.8.0482

Clements, D., and Sarama, J. (2020). *Learning and Teaching Early Math: The Learning Trajectories Approach.* Routledge. doi: 10.4324/9781003083528

Clements, D., Sarama, J., Swaminathan, S., Weber, D., and Trawick-Smith, J. (2018). Teaching and learning geometry: early foundations. *Quadrante* 27, 7–31. doi: 10.48489/quadrante.22970

Clements, D., Swaminathan, S., Hannibal, M. A. Z., and Sarama, J. (1999). Young children's concepts of shape. *J. Res. Math. Educ.* 30, 192–212. doi: 10.2307/749610

Cytowič, R. E. (1993). *The Man Who Tasted Shapes.* New York, NY: G. P. Putnam's Sons.

Delazio, A., Israr, A., and Klatzky, R. L. (2017). "Cross-modal correspondence between vibrations and colors," in *2017 IEEE World Haptics Conference* (Munich: IEEE). doi: 10.1109/whc.2017.7989904

Demattè, M. L., Sanabria, D., and Spence, C. (2006a). Cross-modal associations between odors and colors. *Chem. Senses* 31, 531–538. doi: 10.1093/chemse/bjj057

Demattè, M. L., Sanabria, D., Sugarman, R., and Spence, C. (2006b). Cross-modal interactions between olfaction and touch. *Chem. Senses* 31, 291–300. doi: 10.1093/chemse/bjj031

Deroy, O., and Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychon. Bull. Rev.* 20, 643–664. doi: 10.3758/s13423-013-0387-2

Dreksler, N., and Spence, C. (2019). A critical analysis of colour-shape correspondences: examining the replicability of colour-shape associations. *I-Perception.* 10, 1–34. doi: 10.1177/2041669519834042

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. *KDD.* 96, 226–231.

Fider, N. A., and Komarova, N. L. (2019). Differences in color categorization manifested by males and females: a quantitative World Color Survey study. *Palgrave Commun.* 5, 1–10. doi: 10.1057/s41599-019-0341-7

Gilbert, A. N., Martin, R., and Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: the color of smells. *Am. J. Psychol.* 109, 335–351. doi: 10.2307/1423010

Greene, K. S., and Gynther, M. D. (1995). Blue versus periwinkle: color identification and gender. *Percept. Mot. Skills* 80, 27–32. doi: 10.2466/pms.1995.80.1.27

Hamilton-Fletcher, G., Witzel, C., Reby, D., and Ward, J. (2017). Sound properties associated with equiluminant colours. *Multisens. Res.* 30, 337–362. doi: 10.1163/22134808-00002567

Hanada, M. (2019). Associations of visual forms with colors: the minor role of emotion as the mediator. *Color Res. Appl.* 44, 568–580. doi: 10.1002/col.22382

Hanson-Vaux, G., Crisinel, A.-S., and Spence, C. (2013). Smelling shapes: crossmodal correspondences between odors and shapes. *Chem. Senses* 38, 161–166. doi: 10.1093/chemse/bjs087

International Organization for Standardization. (2018). *Colorimetry—Part 4: CIE 1976 L*a*b* Colour space (ISO Standard No. 11664-4:2018).*

Jacobsen, T. (2002). Kandinsky's questionnaire revisited: fundamental correspondence of basic colors and forms? *Percept. Mot. Skills* 95, 903–913. doi: 10.2466/pms.2002.95.3.903

Jacobsen, T., and Wolsdorff, C. (2007). Does history affect aesthetic preference? Kandinsky's teaching of colour-form correspondence, empirical aesthetics, and the bauhaus. *Design J.* 10, 16–27. doi: 10.2752/146069207789271902

Kandinsky, W. (1912 [1946]). *Wassily Kandinsky on the Spiritual in Art.* New York, NY: Guggenheim Foundation.

Kay, P., and McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language* 54, 610–646. doi: 10.1353/lan.1978.0035

Kharkhurin, A. V. (2012). Is triangle really yellow? An empirical investigation of kandinsky's correspondence theory. *Empir. Stud. Arts* 30, 167–182. doi: 10.2190/EM.30.2.d

Knöferle, K., and Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychon. Bullet. Rev.* 19, 992–1006. doi: 10.3758/s13423-012-0321-z

Lindborg, P. (2014). *Colour association to sound: a perceptual experiment using a CIELab haptic response interface and the Jyväskylä film music set.* SMC Sweden 2014 Sound and Music Computing: Bridging science, art, and industry, Stockholm, Sweden.

Makin, A. D. J., and Wuerger, S. (2013). The IAT shows no evidence for Kandinsky's color-shape associations. *Front. Psychol.* 4, 616–616. doi: 10.3389/fpsyg.2013.00616

Malfatti, M. (2014). *Shape-to-color associations in non-synesthetes: Perceptual, emotional, and cognitive aspects* (Doctoral dissertation). University of Trento, Trento, Italy.

Malfatti, M., Schloss, K. B., Albertazzi, L., and Palmer, S. E. (2014). Shape-to-color associations in non-synesthetes: evidence for emotional mediation. *J. Vis.* 14, 1004. doi: 10.1167/14.10.1004

Mankin, J. L., and Simner, J. (2017). A is for apple: the role of letter-word associations in the development of grapheme-colour synaesthesia. *Multisens. Res.* 30, 409–446. doi: 10.1163/22134808-00002554

Marks, L. E. (1987). On cross-modal similarity: auditory-visual interactions in speeded discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 384–394. doi: 10.1037/0096-1523.13.3.384

Martino, G., and Marks, L. E. (2000). Cross-modal interaction between vision and touch: the role of synesthetic correspondence. *Perception* 29, 745–754. doi: 10.1068/p2984

Mylonas, D., MacDonald, L. W., and Wuerger, S. (2010). "Towards an online color naming model," in *18th Color Imaging Conference* (Springfield).

Mylonas, D., Paramei, G., and MacDonald, L. W. (2014). "Gender differences in colour naming," in *Colour Studies: A Broad Spectrum*. p. 225–239. doi: 10.1075/z.191.15myl

Mylonas, D., Stutters, J., Doval, V., and MacDonald, L. W. (2013). *Colournamer, A Synthetic Observer for Colour Communication (Newcastle Gateshead Convention Bureau Patent).*

Palmer, S. E., and Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8877–8882. doi: 10.1073/pnas.0906172107

Rich, A. N., Bradshaw, J. L., and Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition* 98, 53–84. doi: 10.1016/j.cognition.2004.11.003

Robertson, A. R. (1977). The CIE 1976 Color-Difference Formulae. *Color Res. Appl.* 2, 7–11. doi: 10.1002/j.1520-6378.1977.tb00104.x

Root, N. B., Dobkins, K. R., Ramachandran, V. S., and Rouw, R. (2019). Echoes from the past: synaesthetic colour associations reflect childhood gender stereotypes. *Philos. Trans. R. Soc. B* 374, 20180572–20180572. doi: 10.1098/rstb.2018.0572

Root, N. B., Rouw, R., Asano, M., Kim, C.-Y., Melero, H., Yokosawa, K., et al. (2017). Why is the synesthete's "A" red? Using a five-language dataset to disentangle the effects of shape, sound, semantics, and ordinality on inducer-concurrent relationships in grapheme-color synesthesia. *Cortex* 99, 375–389. doi: 10.1016/j.cortex.2017.12.003

Rosch, E. H. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350. doi: 10.1016/0010-0285(73)90017-0

Rouw, R., and Root, N. B. (2019). Distinct colours in the "synaesthetic colour palette". *Philos. Trans. R. Soc. B* 374, 20190028–20190028. doi: 10.1098/rstb.2019.0028

Schifferstein, H. N. J., and Tanudjaja, I. (2004). Visualising fragrances through colours: the mediating role of emotions. *Perception* 33, 1249–1266. doi: 10.1068/p5132

Schloss, K. B., Witzel, C., and Lai, L. Y. (2020). Blue hues don't bring the blues: questioning conventional notions of color–emotion associations. *J. Opt. Soc. Am. A Optics Image Sci. Vis.* 37, 813–824. doi: 10.1364/JOSAA.383588

Shanon, B. (1982). Colour associates to semantic linear orders. *Psycholo. Res. Psychologische Forschung* 44, 75–83. doi: 10.1007/BF00308557

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., et al. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cogn. Neuropsychol.* 22, 1069–1085. doi: 10.1080/02643290500200122

Simpson, J., and Tarrant, A. W. S. (1991). Sex- and age-related differences in colour vocabulary. *Lang. Speech* 34, 57–62. doi: 10.1177/002383099103400104

Slobodenyuk, N., Jraissati, Y., Kanso, A., Ghanem, L., and Elhajj, I. H. (2015). Cross-modal associations between color and haptics. *Attent. Percept. Psychophys.* 77, 1379–1395. doi: 10.3758/s13414-015-0837-1

Sørensen, T. A., and Kyllingsbæk, S. (2012). Short-term storage capacity for visual objects depends on expertise. *Acta Psychol.* 140, 158–163. doi: 10.1016/j.actpsy.2012.04.004

Spector, F., and Maurer, D. (2008). The colour of Os: naturally biased associations between shape and colour. *Perception* 37, 841–847. doi: 10.1068/p5830

Spector, F., and Maurer, D. (2011). The colors of the alphabet: naturally-biased associations between shape and color. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 484–495. doi: 10.1037/a0021437

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attent. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Valberg, A. (2001). Unique hues: an old problem for a new generation. *Vis. Res.* 41, 1645–1657. doi: 10.1016/S0042-6989(01)00041-4

van Leeuwen, T. M., Dingemanse, M., Todil, B., Agameya, A., and Majid, A. (2016). Nonrandom associations of graphemes with colors in Arabic. *Multisens. Res.* 29, 223–252. doi: 10.1163/22134808-00002511

Velasco, C., Woods, A. T., Deroy, O., and Spence, C. (2015). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Qual. Prefer.* 41, 151–158. doi: 10.1016/j.foodqual.2014.11.010

Velasco, C., Woods, A. T., Petit, O., Cheok, A. D., and Spence, C. (2016). Crossmodal correspondences between taste and shape, and their implications for product packaging: a review. *Food Qual. Prefer.* 52, 17–26. doi: 10.1016/j.foodqual.2016.03.005

Ward, J., Huckstep, B., and Tsakanikos, E. (2006). Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all? *Cortex* 42, 264–280. doi: 10.1016/S0010-9452(08)70352-6

Wilms, L., and Oberfeld, D. (2018). Color and emotion: Effects of hue, saturation, and brightness. *Psychol. Res. Psychol. Forschung* 82, 896–914. doi: 10.1007/s00426-017-0880-8

Witthoft, N., Winawer, J., and Eagleman, D. M. (2015). Prevalence of learned grapheme-color pairings in a large online sample of synesthetes. *PLoS ONE* 10, e0118996. doi: 10.1371/journal.pone.0118996

Witzel, C., and Franklin, A. (2014). Do focal colors look particularly "colorful"? *J. Opt. Soc. Am. A Optics Image Sci. Vis.* 31, A365–A374. doi: 10.1364/JOSAA.31.00A365

Zelazny, A., and Sørensen, T. A. (2020). Shape-color associations across cultures. *J. Vis.* 20, 678. doi: 10.1167/jov.20.11.678

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 June 2023
DOI 10.3389/fpsyg.2023.1172946

Check for updates

OPEN ACCESS

EDITED BY
Na Chen,
Bar-Ilan University, Israel

REVIEWED BY
Vivian Ciaramitaro,
University of Massachusetts Boston,
United States
Ulrich Ansorge,
University of Vienna, Austria

*CORRESPONDENCE
Pi-Chun Huang
✉ pichun_huang@mail.ncku.edu.tw

RECEIVED 24 February 2023
ACCEPTED 16 May 2023
PUBLISHED 05 June 2023

CITATION
Chen Y-C and Huang P-C (2023) Examining the automaticity and symmetry of sound–shape correspondences.
Front. Psychol. 14:1172946.
doi: 10.3389/fpsyg.2023.1172946

Examining the automaticity and symmetry of sound–shape correspondences

Yi-Chuan Chen[1] and Pi-Chun Huang[2]*

[1]Department of Medicine, MacKay Medical College, New Taipei City, Taiwan, [2]Department of Psychology, National Cheng Kung University, Tainan, Taiwan

Introduction: A classic example of sound–shape correspondences is the mapping of the vowel /i/ with angular patterns and the vowel /u/ with rounded patterns. Such crossmodal correspondences have been reliably reported when tested in explicit matching tasks. Nevertheless, it remains unclear whether such sound–shape correspondences automatically occur and bidirectionally modulate people's perception. We address this question by adopting the explicit matching task and two implicit tasks.

Methods: In Experiment 1, we examined the sound–shape correspondences using the implicit association test (IAT), in which the sounds and shapes were both task-relevant, followed by an explicit matching task. In Experiments 2 and 3, we adopted the speeded classification task; when the target was a sound (or shape), a task-irrelevant shape (or sound) that was congruent or incongruent to the target was simultaneously presented. In addition, the participants performed the explicit matching task either before or after the speeded classification task.

Results and Discussion: The congruency effect was more pronounced in the IAT than in the speeded classification task; in addition, a bin analysis of RTs revealed that the congruency effect took time to develop. These findings suggest that the sound–shape correspondences were not completely automatic. The magnitude and onset of visual and auditory congruency effects were comparable, suggesting that the crossmodal modulations were symmetrical. Taken together, the sound–shape correspondences appeared not to be completely automatic, but their modulation was bidirectionally symmetrical once it occurred.

KEYWORDS

crossmodal correspondences, implicit association test, speeded classification task, automatic processing, bidirectional association

1. Introduction

Multisensory signals originating from the same object or event are often correlated rather than arbitrary in some properties, and such crossmodal correspondences provide critical cues for human brains to integrate these signals (e.g., Spence, 2011; Chen and Spence, 2017). For example, a bird's chirping is high-pitched and often associated with its small body, pointy beak, and high elevation on the tree; in contrast, a cow's mooing is low-pitched and often associated with its huge and rounded body and low position on the ground (See Figure 1A; Marks, 1987; Tsur, 2006; Lowe and Haws, 2017). In addition to simple features, crossmodal correspondences between complex stimuli have also been well documented, such that the meaningless speech sounds "kiki" and "bouba" are often matched with angular and rounded shapes, respectively, rather than the reverse (see Figure 1B; Ramachandran and Hubbard, 2001; see Köhler, 1929, 1947; Holland and Wertheimer, 1964 for early studies). In the current study, we aimed to

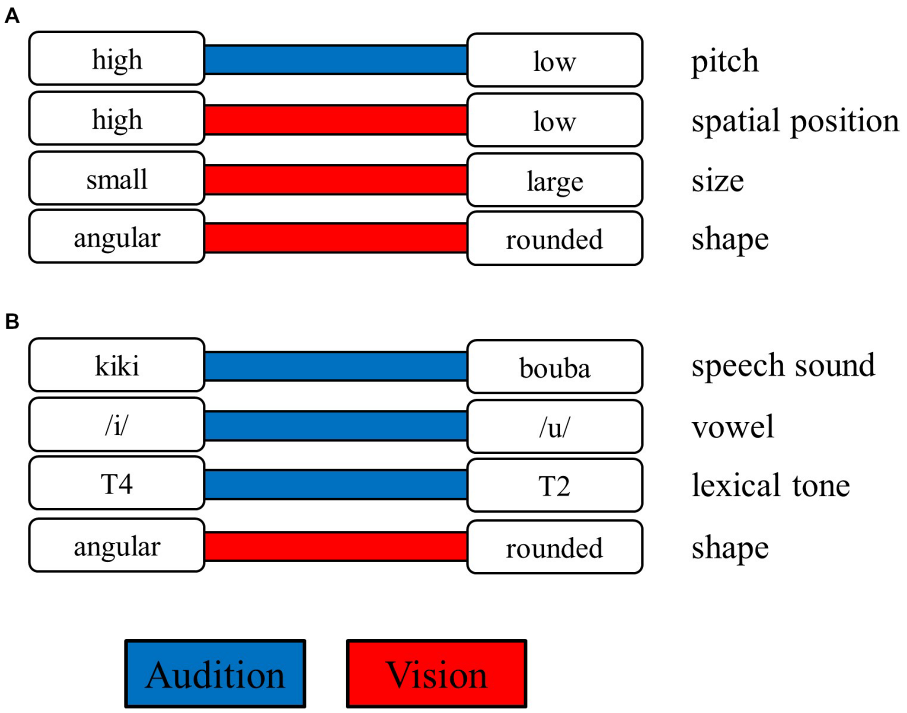

FIGURE 1
Examples of crossmodal correspondences. The contrasting features (in the white rectangles, such as the sound of high vs. low pitch) are connected by a colored bar either in the auditory modality (in blue) or in the visual modality (in red). The features aligned in the left and right columns correspond to each other (e.g., high pitch and high spatial position). **(A)** Illustrates the correspondences associated with pitch, and **(B)** illustrates the correspondence associated with speech sound. T2 and T4 represent lexical tone 2 (consisting of a rising pitch contour) and lexical tone 4 (consisting of a falling pitch contour) in Mandarin (see Figure 2 for details).

investigate whether crossmodal correspondences, specifically the sound–shape correspondences, can occur automatically and symmetrically in influencing the auditory and visual information processing (see Spence and Deroy, 2013, on this issue).

Over the years, researchers have commonly used explicit tasks to explore new crossmodal correspondences. Köhler (1929), for example, explicitly asked participants to match the nonwords "takete" and "baluma" to two line drawings—one with sharp angles and the other with smooth curves. In another study, Sapir (1929) asked participants whether the vowel /a/ symbolized the concept of "large" compared to the vowel /i/. In such *explicit matching tasks*, the participants often only made a single judgment. When a consensual match among this group of participants was reached, the tested correspondences could be confirmed (e.g., Rogers and Ross, 1975; Bremner et al., 2013). Other explicit methods can directly quantify the strength of crossmodal correspondences, such as presenting participants a target stimulus (e.g., sweet liquid) and asking them to rate or adjust a feature along a bipolar dimension in another sensory modality (e.g., a sound's pitch) that best matches the target (e.g., Stevens and Marks, 1965; Crisinel and Spence, 2010; Ngo et al., 2013).

These explicit tasks provide straightforward methods and fruitful outcomes in understanding crossmodal correspondences (e.g., see Spence and Sathian, 2020, for a review). However, the explicit methods have at least three limitations. First, the explicit tasks rely on the participants' introspective ability (see Parise and Spence, 2012; Westbury et al., 2018). Even though Chen et al. (2019) demonstrated that most of their participants who were healthy adults have good introspection when tested with sound–shape correspondences, caution is advised when one explores crossmodal correspondences in people of other ages (such as young children) or special groups (such as those with autistic spectrum disorders, see Spence, 2022, for a review). Second, given that participants are generally being forced to respond in the explicit tasks, the observed associations may be a product of the decisional/response rather than perceptual/cognitive processing. Third, explicit tasks depend on conscious judgments; therefore, the responses are susceptible to individual tendencies, preferences, and intentions. In sum, the responses or judgments in the explicit tasks are prone to be generated from decisional, voluntary, and controlled processing.

To investigate whether crossmodal correspondences occur at the perceptual/cognitive level rather than the decisional/response level, researchers have adopted more sophisticated psychophysical paradigms. One of these tasks is the implicit association test (IAT), which has often been used to measure the strength of associations between two pairs of items or concepts (e.g., the associations between flower–pleasant and insect–unpleasant, Greenwald et al., 1998). In the IAT, a single stimulus is presented in each trial and participants have to categorize the stimulus rapidly and accurately by pressing one of two predesignated response keys (often using their left and right hands). Within an experimental block of trials, the mappings between the stimuli and response keys were either congruent (e.g., "flower" and "pleasant" correspond to the same response key) or incongruent (e.g., "flower" and "unpleasant" correspond to the same response key). It is then assumed that if participants intrinsically associated the items as expected, the response time (RT) would be shorter and/or the accuracy would be higher in the congruent than in the incongruent condition (called the congruency effect). Furthermore, the magnitude of the congruency effect can be a quantitative index of the strength of associations. Taken together, in the IAT, participants judged each item on its merits rather than the associations between items. Therefore, in the IAT, crossmodal correspondences are implicitly encoded in the stimulus–response mappings rather than serving as explicit task goals.

Parise and Spence (2012) utilized the IAT to investigate the core level of processing on five different pairs of crossmodal correspondences. They reported reliable and similar magnitudes of congruency effect regardless of the stimuli's complexity (e.g., pitch–size correspondences rely on simpler stimuli than sound–shape correspondences). By conducting bin analysis of RTs (i.e., the RTs were ordered from the fastest to slowest and then evenly divided into five time bins; see Data Analysis section), Parise and Spence demonstrated that the congruency effect appeared in the initial time bin. Therefore, the rapid onset of the congruency effect suggested that crossmodal correspondences can even occur at the early stage of information processing. Parise and Spence further analyzed the congruency effects in the trials containing a visual or auditory target and demonstrated that their magnitudes were similar, suggesting that the influence in terms of crossmodal correspondences from one to the other modality was bidirectional and symmetrical.

The IAT provides an indirect measure to inspect the role of crossmodal correspondences at the perceptual/cognitive level of information processing. However, the congruency effect observed in the IAT could not unequivocally indicate an *implicit* and *automatic* association. For example, because the congruent and incongruent conditions were tested in different blocks, participants may be aware of the relationships between the stimuli; therefore, the congruency effect reflects implicit and explicit associations or attitudes (see Schimmack, 2021; though see Kurdi et al., 2021). It is also plausible that given the repeated presentations of the stimuli and the well-practiced stimulus–response mappings, certain associations may be learned during the course of the experiment (Hussey and De Houwer, 2018). Therefore, to test whether crossmodal correspondences can occur automatically, adopting other psychophysical paradigms is necessary.

Another implicit task that has been used in the research on crossmodal correspondences is the *speeded classification task* (e.g., Marks, 1987; Melara and O'Brien, 1987; Martino and Marks, 1999; Gallace and Spence, 2006; Evans and Treisman, 2010). In this paradigm, participants quickly respond to the target presented in one modality while ignoring the distractor in the other modality, which is task-irrelevant and lacks an assigned response. The target and distractor are either crossmodally correspondent (i.e., congruent) or not (i.e., incongruent), and the two conditions are intermixed randomly within a block. If the RT is shorter and/or accuracy is higher in the congruent than in the incongruent condition, a typical congruency effect is observed, which suggests that the distractor is processed and facilitates or interferes with the target processing.

Gallace and Spence (2006), for example, asked participants to judge whether the target disc is larger or smaller than the preceding standard disc as rapidly and accurately as possible, and a task-irrelevant tone either in a higher pitch (4,500 Hz) or lower pitch (300 Hz), or else no tone, was simultaneously presented with the target disc. The participants' RT was shorter when the target was paired with a congruent tone (e.g., a smaller disc and the higher-pitched tone) than when paired with an incongruent tone (e.g., a smaller disc and the lower-pitched tone), demonstrating that pitch–size correspondences had occurred. Evans and Treisman (2010) similarly utilized the speeded classification tasks to investigate the crossmodal correspondences' bidirectional influences. In each trial, the researchers presented a tone (either a lower or higher pitch) and a visual disc (either large or small) and asked the participants to discriminate the pitch as either high or low while ignoring the visual disc (i.e., the auditory task) or to discriminate the disc as either large or small while ignoring the pitch (i.e., the visual task). Evans and Treisman found that the congruency effect was significant in the auditory and visual tasks; therefore, the modulation in terms of crossmodal correspondences was bidirectional.

In previous studies, we tested the sound–shape correspondences (e.g., the mapping of "kiki" and "bouba" to angular and rounded patterns) using explicit matching tasks (Chen et al., 2016, 2019, 2021; Chang et al., 2021; Shen et al., 2022). Researchers have also tried to use implicit tasks to investigate the level of processing at which sound–shape correspondences occur. For example, a reliable congruency effect has been reported in the IAT (Parise and Spence, 2012; Shang and Styles, 2023). In addition, the presentation of a congruent as compared to an incongruent sound enhanced the visibility of a masked visual pattern (as indicated by the lower contrast threshold, Hung et al., 2017). These results suggest that the sound–shape correspondences occurred at the perceptual/cognitive level rather than the decisional/response level. In the present study, we aimed to investigate the automaticity and symmetry of the sound–shape correspondences using two implicit tasks, the IAT and speeded classification tasks, and the explicit matching tasks.

According to early research comparing automatic and controlled processing, the former is suggested to be involuntary, capacity-unlimited, and pre-attentive (Schneider and Shiffrin, 1977). Moors and De Houwer (2006) proposed similar criteria for automatic processing—unintentional, efficient, and unconscious, and added the fourth criterion of speed; that is, an automatic process should occur quickly and influence the early stage of information processing. These four criteria are conceptually distinct, allowing them to be investigated individually. In addition, some of the criteria change in degrees rather than qualitatively. According to these propositions, the transition from automatic to controlled processing should occur gradually because there is no single dichotomous criterion (Moors and De Houwer, 2006). Following this gradual approach of automatic processing, we compared the congruency effect in the IAT and speeded classification tasks in terms of the aforementioned criteria.

In Experiments 1, 2, and 3, each participant performed the IAT, sound classification, or shape classification as the main task, respectively, either preceding or following the explicit matching tasks. Considering that RT was the main dependent variable in these implicit tasks, the presentation duration of visual and auditory stimuli should be equalized. Therefore, we utilized the vowels as the auditory stimuli: the vowels /i/ and /u/ are reliably matched to angular and rounded patterns, respectively (Sapir, 1929; Newman, 1933; Taylor and Taylor, 1962; Remez, 2003; Berlin, 2006; Maurer et al., 2006; Kovic et al., 2010; Lockwood and Dingemanse, 2015; Peiffer-Smadja and Cohen, 2019). The *automaticity* of the sound–shape correspondences was examined using the following tests in the implicit tasks: (1) According to the involuntariness (or unintentionality) criterion, the congruency effect should be similar in magnitude in the IAT and speeded classification tasks even though both modalities are task-relevant in the IAT whereas only the target modality was task-relevant in the speeded classification tasks. (2) According to the pre-attentiveness (or unconsciousness) criterion, the congruency effect should be significant irrespective of the order of the explicit matching and implicit tasks. This is because performing the explicit matching tasks would bring the concept of sound–shape correspondences to participants' consciousness; however, an automatic process should occur irrespective of participants' awareness of the correspondences. (3) According to the speed criterion, the congruency effect should emerge when the RT is short (e.g., in the first time bin in the bin analysis of RTs, see Parise and Spence, 2012). The *symmetry* of the sound–shape correspondences was examined using the following tests in the implicit tasks: (1) the magnitude of auditory and visual congruency effect should be similar, and (2) the auditory and visual congruency effect should emerge in similar time bins. These comparisons would provide detailed examinations of the automaticity and symmetry issues.

2. General methods

2.1. Participants

One hundred and twenty participants (24 participants in each experiment) took part in this study. Table 1 provides the participants' demographic information. All of the participants had normal or corrected-to-normal vision and normal hearing by self-report. They were naïve to the purpose of the study and gave their informed consent before the experiment. All of the participants were students at National Cheng Kung University in Taiwan and native Mandarin Chinese speakers. They received monetary compensation for their participation. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki and approved by the National Cheng Kung University research ethics committee for human behavioral sciences (NCKU HREC-E-110-460-2).

To demonstrate the congruency effect of RT in the IAT and speeded classification tasks, we computed the *d-score*, a standardized difference between the RTs in the congruent and incongruent conditions. A d-score greater than zero indicates that a typical congruency effect was significant. An analysis using G*Power (version 3.1.9.4; Faul et al., 2007) for the one-sample *t*-test (one-tailed) suggested that when we test 24 participants with $\alpha = 0.05$ and a power = 0.80, the effect size of the *t*-test would reach 0.52 (a medium effect). The sample (24 participants) was also larger than those in previous studies testing crossmodal correspondences using the IAT ($N = 10$, Parise and Spence, 2012) and speeded classification tasks ($N = 10$ or 15, Gallace and Spence, 2006). We were therefore confident that the number of participants provided adequate effect size to reach statistical significance.

2.2. Apparatus and stimuli

The experimental program was controlled by a computer compatible with the Psychophysics Toolbox (Brainard, 1997) in the MATLAB (The Mathworks, Matick, MA, USA) environment. The visual stimuli were presented on a liquid crystal display monitor

TABLE 1 Demographic information in three experiments.

Experiment	Main task	Explicit matching task (before/ after main task)	Mean age (SD)	Gender (M/F)
1	Implicit association test (IAT)	After	20.4 (1.62)	(12/12)
2A	Sound classification	Before	21.4 (1.86)	(10/14)
2B		After	20.4 (1.57)	(12/12)
3A	Shape classification	Before	21.5 (2.55)	(12/12)
3B		After	20.9 (2.06)	(13/11)

(ASUS XG250) with a resolution of 1,920 × 1,080 pixels, running at 85 Hz. The viewing distance was 100 cm. The auditory stimuli were delivered by closed-ear headphones (Beyer dynamic DT770 Pro). The volume was set to 68 dB. The experiments were conducted in a soundproof, dim room.

The auditory stimuli were two pairs of tonal vowels to increase the variety of auditory stimuli. The first pair was the high front vowel /i/ and the high back vowel /u/, and both were articulated in Tone 1 (i.e., /i1/ and /u1/). The second pair consisted of the same vowels, and /i/ was articulated in Tone 4, and /u/ was articulated in Tone 2 (i.e., /i4/ and /u2/). Each pair of tonal vowels was presumably matched to angular and rounded patterns, respectively (see Figure 1B; Chang et al., 2021). These tonal vowels were produced by a female native speaker who teaches Mandarin Chinese pronunciation at the Chinese Language Center at National Cheng Kung University. These tonal vowels were recorded in a soundproof recording room using a Yeti microphone (Blue, CA, USA) and QuickTime Player (ver. 10.5, Apple Inc., CA, USA) at a 44.1-kHz sampling rate with 16-bit encoding. The audio files were then edited and trimmed using Audacity (ver. 2.3) so that each sound clip's duration was 300 ms. To increase the variety of the acoustic properties while preserving the phonetic features, the Voicemod program was used to transform the female voice into a voice with higher or lower frequencies, resulting in 12 sounds. The volumes of these 12 sounds were equalized to have the same root mean square contrast. The temporal contours of the fundamental frequencies (F0s) of the 12 sounds (/i1/, /u1/, /i4/, and /u2/) as a function of time (ms) are plotted in Figure 2A. Figure 2B shows the mean and standard deviation across the time of the F0, the first formant (F1), the second formant (F2), and the third formant (F3) of the 12 sounds.

Three pairs of angular and rounded visual patterns were adopted from previous studies (Köhler, 1929; Ramachandran and Hubbard, 2001; Bremner et al., 2013). The size of each pattern was 9.3° × 9.3° in visual angle, and each pattern was presented in one of four orientations. The two response keys were labeled using color stickers to ensure that the letters' shapes would not influence the participants' performance.

2.3. Explicit matching tasks

All of the participants conducted the explicit matching tasks either before or after the main experiment. There were two types of matching

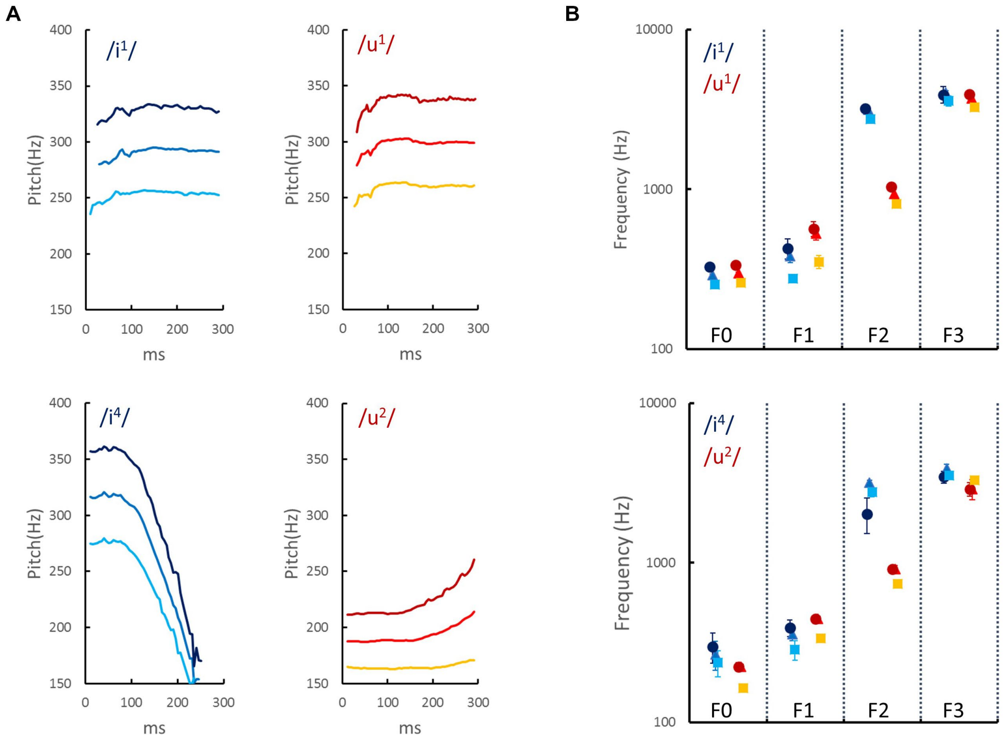

FIGURE 2
Acoustic properties of auditory stimuli used in the current study. **(A)** The temporal contours of the fundamental frequencies (F0s) of two vowels (/i/ and /u/) that were articulated in lexical tone 1 (/i1/ and /u1/ in the upper panels) or in lexical tones 4 and 2 (/i4/ and /u2/ in the lower panels). The bright, intermediate, and dark colors represent voices in the low, medium, and high frequencies, respectively. **(B)** The mean and standard deviation across time of the F0, the first formant (F1), the second formant (F2), and the third formant (F3). The acoustic properties were measured using Praat (Boersma, 2001). The frequencies of formants are associated with the articulation method: F1 increases with lower tongue position and greater jaw opening, and F2 is related to vowel frontness (e.g., the front vowel /i/ has a higher F2 than the back vowel /u/; Ladefoged and Disner, 2012).

tasks. In the sound matching task, in each trial, a tonal vowel and a pair of angular and rounded patterns were presented on the monitor side by side, randomly on the left and right. The participants had to determine which of the two patterns better matched the tonal vowel. The visual patterns disappeared after the participants responded, and the next trial started after 1,000 ms. There were 36 trials (4 tonal vowels * 3 voices * 3 pairs of visual patterns) presented in a randomized order.

In the shape matching task, in each trial, we presented a visual pattern in the center of the monitor and then a pair of sounds (/i1/ and /u1/, or /i4/ and /u2/ in a low, medium, or high frequency) in a random order. The participants had to determine which of the two sounds better matched the visual pattern. The visual pattern disappeared after the participants responded, and the next trial started after 1,000 ms. There were 36 trials (6 visual shapes * 2 pairs of tonal vowels * 3 voices), and their presentation order was randomized. We presented the two matching tasks in a randomized order across participants. The procedure took approximately 10 min to complete.

3. Experiment 1: implicit association test

In Experiment 1, the participants performed the IAT. In every trial, they had to discriminate the sounds as either vowel /i/ or /u/ or the visual patterns as either angular or rounded rapidly and accurately by pressing predesignated keys.

3.1. Design and procedure

The IAT consisted of two sessions, one to test the sound pair /i1/ and /u1/ and the other to test the sound pair /i4/ and /u2/. We conducted these two sessions in an evenly distributed order between the participants. Each session consisted of five blocks: three for training (blocks 1, 2, and 4) and two for the main experiment (blocks 3 and 5). In block 1, the participants practiced to rapidly categorize visual patterns by pressing predesignated keys (e.g., angular shape: left key, rounded shape: right key). In block 2, they practiced to rapidly discriminate the auditory stimuli by pressing the same predesignated keys (e.g., /i1/: left key, /u1/: right key). In block 3, we measured the RTs of congruent audiovisual mappings (e.g., angular shape and /i1/: left key, rounded shape and /u1/: right key). In block 4, participants practiced with the visual patterns again, but we swapped the stimulus–response mappings from block 1 (e.g., rounded shape: left key, angular shape: right key). Finally, we used block 5 to measure the RTs of incongruent audiovisual mappings (e.g., rounded shape and /i1/: left key, angular shape and /u1/: right key). Throughout the experiment, the vowel /i/ always corresponded to the left key and the vowel /u/ always corresponded to the right key regardless of lexical tone. We balanced the presentation order of blocks 1 and 3 and blocks

4 and 5 between participants to control for the potential effect of practicing the IAT.

In the visual and auditory training blocks, we presented six patterns or six sounds (2 tonal vowels * 3 voices) 6 times, resulting in 36 trials. In the congruent and incongruent blocks, we presented each pattern and sound 12 times, resulting in 144 trials. To ensure that the terminology did not prime participants regarding crossmodal association, we showed them two categories of patterns instead of using verbal description when delivering the experimental instruction. At the beginning of each block, the response key (left or right) for each category of patterns and/or sounds was assigned.

In each trial, one stimulus–either a pattern or sound–was presented on the screen or through the headphones. Each stimulus was presented for 300 ms. When the auditory stimulus was presented, a fixation cross was presented on the screen. The participants had to categorize the presented stimulus as quickly and correctly as possible by pressing the corresponding key. The participants were instructed to press the left and right keys (A and L keys on the keyboard) using their left and right index finger, respectively. In the training blocks, an X was presented in the center of the screen when the response was incorrect; however, there was no feedback in the congruent and incongruent blocks. There was no time limit for participants to respond. The IAT task took around 60 min to complete.

3.2. Data analysis

We analyzed the data in the congruent and incongruent blocks. For each participant, we defined the trials with RT less than 100 ms or greater than mean plus three standard deviations (SDs) as outliers and excluded them from further analyses. In each experiment, however, there was no trial with RT shorter than 100 ms, so the excluded trials were all attributed to long responses (i.e., longer than mean + 3SDs). We report the accuracy, RT, and their associated analyses including lexical tones as a within-participant factor in Supplementary materials. In the main text, we combined the two sessions of vowels with different lexical tones.

To quantify the magnitude of the congruency effects in terms of cross-modal correspondences, we computed the *d-score* by taking the RT difference (subtract the mean RT in the congruent condition from the mean RT in the incongruent condition) and then dividing it by the pooled standard deviation. Therefore, the larger the RT difference or the smaller the RT variability, the larger the d-score. Given the prior assumption that the RT should be longer in the incongruent than in the congruent condition, we conducted a one-sample *t*-test (one-tailed) to determine whether the d-score was significantly greater than zero, indicating a significant congruency effect. We computed the d-score for each stimulus modality (visual or auditory), which we would then use to compare the congruency effect across experiments. We compared the visual and auditory d-scores using a paired *t*-test to investigate the symmetry of sound–shape correspondences.

To investigate the stage of information processing at which sound–shape correspondences occurred, we ran a bin analysis of RTs (see De Jong et al., 1994; Vallesi et al., 2005; Parise and Spence, 2012) by separating the RT data in each modality and congruency condition into five time bins, from fastest to slowest, for each participant. We then used the RTs in each time bin to compute the d-score in each time bin. We initially compared these d-scores to zero using one-sample *t*-tests; however, because the d-scores in the five time bins were related rather than independent, we set the criterion of statistical significance at $p < 0.01$ to avoid inflating α. To compare the congruency effect between modalities across time bins, we submitted the d-scores to a two-way repeated-measure ANOVA on the within-participant factors of modality (visual and audition) and bin (1 to 5). We used a Greenhouse–Geisser correction when the sphericity assumption was violated. *Post hoc* comparisons for time bins were pairwise *t*-tests with Bonferroni corrections. The data in Experiments 2 and 3 underwent the same analysis procedure, but we obtained only auditory d-scores and visual d-scores, respectively.

3.3. Results

The overall accuracy was 96.61 ± 0.43% (mean ± SE). We removed the 3.39% incorrect trials and 2.30% trials as RT outliers from the computations of d-scores. The auditory and visual d-scores in the IAT demonstrated reliable congruency effects (Figures 3A,B, respectively): auditory d-score = 0.16 ($SE = 0.05$, $t(23) = 2.99$, $p = 0.003$, *Cohen's* $d = 0.61$) and visual d-score = 0.21 ($SE = 0.08$, $t(23) = 2.52$, $p = 0.01$, *Cohen's* $d = 0.51$). Furthermore, the visual and auditory d-scores were similar ($t(23) = 0.80$, $p = 0.43$, *Cohen's* $d = 0.16$), suggesting a similar magnitude of congruency effect between vision and audition.

When we computed the d-scores individually in each time bin, the results demonstrated that the auditory d-scores were greater than zero in the third, fourth, and fifth bins (all $t(23) > 2.59$, $ps \leq 0.008$, *Cohen's* $d > 0.52$) but not in the first two bins (both $t(23) < 2.05$, $ps \geq 0.025$, *Cohen's* $d < 0.42$); the visual d-scores were greater than zero in the second and third bins (both $t(23) > 2.61$, $ps \leq 0.008$, *Cohen's* $d > 0.53$) but not in the first, fourth, and fifth bins (all $t(23) < 2.34$, $ps \geq 0.014$, *Cohen's* $d < 0.48$). We submitted these d-scores to a two-way ANOVA on the factors of modality and bin. The main effect of bin was significant ($F(2.07, 47.63) = 5.06$, $p = 0.01$, $\eta_p^2 = 0.18$). The *post hoc* tests showed that the d-score was significantly lower in the first bin than in the second and third bins (both $t(23) > 3.12$, $ps \leq 0.048$, *Cohen's* $d > 0.63$). However, the main effect of modality ($F(1, 23) = 1.08$, $p = 0.31$, $\eta_p^2 = 0.05$) and the interaction between modality and bin ($F(1.79, 41.07) = 1.05$, $p = 0.38$, $\eta_p^2 = 0.04$) were not significant. Taken together, the d-score was not significant in the first bin, and it was significantly smaller than in the central bins, and therefore did not support an early onset of the congruency effect. The effect of modality did not reach a significant level, suggesting a symmetrical modulation of crossmodal correspondences (Evans and Treisman, 2010; Parise and Spence, 2012).

4. Experiments 2A and 2B: sound classification task

In Experiments 2A and 2B, the participants had to discriminate the sounds as either the vowel /i/ or /u/ rapidly and accurately irrespective of the lexical tones; meanwhile, a task-irrelevant visual pattern was presented as a distractor. In Experiment 2A, the participants performed the explicit matching tasks *before* the sound classification task (called the $Match_{bef}$ condition hereafter); in

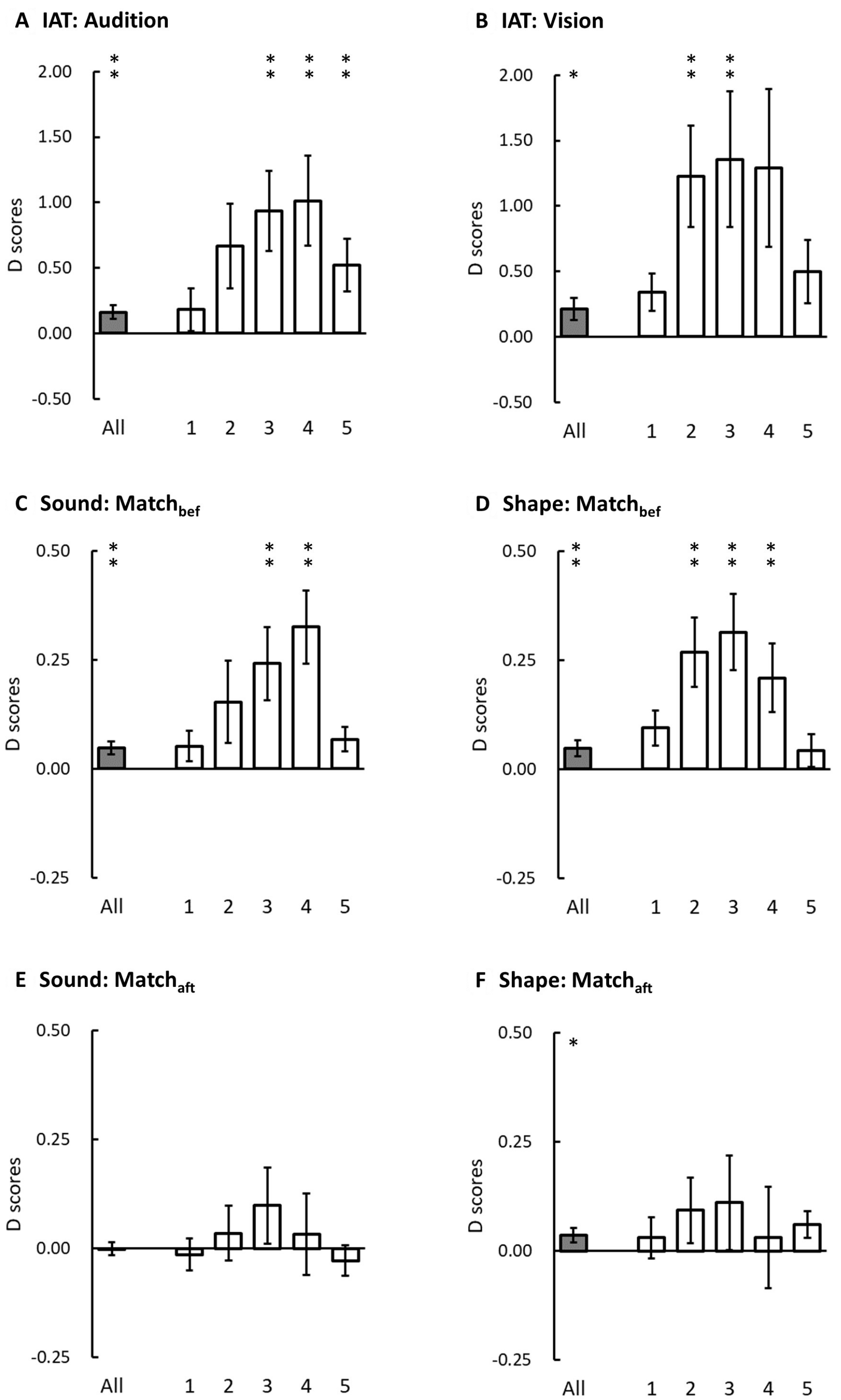

FIGURE 3
Overall d-scores (grey bars) and the d-scores in each time bin (white bars) in three experiments. The left panels are auditory d-scores and the right panels are visual d-scores. **(A,B)** present auditory and visual d-scores in the IAT in Experiment 1. **(C,E)** present auditory d-scores in the sound classification task when the participants performed the explicit matching tasks either before ($Match_{bef}$) or after ($Match_{aft}$) the main task in Experiments 2A and 2B. **(D,F)** present the visual d-scores in the shape classification task when the participants performed the explicit matching tasks before ($Match_{bef}$) or after ($Match_{aft}$) the main task in Experiments 3A and 3B. The error bars represent one standard error of the mean. The asterisks indicate statistical difference from zero (* denotes p<0.05 and ** denotes p<0.01 before Bonferroni corrections; note that the significant criterion for the d-scores in each time bin was set at p<0.01).

Experiment 2B, the participants performed the explicit matching tasks *after* the sound classification task (called $Match_{aft}$ condition hereafter).

4.1. Design and procedure

The sound classification task comprised two sessions, one to test the /i1/ and /u1/ pair and the other to test the /i4/ and /u2/ pair. We evenly distributed the order of these two sessions between participants.

In each trial, the sound was presented for 300 ms; meanwhile, either a congruent visual shape (e.g., an angular shape for /i1/) or an incongruent visual shape (e.g., a rounded shape for /i1/) was presented. Participants had to categorize the sound as /i/ or /u/ (irrespective of the lexical tones) as rapidly and correctly as possible by pressing predesignated keys (A and L keys on the keyboard, counterbalanced between participants) using the left and right index finger, respectively. There was no time limit for participants to respond. After the response, an inter-trial interval (ITI) ranging from 500 ± 50 ms was presented and followed by the next trial. No clue was provided at the beginning of each trial, so the participants had to concentrate on the task continuously. Each session consisted of 400 trials, and the participants had a short break after every 100 trials. The participants practiced a few trials and ensured that they understood the task before the main experiment. The experiment took around 40 min to complete.

4.2. Results

In Experiment 2A (the $Match_{bef}$ condition, Figure 3C), the overall accuracy was 97.68 ± 0.34%. There were 2.32% incorrect trials and 2.13% of trials as RT outliers removed from the computation of d-scores. The auditory d-score was 0.05 ($SE = 0.02$), which was significantly greater than zero ($t(23) = 3.12$, $p = 0.002$, *Cohen's* $d = 0.64$), demonstrating a significant congruency effect induced by visual distractors. When we separately computed the d-scores in the five time bins, the d-score was significantly greater than zero in the third and fourth bins (both $t(23) > 2.88$, $ps \leq 0.004$, *Cohen's* $d > 0.58$) but not in the first, second, or fifth bins (all $t(23) \leq 2.45$, $ps \geq 0.011$, *Cohen's* $d < 0.50$). We submitted the auditory d-scores in the five time bins to a one-way ANOVA on the factor of bin. The main effect was significant ($F(2.29, 52.68) = 5.80$, $p = 0.004$, $\eta_p^2 = 0.20$). The *post hoc* tests showed that the d-score was significantly smaller in the first and fifth bins than in the fourth bin (both $t(23) > 3.44$, $p \leq 0.022$, *Cohen's* $d > 0.70$).

In Experiment 2B (the $Match_{aft}$ condition, Figure 3E), the overall accuracy was 97.43 ± 0.35%. There were 2.57% incorrect trials and 1.89% of trials as RT outliers removed from the computation of d-scores. The overall d-score was −0.0003 ($SE = 0.01$), which was not significantly different from zero ($t(23) = -0.02$, $p = 0.49$, *Cohen's* $d = 0.004$). Consistently, none of the d-scores in the five time bins were significantly different from zero (all $t(23) \leq 1.14$, $ps > 0.13$, *Cohen's* $d < 0.23$).

To investigate the influences of the order of the explicit matching tasks either before or after the sound classification task, we compared the d-scores in Experiments 2A and 2B. The auditory d-score was significantly larger in the $Match_{bef}$ than in the $Match_{aft}$ condition (two-tailed, $t(46) = 2.25$, $p = 0.03$, *Cohen's* $d = 0.65$), suggesting that performing the explicit matching tasks before sound classification task rather than the reversed order induced a larger congruency effect.

5. Experiments 3A and 3B: shape classification task

In Experiments 3A and 3B, the participants had to discriminate a visual pattern rapidly and accurately as belonging to the angular or rounded category; meanwhile, a task-irrelevant sound was presented as a distractor. In Experiment 3A, the participants performed the explicit matching tasks just before the shape classification task (i.e., the $Match_{bef}$ condition); in Experiment 3B, the participants performed the explicit matching tasks after the shape classification task (i.e., the $Match_{aft}$ condition).

5.1. Design and procedure

The shape classification task comprised two sessions. In one session, either the tonal vowel /i1/ or /u1/ was presented as distractor; in the other session, either the tonal vowel /i4/ or /u2/ was presented as distractor. The order of these two sessions was evenly distributed between participants.

In each trial, the shape, either an angular or rounded pattern, was presented in the center of the monitor for 300 ms; meanwhile, either a congruent or incongruent sound (see above) was presented. The participants were required to categorize the shape as an angular or rounded pattern as rapidly and correctly as possible by pressing predesignated keys (A and L keys on the keyboard, evenly distributed between participants) using the left and right index finger, respectively. Other details were the same as in the sound classification task.

5.2. Results

In Experiment 3A (the $Match_{bef}$ condition, Figure 3D), the overall accuracy was 96.10 ± 0.58%. There were 3.90% incorrect trials and 1.94% of trials as RT outliers removed from the computation of d-scores. The visual d-score was 0.05 ($SE = 0.02$), which was significantly greater than zero ($t(23) = 2.60$, $p = 0.008$, *Cohen's* $d = 0.53$). Follow-up bin analysis demonstrated that the visual d-scores in the second, third, and fourth bins were significantly greater than zero (all $t(23) > 2.64$, $ps < 0.007$, *Cohen's* $d > 0.53$) but not in the first and fifth bins (both $t(23) \leq 2.33$, $ps > 0.014$, *Cohen's* $d < 0.48$). The visual d-scores in the five time bins were submitted to a one-way ANOVA on the factor of bin. The main effect was significant ($F(2.58, 59.41) = 4.64$, $p = 0.008$, $\eta_p^2 = 0.17$). The *post hoc* tests showed that the d-score was significantly smaller in the first than in the second bin ($t(23) = 3.18$, $p = 0.04$, *Cohen's* $d = 0.64$).

In Experiment 3B (the $Match_{aft}$ condition, Figure 3F), the overall accuracy was 95.63 ± 0.56%. There were 4.37% incorrect trials and 1.72% of trials as RT outliers removed from the computation of d-scores. The visual d-score was 0.04 ($SE = 0.02$), which was significantly greater than zero ($t(23) = 2.15$, $p = 0.02$, *Cohen's* $d = 0.44$). However, the visual d-scores in the five time bins did not show a significant difference from zero (all $t(23) \leq 2.00$, $ps > 0.029$, *Cohen's* $d < 0.41$). The visual d-scores did not show a significant difference between the $Match_{bef}$ and $Match_{aft}$ conditions (two-tailed, $t(46) = 0.52$, $p = 0.61$, *Cohen's* $d = 0.15$).

6. Comparisons of experiments 1, 2, and 3

6.1. The magnitude of auditory and visual congruency effect

To compare the magnitude of congruency effect in the sound and shape classification tasks, we submitted the d-scores in Experiments 2 and 3 to a two-way ANOVA on the between-participant factors of modality (auditory or visual) and order ($Match_{bef}$ or $Match_{aft}$). The results showed that the d-score was similar in the $Match_{bef}$ and $Match_{aft}$ conditions (0.05 vs. 0.02, $F(1, 92) = 3.46$, $p = 0.066$, $\eta_p^2 = 0.04$). The main effect of modality was not significant (auditory: 0.02, visual: 0.04, $F(1, 92) = 1.24$, $p = 0.27$, $\eta_p^2 = 0.01$), nor was the interaction between modality and order ($F(1, 92) = 1.15$, $p = 0.29$, $\eta_p^2 = 0.01$). Therefore, the magnitude of congruency effect was similar irrespective of the order of explicit and implicit tasks, and it was similar in the two modalities, consistent with the results in IAT (see Results in Experiment 1).

6.2. Modulation of the congruency effect by task relevance

In Experiment 1, the visual and auditory modalities were task-relevant in the IAT; in Experiments 2 and 3, only the auditory or visual modality was task-relevant in the sound and shape classification tasks, respectively. By comparing the d-scores obtained from the IAT to those of the speeded classification tasks, we could determine whether the magnitude of the congruency effect depended on the task relevance of sensory modalities. To be fair, we only compared the experiments in which the explicit matching tasks were performed *after* the main experiment (i.e., Experiments 1, 2B, and 3B). The auditory d-score was higher in the IAT than in the sound classification task (two-tailed, $t(46) = 2.89$, $p = 0.006$, *Cohen's* $d = 0.83$). Similarly, the visual d-score was higher in the IAT than in the shape classification task (two-tailed, $t(46) = 2.06$, $p = 0.045$, *Cohen's* $d = 0.59$). Therefore, the congruency effect was more pronounced when both modalities rather than only one were task-relevant.

6.3. The results of the explicit matching tasks

The matching tendency in the explicit matching tasks, as represented by the proportions that participants matched each tonal vowel to rounded patterns or the proportion that they matched each visual pattern to the vowel /u/, are reported in Figure 4. To examine any consensual matching in each condition, the proportion of matching was compared to 50% using a one-sample *t*-test (two-tailed).

In the sound matching task (upper panels in Figure 4), the participants tended to match the vowel /i/ to angular patterns and the vowel /u/ sound to rounded patterns for most of the tonal vowels ($t(23) \geq 2.07$, $ps \leq 0.05$, *Cohen's* $d > 0.42$) except the /i1/ in Experiments 2A ($t(23) = 1.67$, $p = 0.26$, *Cohen's* $d = 0.34$) and 3A ($t(23) = 0.25$, $p = 0.80$, *Cohen's* $d = 0.05$); here both were the condition where the explicit matching tasks had been performed before the main experiment. Even though the participants' matching between /i1/ and angular pattern was ambiguous, the contrast to the matching between /u1/ and rounded pattern remained significant (Experiment 2A: $t(23) = 7.33$, $p < 0.001$, *Cohen's* $d = 1.50$; Experiment 3A: $t(23) = 4.58$, $p < 0.001$, *Cohen's* $d = 0.93$). The result that the vowel /i1/ was reliably matched to angular patterns only in the $Match_{aft}$ condition (Experiments 1, 2B, and 3B) may suggest that the repeated presentation of the tonal vowels and visual patterns in the preceding main task would improve the reliability of the sound-shape correspondences when the participants responded in the explicit matching tasks.

In the shape matching task (lower panels in Figure 4), the results showed that the participants reliably match rounded patterns to /u/ while matching angular patterns to /i/ (all $t(23) \geq 6.13$, $ps \leq 0.001$, *Cohen's* $d > 1.25$). Compared to the sound matching task, the judgments in the shape matching seemed more reliable and insensitive to the prior exposure of these stimuli.

7. Discussion

In the present study, we examined the automaticity and symmetry of a classic example of crossmodal correspondences, that is, the sound–shape correspondences between vowels /i/ and /u/, and angular and rounded patterns, respectively. We utilized two implicit tasks, the IAT and speeded classification tasks, and the commonly used explicit matching tasks to explore this issue. In the IAT, in which the participants had to categorize a sound or pattern rapidly and accurately, the significant congruency effect was demonstrated in the auditory and visual modalities. In the sound classification task, we only observed the significant congruency effect induced by visual distractors when the explicit matching tasks were performed before rather than after the sound classification task; in contrast, in the shape classification task, we observed the significant congruency effect induced by auditory distractors when the explicit matching tasks were performed either before or after the shape classification task. Further bin analysis of RTs demonstrated that we did *not* observe the congruency effect in the first bin but observed it in the later bins. Finally, when we compared the IAT and speeded classification tasks, the congruency effect was more pronounced in the IAT, suggesting that the task relevance of modalities is critical for the encoding of sound–shape correspondences. Combined, these results of congruency effects suggest that the sound–shape correspondences tested here did not occur automatically.

The congruency effect of sound–shape correspondences did not demonstrate unequivocal evidence of automaticity. The auditory and visual congruency effects were significant in most experiments (except in Experiment 2B), but the congruency effect was never significant in the first time bin. This result contrasts with the observation of Parise and Spence (2012), who demonstrated the congruency effect in the first time bin and suggested that sound–shape associations were encoded automatically according to the speed criterion. In addition, the congruency effect was more pronounced when the auditory and visual modalities were task-relevant in the IAT than when only one

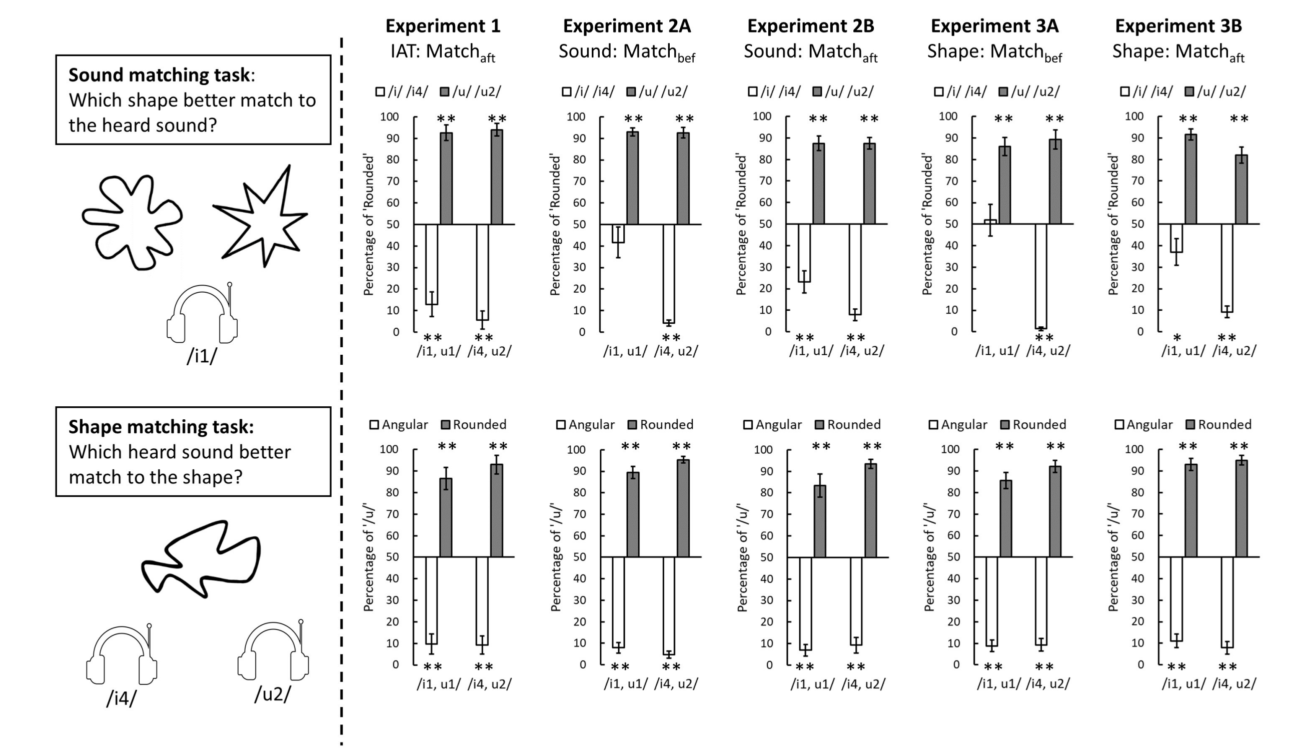

FIGURE 4
The results of explicit matching tasks in three experiments. The sound matching performance is presented in the upper panels and the shape matching performance in the lower panels. The white bars represent the tonal vowels /i1/ or /i4/ in the sound matching task or angular patterns in the shape matching task; the grey bars represent the tonal vowel /u1/ or /u2/ in the sound matching task or rounded patterns in the shape matching task. The error bars represent one standard error of the mean (* denotes $p<0.05$, ** denotes $p<0.01$ of the result of a one-sample t-test compared to 50%).

modality was task-relevant in the speeded classification tasks. This result suggests that encoding the sound–shape correspondences during information processing depended on whether visual and auditory modalities were related to the goal at hand to a certain extent, thereby violating the involuntariness (unintentionality) criterion. These two pieces of evidence therefore do not support the automatic processing of sound–shape correspondences. Note that the auditory congruency effect was only significant when the explicit matching tasks were performed before (Experiment 2A) rather than after (Experiment 2B) the sound classification task; however, we did not observe this order difference in the visual congruency effect (Experiments 3A and 3B). Critically, a further comparison by combining Experiments 2 and 3 did not demonstrate a reliable order difference, either. Therefore, despite the participants' explicit matching of the sound-shape correspondences may facilitate the informativeness of the distractor to target in terms of their congruency, it does not appear to be decisive in the automaticity of the sound-shape correspondences.

Evidence of the symmetrical modulations of sound–shape correspondences in either modality was apparent. When we compared the magnitudes of the auditory and visual congruency effects, we found no significant difference in the IAT or speeded classification tasks. In addition, the results of bin analysis of RTs in the IAT demonstrated a similar pattern of visual and auditory congruency effect across time bins. Taken together, the modulations of sound–shape correspondences were symmetrical in terms of the magnitude and onset of the congruency effects, consistent with Evans and Treisman (2010) and Parise and Spence (2012).

We quantified the congruency effect of sound–shape correspondences by calculating d-scores in the IAT and speeded classification tasks. Even though the magnitude of the congruency effect is presumed to reflect the strength of associations between the sounds and shapes tested in the current study, the underlying mechanism in the two paradigms is substantially different. In the IAT, the participants responded to stimuli in the auditory and visual modalities and tried to associate the sound and shape that were mapped to the same response to speed up the responses and reduce errors. This task-demanded association should be easier to establish in the congruent block, in which the corresponding sounds and shapes were mapped to the same response, than in the incongruent block, in which the contrasting sounds and shapes were mapped to the same response. Therefore, factors that help establish new associations, such as learning and practicing, may influence the magnitude of the congruency effect in the IAT (Greenwald et al., 1998; Hussey and De Houwer, 2018). In the speeded classification task, in contrast, the participants had to respond selectively to the modality in which the target was presented while ignoring (or inhibiting) the other modality in which the distractor was presented; that is, the participants tried to dissociate the simultaneously presented sound and shape. If they failed to inhibit the distractor but processed it to a certain level, it may have facilitated (or interfered with) the processing of the target when they correspond (or contrast).

Therefore, the participant's execution of selective attention or the stimulus's complexity (determining the perceptual load of processing) may influence the congruency effect in the speeded classification tasks (Parise and Spence, 2012; Molloy et al., 2015). Given that the task relevance of sensory modalities and intention to relate the stimuli (associate vs. dissociate) differed in the IAT and speeded classification tasks, utilizing either experimental paradigm alone is perhaps insufficient to examine the automaticity of the crossmodal correspondences.

It should be noted that we do not intend to generalize our conclusion to all types of crossmodal correspondences. In fact, based on our previous studies, we postulate that the sound–shape correspondences occur at the "mid-level" processing, a higher level than the crossmodal correspondences of simple features (such as the pitch–size correspondences in Figure 1A). For example, in the visual modality, the process of perceptual grouping from elements to a global contour is essential in the sound–shape correspondences (Chen et al., 2021). In addition, not only the acoustic features (Knoeferle et al., 2017) but also the phonemes that are perceptually invariant (Shen et al., 2022) seem to drive the sound–shape correspondences. It has been suggested that crossmodal correspondences can occur at multiple levels of information processing (Spence, 2011), and it is plausible that higher-level correspondences may be more susceptible to cognitive factors and therefore more controllable than lower-level correspondences.

In conclusion, we utilized two implicit tasks—IAT and speeded classification tasks—to examine the automaticity and symmetry of sound-shape correspondences in modulating people's speeded behavioral performance. Given that the congruency effects occurred with later responses and they were susceptible to the task relevance of sensory modalities, we suggest that the associations between sounds and shapes were not encoded automatically. In addition, the magnitude and onset of the visual and auditory congruency effects were comparable, suggesting a symmetrical modulation of sound-shape correspondence between sensory modalities once it occurred. The current study demonstrates a methodological solution when one aims to examine the automaticity and symmetry of cross-modal correspondences carefully (see Spence and Deroy, 2013, for a review and discussion).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by National Cheng Kung University research ethics committee for human behavioral sciences (NCKU HREC-E-110-460-2). The patients/participants provided their written informed consent to participate in this study.

Author contributions

P-CH and Y-CC: conception and design of study, interpretation of data, and writing the article. P-CH: acquisition and analysis of data. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Ministry of Science and Technology in Taiwan (MOST 108-2410-H-006-041-MY2 to P-CH and MOST 110-2423-H-715-001-MY3 to Y-CC), and MacKay Medical College (MMC-RD-110-1B-P024 to Y-CC).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1172946/full#supplementary-material

References

Berlin, B. (2006). The first congress of Ethnozoological nomenclature. *J. R. Anthropol. Inst.* 12, S23–S44. doi: 10.1111/j.1467-9655.2006.00271.x

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot Int.* 5, 341–345.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., and Spence, C. (2013). "Bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to westerners. *Cognition* 126, 165–172. doi: 10.1016/j.cognition.2012.09.007

Chang, Y. H., Zhao, M., Chen, Y.-C., and Huang, P. C. (2021). The effects of mandarin Chinese lexical tones in sound-shape and sound-size correspondences. *Multisens. Res.* 35, 243–257. doi: 10.1163/22134808-bja10068

Chen, Y.-C., Huang, P. C., and Spence, C. (2021). Global shape perception contributes to crossmodal correspondences. *J. Exp. Psychol. Hum. Percept. Perform.* 47, 357–371. doi: 10.1037/xhp0000811

Chen, Y.-C., Huang, P. C., Woods, A., and Spence, C. (2016). When "Bouba" equals "Kiki": cultural commonalities and cultural differences in sound-shape correspondences. *Sci. Rep.* 6:26681. doi: 10.1038/srep26681

Chen, Y.-C., Huang, P. C., Woods, A., and Spence, C. (2019). I know that "Kiki" is angular: the metacognition underlying sound–shape correspondences. *Psychon. Bull. Rev.* 26, 261–268. doi: 10.3758/s13423-018-1516-8

Chen, Y.-C., and Spence, C. (2017). Assessing the role of the 'Unity assumption' on multisensory integration: a review. *Front. Psychol.* 8:445. doi: 10.3389/fpsyg.2017.00445

Crisinel, A. S., and Spence, C. (2010). As bitter as a trombone: synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Atten. Percept. Psychophys.* 72, 1994–2002. doi: 10.3758/app.72.7.1994

De Jong, R., Liang, C. C., and Lauber, E. (1994). Conditional and unconditional automaticity: a dual-process model of effects of spatial stimulus-response correspondence. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 731–750. doi: 10.1037/0096-1523.20.4.731

Evans, K. K., and Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *J. Vis.* 10, 6.1–6.12. doi: 10.1167/10.1.6

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Gallace, A., and Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Percept. Psychophys.* 68, 1191–1203. doi: 10.3758/bf03193720

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Holland, M. K., and Wertheimer, M. (1964). Some physiognomic aspects of naming, or, maluma and takete revisited. *Percept. Mot. Skills* 19, 111–117. doi: 10.2466/pms.1964.19.1.111

Hung, S. M., Styles, S. J., and Hsieh, P. J. (2017). Can a word sound like a shape before you have seen it? Sound-shape mapping prior to conscious awareness. *Psychol. Sci.* 28, 263–275. doi: 10.1177/0956797616677313

Hussey, I., and De Houwer, J. (2018). Implicit association test as an analogical learning task. *Exp. Psychol.* 65, 272–285. doi: 10.1027/1618-3169/a000416

Knoeferle, K., Li, J., Maggioni, E., and Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Sci. Rep.* 7:5562. doi: 10.1038/s41598-017-05965-y

Köhler, W. (1929). *Gestalt psychology*. New York: H. Liveright.

Köhler, W. (1947). *Gestalt psychology, an introduction to new concepts in modern psychology*, *2nd edition*. New York: H. Liveright.

Kovic, V., Plunkett, K., and Westermann, G. (2010). The shape of words in the brain. *Cognition* 114, 19–28. doi: 10.1016/j.cognition.2009.08.016

Kurdi, B., Ratliff, K. A., and Cunningham, W. A. (2021). Can the implicit association test serve as a valid measure of automatic cognition? A response to Schimmack (2021). *Perspect. Psychol. Sci.* 16, 422–434. doi: 10.1177/1745691620904080

Ladefoged, P., and Disner, S. F. (2012). *Vowels and consonants*. Hoboken, NJ: John Wiley & Sons.

Lockwood, G., and Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Front. Psychol.* 6:1246. doi: 10.3389/fpsyg.2015.01246

Lowe, M. L., and Haws, K. L. (2017). Sounds big: the effects of acoustic pitch on product perceptions. *J. Mark. Res.* 54, 331–346. doi: 10.1509/jmr.14.0300

Marks, L. E. (1987). On cross-modal similarity: auditory-visual interactions in speeded discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 384–394. doi: 10.1037//0096-1523.13.3.384

Martino, G., and Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: tests of the semantic coding hypothesis. *Perception* 28, 903–923. doi: 10.1068/p2866

Maurer, D., Pathman, T., and Mondloch, C. J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Dev. Sci.* 9, 316–322. doi: 10.1111/j.1467-7687.2006.00495.x

Melara, R. D., and O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *J. Exp. Psychol. Gen.* 116, 323–336. doi: 10.1037/0096-3445.116.4.323

Molloy, K., Griffiths, T. D., Chait, M., and Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *J. Neurosci.* 35, 16046–16054. doi: 10.1523/jneurosci.2931-15.2015

Moors, A., and De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychol. Bull.* 132, 297–326. doi: 10.1037/0033-2909.132.2.297

Newman, S. S. (1933). Further experiments in phonetic symbolism. *Am. J. Psychol.* 45, 53–75. doi: 10.2307/1414186

Ngo, M. K., Velasco, C., Salgado, A., Boehm, E., O'Neill, D., and Spence, C. (2013). Assessing crossmodal correspondences in exotic fruit juices: the case of shape and sound symbolism. *Food Qual. Prefer.* 28, 361–369. doi: 10.1016/j.foodqual.2012.10.004

Parise, C. V., and Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Exp. Brain Res.* 220, 319–333. doi: 10.1007/s00221-012-3140-6

Peiffer-Smadja, N., and Cohen, L. (2019). The cerebral bases of the bouba-kiki effect. *NeuroImage* 186, 679–689. doi: 10.1016/j.neuroimage.2018.11.033

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

Remez, R. E. (2003). Establishing and maintaining perceptual coherence: unimodal and multimodal evidence. *J. Phon.* 31, 293–304. doi: 10.1016/S0095-4470(03)00042-1

Rogers, S. K., and Ross, A. S. (1975). A cross-cultural test of the Maluma-Takete phenomenon. *Perception* 4, 105–106. doi: 10.1068/p040105

Sapir, E. (1929). A study in phonetic symbolism. *J. Exp. Psychol.* 12, 225–239. doi: 10.1037/h0070931

Schimmack, U. (2021). The implicit association test: a method in search of a construct. *Perspect. Psychol. Sci.* 16, 396–414. doi: 10.1177/1745691619863798

Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. detection, search, and attention. *Psychol. Rev.* 84, 1–66. doi: 10.1037/0033-295X.84.1.1

Shang, N., and Styles, S. J. (2023). Implicit association test (IAT) studies investigating pitch-shape audiovisual cross-modal associations across language groups. *Cogn. Sci.* 47:e13221. doi: 10.1111/cogs.13221

Shen, Y. C., Chen, Y.-C., and Huang, P. C. (2022). Seeing sounds: the role of vowels and consonants in Crossmodal correspondences. *Iperception* 13:20416695221084724. doi: 10.1177/20416695221084724

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2022). Exploring group differences in the Crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079

Spence, C., and Deroy, O. (2013). How automatic are crossmodal correspondences? *Conscious. Cogn.* 22, 245–260. doi: 10.1016/j.concog.2012.12.006

Spence, C., and Sathian, K. (2020). "Chapter 11 - audiovisual crossmodal correspondences: behavioral consequences and neural underpinnings" in *Multisensory perception*. eds. K. Sathian and V. S. Ramachandran (Cambridge, MA: Academic Press), 239–258.

Stevens, J. C., and Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proc. Natl. Acad. Sci. U. S. A.* 54, 407–411. doi: 10.1073/pnas.54.2.407

Taylor, I. K., and Taylor, M. M. (1962). Phonetic symbolism in four unrelated languages. *Can. J. Psychol.* 16, 344–356. doi: 10.1037/h0083261

Tsur, R. (2006). Size–sound symbolism revisited. *J. Pragmat.* 38, 905–924. doi: 10.1016/j.pragma.2005.12.002

Vallesi, A., Mapelli, D., Schiff, S., Amodio, P., and Umiltà, C. (2005). Horizontal and vertical Simon effect: different underlying mechanisms? *Cognition* 96, B33–B43. doi: 10.1016/j.cognition.2004.11.009

Westbury, C., Hollis, G., Sidhu, D. M., and Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: distributional properties predict cue strength. *J. Mem. Lang.* 99, 122–150. doi: 10.1016/j.jml.2017.09.006

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 15 June 2023
DOI 10.3389/fpsyg.2023.1190103

OPEN ACCESS

EDITED BY
Philipp Beckerle,
University of Erlangen-Nuremberg, Germany

REVIEWED BY
Merle Theresa Fairhurst,
Dresden University of Technology, Germany
Ioannis Delis,
University of Leeds, United Kingdom

*CORRESPONDENCE
Mattia Pinardi
✉ m.pinardi@unicampus.it

RECEIVED 20 March 2023
ACCEPTED 30 May 2023
PUBLISHED 15 June 2023

CITATION
Pinardi M, Di Stefano N, Di Pino G and Spence C (2023) Exploring crossmodal correspondences for future research in human movement augmentation.
Front. Psychol. 14:1190103.
doi: 10.3389/fpsyg.2023.1190103

Exploring crossmodal correspondences for future research in human movement augmentation

Mattia Pinardi[1]*, Nicola Di Stefano[2], Giovanni Di Pino[1] and Charles Spence[3]

[1]NeXT Lab, Neurophysiology and Neuroengineering of Human-Technology Interaction Research Unit, Università Campus Bio-Medico di Roma, Rome, Italy, [2]Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy, [3]Crossmodal Research Laboratory, University of Oxford, Oxford, United Kingdom

"Crossmodal correspondences" are the consistent mappings between perceptual dimensions or stimuli from different sensory domains, which have been widely observed in the general population and investigated by experimental psychologists in recent years. At the same time, the emerging field of human movement augmentation (i.e., the enhancement of an individual's motor abilities by means of artificial devices) has been struggling with the question of how to relay supplementary information concerning the state of the artificial device and its interaction with the environment to the user, which may help the latter to control the device more effectively. To date, this challenge has not been explicitly addressed by capitalizing on our emerging knowledge concerning crossmodal correspondences, despite these being tightly related to multisensory integration. In this perspective paper, we introduce some of the latest research findings on the crossmodal correspondences and their potential role in human augmentation. We then consider three ways in which the former might impact the latter, and the feasibility of this process. First, crossmodal correspondences, given the documented effect on attentional processing, might facilitate the integration of device status information (e.g., concerning position) coming from different sensory modalities (e.g., haptic and visual), thus increasing their usefulness for motor control and embodiment. Second, by capitalizing on their widespread and seemingly spontaneous nature, crossmodal correspondences might be exploited to reduce the cognitive burden caused by additional sensory inputs and the time required for the human brain to adapt the representation of the body to the presence of the artificial device. Third, to accomplish the first two points, the benefits of crossmodal correspondences should be maintained even after sensory substitution, a strategy commonly used when implementing supplementary feedback.

KEYWORDS

crossmodal correspondence, augmentation, multisensory integration, sensory feedback, embodiment

1. Introduction

1.1. Crossmodal correspondences

"Crossmodal correspondences" are the consistent matchings between perceptual dimensions or stimulus attributes from different sensory domains that are observed in normal observers (i.e., in non-synesthetes; see Spence, 2011, for a review). One of the most famous audiovisual associations was discovered almost a century ago by Köhler (1929). In particular, the German psychologist observed that people tend to associate the term "baluba" with curved lines, while the term "takete" is associated with angular shapes instead. To date, the existence of a very wide range of such crossmodal associations have been demonstrated, involving most, if not all, combinations of senses (see Spence, 2011, for a review). Associations involving auditory and visual stimuli have been predominantly addressed. That said, combinations involving haptics (Slobodenyuk et al., 2015; Wright et al., 2017; Lin et al., 2021) and/or the chemical senses (i.e., taste and gustation) have increasingly been investigated over the last decade, also in relation to the growing interest in research on food and wine pairing, and food marketing/packaging design (North, 2012; Spence and Wang, 2015; Pramudya et al., 2020; Spence, 2020).

Crossmodal associations have primarily been investigated in the field of experimental psychology/perception research, in which it has, on occasion, been related to the wider phenomenon of multisensory integration (Parise and Spence, 2009), conceived of as the processing and organization of sensory inputs received from different pathways into a unified whole (Tong et al., 2020). A number of intriguing questions concerning the nature of crossmodal correspondences and the relative contribution of different factors remains open, with some researchers putting forward the existence of intersensory, or suprasensory stimulus qualities, such as brightness, intensity, roughness, that can be picked-up by multiple senses (e.g., see, von Hornbostel, 1931; Walker-Andrews et al., 1994), and others focusing more on the integration of information occurring in the brain, as a result of crossmodal binding mechanisms (Calvert and Thesen, 2004).

1.2. Multisensory processing and feedback in the context of human augmentation

The mechanisms behind the processing of multisensory stimuli are also of great interest to scientists seeking feedback strategies to be implemented in the context of human motor control. Indeed, multisensory feedback has been shown to improve an individual's ability to command their body significantly, especially in complex interactions requiring constant adjustment and recalibration (Wolpert et al., 1995; Diedrichsen et al., 2010; Shadmehr et al., 2010). The same benefit can potentially also be extended to the control of artificial devices, such as teleoperated robots or prostheses. In the field of prosthetics research, researchers have endowed prostheses with artificial senses, so that users can retrieve sensory information from these artificial devices as they would with their natural limbs, ultimately leading to better control and higher acceptability (Di Pino et al., 2009; Di Pino et al., 2012, 2020; Raspopovic et al., 2014; Marasco et al., 2018; Page et al., 2018; Zollo et al., 2019).

In recent years, the role of multisensory feedback has been studied within the framework of human augmentation, an emerging field that aims at enhancing human abilities beyond the level that is typically attainable by able-bodied users (Di Pino et al., 2014; Eden et al., 2021). Movement augmentation can, for instance, be achieved through the use of supernumerary robotic limbs (SRLs), artificial devices that can be controlled simultaneously with, but independently from, the natural limbs, thus opening up new possibilities in the interaction with the environment. SRLs can be controlled through the movement of body parts that may not be directly involved in the task, actuated by means of joysticks, trackers, or retrieved from the electrical muscle activity (Eden et al., 2021). It turns out that the quality of the motor interface is of pivotal importance as far as achieving proficient performance is concerned. However, proficient control can only be achieved by closing the sensorimotor loop between user and robotic device by means of reliable feedback, possibly covering multiple senses. Indeed, it has recently been demonstrated how supplementary sensory feedback improved the regulation of the force exerted by the SRL's end-effector when interacting with an object (Hussain et al., 2015; Hernandez et al., 2019; Guggenheim and Asada, 2021), as well as increasing the accuracy in reaching a target or replicating the end-effector position (Segura Meraz et al., 2017; Aoyama et al., 2019; Pinardi et al., 2021, 2023). It can also reduce the amount of time required to complete behavioral tasks (Hussain et al., 2015; Sobajima et al., 2016; Noccaro et al., 2020).

Multisensory integration plays a key role in building body representations, the map that our brain uses to recognize and identify our body and the model the brain uses to control its movement (i.e., body image and body schema; de Vignemont, 2010; Blanke, 2012; Moseley et al., 2012). Sensory feedback continuously updates these representations, allowing the brain to accept artificial limbs, as has frequently been demonstrated in the famous Rubber Hand Illusion (Botvinick and Cohen, 1998), where a visible rubber hand and the participant's own hidden hand are brushed synchronously (though see Pavani et al., 2000, for evidence that synchronous stroking is not always required to induce the illusion). As a result of this visuotactile congruency, the artificial limb is included in the body representation and thus perceived as belonging to the body of the person experiencing the illusion. This process, often labeled "embodiment" (Botvinick and Cohen, 1998; Pinardi et al., 2020a), has been reported for several types of artificial limbs, including prostheses, and has been linked to an improvement in artificial limb control and acceptability (Marasco et al., 2011; Gouzien et al., 2017). Recent works suggests that embodiment might be perceived for SRLs as well, making it a relevant topic in the field of human augmentation (Di Pino et al., 2014; Arai et al., 2022; Di Stefano et al., 2022; Umezawa et al., 2022).

1.3. Perspectives for future research in human augmentation

In this paper, we present three novel considerations concerning how future research on human augmentation might benefit from the vast knowledge concerning crossmodal correspondences. First, crossmodal correspondences could make multisensory feedback from the augmenting device easier to process and understand. Second, the spontaneous and widespread mechanisms that characterize crossmodal correspondences might well be expected to reduce the amount of time that is needed for users to learn and decode the supplementary feedback of SRLs and to embody it faster. Third, the benefits of crossmodal correspondences involving two sensory modalities should persist if one modality is transduced but the informative content is maintained (see **Figure 1**).

2. Improving supplementary feedback integration

Crossmodal correspondences modulate multisensory integration (see Spence, 2011, for a review). More specifically, crossmodally congruent multisensory stimuli possess a perceptual bond that leads to increased strength/likelihood of integration thus potentially reducing spatiotemporal discrepancy. Indeed, when participants had to judge the temporal order of auditory and visual stimuli presented in pairs, worse performance was observed with pairs of stimuli that were crossmodally congruent as compared to those that were incongruent (Parise and Spence, 2009).

Crossmodally congruent pairs of audiovisual stimuli give rise to larger spatial ventriloquism effects (i.e., mislocalization of the source of a sound in space toward an irrelevant visual stimulus) compared to incongruent pairs (Parise and Spence, 2008). Crossmodal correspondences can also impact the perceived direction of moving visual stimuli: participants perceived gratings with ambiguous motion as moving upward when coupled with an ascending pitch sound and vice versa (Maeda et al., 2004).

Additional research findings have also demonstrated that crossmodal correspondences can improve behavioral performance in sensorimotor tasks. For instance, in a speeded target discrimination task in which participants were asked whether the pairs of color and sound stimuli that were presented appropriately matched, higher accuracy and faster reaction times were obtained with crossmodally congruent pairs of stimuli as compared to incongruent ones (Marks, 1987; Klapetek et al., 2012; Sun et al., 2018). Recently, audiovisual correspondences were shown to promote higher accuracy and faster responses during a reaction time task when stimuli were congruent rather than incongruent (Ihalainen et al., 2023). Even complex motor tasks, such as drawing, are affected by crossmodal correspondences, with participants obtaining higher accuracy in representing graphically certain features of auditory stimuli (e.g., pitch) compared to others (e.g., loudness) (Küssner and Leech-Wilkinson, 2014). Therefore, a first suggestion is that crossmodal correspondences can be exploited to improve the integration of multisensory information concerning the status of the SRL. According to the Bayesian integration framework (Körding and Wolpert, 2004; Chandrasekaran, 2017), supplementary feedback is particularly useful when it carries information with lower "noise" (i.e., less disturbance in

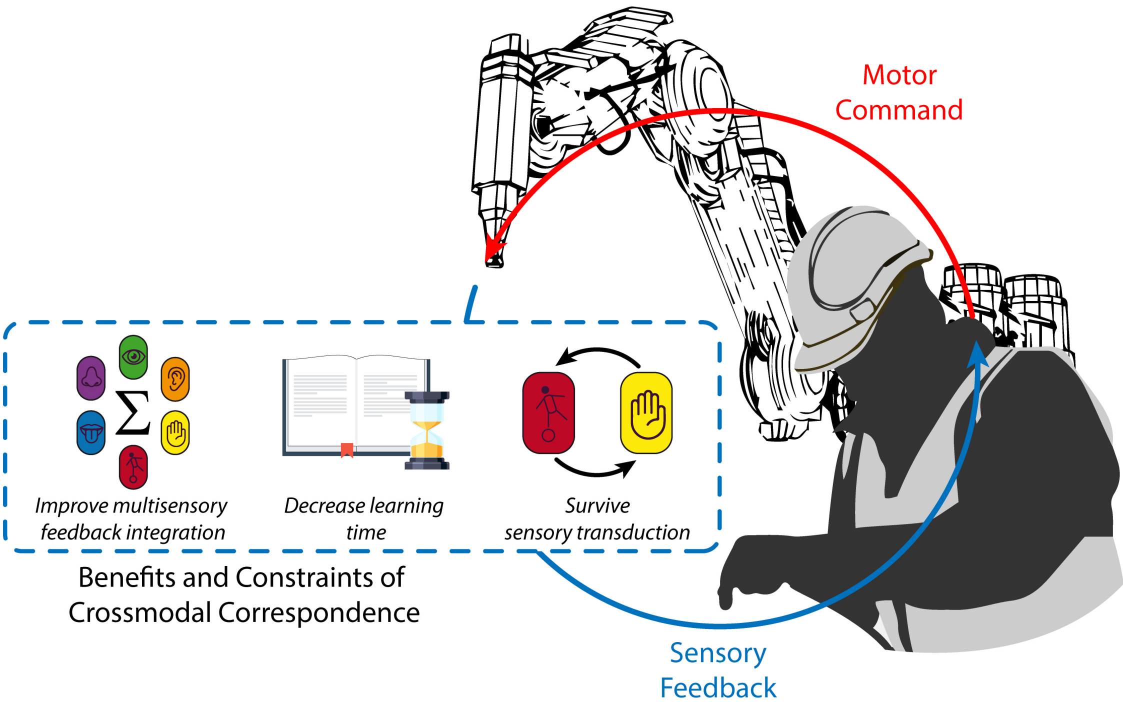

FIGURE 1
Crossmodal correspondences relevant to human augmentation. Relying on crossmodal correspondences when designing the SRL supplementary feedback encoding (blue dotted line) could make a robot status parameter (e.g., its position) relayed through multiple sensory modalities (e.g., somatosensation and vision), more useful and easier to integrate. This, in turn, might help to decrease the amount of time required to learn to efficiently decode SRL feedback. Existing evidence suggests that these benefits should be resistant to sensory modality transduction. As a result, supplementary sensory feedback from the SRL (solid blue line) should be more efficient and motor control (solid red line) and embodiment of the augmenting device should improve.

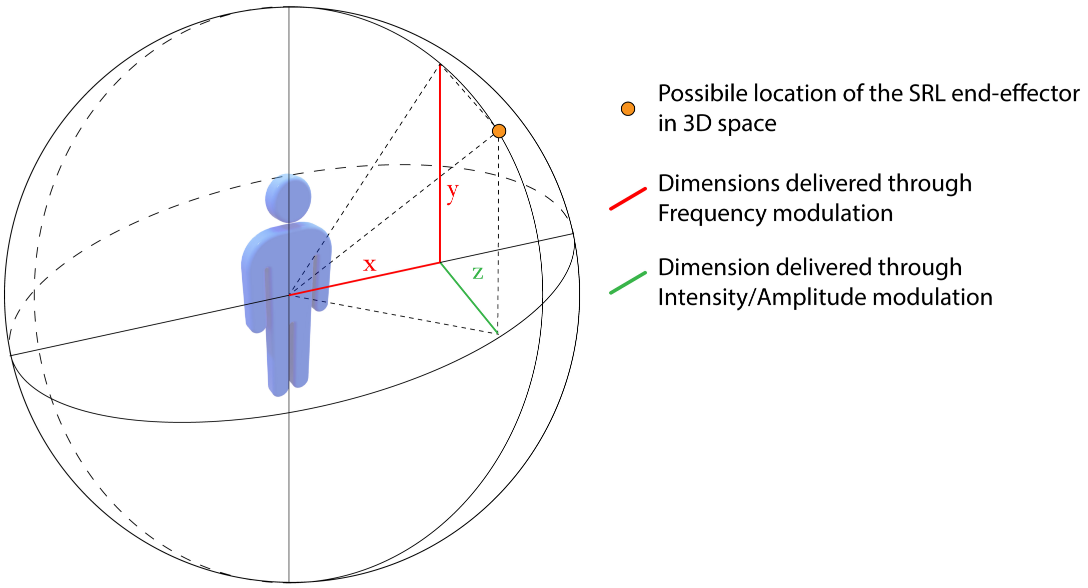

FIGURE 2
Possible encoding strategy that exploit visuo/haptic/auditory crossmodal correspondences to facilitate the localization of an SRL end-effector. 3D space is represented as a sphere surrounding the user, while the possible location of the SRL end-effector is shown as an orange dot. Dimensions that could be delivered by modulating sound (or possibly vibrotactile) frequency are shown in red. Since pitch modulation has been associated with localization on both the *x* and, especially, *y* axis (red), disentangling them while using the same encoding strategy (i.e., frequency modulation) remains an open issue that could potentially be addressed initially by simplifying the scenario and considering only the *y* axis. Dimension that could be delivered by modulating either sound intensity or possibly vibrotactile amplitude is shown in green.

the signal) compared to other sensory cues. Hence, if the user is operating the SRL in an environment with reduced visibility (e.g., presence of smoke or dust, or darkness for workers operating under water), supplementary feedback carrying information on the motion of the robot would become particularly useful. This usefulness could be increased even more by capitalizing on the correspondence between perceived motion and acoustic pitch. For instance, playing an ascending pitch when the robot moves upward would easily bias the attention of the user toward that direction, because of the integration guided by crossmodal correspondence, thus improving the interaction under low visibility conditions.

3. Reducing learning time

A striking feature of crossmodal correspondences is how spontaneous and widespread they appear to be. Whether this perceptual phenomenon reflects an innate and universal mechanism is not yet clear, since research has produced conflicting evidence (Ludwig et al., 2011; Spence and Deroy, 2012). Despite an experience-based explanation for certain crossmodal correspondences having been proposed (Speed et al., 2021), and possible cultural differences between groups occasionally being reported (Spence, 2022), the empirical evidence that has been published to date shows that most healthy adults perceive crossmodal correspondence spontaneously, consistently, and without the need for any specific training (though see also Klapetek et al., 2012). The possibility of associating information from different sensory modalities without a dedicated training is of great interest for the field of human augmentation.

Indeed, learning to use a sensory augmentation tool requires the user to become familiar with its physical features (e.g., dimensions, weight, joint stiffness, and degrees of freedom) in order to interact with it proficiently. Rich supplementary feedback can greatly facilitate this process, but learning to understand and use such feedback often requires extensive training, especially when supplementary feedback carries information that is not available to the other senses. This sensory substitution requires being exposed to both the original sensory feedback (e.g., vision) and the supplementary one (e.g., vibrotactile stimulation) in order to establish an association that is sufficiently strong to survive the removal of the former and thus allow the latter to provide useful information. The latest research demonstrates that this is possible in the specific context of human augmentation (Pinardi et al., 2021; Umezawa et al., 2022). However the long training sessions required to fully benefit from supplementary feedback, can all too easily result in participant exhaustion and demotivation, thus limiting the feasibility of the experimental approach. To avoid this, supplementary feedback usually relies on a simple encoding strategy that can be learned quickly. For instance, it is useful to respect the spatial distribution between workspace organization and feedback interface on the participant's body (i.e., feedback of proximal joints is delivered proximally on the participant's body; Noccaro et al., 2020; Pinardi et al., 2021), or to couple a vibrotactile stimulation frequency that is higher with a higher force exerted by the robotic limb (Hussain et al., 2015). By removing the need for feedback decoding, the effort to facilitate feedback learning could be pushed toward a more efficient application by exploiting the spontaneous associations that characterize crossmodal correspondences, such as the well-documented space-pitch/loudness correspondence.

3.1. Application scenario: exploiting space-pitch/loudness correspondence for SRL localization

Pitch is consistently associated with spatial locations on the vertical (Bordeau et al., 2023), and to a lesser extent horizontal, axis, such that a low-pitched sound is more likely to be associated with a left/low location whereas a high pitch is associated with a right/high location (Mudd, 1963; Rusconi et al., 2006; Spence, 2011; Guilbert, 2020). At the same time, however, loudness is known to be an effective indicator of distance, such that, when presented with two sounds which differ only in loudness, people tend to associate the louder sound with a closer sound source, and *vice versa* (Eitan, 2013; Di Stefano, 2022).

Sonification is a strategy of heteromodal sensory substitution which relays kinematic and dynamic features of a movement with sounds (Castro et al., 2021a,b); however, its effectiveness is dependent on the specific rules used for encoding. Sonification can also be exploited to relay feedback related to an SRL. In light of enhancing the intuitiveness and hence usability of sensory substitution devices by means of the incorporation of crossmodal correspondence, we propose to use sound frequency and intensity (which determines perceived pitch and loudness, respectively) to help users to more easily localize the SRL in tri-dimensional space with minimal need for dedicated training, since pitch/space correspondence has been shown to occur spontaneously (Guilbert, 2020), even in new-borns (Walker et al., 2018), and space/loudness is a reliable and consistent association which reflects the regularities in the environment (Eitan, 2013) (see **Figure 2**). Given that the spatial resolution of pitch/space correspondences has yet to be determined, it is hard to provide clear indications on the effective contribution of this association to the spatial localization of SRLs. However, a rough indication on the region of space occupied by the SRL in terms of the two principal axes (i.e., up/down, and right/left) would be probably enough informative for these kinds of tasks, and we expect that such information can be likely provided intuitively, thus potentially reducing the learning time.

The association between pitch and the vertical/horizontal axes could also be used in human augmentation to generate more “instinctive” warning signals that might direct the operators’ spatial attention in an intuitive way, i.e., with no need for dedicated training (Ho and Spence, 2008). Finally, sonification has been shown to improve the behavioral performance of sportsmen (e.g., Maes et al., 2016; Lorenzoni et al., 2019; O’Brien et al., 2020), and thus coupling it with crossmodal correspondence mechanisms might further facilitate this beneficial effect. Since pitch is associated at the same time with both horizontal and vertical direction, it would be important to disentangle the modulation of these axis. This could be addressed initially by simply considering one of the two axes (2D workspace), and later by acting on how feedback is delivered (e.g., deliver two pitches sequentially, the first coding the x axis and the second coding the y axis). Additionally, while the loudness/distance association has been investigated primarily on the z axis, it should be tested whether it also consistently present on the others.

This perspective is further corroborated by studies involving blind participants, that show that the use of sensory substitution devices, despite not being devoid of challenges (Barilari et al., 2018; Hamilton-Fletcher et al., 2018; Sourav et al., 2019), is somewhat easier to learn if they are based on crossmodal correspondences (Hamilton-Fletcher et al., 2016) and a similar correspondence-based facilitation has also emerged in the case of language learning as well (Imai et al., 2008, 2015).

Moreover, it can be predicted that the haptic feedback might be used similarly to pitch, assuming that frequency and amplitude of a vibrotactile stimulation are associated with vertical/horizontal localization and with distance, or depth, respectively. Should this be confirmed empirically, it would represent a useful principle to guide the design of supplementary feedback. Additionally, given the multifaceted nature of the somatosensory modality, haptic feedback could potentially convey several informative contents simultaneously, so that multiple stimulus dimensions can be captured more clearly compared to information conveyed through auditory feedback. Indeed, somatosensation includes a great variety of sensory stimuli, such as, vibration, pressure, and skin stretch. Combining different types of stimuli, such as high frequency vibrotactile stimulation and skin stretch, to encode different information (e.g., coordinates on different axis), could reduce the risk of ambiguous encoding while maintaining a high informative value. Though, that being said, one should never forget the inherent limitations associated with the skin as a means of information transfer (see Spence, 2014).

4. Sensory transduction

Studies on human motor control have proven that proprioception is an unobtrusive yet rich sensory modality allowing people to complete dextrous motor tasks with minimal attentional effort (Proske and Gandevia, 2012). This is further confirmed in patients who lost proprioception, and must continuously rely on visual monitoring to carry out daily life actions, with a huge attentional effort and lower learning performance (Danna and Velay, 2017; Miall et al., 2019, 2021). Hence, relaying proprioceptive-like information concerning SRLs (i.e., joint angle, end effector acceleration or exerted force) to the user can result in improved SRL control. However, in the case of supplementary feedback, it is often difficult to adopt a homomodal coding approach (information related to a given sensory modality are fed back to the user exploiting the same sensory modality e.g., pressure on the robot surface with tactors upon the user skin). This is due to the fact that SRL’s proprioceptive parameters should be delivered using proprioceptive stimulation, and this is sometimes impossible or impractical (e.g., delivering joint angles by means of a kinesthetic illusion (Pinardi et al., 2020b) imposes heavy constraints on experimental protocols). Hence, translating a proprioceptive parameter of the robot (e.g., joint angle) into a pattern of tactile stimulation (e.g., vibratory stimulation) is easier to implement, resulting in heteromodal feedback. This approach solves a number of technical issues but introduces the need to translate supplementary sensory feedback from one modality (e.g., proprioception) to another (e.g., touch). Hence, the possibility of maintaining the benefits of crossmodal correspondence across different sensory modalities would be an empirical direction worth investigating. For instance, higher auditory pitch is associated with faster motion (Zhang et al., 2021) and higher physical acceleration (Eitan and Granot, 2006). However, in a human augmentation paradigm, participants would not be directly informed about speed or acceleration through supplementary feedback, but rather with vibratory stimulation that codes those parameters (e.g., higher vibrotactile frequency codes for higher acceleration). Hence, would the association between acceleration (or speed) and auditory pitch still be valid, if acceleration were to be substituted with a vibrotactile stimulation coding the same information?

To the best of our knowledge, there are no studies that directly address this question, however, considering that crossmodal correspondences have been demonstrated between every pair of sensory modalities (North, 2012; Nava et al., 2016; Wright et al., 2017; Speed et al., 2021) and are, by definition, cross-modal, it is likely that the underlying mechanisms would not be disrupted by sensory modality transduction.

5. Discussion and conclusion

In the present perspective paper, possible intersections between two different and, to the best of our knowledge, previously unrelated research fields, namely human augmentation and crossmodal correspondences have been explored. We propose that future research on the control and embodiment of SRLs could benefit from exploiting the spontaneous and widespread mechanisms of crossmodal correspondences, which have been shown to impact on spatiotemporal features and constraints of multisensory integration. The potential improvement of the control of SRLs would likely result as a by-product of the well-documented positive effect of crossmodal correspondence on attentional processing (Spence, 2011; Klapetek et al., 2012; Orchard-Mills et al., 2013; Chiou and Rich, 2015). Indeed, attention is considered a crucial resource when it comes to learning new motor skills and improving motor performance, especially in complex environments or during the performance of difficult tasks (Song, 2019), such as controlling an SRL. The link between motor control, attention and sensory processing is further demonstrated by the fact that multisensory contingencies can modulate the coupling between attention and motor planning (Dignath et al., 2019). However, further evidence is required to assess and disentangle the specific impact of crossmodal correspondences on motor control or attentional processing in the framework of human augmentation.

While the insights suggested here could be extended to any robotic teleoperated device, they are particularly fit for the field of human augmentation. Indeed, SRL users are supposed to control robotic limbs and natural ones at the same, while receiving feedback from both. This produces an additional challenge that is not present in prosthetics or when dealing with robotic teleoperation. The added sensorimotor processing required in human augmentation, and the absence of residual neural resources that can be repurposed for conveying feedback (Makin et al., 2017; Dominijanni et al., 2021) justifies the need for new strategies that minimize the effort required for sensory feedback learning and maximize its usefulness, and even though benefits of crossmodal correspondence have usually been observed in simple tasks (i.e., speeded target discrimination) and control and embodiment of an augmenting device are much more complex tasks, they ultimately rely on the same mechanisms of multisensory integration.

Finally, we argue that crossmodal correspondences could be exploited to deliver even highly sophisticated feedback with reduced cognitive burden. For instance, by capitalizing on the correspondence between tactile and auditory roughness (see Di Stefano and Spence, 2022, for a review), it might be possible to deliver the sensation of a rough texture through a rough sound (e.g., equal intensity for all frequencies, such as white noise, Pellegrino et al., 2022). More specifically, simply using a camera located on the tip of the end-effector, it would be possible to analyze surface texture, determine its roughness level through a machine-learning algorithm, and modulate the features of an auditory cue to provide information to the user on the surface tactile quality (Guest et al., 2002).

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

MP and NDS conceived the work and wrote the first draft of the manuscript. CS and GDP revised the entire manuscript and provided the insightful comments on all sections. All authors read and approved the submitted version.

Funding

This work was supported by the Italian Ministry of Education, University and Research under the “FARE: Framework Attrazione e Rafforzamento Eccellenze Ricerca in Italia” research program (ENABLE, no. R16ZBLF9E3) and by the European Commission under the “NIMA: Non-invasive Interface for Movement Augmentation” project (H2020-FETOPEN-2018-2020, ID: 899626).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

References

Aoyama, T., Shikida, H., Schatz, R., and Hasegawa, Y. (2019). Operational learning with sensory feedback for controlling a robotic thumb using the posterior auricular muscle. *Adv. Robot.* 33, 243–253. doi: 10.1080/01691864.2019.1566090

Arai, K., Saito, H., Fukuoka, M., Ueda, S., Sugimoto, M., Kitazaki, M., et al. (2022). Embodiment of supernumerary robotic limbs in virtual reality. *Sci. Rep.* 12:9769. doi: 10.1038/s41598-022-13981-w

Barilari, M., de Heering, A., Crollen, V., Collignon, O., and Bottini, R. (2018). Is red heavier than yellow even for blind? *i-Perception* 9:2041669518759123. doi: 10.1177/2041669518759123

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13, 556–571. doi: 10.1038/nrn3292

Bordeau, C., Scalvini, F., Migniot, C., Dubois, J., and Ambard, M. (2023). Cross-modal correspondence enhances elevation localization in visual-to-auditory sensory substitution. *Front. Psychol.* 14:1079998. doi: 10.3389/fpsyg.2023.1079998

Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391:756. doi: 10.1038/35784

Calvert, G. A., and Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *J. Physiol. Paris* 98, 191–205. doi: 10.1016/j.jphysparis.2004.03.018

Castro, F., Bryjka, P. A., Di Pino, G., Vuckovic, A., Nowicky, A., and Bishop, D. (2021a). Sonification of combined action observation and motor imagery: Effects on corticospinal excitability. *Brain Cogn.* 152:105768. doi: 10.1016/j.bandc.2021.105768

Castro, F., Osman, L., Di Pino, G., Vuckovic, A., Nowicky, A., and Bishop, D. (2021b). Does sonification of action simulation training impact corticospinal excitability and audiomotor plasticity? *Exp. Brain Res.* 239, 1489–1505. doi: 10.1007/s00221-021-06069-w

Chandrasekaran, C. (2017). Computational principles and models of multisensory integration. *Curr. Opin. Neurobiol.* 43, 25–34. doi: 10.1016/j.conb.2016.11.002

Chiou, R., and Rich, A. N. (2015). Volitional mechanisms mediate the cuing effect of pitch on attention orienting: The influences of perceptual difficulty and response pressure. *Perception* 44, 169–182. doi: 10.1068/p7699

Danna, J., and Velay, J. L. (2017). On the auditory-proprioception substitution hypothesis: Movement sonification in two deafferented subjects learning to write new characters. *Front. Neurosci.* 11:137. doi: 10.3389/fnins.2017.00137

de Vignemont, F. (2010). Body schema and body image–pros and cons. *Neuropsychologia* 48, 669–680. doi: 10.1016/j.neuropsychologia.2009.09.022

Di Pino, G., Guglielmelli, E., and Rossini, P. M. (2009). Neuroplasticity in amputees: Main implications on bidirectional interfacing of cybernetic hand prostheses. *Prog. Neurobiol.* 88, 114–126. doi: 10.1016/j.pneurobio.2009.03.001

Di Pino, G., Maravita, A., Zollo, L., Guglielmelli, E., and Di Lazzaro, V. (2014). Augmentation-related brain plasticity. *Front. Syst. Neurosci.* 8:109. doi: 10.3389/fnsys.2014.00109

Di Pino, G., Porcaro, C., Tombini, M., Assenza, G., Pellegrino, G., Tecchio, F., et al. (2012). A neurally-interfaced hand prosthesis tuned inter-hemispheric communication. *Restor. Neurol. Neurosci.* 30, 407–418. doi: 10.3233/RNN-2012-120224

Di Pino, G., Romano, D., Spaccasassi, C., Mioli, A., D'Alonzo, M., Sacchetti, R., et al. (2020). Sensory- and action-oriented embodiment of neurally-interfaced robotic hand prostheses. *Front. Neurosci.* 14:389. doi: 10.3389/fnins.2020.00389

Di Stefano, N. (2022). The spatiality of sounds. From sound-source localization to musical spaces. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico* 15, 173–185.

Di Stefano, N., and Spence, C. (2022). Roughness perception: A multisensory/crossmodal perspective. *Atten. Percept. Psychophys.* 84, 2087–2114. doi: 10.3758/s13414-022-02550-y

Di Stefano, N., Jarrassé, N., and Valera, L. (2022). The ethics of supernumerary robotic limbs. An enactivist approach. *Sci. Eng. Ethics.* 28:57. doi: 10.1007/s11948-022-00405-1

Diedrichsen, J., Shadmehr, R., and Ivry, R. B. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends Cogn Sci.* 14, 31–39. doi: 10.1016/j.tics.2009.11.004

Dignath, D., Herbort, O., Pieczykolan, A., Huestegge, L., and Kiesel, A. (2019). Flexible coupling of covert spatial attention and motor planning based on learned spatial contingencies. *Psychol. Res.* 83, 476–484. doi: 10.1007/s00426-018-1134-0

Dominijanni, G., Shokur, G., Salvietti, G., Buehler, S., Palmerini, E., Rossi, S., et al. (2021). Enhancing human bodies with extra robotic arms and fingers: The neural resource allocation problem. *Nat. Mach. Intell.* 3, 850–860.

Eden, J., Bräcklein, M., Pereda, J. I., Barsakcioglu, D. Y., Di Pino, G., Farina, D., et al. (2021). Human movement augmentation and how to make it a reality. *arXiv preprint. arXiv:2106.08129*.

Eitan, Z. (2013). "How pitch and loudness shape musical space and motion," in *The psychology of music in multimedia*, eds S.-L. Tan, A. J. Cohen, S. D. Lipscomb, and R. A. Kendall (Oxford: Oxford University Press), 165–191.

Eitan, Z., and Granot, R. Y. (2006). How music moves: Musical parameters and listeners images of motion. *Music Percept.* 23, 221–248.

Gouzien, A., de Vignemont, F., Touillet, A., Martinet, N., De Graaf, J., Jarrassé, N., et al. (2017). Reachability and the sense of embodiment in amputees using prostheses. *Sci. Rep.* 7:4999. doi: 10.1038/s41598-017-05094-6

Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. *Exp. Brain Res.* 146, 161–171. doi: 10.1007/s00221-002-1164-z

Guggenheim, J. W., and Asada, H. H. (2021). Inherent haptic feedback from supernumerary robotic limbs. *IEEE Trans. Haptics*. 14, 123–131. doi: 10.1109/TOH.2020.3017548

Guilbert, A. (2020). About the existence of a horizontal mental pitch line in non-musicians. *Laterality*. 25, 215–228. doi: 10.1080/1357650X.2019.1646756

Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefanczyk, Ward, J., and Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition* 175, 114–121. doi: 10.1016/j.cognition.2018.02.023

Hamilton-Fletcher, G., Wright, T. D., and Ward, J. (2016). Cross-modal correspondences enhance performance on a colour-to-sound sensory substitution device. *Multisens. Res.* 29, 337–363. doi: 10.1163/22134808-00002519

Hernandez, J. W., Haget, A, Bleuler, H., Billard, A., and Bouri, M. (2019). Four-arm manipulation via feet interfaces. *arXiv* [Preprint] doi: 10.48550/arXiv.1909.04993

Ho, C., and Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. Farnham: Ashgate Publishing, Ltd.

Hussain, I., Meli, L., Pacchierotti, C., Salvietti, G., and Prattichizzo, D. (2015). "Vibrotactile haptic feedback for intuitive control of robotic extra fingers," in *IEEE World Haptics Conference, WHC 2015*, Evanston, IL, 394–399. doi: 10.1109/WHC.2015.7177744

Ihalainen, R., Kotsaridis, G., Vivas, A. B., and Paraskevopoulos, E. (2023). *The* relationship between musical training and the processing of audiovisual correspondences: Evidence from a reaction time task. *PLoS One* 18:e0282691. doi: 10.1371/journal.pone.0282691

Imai, M., Kita, S., Nagumo, M., and Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition* 109, 54–65. doi: 10.1016/j.cognition.2008.07.015

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., et al. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS One* 10:e0116494. doi: 10.1371/journal.pone.0116494

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

Köhler, W. (1929). An old pseudoproblem. *Naturwissenschaften* 17, 395–401.

Körding, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247. doi: 10.1038/nature02169

Küssner, M. B., and Leech-Wilkinson, D. (2014). Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing paradigm. *Psychol. Music* 42, 448–469. doi: 10.1177/0305735613482022

Lin, A., Scheller, M., and Feng, F. (2021). "Feeling colours: Crossmodal correspondences between tangible 3D objects, colours and emotions," in *Proceedings of the Conference on human factors in computing systems*, Yokohama, doi: 10.1145/3411764.3445373

Lorenzoni, V., Van den Berghe, P., Maes, P. J., De Bie, T., De Clercq, D., and Leman, M. (2019). Design and validation of an auditory biofeedback system *for* modification of running parameters. *J. Multim. User Interf.* 13, 167–180. doi: 10.1007/s12193-018-0283-1

Ludwig, V. U., Adachi, I., and Matsuzawa, T. (2011). Visuoauditory mappings between high luminance *and* high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20661–20665. doi: 10.1073/pnas.1112605108

Maeda, F., Kanai, R., and Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Curr. Biol.* 14, R990–R991. doi: 10.1016/j.cub.2004.11.018

Maes, P. J., Buhmann, J., and Leman, M. (2016). 3Mo: A model for music-based biofeedback. *Front. Neurosci.* 10:548. doi: 10.3389/fnins.2016.00548

Makin, T. R., de Vignemont, F., and Faisal, A. A. (2017). Neurocognitive barriers to the embodiment of technology. *Nat. Biomed. Eng.* 1:14. doi: 10.1038/s41551-016-0014

Marasco, P. D., Hebert, J. S., Sensinger, J. W., Shell, C. E., Schofield, J. S., Thumser, Z. C., et al. (2018). Illusory movement perception improves motor control for prosthetic hands. *Sci. Transl. Med.* 10:eaao6990. doi: 10.1126/scitranslmed.aao6990

Marasco, P. D., Kim, K., Colgate, J. E., Peshkin, M. A., and Kuiken, T. A. (2011). Robotic touch shifts perception of embodiment to a prosthesis in targeted reinnervation amputees. *Brain* 134(Pt 3), 747–758. doi: 10.1093/brain/awq361

Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 384–394. doi: 10.1037//0096-1523.13.3.384

Miall, R. C., Afanasyeva, D., Cole, J. D., and Mason, P. (2021). The role of somatosensation in automatic visuo-motor control: A comparison of congenital and acquired sensory loss. *Exp. Brain Res.* 239, 2043–2061. doi: 10.1007/s00221-021-06110-y

Miall, R. C., Rosenthal, O., Ørstavik, K., Cole, J. D., and Sarlegna, F. R. (2019). Loss of haptic feedback impairs control of hand posture: A study in chronically deafferented individuals when grasping and lifting objects. *Exp. Brain Res.* 237, 2167–2184. doi: 10.1007/s00221-019-05583-2

Moseley, G. L., Gallace, A., and Spence, C. (2012). Bodily illusions in health and disease: Physiological and clinical perspectives and the concept of a cortical 'body matrix'. *Neurosci. Biobehav. Rev*. 36, 34–46. doi: 10.1016/j.neubiorev.2011.03.013

Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *J. Exp. Psychol.* 66, 347–352. doi: 10.1037/h0040045

Nava, E., Grassi, M., and Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisens. Res.* 29, 93–111. doi: 10.1163/22134808-00002493

Noccaro, A. L., Raiano, M., Pinardi, D., Formica, G., and Pino, D. (2020). "A novel proprioceptive feedback system for supernumerary robotic limb," in *Proceedings of the IEEE RAS and EMBS international conference on Biomedical robotics and biomechatronics*, 2020-Novem, New York, NY, 1024–1029. doi: 10.1109/BioRob49111.2020.9224450

North, A. C. (2012). The effect of background music on the taste of wine. *Br. J. Psychol.* 103, 293–301. doi: 10.1111/j.2044-8295.2011.02072.x

O'Brien, B., Hardouin, R., Rao, G., Bertin, D., and Bourdin, C. (2020). Online sonification improves cycling performance through kinematic *and* muscular reorganisations. *Sci. Rep.* 10, 20929. doi: 10.1038/s41598-020-76498-0

Orchard-Mills, E., Alais, D., and Van der Burg, E. (2013). Cross-modal associations between vision, touch, *and* audition influence visual search through top-down attention, not bottom-up capture. *Atten. Percept. Psychophys.* 75, 1892–1905. doi: 10.3758/s13414-013-0535-9

Page, D. M., George, J. A., Kluger, D. T., Duncan, C., Wendelken, S., Davis, T., et al. (2018). Motor control and sensory feedback enhance prosthesis embodiment and reduce phantom pain after long-term hand amputation. *Front. Hum. Neurosci.* 12:352. doi: 10.3389/fnhum.2018.00352

Parise, C. V., and Spence, C. (2009). 'When birds of a feather flock together': Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One* 4:e5664. doi: 10.1371/journal.pone.0005664

Parise, C., and Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neurosci. Lett.* 442, 257–261. doi: 10.1016/j.neulet.2008.07.010

Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychol. Sci.* 11, 353–359. doi: 10.1111/1467-9280.00270

Pellegrino, G., Pinardi, M., Schuler, A. L., Kobayashi, E., Masiero, S., Marioni, G., et al. (2022). Stimulation with acoustic white noise enhances motor excitability and sensorimotor integration. *Sci. Rep.* 12:13108. doi: 10.1038/s41598-022-17055-9

Pinardi, M., Noccaro, A., Raiano, L., Formica, D., and Di Pino, G. (2023). Comparing end-effector position and joint angle feedback for online robotic limb tracking. *PLoS One* (in press). doi: 10.1371/journal.pone.0286566

Pinardi, M. L., Raiano, A., Noccaro, D., Formica, G., and Pino, D. (2021). "Cartesian space feedback for real time tracking of a supernumerary robotic limb: A pilot study," in *International IEEE/EMBS conference on neural engineering, NER*, 2021-May, Rome, 889–892. doi: 10.1109/NER49283.2021.9441174

Pinardi, M., Ferrari, F., D'Alonzo, M., Clemente, F., Raiano, L., Cipriani, C., et al. (2020a). Doublecheck: A sensory confirmation is required to own a robotic hand, sending a command to feel in charge *of* it. *Cogn. Neurosci.* 11, 216–228. doi: 10.1080/17588928.2020.1793751

Pinardi, M., Raiano, L., Formica, D., and Di Pino, G. (2020b). Altered proprioceptive feedback influences movement kinematics in a lifting task. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2020, 3232–3235. doi: 10.1109/EMBC44109.2020.9176252

Pramudya, R. C., Choudhury, D., Zou, M., and H-S, Seo (2020). "Bitter touch": Cross-modal associations between hand-feel touch and gustatory cues in the context of coffee consumption experience. *Food Qual. Preferen.* 83:103914. doi: 10.1016/j.foodqual.2020.103914

Proske, U., and Gandevia, S. C. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiol. Rev.* 92, 1651–1697. doi: 10.1152/physrev.00048.2011

Raspopovic, S., Capogrosso, M., Petrini, F. M., Bonizzato, M., Rigosa, J., Di Pino, G., et al. (2014). Restoring natural sensory feedback in real-time bidirectional hand prostheses. *Sci. Transl. Med.* 6:222ra19. doi: 10.1126/scitranslmed.3006820

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., and Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition* 99, 113–129. doi: 10.1016/j.cognition.2005.01.004

Segura Meraz, N., Shikida, H., and Hasegawa, Y. (2017). "Auricularis muscles based control interface for robotic extra thumb," in *Proceedings of the international symposium on micro-nanomechatronics and human science (MHS)*, Nagoya. doi: 10.3233/RNN-150579

Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* 33, 89–108. doi: 10.1146/annurev-neuro-060909-153135

Slobodenyuk, N., Jraissati, Y., Kanso, A., Ghanem, L., and Elhajj, I. (2015). Cross-modal associations between color and haptics. *Atten. Percept. Psychophys.* 77, 1379–1395. doi: 10.3758/s13414-015-0837-1

Sobajima, M., Sato, Y., Xufeng, W., and Hasegawa, Y. (2016). "Improvement *of* operability of extra robotic thumb using tactile feedback by electrical stimulation," in *2015 International Symposium on Micro-NanoMechatronics and Human Science, MHS 2015*, (Nagoya), 3–5. doi: 10.1109/MHS.2015.7438269

Song, J. H. (2019). The role of attention in motor control and learning. *Curr Opin Psychol.* 29, 261–265. doi: 10.1016/j.copsyc.2019.08.002

Sourav, S., Kekunnaya, R., Shareef, I., Banerjee, S., Bottari, S., and Röder, B. (2019). A protracted sensitive period regulates the development of cross-modal sound–shape associations in humans. *Psychol. Sci.* 30, 1473–1482. doi: 10.1177/0956797619866625

Speed, L. J., Croijmans, I., Dolscheid, S., and Majid, A. (2021). Crossmodal associations with olfactory, auditory, and tactile stimuli in children and adults. *i-Perception* 12:20416695211048513. doi: 10.1177/20416695211048513

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2014). The skin as a medium for sensory substitution. *Multisens. Res.* 27, 293–312. doi: 10.1163/22134808-00002452

Spence, C. (2020). Food and beverage flavour pairing: A critical review of the literature. *Food Res. Int.* 133:109124. doi: 10.1016/j.foodres.2020.109124

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079

Spence, C., and Deroy, O. (2012). Crossmodal correspondences: Innate or learned?. *Iperception* 3, 316–318. doi: 10.1068/i0526ic

Spence, C., and Wang, Q. (2015). Wine and music (I): On the crossmodal matching of wine and music. *Flavour* 4, 1–14. doi: 10.1186/S13411-015-0045-X

Sun, X., Li, X., Ji, L., Han, F., Wang, H., Liu, Y., et al. (2018). An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. *PeerJ* 6:e4443. doi: 10.7717/peerj.4443

Tong, J., Li, L., Bruns, P., and Röder, B. (2020). Crossmodal associations modulate multisensory spatial integration. *Atten. Percept. Psychophys.* 82, 3490–3506. doi: 10.3758/s13414-020-02083-2

Umezawa, K., Suzuki, Y., Ganesh, G., and Miyawaki, Y. (2022). Bodily ownership of an independent supernumerary limb: An exploratory study. *Sci. Rep.* 12:2339. doi: 10.1038/s41598-022-06040-x

von Hornbostel, E. M. (1931). Ueber geruchshelligkeit. *Pflügers Archiv Gesamte Physiol. Menschen Tiere* 228, 517–537. doi: 10.1007/BF01755351

Walker, P., Bremner, J. G., Lunghi, M., Dolscheid, S., Barba, D., and Simion, F. (2018). Newborns are sensitive to the correspondence between auditory pitch and visuospatial elevation. *Dev. Psychobiol.* 60, 216–223. doi: 10.1002/dev.21603

Walker-Andrews, A., Lewkowicz, D. J., and Lickliter, R. (1994). "Taxonomy for intermodal relations," in *The development of intersensory perception: Comparative perspectives*, eds D. J. Lewkowicz and R. Lickliter (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 39–56.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931

Wright, O., Jraissati, Y., and Özçelik, D. (2017). Cross-modal associations between color and touch: Mapping haptic and tactile terms to the surface of the munsell color solid. *Multisens. Res.* 30, 691–715. doi: 10.1163/22134808-00002589

Zhang, G., Wang, W., Qu, J., Li, H., Song, X., and Wang, Q. (2021). Perceptual influence *of* auditory pitch on motion speed. *J. Vis.* 21:11. doi: 10.1167/jov.21.10.11

Zollo, L., Di Pino G, Ciancio, A. L., Ranieri, F., Cordella, F., Gentile, C., et al. (2019). Restoring tactile sensations via neural interfaces for real-time force-and-slippage closed-loop control of bionic hands. *Sci. Robot.* 4:eaau9924. doi: 10.1126/scirobotics.aau9924

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 15 June 2023
DOI 10.3389/fpsyg.2023.1170674

OPEN ACCESS

EDITED BY
Charles Spence,
University of Oxford, United Kingdom

REVIEWED BY
Magdalena Szubielska,
The John Paul II Catholic University of Lublin, Poland
Nicola Di Stefano,
National Research Council (CNR), Italy

*CORRESPONDENCE
Shigeo Yoshida
shigeo.yoshida@sinicx.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 21 February 2023
ACCEPTED 01 June 2023
PUBLISHED 15 June 2023

CITATION
Ogata K, Gakumi R, Hashimoto A, Ushiku Y and Yoshida S (2023) The influence of Bouba- and Kiki-like shape on perceived taste of chocolate pieces. *Front. Psychol.* 14:1170674. doi: 10.3389/fpsyg.2023.1170674

The influence of Bouba- and Kiki-like shape on perceived taste of chocolate pieces

Kazuhiro Ogata[1,2†], Reo Gakumi[2,3†], Atsushi Hashimoto[2], Yoshitaka Ushiku[2] and Shigeo Yoshida[2]*

[1]Department of Design, Kyoto Institute of Technology, Kyoto, Japan, [2]OMRON SINIC X Corporation, Tokyo, Japan, [3]Graduate School of Informatics, Kyoto University, Kyoto, Japan

In this paper, we present the findings of a study investigating the impact of shape on the taste perception of chocolate. Previous research has explored the influence of various sensory information on taste perception, but there has been little focus on the effect of food shape being eaten on taste perception. To explore this, we focused on the Bouba-Kiki effect, illustrating an interaction between shape and several modalities, and investigated the effect of Bouba- and Kiki-shaped (rounded and angular) foods eaten on taste perception. We utilized a 3D food printer to produce four different shapes of chocolate pieces based on the Bouba-Kiki. Participants tasted each piece and completed a chocolate flavor questionnaire. With Bayesian analysis, we determined that the Bouba-shaped chocolate pieces were perceived as sweeter than the Kiki-shaped ones, supporting earlier studies on crossmodal correspondences between shape and taste perception. However, there were no significant differences in ratings of other tastes, such as sourness and bitterness. Our research indicates that shape can affect taste perception during consumption and suggests that 3D food printers offer an opportunity to design specific shapes that influence taste experiences.

KEYWORDS
crossmodal correspondences, Bouba-Kiki effect, chocolate, shape-taste correspondences, 3D food printer

1. Introduction

The experience of food results from the integration of multisensory information. Previous studies have shown that the visual and tactile characteristics of food [e.g., color: Spence et al. (2015), Velasco et al. (2016a), shape: Gal et al. (2007), Spence and Gallace (2011), Spence and Deroy (2014), and its texture: Slocombe et al. (2016)] affect taste perception. Extrinsic or environmental characteristics when eating also influence taste perception. For example, studies of taste perception and related experiences have demonstrated that the taste of food influenced by the characteristics of the cutlery (e.g., weight, size, shape, and color: Harrar and Spence, 2013; Michel et al., 2015b), how food is arranged on the plates (Michel et al., 2015a), environmental sound (Mesz et al., 2011; Bronner et al., 2012), atmosphere, and room lighting (Velasco et al., 2013; Spence et al., 2014; Nygård and Lie, 2020), product packages and labels (Velasco et al., 2016b), and congruency between content and glassware (Wan et al., 2015; Van Doorn et al., 2017). Such non-arbitrary associations among different modalities are called "crossmodal correspondences", and various associations among modalities have been demonstrated for senses other than taste (Spence, 2011).

The Bouba-kiki effect is a well-known example of crossmodal correspondences that illustrate an interaction between shape and speech sound. Bouba-Kiki was derived from an earlier study (Köhler, 1929) that presented many people with similar associations between two non-sense words or pseudowords (Maluma/Taketa) and two types of geometric shapes (rounded/angular). Subsequently, Ramachandran and Hubbard (2001) mentioned that 95% of their participants responded to the same associations between the two types of rounded or angular shapes and speech sounds. Studies have extended the Bouba-Kiki effect to reveal correspondences between foods/flavors and words (Gallace et al., 2011), flavors of liquids (Deroy and Valentin, 2011), and between rounded/angular shapes and cheese taste (Spence et al., 2013). Moreover, some other studies have investigated the Bouba-Kiki effect in terms of different transparent window shapes in product packaging (Simmonds et al., 2019) and the effect of different bowl shapes on the expected taste for a dessert (Chen et al., 2018). The results showed correspondences between sweetness and rounded shapes and bitterness/sourness and angular shapes. Several reports have also demonstrated similar results (Ngo and Spence, 2011; Ngo et al., 2011; Wang et al., 2017). In addition, Salgado-Montejo et al. (2015) conducted an experiment on the Bouba-Kiki effect using two additional properties of the shapes and shape symmetry (Makin et al., 2012), and number of elements (Jacobsen et al., 2006; Palmer et al., 2013).

Many studies have demonstrated a relationship between the appearance and shape of food and its taste, including the Bouba-Kiki, but it remains unclear how taste changes when such shaped foods are consumed. For example, Lenfant et al. (2013) demonstrated that different shapes (e.g., rectangle, triangle, round) of dark chocolate pieces affect flavor and texture. In contrast, Baptista et al. (2022) revealed that differences in shape have an impact on expected/perceived creaminess but not on sweetness, bitterness, and likability through a large-scale experiment. Wang et al. (2017) suggested there are no differences in the taste ratings of different shaped (rounded and angular) chocolate pieces when consumed, although ratings of expected taste before eating did differ by shape. In contrast, a recent study reported that the 3D Bouba-shaped edible samples were perceived as sweeter than the Kiki-shaped samples (Cornelio et al., 2022). However, the edible sample was prepared and designed only for the experiments (the ingredients were water, pork gelatin, agar, and sugar); unlike foods normally consumed, the taste was relatively neutral. Hence, it is unclear whether the Bouba-Kiki effect applies to the taste perception of foods usually consumed.

In this paper, we investigated the effects of Bouba-Kiki-shaped edible samples of a commonly consumed food item (i.e., chocolate, Figure 1) on taste perception. Chocolate was chosen as a material for the edible samples because chocolate has some inherent flavors and aromas, and has been used in previous taste perception studies, as described above. Chocolate is a common material used for 3D food printing (Hao et al., 2010; Lanaro et al., 2019; Mantihal et al., 2019; Souto et al., 2022). Chocolate as a 3D food printer material enables easy to control of shape and permits the production of a large number of edible samples for experiment.

As discussed above, several studies of crossmodal correspondences have demonstrated that Bouba (rounded) shapes are associated with sweetness and Kiki (angular) shapes are associated with sourness/bitterness. Thus, we hypothesized that Bouba-shaped chocolate pieces would be rated as sweeter than Kiki-shaped pieces (H_1), whereas ratings of sourness and bitterness would be greater for Kiki- than for Bouba-shaped pieces (H_2:sourness/H_3:bitterness). In addition, based on the premise that rounder, symmetrical, and those shapes with fewer elements are more likely to be judged as sweeter and angular, asymmetrical, and those shapes with a greater number of elements are more likely to be judged as more sour (Salgado-Montejo et al., 2015), we assumed that the symmetrical Bouba-shaped chocolate piece was more likely to be perceived as sweeter than the asymmetrical Bouba-shaped sample (H_4) and the Kiki-shaped sample with a greater number of elements would be perceived as more sour than that with fewer elements (H_5). Similarly, the effect of shape (an asymmetric shape with more elements) on bitterness was also investigated (H_6).

FIGURE 1
Chocolate pieces used in the study. The alphabetical letter following the hyphen means symmetrical (S) and asymmetrical (A), and the number means the number of elements (6 or 10 in this study).

2. Materials and methods

2.1. Taste stimuli

2.1.1. Stimulus design

We designed four Bouba/Kiki-shaped stimuli for this study (Figure 1). These shapes are derived from geometric forms, as shown in Köhler (1929). Each shape we designed was named according to its Bouba/Kiki form (i.e., rounded/angular), symmetry/asymmetry, and number of elements. For example, Bouba-S6 denotes a symmetric Bouba-shaped chocolate piece with six elements, whereas Kiki-A10 denotes an asymmetric Kiki-shaped chocolate piece with 10 elements.

First, two shapes with rounded or angular tips of the same basic shape (Bouba-A6 and Kiki-A6 in Figure 1) were chosen as a standard edible samples. This was consistent with our basic hypothesis, namely that a Bouba-shaped edible sample would be perceived as sweeter than an equivalent Kiki-shaped sample (conversely, a Kiki-shaped sample was hypothesized to be perceived as more sourer/bitter than the equivalent Bouba-shaped one). In addition, stimuli with modified symmetry and number of elements were designed (Bouba-S6 and Kiki-A10) because prior work demonstrated differences in taste expectation by changing the symmetry/asymmetry and number of elements of Bouba-Kiki-shaped stimuli (Salgado-Montejo et al., 2015). Therefore, we focused on these parameters, which had a particular influence on sweetness and sourness/bitterness in prior work, and designed

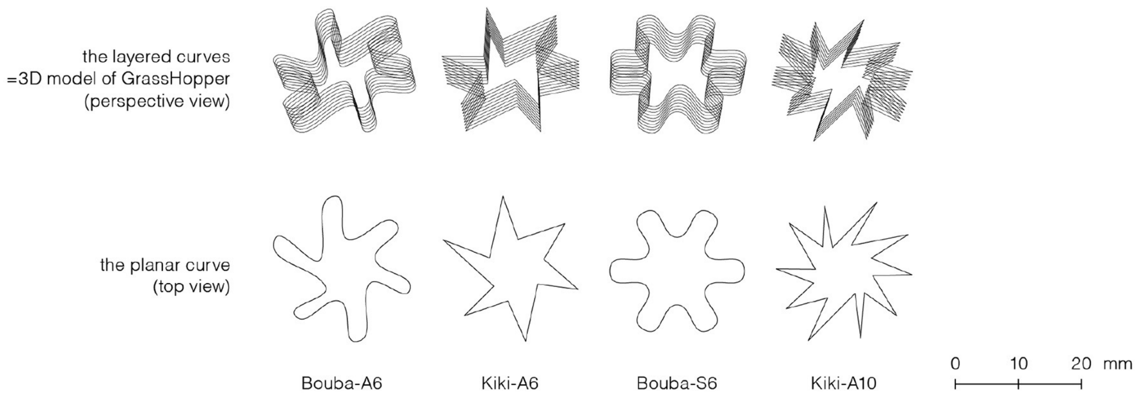

FIGURE 2
3D model and its planar curve used to print the chocolate pieces.

symmetric Bouba-shaped stimuli with few elements (6 in our study), and asymmetric Kiki-shaped stimuli with a large number of elements (10 in our study).

Prior studies found no differences in participants' reports of taste differences following actual consumption of round and angular chocolate pieces (Wang et al., 2017). We assumed that the amount of chocolate eaten in this prior study was too large, which caused the actual taste to dominate any effect of perceived shape on taste. Therefore, we designed a ring-shaped stimulus that was not filled with chocolate at its center to avoid the need to consume excessive chocolate taste/flavor, while maintaining perceived differences in shape.

2.1.2. 3D modeling and printing

Here, we briefly explain the conditions for 3D modeling and chocolate printing.

We used 3D CAD software (Rhinoceros[1]) and its plug-in (Grasshopper[2]) to design the stimuli. The edible samples were 24 mm in width and depth, and 3–4 mm in height (Figure 2). This volume was determined based on the results of a pilot study. Participants found it easy to place the samples in their mouth in one bite. The height was designed to be greater than the minimum discriminable distance between two points on the tongue, namely 1.650 ± 0.433 mm at the tip of the tongue and 2.650 ± 0.856 mm and 1.675 ± 0.269 mm at the sides of the anterior margins of each side of the tongue (Maeyama and Plattig, 1989); thus differences in texture could be perceived. The minimum two-point discrimination distance was also used to determine the distance between the tips of the shape, and the number of elements. Additionally, to equalize the weight of each sample, the height of each sample was controlled by varying the number of layers from 7 to 10. The symmetry/asymmetry and number of elements of the Bouba-Kiki shapes were partly determined with reference to a previous study (Salgado-Montejo et al., 2015).

To obtain precise and flexible control over the shape, mass, and volume when producing ring-shaped edible samples, a 3D food printer (Foodbot[3]) was used. This printer is designed for printing edible ingredients using a temperature-controllable food-grade syringe. The nozzle can also be changed to permit the printing of food objects with subtle textures and smooth surfaces. We used the nozzle of diameter of 0.6 mm in this study to make the edge of the Kiki-shaped samples precisely angular. The printer can directly read G-code files, which is a geometric code that controls printer direction, speed, and other relevant parameters (Derossi et al., 2019). In many cases, G-code is generated using slicing software. However, 3D objects can also be created and G-code generated using 3D CAD software and plug-ins. Printing conditions were determined based on previous studies that fabricated chocolate using 3D printers (Mantihal et al., 2017; Lanaro et al., 2019). Our printing conditions were as follows: print speed from 12.0 to 15.0 mm/s, extrusion multiplier of 0.8%, and layer pitch of 0.42 mm. Room temperature during printing maintained below 25.0°C to prevent the chocolate from melting.

2.1.3. Ingredients

The chocolate used as the edible sample was a commercially available 74% cacao couverture chocolate (ingredients: organic cacao mass, organic cane sugar, organic cacao butter, organic cacao powder, organic vanilla powder). This chocolate was used because lower-purity cacao complicates the tempering and printing process (Souto et al., 2022). That is, it is necessary to perform tempering to achieve crystallization of a stable form of chocolate that can be precisely printed. First, chocolate was melted at approximately 45°C, after which the temperature was reduced to 34°C. The chocolate was mixed with 1% (by weight of chocolate) cocoa butter powder as a seed crystal to produce stable form of chocolate (Lanaro et al., 2019). The temperature was maintained at 33°C at the end of this process. Mixing cocoa butter simplifies the tempering process and helps promote the growth of Form V crystal

1 https://www.rhino3d.com
2 https://www.grasshopper3d.com
3 https://www.3dfoodbot.com

nuclei. The tempered chocolate was poured into a plastic syringe and printed using the 3D food printer.

2.2. Experimental setup and design

The experiment was conducted in a quiet room with closed blinds. The air conditioning was set at 23°C, and a fan was running quietly in the background. The participants were seated in the room and tasted one edible sample at a time and rated the samples using a questionnaire presented on a tablet. We collected demographic information (age and sex) and taste ratings for each sample. Participants rated the following questions in a visual analogue scale (VAS) ranging from 0 to 100:

1. How much sweetness did you perceive in the chocolate?
2. How much sourness did you perceive in the chocolate?
3. How much bitterness did you perceive in the chocolate?

The experimenter placed the edible samples individually on a disposable plastic spoon (white, 160 mm long, 38 mm wide, for hygienic reasons) on a paper plate in front of the participant. The edible samples were presented at room temperature. The experiment was conducted using a within-participants design. Each participant consumed four samples with different shapes. The order in which the participants consumed the samples was predetermined by Latin square design, resulting in 24 possible sequences (factorial of 4 = 24).

2.3. Participants

Twenty-four Japanese participants (12 male, 12 female; age range 22–47 years, average age 33 years) without chocolate or dairy product allergies were recruited for this study. They were requested not to consume spicy food 24 h prior to the experiment, alcohol 6 h prior, or any food or drink (except water) 1 h prior (Obrist et al., 2014). Ethical approval was obtained from the Research Ethics and Review Committee of our company and written informed consent was obtained from all participants.

2.4. Procedure

The procedure of the experiment was as follows (Figure 3):

1. The experimenter put one piece chocolate sample on the spoon.
2. The participant was instructed to bring the sample to their mouth with the spoon and put it on their tongues. They were instructed to let the sample melt in their mouth and avoid biting. They were allowed to move the sample into their mouths, and the movement of their tongue was not restricted.
3. After the sample was melted and the participant swallowed it, they were instructed to wait 30 s to perceive the aftertaste of the sample.
4. The participant was then asked to complete a questionnaire that was presented on a tablet. They rated each sample in terms of three characteristics (sweetness, sourness, and bitterness).
5. The participant was asked to clear the aftertaste from their mouth by drinking room-temperature mineral water. After drinking the water, they remained in their seats and rested for 3 min each time to avoid any carry-over effect. In this study, we used room-temperature water, which is effective for cleaning the taste of astringent foods (Lee and Vickers, 2010).

The experiment consisted of repeating steps 1–5 four times and was completed in approximately 30 min. The study was conducted in Japanese, and the questionnaires used in the study were written in Japanese.

2.5. Analysis

We performed a statistical analysis[4] of the experimental results using Bayesian analysis with Markov chain Monte Carlo methods (MCMC). We employed Bayesian analysis as it facilitates a natural interpretation of the probability that the research hypothesis holds true, and it directly provides the probabilistic range of the parameters, enabling flexible hypothesis evaluation.

We used Python3 (v3.8) for data preprocessing and Bayesian analysis via PyMC (v4.1),[5] which is a Python package for the Bayesian statistical modeling of advanced MCMC. Five chains of 5,000 samples were generated using PyMC. The burn-in period was set to 1,000, 20,000 random numbers obtained using the No-U-Turn Sampler (NUTS) method were used to approximate the posterior and predictive distributions. The convergence in sampling was diagnosed using the Gelman-Rubin statistic, $|\hat{R} - 1| < 0.01$ for all parameters, and visually confirming indicate that the each Markov chain was converging to a stationary state properly.

It was assumed that each edible sample has its estimated value of population mean and population standard deviation denoted as μ_i, σ_i, and each of them follows a uniform distribution as a prior distribution.

$$\mu_i \sim \text{Uniform}\,(0, 100)$$
$$\sigma_i \sim \text{Uniform}\,(0, 50)$$

The calculation of the multivariate normal log-likelihood that we used as the likelihood in our analysis requires a vector of means and a variance-covariance matrix. Since the variance-covariance matrix can be generated from the correlation matrix and the variance of each variable, we used the LKJ (Lewandowski-Kurowicka-Joe) correlation distribution (Lewandowski et al., 2009) as a prior distribution for the correlation matrix.

$u^{(t)}_{\mu_i - \mu_j} > 0$ was set to compare the population means of the two edible samples. The probability that the hypothesis is true,

4 The code is available here: https://github.com/omron-sinicx/ComputationalMouthfeel.

5 https://www.pymc.io/welcome.html

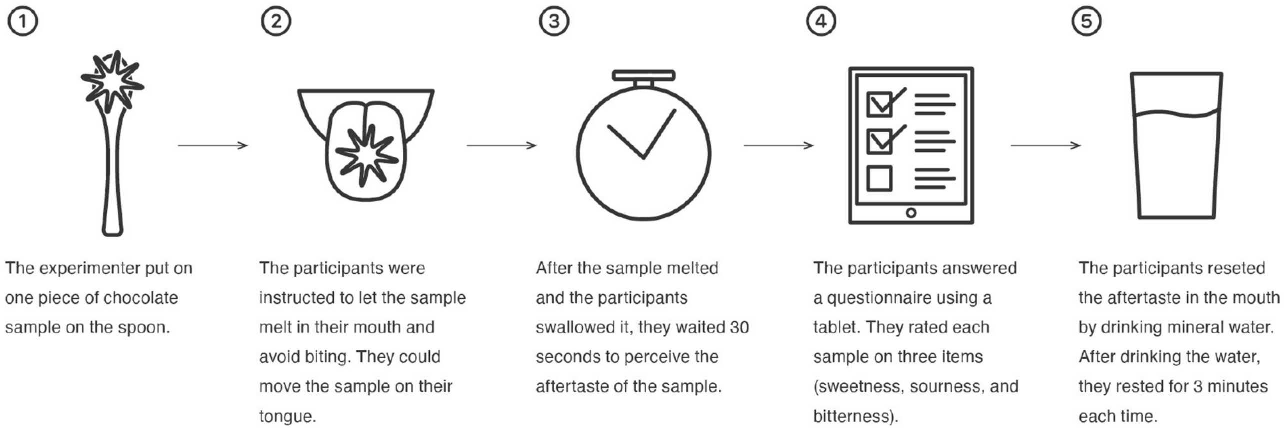

FIGURE 3
The experimental procedure followed by the participants. Step 1 to 5 were repeated four times.

$p(\mu_i - \mu_j > 0)$, was evaluated by the EAP (expected a posteriori) of the following generated quantity:

$$u^{(t)}_{\mu_i-\mu_j>0} = u^{(t)}_{\mu_i>\mu_j} = \begin{cases} 1 & \mu_i^{(t)} - \mu_j^{(t)} > 0 \\ 0 & \text{otherwises} \end{cases} \tag{1}$$

3. Results

3.1. Perceived sweetness

Figure 4 shows the box-plot of participants' sweetness ratings for each edible sample.

Point estimates of the population mean of perceived sweetness using EAP for Bouba-A6, Kiki-A6, Bouba-S6, and Kiki-A10 were as follows: 38.52 (3.59) [31.21, 45.27], 28.90 (3.39) [21.96, 35.25], 40.63 (4.80) [31.58, 50.29], and 30.67 (3.73) [23.04, 37.76], respectively. Note that "( )" denotes the posterior standard deviation and "[ ]" denotes the 95% HDI (highest density interval).

Table 1 lists the probability that the sample in row *i* was perceived as sweeter than that in column *j*. Comparison of the population mean of perceived sweetness between the basic Bouba-shaped sample (Bouba-A6) and Kiki-shaped sample (Kiki-A6) (H_1: Bouba-A6 > Kiki-A6) shows the Bouba-A6 was perceived as sweeter than the Kiki-A6 with a probability of 98.7%.[6] Regarding H_4 (Bouba-S6 > Bouba-A6), the probability of differences in the perceived sweetness due to minor shape differences was 65.5%.[7] The posterior distribution of the difference in population means between each sample and the 95% HDI is shown in Figure 5.

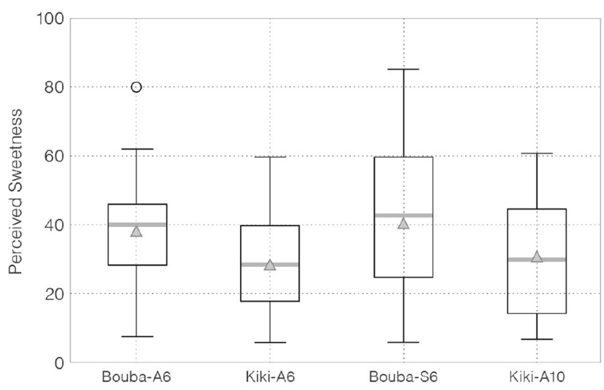

FIGURE 4
VAS (Visual Analog Scale) scores in participants' sweetness ratings. The boxes represent the middle 50% of the data, and the inside thick gray lines indicate the median value while the triangles indicate the mean value. The whiskers represent the maximum or minimum value of the data within 1.5 times the length of the box. Values not within this range were considered outliers (shown as individual data points).

3.1.1. Perceived sourness

Figure 6 shows a box-plot of participants' ratings of sourness for each edible sample.

Point estimation of population mean on perceived sourness by EAP regarding Bouba-A6, Kiki-A6, Bouba-S6, and Kiki-A10 are as follows: 28.49 (4.54) [19.20, 37.75], 33.45 (5.87) [22.41, 45.23], 25.75 (5.02) [15.88, 35.52], and 28.91 (5.40) [17.81, 39.36], respectively.

Table 2 lists the probability that the sample in row *i* is perceived as more sour than that in column *j*. A comparison of the population mean for perceived sourness between the basic Kiki-shaped sample (Kiki-A6) and Bouba-shaped sample (Bouba-A6) (H_2: Kiki-A6 >

6 For reference, the results of a *t*-test instead of a Bayesian analysis are presented here. Cohen's d is used to represent the effect size. The *t*-test results, $t_{(23)} = 2.22$, $p = 0.018$, $d = 0.45$, also show a significant difference between the two samples, supporting hypothesis H_1.

7 The *t*-test results, $t_{(23)} = 0.47$, $p = 0.321$, $d = 0.10$, suggest that the *p*-value is greater than the conventional significance level of 0.05. This value indicates a failure to reject the null hypothesis, as it does not provide significant statistical evidence to support the alternative hypothesis H_4.

TABLE 1 Probability that the sample in row *i* was perceived as sweeter than the sample in column *j*.

Sample	Sample			
	Bouba-A6 (μ_1)	Kiki-A6 (μ_2)	Bouba-S6 (μ_3)	Kiki-A10 (μ_4)
Bouba-A6 (μ_1)	0	0.987	0.345	0.956
Kiki-A6 (μ_2)	0.013	0	0.013	0.349
Bouba-S6 (μ_3)	0.655	0.987	0	0.966
Kiki-A10 (μ_4)	0.044	0.651	0.034	0

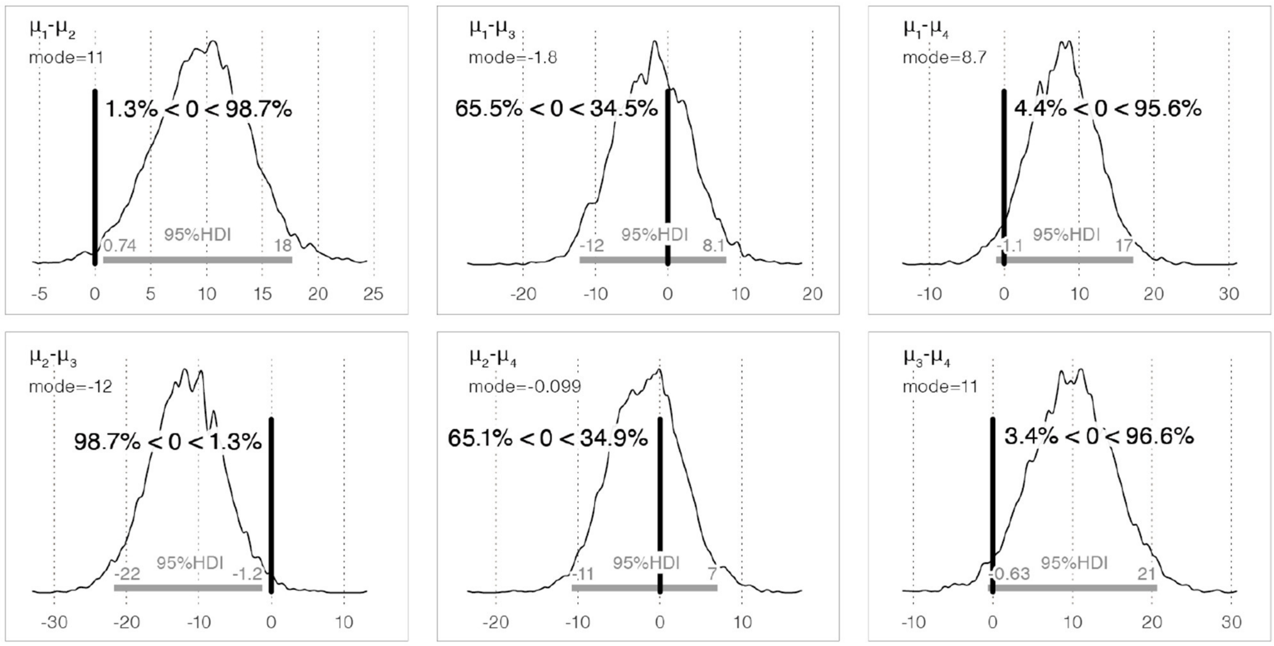

FIGURE 5
The posterior distribution of the difference in the population means for perceived sweetness between each sample, and the 95% HDI. Here, μ_i denotes the population mean of each edible sample (μ_1: Bouba-A6, μ_2: Kiki-A6, μ_3: Bouba-S6, μ_4: Kiki-A10). The vertical thick bar indicates the position of 0 and means the reference value. Note that $a\% < 0 < b\%$ of $\mu_i - \mu_j$ next to the vertical thick bar denotes that the research hypothesis $\mu_i > \mu_j$ holds the probability at $b\%$ and $\mu_i < \mu_j$ holds at $a\%$.

Bouba-A6) indicates that Kiki-A6 was perceived as more sour than Bouba-A6 with a probability of 82.0%.[8] Regarding H_5 (Kiki-A10 > Kiki-A6), the probability of differences in perceived sourness due to minor shape differences was 22.1%.[9] The posterior distribution of the differences in the population means of each sample and the 95% HDI are shown in Figure 7.

3.2. Perceived bitterness

Figure 8 shows a box-plot of participants' ratings of bitterness for each edible sample.

Point estimates of population mean for perceived bitterness by EAP for Bouba-A6, Kiki-A6, Bouba-S6, and Kiki-A10 were as follows: 54.29 (3.84) [46.62, 61.74], 55.21 (4.06) [47.07, 62.94], 49.14 (5.11) [38.94, 58.90], and 54.40 (5.24) [44.39, 65.04], respectively.

Table 3 lists the probabilities that the sample in row *i* was perceived as more bitter than that in column *j*. Comparison of the population mean of perceived bitterness between the basic Kiki-shaped sample (Kiki-A6) and Bouba-shaped sample (Bouba-A6) (H_3: Kiki-A6 > Bouba-A6) showed that Kiki-A6 was perceived as more bitter than Bouba-A6 with a probability of 57.3%.[10] Regarding H_6 (Kiki-A10 > Kiki-A6), the probability of differences in perceived bitterness owing to minor differences in shape was 43.7%.[11]

8 The *t*-test results, $t_{(23)} = 1.21$, $p = 0.120$, $d = 0.25$, indicate that the *p*-value is greater than the conventional significance level of 0.05. The results do not provide substantial statistical evidence to support the alternative hypothesis H_2.

9 The *t*-test results, $t_{(23)} = -0.99$, $p = 0.834$, $d = -0.20$, indicate that the *p*-value is considerably higher than the conventional significance level of 0.05. Consequently, the results do not provide substantial statistical evidence to support the alternative hypothesis H_5.

10 The *t*-test results, $t_{(23)} = 0.22$, $p = 0.414$, $d = 0.05$, indicate that the results do not provide substantial statistical evidence to support the alternative hypothesis H_3.

11 The *t*-test results, $t_{(23)} = -0.17$, $p = 0.568$, $d = -0.04$, indicate that the results do not provide substantial statistical evidence to support the alternative hypothesis H_6.

The posterior distribution of the difference in population means between each sample and the 95% HDI is shown in Figure 9.

4. Discussion

4.1. Reflections on study results

This study investigated crossmodal correspondences between food shape and taste. We found that Bouba-shaped (round) chocolates were perceived as sweeter than Kiki-shaped (angular) chocolates. The results supported the first hypothesis (H_1) and did not support the other hypotheses.

Regarding sweetness, our results supported H_1 and did not support H_4. The result for H_1 is consistent with previous studies on crossmodal correspondences between shape and taste perception, in other words, basic Bouba-shaped (rounded) stimuli are perceived as sweeter than Kiki-shaped (angular) stimuli. In particular, our results are consistent with previous studies showing crossmodal correspondences between round shape and sweetness both in visual food perception (Deroy and Valentin, 2011; Ngo et al., 2011; Salgado-Montejo et al., 2015) and in actual consumption experiments (Lenfant et al., 2013; Cornelio et al., 2022). However, it is possible that the participants in our experiment did not perceive a sufficient difference in the tactile stimulation provided by the different edible samples, even though we designed the samples considering the tip distances at the shape extremities.

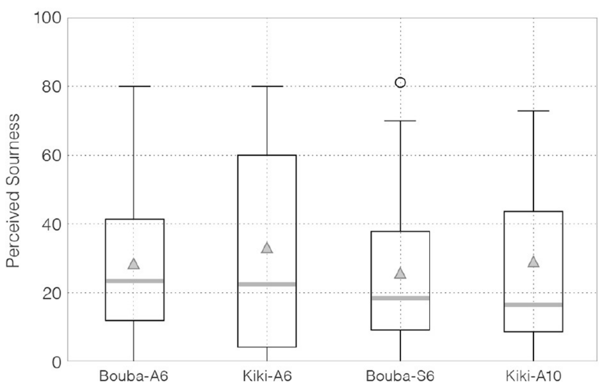

FIGURE 6
VAS scores in participants' sourness ratings.

Regarding sourness, the results did not support H_2 and H_5. A closer look at the ratings of sourness (Figure 6) showed that the average values of each sample were lower than the other two tastes. Based on these results, the following hypothesis is proposed: it was unfamiliar or unclear for participants to rate the sourness of the chocolate samples, unlike ratings of sweetness and bitterness. This will play a role in constructing a further studies to understand correspondences between shape and sourness: one might use a standard sample as a sour taste reference, with participants rating the difference between each sample and this reference, so that participants can more easily estimate the range of differences.

Finally, regarding bitterness, the results did not support H_3 and H_6. Bitterness is the most complex of the basic tastes (sweet, sour, bitter, salty and umami) (Trivedi, 2012) and the ability to detect bitter tastes decreases with age (Drewnowski, 2001). Thus, it is possible that the age range of the participants in our study caused the lack of shape effects on bitterness.

To further explore the relationship between shape and taste, the four types of chocolate used in the experiment were analyzed separately for Bouba- and Kiki-shaped samples. Assuming $(\mu_i, \mu_j) > (\mu_k, \mu_l)$, in which there is no up-down relationship between i and j or between k and l, but there is an up-down relationship between the two groups, the probability of this hypothesis being true can be obtained by calculating the EAP of the generation of the following Equation 2.

$$u^{(t)}_{\mu_i > \mu_k} \times u^{(t)}_{\mu_i > \mu_l} \times u^{(t)}_{\mu_j > \mu_k} \times u^{(t)}_{\mu_j > \mu_l} \tag{2}$$

The overall probability of Bouba-shaped samples being perceived as sweeter than the Kiki-shaped samples, regardless of shape differences, was 90.0% (0.900 = 0.956 × 0.987 × 0.987 × 0.966). For some of those sample pairs, the 95% HDI included zero; however, the probability mass of the posterior distribution was concentrated in the positive or negative range. Therefore, a small effect might have been present. Regarding sourness and bitterness, the results showed no significant difference between the taste of Bouba-shaped and Kiki-shaped samples. The overall probability of Kiki-shaped samples being perceived as more sour than the Bouba-shaped samples, regardless of shape differences, was 29.2% (0.292 = 0.820 × 0.917 × 0.535 × 0.728). The overall probability of Kiki-shaped samples being perceived as more bitter than the Bouba-shaped samples, regardless of shape differences, was 19.6% (0.196 = 0.573 × 0.851 × 0.506 × 0.794).

TABLE 2 Probability that the sample in row *i* was perceived as more sour than the sample in column *j*.

Sample	Sample			
	Bouba-A6 (μ_1)	Kiki-A6 (μ_2)	Bouba-S6 (μ_3)	Kiki-A10 (μ_4)
Bouba-A6 (μ_1)	0	0.180	0.705	0.465
Kiki-A6 (μ_2)	0.820	0	0.917	0.779
Bouba-S6 (μ_3)	0.295	0.083	0	0.272
Kiki-A10 (μ_4)	0.535	0.221	0.728	0

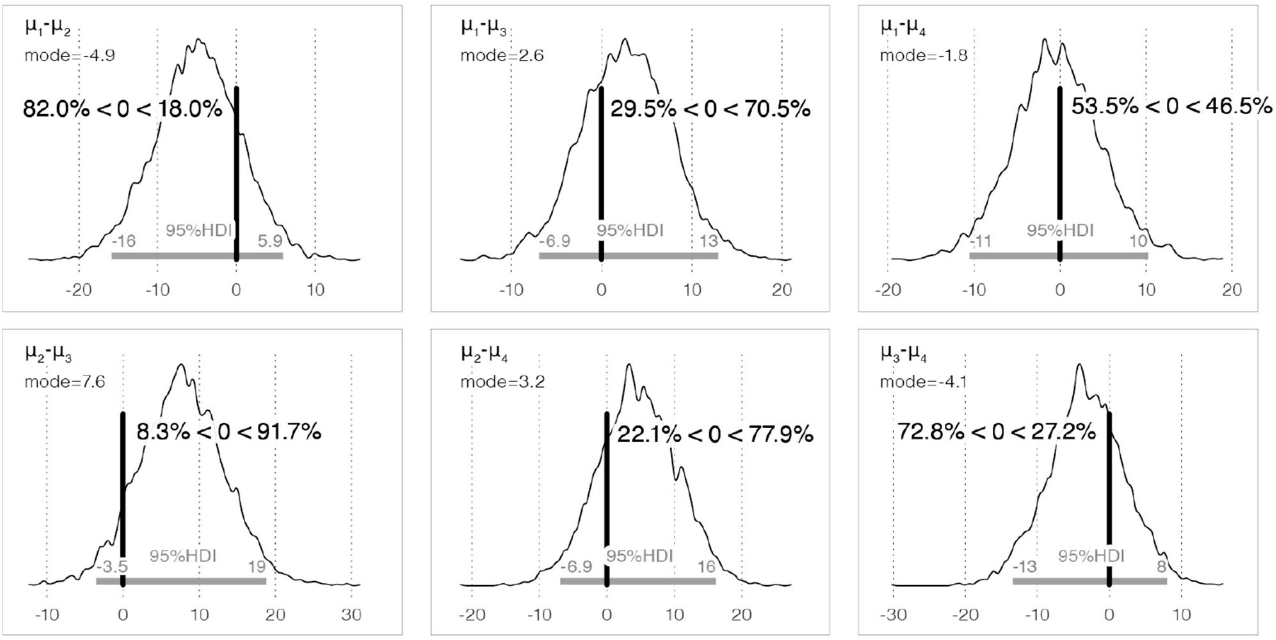

FIGURE 7
The posterior distribution of the difference in the population means for perceived sourness between each sample, and the 95% HDI.

We should note that generalizing the above results, namely the overall probability regardless of shape differences, may be challenging due to the limited samples. This limitation stems from our adoption of an asymmetric experimental design with a reduced sample size, which was intended to focus on examining the effects of the most likely effective parameters. Further study would be needed to test the influence of shape type, (a)symmetry, and a number of elements. In addition, we are still uncertain how much the change in taste/flavor itself affects generalization. In future work, we need to conduct experiments to see if the same round shape but different taste/flavor chocolate samples (e.g., milk chocolate, white chocolate) are perceived as sweeter than the angular shape.

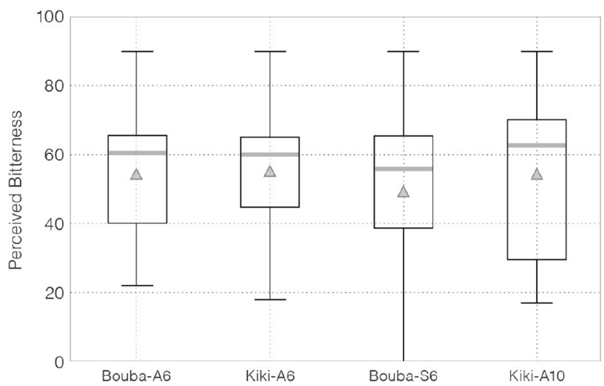

FIGURE 8
VAS scores in participants' bitterness ratings.

4.2. Consideration for actual eating experiments

Although we have demonstrated the possibility of crossmodal correspondences between shape and taste, some limitations such as visual stimuli, food structure, individual differences, and expansion to other modalities and taste characteristics remain uncertain in our study. We briefly discuss them as below.

4.2.1. Visual influence and tactile influence

One limitation of our study is that we could not distinguish between the effects of visual and tactile (mouthfeel) sensations of food shape on taste perception. It is known that visually perceived shape influences expectations regarding food taste. For example, consumers make judgments of food taste based on its packaging (Schifferstein, 2001). It is also known that haptic information through a hand plays a significant role in modulating food perception (Barnett-Cowan, 2010; Biggs et al., 2016). We focused on the effect of food shape on taste perception, and conducted an experiment in which participants ate four chocolate pieces with difference shapes. However, in the experiment, the participants could see the samples before eating to ensure that they were putting something edible into their mouths. Thus, we did not investigate the independent effects of visually perceived shape and tactile perceived shape. Further experiments are required to compare the results with and without observing stimuli. This would contribute to investigating crossmodal effects on taste, focusing only on shape stimuli perceived through the sense of touch.

TABLE 3 Probability that the sample in row *i* was perceived as more bitter than the sample in column *j*.

Sample	Sample			
	Bouba-A6 (μ_1)	Kiki-A6 (μ_2)	Bouba-S6 (μ_3)	Kiki-A10 (μ_4)
Bouba-A6 (μ_1)	0	0.427	0.822	0.494
Kiki-A6 (μ_2)	0.573	0	0.851	0.563
Bouba-S6 (μ_3)	0.178	0.149	0	0.206
Kiki-A10 (μ_4)	0.506	0.437	0.794	0

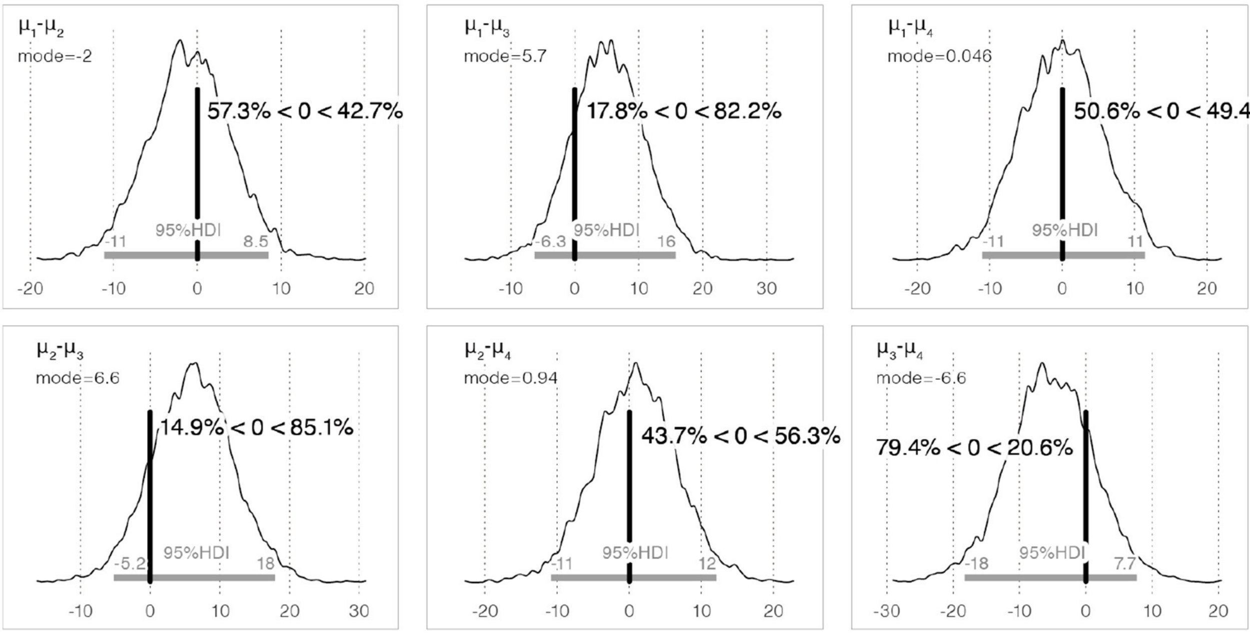

FIGURE 9
The posterior distribution of the difference in population means for perceived bitterness between each sample, and the 95% HDI.

Besides, we can also present a standard sample shape to clarify a taste evaluation baseline. Such an experimental design will help distinguish whether our results were perceived through food consumption or attributed through crossmodal correspondence, which our study did not clarify.

4.2.2. Food structure, fragility, and amount

In our experiment, we used two main characteristic shapes (round and angular) to measure the influence of food shape. The weights of the edible samples were balanced, and the influence of differences in the amount of material was controlled. However, structural engineering properties such as brittleness, fragility, and elasticity of the edible samples, were not included in the design parameters. In their open-ended responses, several participants stated that the Kiki-shaped (angular) samples crumbled more easily and quickly transformed into small particles that melted more quickly. There is a possibility that the Kiki-shaped samples had areas of greater structural stress and quickly crumbled into smaller particles, even though the edible samples were not bitten.

On the other hand, a previous study found that the ease of crumbling depended on the direction in which the mechanical metamaterial chocolate was bitten, and that there was a positive correlation with the number of cracks and sensory ratings (e.g., how crunchy and easy to bite) of the overall taste (Souto et al., 2022). Another study showed that food with a honeycomb infill pattern requires longer to chew than food with a rectilinear pattern (Lin et al., 2020). A longer chewing time indicates that the food is structurally stronger. If this structure is related to the ease of melting, it could be utilized when studying how taste perception may be related to changes in structure.

Finally, we believe the ring-shaped samples used in our study could be generalized for comparing other shapes, such as circles vs. triangles. However, we assume that if the surface area expanded, the difference in taste perception would be smaller. As we pointed out in the Introduction and Section 2.1.1, we assumed that the filled samples had too much chocolate to be sufficiently discriminable in a previous study (Wang et al., 2017). Although we recommend using ring-shaped samples, the optimal amount of chocolate needs more exploration to be addressed in future research.

4.2.3. Relationships between shapes and other taste characteristics

In this study, we focused on sweetness, bitterness, and sourness, which are basic human taste characteristics and tastes/flavors that chocolate typically has. However, there are other taste/flavor expressions such as spicy, oily, bland, astringency, and so on. Especially, astringency is a term that implies a characteristic of chocolate, wine, coffee, and so forth. When it comes to taste characteristics associated with chocolate, it may be possible to influence astringency by controlling the shape itself. Di Stefano and Spence (2022) noted that tactile stimuli produced by physical/mechanical contact between the food and the palate contribute to texture determination. These tactile stimuli include hardness, viscosity, elasticity and roughness. The study also discussed the relationship between roughness and astringency. In addition, Breen et al. (2019) showed that participants in the higher roughness sensitivity group responded to more subtle differences in chocolate roughness than those in the lower sensitivity group, suggesting that texture perception may be controllable. Roughness (which is angular), which tends to be perceived as unpleasant, like auditory and some visual and tactile stimuli, affects sensory and aesthetic responses in taste. Although our study highlighted results related to roundness and sweetness, it may be a good topic for future work to examine the possibility of manipulating the astringency evaluation by considering angularity as roughness.

Roughness, synonymous with angularity in this context, tends to be perceived as unpleasant and affects sensory and aesthetic responses to taste. Juravle et al. (2022) found that angular shapes, such as the Platonic solids, were typically associated with sour and bitter tastes. On the other hand, the study revealed a preference for sweetness and umami tastes with rounded shapes, such as spheres. This result suggests that the perception of different taste characteristics, including astringency, could be manipulated through the shape's angularity or roundness. Although our study primarily highlighted results related to roundness and sweetness, future work could consider the possibility of influencing astringency evaluations by manipulating the shape's roughness or angularity.

Moreover, regarding melting, Szubielska (2019) showed that chocolate was rated as tastier and more melt-in-the-mouth when eaten barefoot on a soft and smooth surface than a rough and hard surface. This study showed that taste perception also depends on the tactile properties of a contextual stimulus that occurs in the eating situation but is not related to the food itself, and this is in contrast to our finding that rough stimuli of the food itself influence melt-in-the-mouth. It will be interesting in future work to test which variable affects taste perception more between tactile contextual indicators or the tactile stimuli of the food itself.

4.2.4. Individual and cultural differences

The other consideration is individual differences in taste perception. Individual differences, e.g., gender/age/sweetness preference, could influence taste perception, which Bertelsen et al. (2020) demonstrated through their large-scale experiment. The ratings of chocolate samples might have changed slightly during the experiment depending on whether one prefers sweet or bitter foods. The chocolate used in our experiment was 74% cacao, which is relatively bitter. Those who prefer sweeter chocolate may have perceived a stronger bitter taste when they tasted the edible samples. One of the participants stated, "*I found the sourness was not my thing because I like sweet chocolate*", in the open-ended responses. In future work, we would like to investigate a correlation between individual differences and their evaluation of taste.

Geographical and cultural diversity should also be considered. A study focusing on the shape-taste matching of Bouba-Kiki showed that the effect size of each factor affecting shape-taste matching differed between the experimental participants from two different countries, namely the United Kingdom and Colombia (Salgado-Montejo et al., 2015), Taiwan and North America (Chen et al., 2016). Their study showed that shape-taste matching was similar across countries. On the other hand, Bremner et al. (2013) showed that the Himba of northern Namibia tended to associate less bitter chocolate samples with angular rather than rounded shapes, which is the opposite mapping to that shown by Westerners. It is prudent to continue to consider the possibility of cross-cultural differences in taste preferences (Prescott and Bell, 1995).

4.2.5. Crossmodal harmony

Besides, there is a recent study that argues about the "harmony" of multiple modalities. Various food scientists have used the term "crossmodal harmony" to represent this notion. Some of them work on the combination of fragrances and flavors and use it to "*refer to fused, or united, percepts while at the same time talking of the balanced composition of a mixture of elements*" in a dish (Spence and Di Stefano, 2022). A congruent crossmodal will be associated with an affectively more positive value. As we obtained from the open-ended responses, the participants stated, "*I was curious to see unusual star-shaped chocolates.*" or "*This is the first time I had ever eaten chocolate in such a unique shape, and it is interesting.*" They enjoyed the opportunity to eat unique-shaped chocolate, which might positively affect them. It would be highly insightful as future work to comprehensively capture variables such as visual appearance, taste, tactile, and shape to consider evaluating their balance and perceptions. Such a comprehensive perspective regarding harmony will be meaningful for studying taste perception through actual food consumption.

5. Conclusion

This study focused on the crossmodal correspondence between the food shape and taste perception. Accordingly, in an empirical experiment, participants ate Bouba- and Kiki-shaped chocolate pieces that were created using a 3D food printer and sensory evaluations were performed. A Bayesian analysis of the survey results revealed that participants perceived the Bouba-shaped (round) chocolate pieces as sweeter than the Kiki-shaped (angular) pieces. Our findings follow previous research on the crossmodal correspondences that occur between visual stimuli and taste perception. The results of this study will contribute to broadening the scope of food shape design and the proposal of new dining methods and experiences, as well as developing further crossmodal research.

Our results may enrich the dining experiences in the metaverse, which is a line of research that has gained momentum in recent years (Jaller et al., 2021; Cornelio et al., 2022). Research exploring dining experiences in VR spaces has already begun and is expected to become an area that will increasingly attract in the future. Similarly, some researchers are exploring the association between 3D shapes and expected taste, liking, and so on (Chen et al., 2019), which could lead directly to the dining experiences in the metaverse. The results of the present study suggest that taste perception can be controlled by changing the shape of food. In addition to visually manipulating taste (Narumi et al., 2011; Narumi, 2016), our study may contribute to approaches that use non-digital elements (Harley et al., 2018) and augment the visual experience with other sensory stimuli (Velasco et al., 2018).

In addition, we believe that it is worth exploring the effects of such a multi-sensory tasting to enrich the dining experiences and contribute to providing sustainable food experiences (Pedersen et al., 2021) in a way that reveals not only what to eat but also how to eat. In fact, chocolate has the highest carbon footprint of any plant-based food, emitting 19 kg of CO_2 equivalents per kilogram, compared to 17 kg for coffee and 8 kg for palm oil.[12] Given that the shape of chocolate can affect the perception of sweetness, it may be possible to deliver the same or much more sweetness with a smaller amount of chocolate by changing its shape, as demonstrated by the ring-shaped pieces used in this study. From the perspective of 3D printing, or additive manufacturing, this production process has potential environmental benefits (Jambrak et al., 2021). It is possible to decentralize the production site, to produce the designed chocolate locally at the desired time/unit, and to reduce CO_2 emissions during the production process, especially during transportation. If we can enjoy eating experiences with less chocolate, this will lead to a reduction in the ingredients needed and greenhouse gases from land use change and farm stage of the supply chain, which are the main contributors to emissions.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

12 https://ourworldindata.org/food-choice-vs-eating-local

Ethics statement

The studies involving human participants were reviewed and approved by OMRON Corporation Technology and Intellectual Property H.Q. Research Ethics Review Committee. The participants provided their written informed consent to participate in this study.

Author contributions

KO, RG, AH, YU, and SY designed the study. KO, RG, and SY performed the experiment. KO designed and prepared the experimental materials with RG. RG analyzed the data with SY. KO and RG wrote the first draft of the manuscript under the supervision of SY with further review by AH and YU. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by JST PRESTO Grant Number JPMJPR18JA and JSPS KAKENHI Grant Number JP21H04910.

Acknowledgments

We would like to thank everyone who participated in our preliminary pilot study and provided feedback. We also thank Sonoda Yasuko for tempering the chocolate for the experiment.

Conflict of interest

KO, RG, AH, YU, and SY were employed by OMRON SINIC X Corporation.

Publisher's note

References

Baptista, I. Y. F., Carvalho, F., Efraim, P., de Souza Silveira, P. T., and Behrens, J. (2022). The shape of creaminess: consumers expected and perceived rounded chocolates as creamier than squared. *Brit. Food J.* 124, 1697–1711. doi: 10.1108/BFJ-06-2021-0675

Barnett-Cowan, M. (2010). An illusion you can sink your teeth into: haptic cues modulate the perceived freshness and crispness of pretzels. *Perception* 39, 1684–1686. doi: 10.1068/p6784

Bertelsen, A. S., Mielby, L. A., Alexi, N., Byrne, D. V., and Kidmose, U. (2020). Individual differences in sweetness ratings and cross-modal aroma-taste interactions. *Foods* 9, 146. doi: 10.3390/foods9020146

Biggs, L., Juravle, G., and Spence, C. (2016). Haptic exploration of plateware alters the perceived texture and taste of food. *Food Qual. Prefer.* 50, 129–134. doi: 10.1016/j.foodqual.2016.02.007

Breen, S. P., Etter, N. M., Ziegler, G. R., and Hayes, J. E. (2019). Oral somatosensory function and the perception of chocolate particle size. *Sci. Rep.* 9, 7437. doi: 10.1038/s41598-019-43944-7

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., and Spence, C. (2013). "Bouba" and "Kiki" in namibia? a remote culture make similar shape-sound matches, but different shape-taste matches to westerners. *Cognition* 126, 165–172. doi: 10.1016/j.cognition.2012.09.007

Bronner, K., Frieler, K., Bruhn, H., Hirt, R., and Piper, D. (2012). "What is the sound of citrus? Research on the correspondences between the perception of sound and flavour," in *Proceedings of the ICMPC*. Available online at: http://icmpc.org/icmpc_conferences.html

Chen, Y.-C., Huang, P.-C., Woods, A., and Spence, C. (2016). When "Bouba" equals "Kiki": Cultural commonalities and cultural differences in sound-shape correspondences. *Sci. Rep.* 6, 26681. doi: 10.1038/srep26681

Chen, Y.-C., Huang, P.-C., Woods, A., and Spence, C. (2019). I know that "Kiki" is angular: the metacognition underlying sound-shape correspondences. *Psychon. Bull. Rev.* 26, 261–268. doi: 10.3758/s13423-018-1516-8

Chen, Y.-C., Woods, A. T., and Spence, C. (2018). Sensation transference from plateware to food: the sounds and tastes of plates. *Int. J. Food Design* 3, 41–62. doi: 10.1386/ijfd.3.1.41_1

Cornelio, P., Dawes, C., Maggioni, E., Bernardo, F., Schwalk, M., Mai, M., Pawlizak, S., et al. (2022). Virtually tasty: an investigation of the effect of ambient lighting and 3 D-shaped taste stimuli on taste perception in virtual reality. *Int. J. Gastronomy Food Sci.* 30, 100626. doi: 10.1016/j.ijgfs.2022.100626

Derossi, A., Caporizzi, R., Ricci, I., and Severini, C. (2019). "Critical variables in 3D food printing," in *Fundamentals of 3D food Printing and Applications*, eds F. C. Godoi, B. R. Bhandari, S. Prakash, and M. Zhang (Cambridge, MA: Academic Press), 41–91.

Deroy, O., and Valentin, D. (2011). Tasting liquid shapes: investigating the sensory basis of cross-modal correspondences. *Chemosens. Percept.* 4, 80–90. doi: 10.1007/s12078-011-9097-1

Di Stefano, N., and Spence, C. (2022). Roughness perception: a multisensory/crossmodal perspective. *Attent. Percept. Psychophys.* 84, 2087–2114. doi: 10.3758/s13414-022-02550-y

Drewnowski, A. (2001). The science and complexity of bitter taste. *Nutr. Rev.* 59, 163–169. doi: 10.1111/j.1753-4887.2001.tb07007.x

Gal, D., Wheeler, S. C., and Shiv, B. (2007). *Cross-Modal Influences on Gustatory Perception*. Available online at: https://ssrn.com/abstract=1030197

Gallace, A., Boschin, E., and Spence, C. (2011). On the taste of "Bouba" and "Kiki": an exploration of word–food associations in neurologically normal participants. *Cogn. Neurosci.* 2, 34–46. doi: 10.1080/17588928.2010.516820

Hao, L., Mellor, S., Seaman, O., Henderson, J., Sewell, N., and Sloan, M. (2010). Material characterisation and process development for chocolate additive layer manufacturing. *Virtual Phys. Protot.* 5, 57–64. doi: 10.1080/17452751003753212

Harley, D., Verni, A., Willis, M., Ng, A., Bozzo, L., and Mazalek, A. (2018). "Sensory VR: smelling, touching, and eating virtual reality," in *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction* (New York, NY: Association for Computing Machinery), 386–397. doi: 10.1145/3173225.3173241

Harrar, V., and Spence, C. (2013). The taste of cutlery: how the taste of food is affected by the weight, size, shape, and colour of the cutlery used to eat it. *Flavour* 2, 1–13. doi: 10.1186/2044-7248-2-21

Jacobsen, T., Schubotz, R. I., Höfel, L., and Cramon, D. Y. (2006). Brain correlates of aesthetic judgment of beauty. *Neuroimage* 29, 276–285. doi: 10.1016/j.neuroimage.2005.07.010

Jaller, C., Andersen, N. B. S., Nilsson, N. C., Paisa, R., Damsbo, M., and Serafin, S. (2021). "MARTYR: exploring ingredients of virtual dining experiences," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (IEEE Communications Society), 226–231. doi: 10.1109/VRW52623.2021.00049

Jambrak, A. R., Nutrizio, M., Djekic, I., Pleslic, S., and Chemat, F. (2021). Internet of nonthermal food processing technologies (IONTP): food industry 4.0 and sustainability. *Appl. Sci.* 11, 666. doi: 10.3390/app11020686

Juravle, G., Olari, E.-L., and Spence, C. (2022). A taste for beauty: on the expected taste, hardness, texture, and temperature of geometric shapes. *i-Perception* 13, 1–19. doi: 10.1177/20416695221120948

Köhler, W. (1929). *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York, NY: Liveright.

Lanaro, M., Desselle, M. R., and Woodruff, M. A. (2019). "3D printing chocolate: properties of formulations for extrusion, sintering, binding and ink jetting," in *Fundamentals of 3D Food Printing and Applications*, eds F. C. Godoi, B. R. Bhandari, S. Prakash, and M. Zhang (Cambridge, MA: Academic Press), 151–173.

Lee, C. A., and Vickers, Z. M. (2010). Discrimination among astringent samples is affected by choice of palate cleanser. *Food Qual. Prefer.* 21, 93–99. doi: 10.1016/j.foodqual.2009.08.003

Lenfant, F., Hartmann, C., Watzke, B., Breton, O., Loret, C., and Martin, N. (2013). Impact of the shape on sensory properties of individual dark chocolate pieces. *LWT Food Sci. Technol.* 51, 545–552. doi: 10.1016/j.lwt.2012.11.001

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *J. Multivar. Anal.* 100, 1989–2001. doi: 10.1016/j.jmva.2009.04.008

Lin, Y.-J., Punpongsanon, P., Wen, X., Iwai, D., Sato, K., Obrist, M., et al. (2020). "Foodfab: creating food perception illusions using food 3 D printing," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY: Association for Computing Machinery), 1–13. doi: 10.1145/3313831.3376421

Maeyama, T., and Plattig, K.-H. (1989). Minimal two-point discrimination in human tongue and palate. *Am. J. Otolaryngol.* 10, 342–344.

Makin, A. D. J., Pecchinenda, A., and Bertamini, M. (2012). Implicit affective evaluation of visual symmetry. *Emotion* 12, 1021. doi: 10.1037/a0026924

Mantihal, S., Prakash, S., and Bhandari, B. (2019). Texture-modified 3 D printed dark chocolate: sensory evaluation and consumer perception study. *J. Texture Stud.* 50, 386–399. doi: 10.1111/jtxs.12472

Mantihal, S., Prakash, S., Godoi, F. C., and Bhandari, B. (2017). Optimization of chocolate 3 D printing by correlating thermal and flow properties with 3 D structure modeling. *Innovat. Food Sci. Emerg. Technol.* 44, 21–29. doi: 10.1016/j.ifset.2017.09.012

Mesz, B., Trevisan, M., and Sigman, M. (2011). The taste of music. *Perception* 40, 209–219. doi: 10.1068/p6801

Michel, C., Velasco, C., Fraemohs, P., and Spence, C. (2015a). Studying the impact of plating on ratings of the food served in a naturalistic dining context. *Appetite* 90, 45–50. doi: 10.1016/j.appet.2015.02.030

Michel, C., Velasco, C., and Spence, C. (2015b). Cutlery matters: heavy cutlery enhances diners' enjoyment of the food served in a realistic dining environment. *Flavour* 4, 1–8. doi: 10.1186/s13411-015-0036-y

Narumi, T. (2016). "Multi-sensorial virtual reality and augmented human food interaction," in *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction* (New York, NY: Association for Computing Machinery), 1–6. doi: 10.1145/3007577.3007587

Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M. (2011). "Meta cookie+: an illusion-based gustatory display," in *Virtual and Mixed Reality - New Trends* (Berlin; Heidelberg: Springer), 260–269.

Ngo, M. K., Misra, R., and Spence, C. (2011). Assessing the shapes and speech sounds that people associate with chocolate samples varying in cocoa content. *Food Qual.Prefer.* 22, 567–572. doi: 10.1016/j.foodqual.2011.03.009

Ngo, M. K., and Spence, C. (2011). Assessing the shapes and speech sounds that consumers associate with different kinds of chocolate. *J. Sens. Stud.* 26, 421–428. doi: 10.1111/j.1745-459X.2011.00359.x

Nygård, M., and Lie, C. L. (2020). *How changes in environmental colour hue affect taste expectations, perceptions, and product preferences at different levels of attention towards atmospheric cues: a mixed experimental design* (Master's thesis). Handelshøyskolen BI, Oslo, Norway..

Obrist, M., Comber, R., Subramanian, S., Piqueras-Fiszman, B., Velasco, C., and Spence, C. (2014). "Temporal, affective, and embodied characteristics of taste experiences: a framework for design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY: Association for Computing Machinery), 2853–2862.

Palmer, S. E., Schloss, K. B., and Sammartino, J. (2013). Visual aesthetics and human preference. *Annu. Rev. Psychol.* 64, 77–107. doi: 10.1146/annurev-psych-120710-100504

Pedersen, M. T., Hansen, P. L., and Clausen, M. P. (2021). Gastronomy unravelled by physics: gastrophysics. *Int. J. Food Design* 6, 153–180. doi: 10.1386/ijfd_00029_1

Prescott, J., and Bell, G. (1995). Cross-cultural determinants of food acceptability: recent research on sensory perceptions and preferences. *Trends Food Sci. Technol.* 6, 201–205.

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

Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., and Spence, C. (2015). The sweetest thing: the influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Front. Psychol.* 6, 1382. doi: 10.3389/fpsyg.2015.01382

Schifferstein, H. N. J. (2001). "Effects of product beliefs on product perception and liking," in *Food, People and Society: A European Perspective of Consumers' Food Choices*, L. J. Frewer, E. Risvik, E., and H. Schifferstein (Berlin; Heidelberg: Springer), 73–96.

Simmonds, G., Woods, A. T., and Spence, C. (2019). 'Shaping perceptions': exploring how the shape of transparent windows in packaging designs affects product evaluation. *Food Qual.Prefer.* 75, 15–22. doi: 10.1016/j.foodqual.2019.02.003

Slocombe, B., Carmichael, D., and Simner, J. (2016). Cross-modal tactile–taste interactions in food evaluations. *Neuropsychologia* 88, 58–64. doi: 10.1016/j.neuropsychologia.2015.07.011

Souto, A., Zhang, J., Aragón, A. M., Velikov, K. P., and Coulais, C. (2022). Edible mechanical metamaterials with designed fracture for mouthfeel control. *Soft Matter* 18, 2910–2919. doi: 10.1039/D1SM01761F

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attent. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C., and Deroy, O. (2014). On the shapes of flavours: a review of four hypotheses. *Theor. Histor. Sci.* 10, 207–238. doi: 10.12775/ths-2013-0011

Spence, C., and Di Stefano, N. (2022). Crossmodal harmony: looking for the meaning of harmony beyond hearing. *i-Perception* 13, 20416695211073817. doi: 10.1177/20416695211073817

Spence, C., and Gallace, A. (2011). Tasting shapes and words. *Food Qual. Prefer.* 22, 290–295. doi: 10.1016/j.foodqual.2010.11.005

Spence, C., Ngo, M. K., Percival, B., and Smith, B. (2013). Crossmodal correspondences: assessing shape symbolism for cheese. *Food Qual. Prefer.* 28, 206–212. doi: 10.1016/j.foodqual.2012.08.002

Spence, C., Velasco, C., and Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour* 3, 1–12. doi: 10.1186/2044-7248-3-8

Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., et al. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour* 4, 1–17. doi: 10.1186/s13411-015-0033-1

Szubielska, M. (2019). Wpływ kontekstowego bodźca dotykowego na percepcję smaku czekolady [Effect of a contextual tactile stimulus on the chocolate taste perception]. *Rocznik Kognitywistyczny* 11, 45–54. doi: 10.4467/20843895RK.18.004.9758

Trivedi, B. P. (2012). Gustatory system: the finer points of taste. *Nature* 486, S2–S3. doi: 10.1038/486S2a

Van Doorn, G., Woods, A., Levitan, C. A., Wan, X., Velasco, C., Bernal-Torres, C., et al. (2017). Does the shape of a cup influence coffee taste expectations? A cross-cultural, online study. *Food Qual. Prefer.* 56, 201–211. doi: 10.1016/j.foodqual.2016.10.013

Velasco, C., Jones, R., King, S., and Spence, C. (2013). Assessing the influence of the multisensory environment on the whisky drinking experience. *Flavour* 2, 1–11. doi: 10.1186/2044-7248-2-23

Velasco, C., Michel, C., Youssef, J., Gamez, X., Cheok, A. D., and Spence, C. (2016a). Colour–taste correspondences: designing food experiences to meet expectations or to surprise. *Int. J. Food Design* 1, 83–102. doi: 10.1386/ijfd.1.2.83_1

Velasco, C., Tu, Y., and Obrist, M. (2018). "Towards multisensory storytelling with taste and flavor," in *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction* (New York, NY: Association for Computing Machinery).

Velasco, C., Woods, A. T., Petit, O., Cheok, A. D., and Spence, C. (2016b). Crossmodal correspondences between taste and shape, and their implications for product packaging: a review. *Food Qual. Prefer.* 52, 17–26. doi: 10.1016/j.foodqual.2016.03.005

Wan, X., Zhou, X., Woods, A. T., and Spence, C. (2015). Influence of the glassware on the perception of alcoholic drinks. *Food Qual. Prefer.* 44, 101–110. doi: 10.1016/j.foodqual.2015.03.018

Wang, Q. J., Reinoso Carvalho, F., Persoone, D., and Spence, C. (2017). Assessing the effect of shape on the evaluation of expected and actual chocolate flavour. *Flavour* 6, 1–6. doi: 10.1186/s13411-017-0052-1

TYPE Original Research
PUBLISHED 23 June 2023
DOI 10.3389/fpsyg.2023.1166543

OPEN ACCESS

EDITED BY
Na Chen,
Bar-Ilan University, Israel

REVIEWED BY
Frederique Robin,
Université de Nantes, France
Antonio Aquino,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

*CORRESPONDENCE
Michio Nomura
✉ nomura.michio.8u@kyoto-u.ac.jp

RECEIVED 15 February 2023
ACCEPTED 08 June 2023
PUBLISHED 23 June 2023

CITATION
Hitsuwari J and Nomura M (2023) Developing and validating a Japanese version of the Plymouth Sensory Imagery Questionnaire. *Front. Psychol.* 14:1166543.
doi: 10.3389/fpsyg.2023.1166543

Developing and validating a Japanese version of the Plymouth Sensory Imagery Questionnaire

Jimpei Hitsuwari and Michio Nomura*

Department of Education, Kyoto University, Kyoto City, Japan

Mental imagery refers to the representation of stimuli that are not physically present and has long been a subject of interest in psychology. However, most research on mental imagery has been limited to visual images, with other types of imagery, such as sound and smell, receiving little attention. A possible reason for this is the lack of appropriate scales to measure the vividness of multisensory imagery. The Plymouth Sensory Imagery Scale (Psi-Q) has been developed to address this issue and has been used in several studies to measure the vividness of seven imageries: vision, sound, smell, taste, touch, body, and feeling. In this study of 400 participants in Japan, the Psi-Q was translated into Japanese and tested for reliability and validity. The results showed good internal reliability and retest reliability and moderate to high correlations with other measures of construct validity, including mindfulness, Big Five, and life satisfaction. Additionally, there is no significant difference in total Psi-Q scores between the Japanese and British samples, although some differences are found in individual sensory imagery abilities. This study provides valuable insights into multisensory mental imagery, and it is expected that research dealing simultaneously with the responses of multisensory modalities will further accumulate.

KEYWORDS
multi-sensory imagery, mental imagery, scale development, individual differences, cultural comparison, assessment, mindfulness

1. Introduction

Mental imagery is defined as "an internal representation of stimuli that are not physically present (Shaw, 2008, p. 175)" and has been a central theme in psychology over the past several decades (Pylyshyn, 1973; Pearson, 2019). The range of topics utilizing mental imagery includes physiology, perception, learning, memory, and exercise (Hatakeyama, 2019), and applied research has been conducted in areas such as marketing and medicine (e.g., Pearson et al., 2015; Zaleskiewicz et al., 2020). However, most of these studies regarding mental imagery have been limited to vision; other mental imagery, such as sound and smell, have been studied less actively (Andrade et al., 2014). This is contrary to the growing interest in the topic of crossmodal correspondence, which attempts to examine the interaction of multiple sensory perceptions rather than just one (Uno and Yokosawa, 2022). One possible reason for this is that there is a lack of a scale with sufficient validity and reliability available to measure the vividness of multisensory imagery simultaneously. In reality, some scales such as the Questionnaire upon Mental Imagery (QMI; Betts, 1909; Sheehan, 1967 for short version) and the Survey of Mental Imagery (SMI; Switras, 1978; Grebot, 2003 for French version), which can measure multisensory imagery ability, have been developed, but these measures are problematic in three aspects. First, the scales were not created through adequate psychometric testing (McAvinue and Robertson, 2007; Andrade et al., 2014). For example, the SMI does not examine correlations with existing scales to determine construct validity, and there are no reports of test–retest reliability. Second, the QMI and SMI are long scales with 150 and 86 items, respectively, which has the disadvantage of placing a high burden on respondents. Third, these scales have been in development for a long time, and the wordings of some items are not suitable today. For example, the QMI uses items that reference the whistle of a locomotive (sound), velvet (touch), and smoke from a train (smell), which are no longer common experiences. Based on these observations, the Plymouth Sensory Imagery Scale (Psi-Q; Andrade et al., 2014) was developed to overcome these problems and simultaneously measure the vividness of seven modality imageries: vision, sound, smell, taste, touch, body sensation, and feeling. Andrade et al. (2014) sampled extensively across multisensory modalities, reviewing the QMI and the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). Consequently, two items were retained and eight items were rephrased from the short version of the QMI, and 25 new items were added. The scale was developed through three experiments ($N=854$). This has demonstrated construct validity ($r=0.18$–0.40 with the Spontaneous Use of Imagery Scale (SUIS); Reisberg et al., 2003), internal reliability ($\alpha=0.93$–0.96 for the Psi-Q; $\alpha=0.80$–0.97 for the subscales), and retest reliability ($r=0.43$–0.84). Since then, the Psi-Q has been translated worldwide, including a Spanish version (Pérez-Fabello and Campos, 2020) and a Dutch version (Woelk et al., 2022), and has become a leading scale for measuring multisensory imagery ability. The Spanish version showed a seven-factor structure as in the original version, despite having removed four items, with adequate to high internal consistency ($\alpha=0.92$ for the Psi-Q; $\alpha=0.68$–0.77 for the subscales) and correlation with the QMI ($r=0.40$–0.56), confirming its construct validity. The Dutch version showed a seven-factor structure, similar to the original version, with adequate to good internal consistency ($\alpha=0.94$–0.96 for Psi-Q; $\alpha=0.76$–0.88 for subscales) and test–retest reliability ($r=0.83$ for Psi-Q; $r=0.67$–0.75 for subscales), and construct validity ($r=0.32$ with SUIS) using two surveys of students and a general sample. Translating and developing the Psi-Q in Japan will ensure the measurement of multisensory imagery ability simultaneously and confirm that the same factor structure can be found across cultures.

The Psi-Q has been employed in a variety of surveys and experiments (e.g., Kharlas and Frewen, 2016; Clark et al., 2020; Koivisto and Grassini, 2022). Kharlas and Frewen (2016) have examined a trait mindfulness scale's association with the Psi-Q and found moderate positive correlations between its five subscales and each of the Psi-Q's sensory imagery abilities. In particular, the "Observe" factor of the trait mindfulness scale was significantly correlated with mental imagery in all seven sensory organs ($r=0.23$–0.47, $ps<0.01$), while only bodily and feeling imagery was not correlated with "Act with awareness" ($r=0.06$–0.18, $ps>0.05$), suggesting that there were differences across sensory organs. In addition, the Psi-Q has been used in behavioral experiments; for example, Koivisto and Grassini (2022) have shown that images of nature (vs. urban environment and architecture) are associated with positive affect and relaxation. They further have found that visual imagery ability in the Psi-Q predicted the vividness of imagery most successfully ($B=0.34$, $p<001$), while bodily imagery ability predicted relaxation most successfully ($B=26$, $p=0.002$). Thus, it is important to simultaneously measure individual differences in multisensory imagery ability and identify commonalities and differences among sensory organs. In this study, in addition to mindfulness, we further aim to deepen our understanding of multisensory mental imagery by examining its relevance to other scales, such as the Big Five and Satisfaction with Life Scale, because these measures have not yet been examined in relation to multisensory imagery abilities but have been shown to be associated with single modalities such as visual imagery (e.g., Wilson et al., 2018; Budnik-Przybylska et al., 2023). We hypothesized, with reference to these previous studies, that multisensory imagery ability would correlate positively with mindfulness, negatively with neuroticism, positively with openness and extraversion, and positively with life satisfaction.

Scales measuring the individual differences in mental imagery have developed for each modality rather than multiple modalities simultaneously. For example, the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973), Vividness of Olfactory Imagery Questionnaire (VOIQ; Gilbert et al., 1998), and Auditory Imagery Questionnaire (AIQ; Hishitani, 2009), which measure visual, olfactory, and auditory imagery abilities, respectively, have been developed. Japanese versions of these scales likewise exist and have been widely used (Hishitani, 2005; Hishitani, 2011; Yamamoto et al., 2018). However, there is a growing need to simultaneously measure multisensory mental imagery abilities for use with other scales and behavioral experiments; therefore, it is important to develop a Japanese version of the Psi-Q.

To our knowledge there are no studies examining the cultural differences in multisensory imagery ability. While not comparing countries, Talamini et al. (2022) has examined musicians' and non-musicians' questionnaire responses regarding visual and auditory imagery abilities and found that musicians had higher auditory imagery abilities than non-musicians, however, there was no difference in visual imagery abilities. The results suggest that mental imagery is influenced by culture and environment and that this influence differs from one sensory system to another. These differences are not only in the concepts that have been treated in psychology, such as the cultural self (Markus and Kitayama, 1991) and genotypes associated with depression (Chiao and Blizinsky, 2010), but also in the customs and food culture. Therefore, it is vital to exploratively examine the cultural differences of multisensory imagery.

1.1. Aims

In this study, we first translated the original version of the Psi-Q (Andrade et al., 2014) to develop a Japanese version, and then examined its validity and reliability. To examine construct validity, we measured existing imagery ability scales, such as the VOIQ, and examined relationships with individual differences in the Big Five, trait mindfulness, and life satisfaction. Finally, an exploratory examination of cultural differences in multisensory imagery ability between the United Kingdom and Japan was conducted using original data (Andrade et al., 2014). This study can inspire research on cultural comparisons of multisensory imagery abilities which have not yet been directly compared, despite the repeated implication that they are influenced by culture and environment.

2. Methods

This study was approved by the Ethics Committee of Kyoto University (CPE-496). All data and scripts are available online.[1]

1 https://osf.io/u8cpd/

2.1. Participants

Four hundred people were recruited through the Japanese crowdsourcing platform CrowdWorks.[2] Of these, 389 (*M* = 40.72, *SD* = 10.73; 155 men, 234 women) were analyzed, excluding those with extremely short response times and those who missed the attention check. To examine retest reliability, the same participants were invited to respond again 1 week later. Data from 344 participants (*M* = 40.85, *SD* = 10.65; 137 men, 207 women), whose ID matched their first completed questionnaire, were used for the retest reliability analysis.

2.2. Materials

2.2.1. The Plymouth sensory imagery scale (Psi-Q)

After obtaining permission from the original authors, the first author, a fluent and native Japanese speaker, translated the Psi-Q into Japanese (Andrade et al., 2014). All items used in the original version of the Psi-Q were not changed to consider cultural influences because they all are familiar to Japanese. For the Japanese version, we conducted three preliminary surveys with a sample size of 100–150 participants and made minor revisions to the items. Back-translation was then performed using a translation service (NAI Inc.[3]). The original author was then asked to confirm whether there were any differences in meaning or intent from the original version. The original version of the scale consists of seven subscales, five items each, for "vision," "sound," "smell," "taste," "touch," "body," and "feeling," for a total of 35 items. A shortened version, with three items each for a total of 21 items, has also been developed. For example, in the case of "sense of smell," the participants are asked to respond to the question, "Imagine the smell of a rose," using a 7-point scale from 1, "I cannot imagine it at all," to 7, "It is as vivid as if it were right in front of my eyes." The completed Japanese version of Psi-Q is described in Table 1.

2.2.2. The vividness of olfactory imagery questionnaire (VOIQ)

The VOIQ is a 14-item scale measuring olfactory imagery ability (Gilbert et al., 1998; Yamamoto et al., 2018). The responses are provided using a 5-point scale, ranging from 1, "I cannot smell anything at all, I just know that I am thinking about the smell that I am told," to 5, "I can smell the smell completely clearly, as if I am smelling a real object."

2.2.3. The auditory imagery questionnaire (AIQ)

The AIQ is a 12-item scale measuring auditory imagery ability (Hishitani, 2009), using a 5-point scale ranging from 1, "I have no image at all, I just 'know' that I am thinking about what I am told," to 5, "It is completely clear, like I am hearing a real thing."

2 https://crowdworks.jp/

3 https://www.nai.co.jp/

2.2.4. Short version of the Japanese Big Five Scale

This shortened version of the Big Five Scale is a 29-item scale measuring the Big Five of extraversion, agreeableness, neuroticism, openness, and conscientiousness (Namikawa et al., 2012). The responses to the personality adjectives are provided on a 7-point scale ranging from 1, "not at all true," to 7, "very true." Although no studies have examined the relationship between multisensory mental imagery ability and Big Five personality, visual imagery ability and Big Five extraversion have been found to be positively related (McDougall and Pfeifer, 2012), and thus construct validity can be examined. In addition, we also examine the correlations between other personalities and mental imagery abilities in an exploratory manner.

2.2.5. Short version of the five facet mindfulness questionnaire (FFMQ)

The shortened version of the Five Facet Mindfulness Questionnaire is a 24-item scale consisting of five factors (Observing, Non-reactivity, Non-judging, Describing, and Acting with awareness) measuring trait mindfulness (Takahashi et al., 2022). A 5-point scale ranged from 1, "not at all true," to 5, "always true." It has shown a moderate correlation with the Psi-Q (Kharlas and Frewen, 2016) and serves as an index for examining construct validity.

2.2.6. The satisfaction with life scale (SWLS)

The Satisfaction With Life Scale is a one-factor, five-item scale measuring life satisfaction (Diener et al., 1985; Sumino, 1994). A 7-point scale ranging from 1, "not at all agree," to 7, "very much agree," is used to answer the questions. To our knowledge, there are no studies that examine the correlation between multi-sensory mental imagery ability and life satisfaction. However, well-being and mental imagery are related, as people with higher depression experience more negative imagery (Holmes et al., 2016); therefore, it will be interesting to look at the relationship between other-sensory imagery ability and life satisfaction in this study.

2.3. Procedure

The participants, recruited through Crowdworks, completed a web-based questionnaire created by Qualtrics. The participants were briefed on the survey and signed an informed consent form. They then began with responding to the Psi-Q, followed by the other five scales, which were presented in a randomized order. The items were also presented in a randomized order. The survey took approximately 10 min to complete. One week later, the same participants responded to the Psi-Q as described above.

2.4. Data analysis

First, using the *fa* function of the *psych* package (Revelle, 2022) in R (ver. 4.2.2; R Core Team, 2022), an exploratory factor analysis was conducted. A confirmatory factor analysis as well as a comparison of the models was then conducted using the *cfa* function from the *lavaan* package (Rosseel, 2012). Based on previous research (Yong and Pearce, 2013), we set a cutoff criterion of 0.4. To compare the model fit, we used chi square (χ^2), goodness of fit index (GFI), adjusted GFI (AGFI), normed fit index (NFI), comparative fit index (CFI), root

TABLE 1 Factor loadings for each item in the factor analysis with a six-factor structure (only items with factor loadings of 0.4 or higher are shown).

	English	Japanese	*M*	SD	Vision	Sound	Smell	Taste	Touch	Body
Imagine the appearance of...										
1	*a bonfire	焚き火	5.67	1.19	0.75					
2	*a sunset	夕焼け	5.84	1.08	0.69					
3	*a cat climbing a tree	木に登る猫	5.35	1.39	0.52					
4	a friend you know well	よく知っている友人	5.74	1.26	0.51					
5	the front door of your house	自宅の玄関扉	6.16	1.02						
Imagine the sound of...										
6	*the sound of a car horn	車のクラクションの音	5.81	1.12		0.70				
7	*hands clapping in applause	割れんばかりの拍手	5.58	1.24		0.45				
8	*an ambulance siren.	救急車のサイレン	6.06	1.02		0.74				
9	the sound of children playing	子どもの遊ぶ声	5.71	1.21		0.57				
10	the mewing of a cat	猫の鳴き声	5.87	1.13		0.57				
Imagine the smell of...										
11	*newly cut grass	刈りたての草	4.66	1.55						
12	*burning wood	燃えている木	3.93	1.65			0.64			
13	*a rose	バラの花	4.21	1.83			0.49			
14	fresh paint	塗りたてのペンキ	4.43	1.55			0.52			
15	a stuffy room	むっとする部屋	4.27	1.53						
Imagine the taste of...										
16	*black pepper	ブラックペッパー	5.09	1.52				0.68		
17	*lemon	レモン	5.87	1.13				0.57		
18	*mustard	マスタード	4.88	1.50				0.74		
19	toothpaste	歯磨き粉	5.88	1.16				0.51		
20	sea water	海水	4.63	1.59						
Imagine touching...										
21	*fur	ふわふわとした毛皮	5.10	1.50					0.52	
22	*warm sand	暖かさをもった砂	4.83	1.50					0.64	
23	*a soft towel	柔らかいタオル	5.89	1.05						
24	icy water	氷水	5.84	1.13					0.45	
25	the point of a pin	ピンの先	4.82	1.65					0.53	
Imagine the bodily sensation of...										
26	*relaxing in a warm bath	温かいお風呂につかってリラックスする	5.98	1.09						0.68
27	*walking briskly in the cold	真冬に外で足早に歩く	4.86	1.53						0.46
28	*jumping into a swimming pool	プールの水面に飛び込む	4.64	1.57						0.48
29	having a sore throat	のどが痛む	5.59	1.21						0.55
30	threading a needle	針に糸を通す	5.50	1.26						0.50

M, mean; SD, standard deviation. * Indicates items used in the short version.

mean square error of approximation (RMSEA), and Akaike's information criterion (AIC). To examine internal reliability, we used the *alpha* function from the *psych* package, and retest reliability was confirmed by calculating correlations with data obtained from the same sample 1 week later. For the British sample, we used data from Study 2 of Andrade et al. (2014), using the *t.test* function and the *mes* function from the *compute.es* package (Del Re, 2013) for effect size calculations.

3. Results

3.1. Factor structure

First, exploratory factor analysis was conducted to examine the eight-factor structure indicated by the scree plot, which resulted in only two items loading on "body" and two "feeling" factors (Supplementary Table S1). Alternatively, when the factor analysis was conducted assuming a six-factor structure, excluding the "feeling" factor from the original version, all factors loaded three or more items, resulting in a cohesive result (Table 1 and Supplementary Table S2 showing all factor loads).

As is clear from the model fit indices such as GFI and AIC in Table 2, the six-factor structure excluding the "feeling" factor is a better model than the original seven-factor structure.

Following the original study (Andrade et al., 2014), we also examined the factor structure of the short version of the Psi-Q, excluding two items from each of the six factors. The results show that the factor loadings for the three items did not exceed 0.4 but generally showed good coherence (Supplementary Table S2).

3.2. Descriptive statistics and reliability

As a result of the factor analysis, a six-factor Psi-Q structure was employed, and its descriptive statistics and internal reliability for the entire scale and each factor were calculated (Table 3). The results show good internal reliability, $\alpha = 0.78 \sim 0.94$, and good retest reliability, $r = 0.67 \sim 0.82$ ($ps = 0.00$), when retests were conducted 1 week apart.

3.3. Examination of construct validity

To examine construct validity, correlations with other measures were calculated (Table 4).

The results show moderate to high correlations and validity with existing measures of imagery ability, with total Psi-Q scores correlating with the AIQ ($r = 0.68$, $p = 0.00$), which measures auditory imagery ability, and the VOIQ ($r = 0.59$, $p = 0.00$), which measures olfactory imagery ability. There are also positive correlations with the Big Five factors of extraversion ($r = 0.27$, $p = 0.00$), openness ($r = 0.19$, $p = 0.00$), and agreeableness ($r = 0.14$, $p = 0.00$), and a negative correlation with neuroticism ($r = -0.13$, $p = 0.01$). In addition to the positive correlation between the total FFMQ scores and the Psi-Q ($r = 0.24$, $ps = 0.00$), the Observe factor and each subfactor of the Psi-Q shows moderate correlations ($r = 0.18 \sim 0.32$, $p = 0.00$). Life satisfaction and total scores on the Psi-Q and each subscale are positively correlated ($r = 0.16 \sim 0.24$, $p = 0.00$). Age ($r = 0.21$, $p = 0.00$) and gender ($r = 0.14$, $p = 0.01$) also correlated positively with the Psi-Q. This indicates that older participants and female participants have higher imagery ability. These correlations indicating construct validity were significant, but the effect sizes were week to moderate.

3.4. Cultural comparison

Finally, we compare the Psi-Q scores of the Japanese sample in this study with the British sample in the original study (Table 1 and Figure 1). ANOVA results showed a non-significant main effect of culture [$F(1, 596) = 1.06$, $p = 0.30$, $\eta^2 = 0.001$], a significant main effect of modality [$F(6, 3{,}576) = 243.31$, $p = 0.00$, $\eta^2 = 0.13$], and a significant interaction [$F(6, 3{,}576) = 28.17$, $p < 0.001$, $\eta^2 = 0.02$]. Multiple comparison showed significant differences of sound [$F(1, 596) = 25.50$, $p < 0.001$, $\eta^2 = 0.04$], smell [$F(1, 596) = 8.81$, $p = 0.003$, $\eta^2 = 0.01$], taste [$F(1, 596) = 25.75$, $p < 0.001$, $\eta^2 = 0.04$], and touch [$F(1, 596) = 5.18$, $p = 0.02$, $\eta^2 = 0.01$] between cultures. For sound and taste, Japanese scores were higher than British scores. For smell and taste, British scores were higher than Japanese scores. The culture differences for overall Psi-Q score [$F(1, 596) = 1.06$, $p = 0.30$, $\eta^2 = 0.002$], vision [$F(1, 596) = 0.88$, $p = 0.35$, $\eta^2 = 0.002$], and body [$F(1, 596) = 1.02$, $p = 0.31$, $\eta^2 = 0.002$] were non-significant.

4. Discussion

4.1. The Japanese version of Plymouth Sensory Imagery Questionnaire

In this study, we developed a Japanese version of the Psi-Q (Andrade et al., 2014), which measures individual differences in multisensory imagery abilities. Internal reliability and retest reliability for this scale were high. The original version measured imagery ability on seven subscales: vision, sound, smell, taste, touch, body, and emotion; however, the Japanese version, which was administered to a Japanese sample, showed the best coherence on six subscales, excluding the emotion factor. When considering the sensory organs, the emotional factor is rarely included, and it is thought to work well as a scale measuring multisensory imagery ability.

TABLE 2 Comparison of the goodness of fit for each model.

	χ^2	df	*P*-value	GFI	AGFI	NFI	CFI	RMSEA	AIC
7 factors, 35 items	1245.44	539.00	0.00	0.84	0.81	0.81	0.88	0.06	41290.04
6 factors, 30 items	806.49	390.00	0.00	0.87	0.85	0.85	0.92	0.05	35104.72
6 factors, 25 items	530.35	260.00	0.00	0.90	0.87	0.88	0.93	0.05	29281.00

GFI, goodness of fit index; AGFI, adjusted GFI; NFI, normed fit index; CFI, comparative fit index; RMSEA, root mean square error of approximation; AIC, Akaike's information criterion. 25-items version excludes items that do not load on each factor in Table 1.

TABLE 3 Descriptive statistics and retest reliability of the Plymouth Sensory Imagery Questionnaire, compared to United Kingdom data.

	Day-1 data (*N*=389)			Day-2 data (*N*=344)				UK data from study 2 of Andrade et al. (2014) (*N*=209)				
	M	SD	*α*	*M*	SD	*r*	*p*	M	SD	*t*	*p*	*d*
Psi-Q	5.29	0.80	0.94	5.14	0.82	0.81	0.00	5.22	0.75	1.00	0.30	0.09
Vision	5.75	0.87	0.78	5.58	0.78	0.69	0.00	5.82	0.65	−1.02	0.31	−0.08
Sound	5.81	0.90	0.84	5.58	0.93	0.71	0.00	5.41	0.98	4.91	0.00	0.43
Smell	4.30	1.22	0.81	4.22	1.17	0.68	0.00	4.60	1.16	−3.01	0.00	−0.25
Taste	5.27	1.07	0.83	5.16	1.05	0.72	0.00	4.78	1.25	4.85	0.00	0.44
Touch	5.29	1.02	0.80	5.20	1.02	0.71	0.00	5.49	0.97	−2.31	0.02	−0.20
Body	5.31	0.98	0.79	5.10	1.00	0.67	0.00	5.23	0.96	1.02	0.31	0.09

Psi-Q, Plymouth Sensory Imagery Questionnaire; M, mean; SD, standard deviation.

TABLE 4 Correlations between subfactors of the Plymouth Sensory Imagery Questionnaire and correlations with other scales.

		M	SD	Psi-Q	Vision	Sound	Smell	Taste	Touch	Body
Psi-Q		5.29	0.80	–	**0.70**	**0.79**	**0.79**	**0.82**	**0.82**	**0.81**
	Vision	5.75	0.87		–	**0.63**	**0.41**	**0.45**	**0.44**	**0.51**
	Sound	5.81	0.90			–	**0.50**	**0.53**	**0.60**	**0.58**
	Smell	4.30	1.22				–	**0.63**	**0.56**	**0.54**
	Taste	5.27	1.07					–	**0.63**	**0.60**
	Touch	5.29	1.02						–	**0.62**
	Body	5.31	0.98							–
AIQ		3.40	0.70	**0.68**	**0.48**	**0.54**	**0.53**	**0.51**	**0.59**	**0.57**
VOIQ		3.10	0.80	**0.59**	**0.33**	**0.39**	**0.56**	**0.49**	**0.47**	**0.49**
Big Five										
	Extraversion	3.74	1.23	**0.27**	**0.21**	**0.23**	**0.22**	**0.21**	**0.23**	**0.18**
	Neuroticism	4.92	1.25	**−0.13**	−0.08	−0.09	**−0.14**	−0.06	**−0.13**	−0.09
	Openness	4.01	1.05	**0.19**	**0.16**	**0.13**	**0.15**	**0.12**	**0.17**	**0.16**
	Conscientiousness	4.25	1.09	0.10	0.05	0.08	**0.13**	**0.12**	0.07	0.02
	Agreeableness	4.26	1.04	**0.14**	**0.14**	**0.11**	**0.14**	0.08	**0.15**	0.05
FFMQ		3.15	0.38	**0.24**	**0.16**	**0.15**	**0.21**	**0.16**	**0.27**	**0.16**
	Observe	3.47	0.72	**0.32**	**0.21**	**0.24**	**0.28**	**0.18**	**0.30**	**0.28**
	Non-react	2.90	0.67	**0.13**	**0.12**	0.07	**0.11**	0.09	**0.17**	0.05
	Non-judge	2.99	0.65	−0.04	−0.03	−0.01	−0.06	−0.04	−0.02	−0.03
	Describe	2.88	0.73	**0.16**	0.09	0.08	**0.19**	0.09	**0.20**	0.08
	Actaware	3.49	0.72	0.06	0.04	0.00	0.05	0.10	0.07	0.03
SWLS		3.64	1.41	**0.24**	**0.24**	**0.24**	**0.17**	**0.16**	**0.17**	**0.18**
Age		40.72	10.73	**0.21**	**0.10**	**0.16**	**0.28**	**0.17**	**0.15**	**0.12**
Gender		1.60	0.49	**0.14**	**0.12**	**0.18**	0.06	0.08	**0.15**	0.10
Education		3.33	0.95	−0.01	−0.03	−0.05	0.01	0.03	0.03	−0.04

M, mean; SD, standard deviation; Psi-Q, Plymouth Sensory Imagery Questionnaire; AIQ, Auditory Imagery Questionnaire; VOIQ, Vividness of Olfactory Imagery Questionnaire; FFMQ, Five Facet Mindfulness Questionnaire; SWLS, Satisfaction With Life Scale. Bolded figures have a value of *p* of less than 0.05. Gender is numbered 1 for men and 2 for women.

4.2. Construct validity of the Japanese version of Psi-Q

To examine construct validity, we used scales that have already been shown to be related to existing imagery ability scales and the Psi-Q, as well as scales that are newly examined in the present study. First, the validity of the Japanese version of the Psi-Q was confirmed by its moderate to high positive correlations with olfactory imagery ability (Gilbert et al., 1998) and auditory imagery ability (Hishitani, 2009). In particular, the correlation between olfactory imagery ability, as measured by the VOIQ, and the smell subfactor, measured on the Psi-Q, was higher than the correlations for the other modalities,

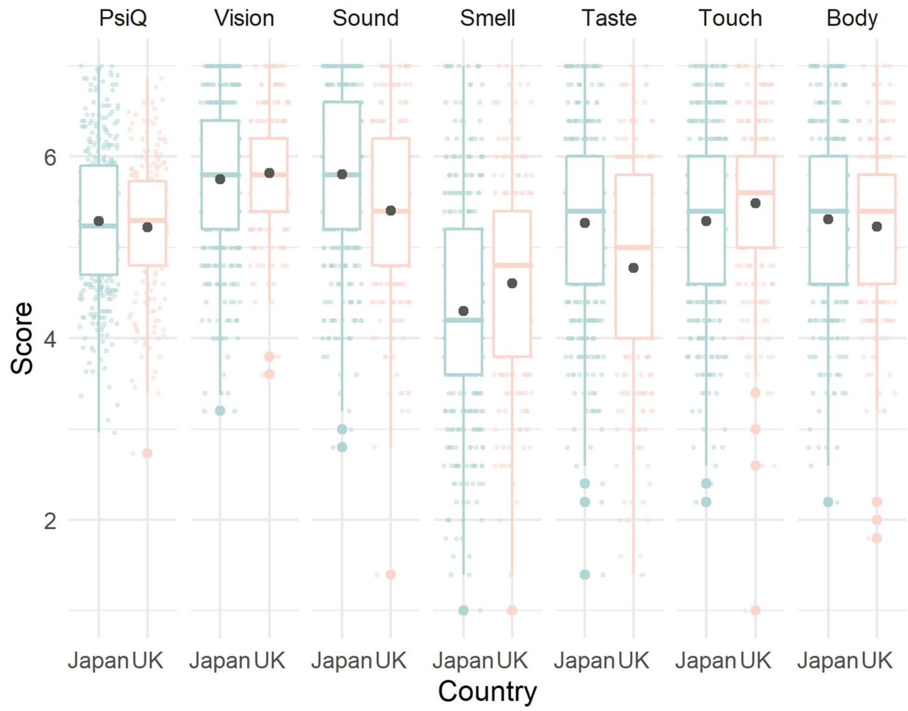

FIGURE 1
Box plot showing the Japanese-United Kingdom comparison of Plymouth Sensory Imagery Questionnaire.

suggesting that it is correctly measuring what it should measure. We also replicated previous research (Kharlas and Frewen, 2016) and found correlations with trait mindfulness, which has been shown to be related to multisensory imagery ability. Individual differences in mental imagery vividness are thought to be related to actual perceptual, emotional, and physical experiences (Cui et al., 2007; Kharlas and Frewen, 2016). Indeed, a positive correlation between participants' ages and Psi-Q total and subfactor scores was found in this study and suggests that an increase in the accumulation of real experience makes images more vivid until the middle age of life. In this way, it can be interpreted that mental imagery vividness is related to methods of self-emotional and physical regulation, such as mindfulness. We further found a relationship between multisensory imagery ability and the Big Five personality factors. Specifically, extraversion was associated with imagery vividness in all modalities. Again, this could be explained similarly to mindfulness and age, as individuals with high extraversion are likely to have more sense or opportunity to have a greater variety of perceptual and emotional experiences than those with low extraversion, which may be linked to mental imagery ability. Similarly, in this study, we found, for the first time, a positive correlation between multisensory imagery ability and life satisfaction. Individuals with vivid multisensory imagery may lead more fulfilling lives, both in their everyday and non-everyday life. For example, it is known that the higher their visual imagery ability, the more people appreciate the beauty of poetry (Hitsuwari and Nomura, 2021), which may potentially enhance their art experience. These relationships between trait mindfulness, the Big Five, life satisfaction, and multisensory imagery ability are correlational, and causal relationships are unresolved and should be explored in future research.

4.3. Cultural comparison of multisensory imagery ability

This study was the first to examine cultural differences in multisensory imagery ability. The results reveal that while there is no difference in general imagery ability between the Japanese and British samples, however, auditory and gustatory imagery is higher among the Japanese sample, and olfactory imagery is higher among the British sample. By measuring multisensory imagery simultaneously, we can show these differences by modality. First, regarding differences in sound imagery, there are cultural differences in the occurrence of the McGurk effect,[4] which is less pronounced in the Japanese sample (Sekiyama, 1994). Sekiyama (1994) has argued that Japanese people are less likely to engage in face-to-face communication and the simplicity of the Japanese phonological structure may allow language interaction based on auditory information alone, without reference to visual information. Second, regarding smell imagery, the influence of COVID-19 must be considered (c.f., Olofsson and Pierzchajlo, 2021). In 2022, the year in which data were collected from the Japanese sample, people were permanently wearing masks to prevent the spread of infection; however, this was not so in 2014 when the data were collected from the British sample. This continued wearing of masks (Ataka, 2022) may have caused reduced smell sensitivity and imagery for the Japanese sample. Third, regarding taste imagery, it is thought that there are cultural

4 A phenomenon in which a third phoneme is perceived when a video of one phonological speech utterance is viewed in combination with another (McGurk and MacDonald, 1976).

differences regarding sensitivity to taste. For example, umami is the fifth primary taste in European taste tests and tends to be defined as salty or sweet (Mueller et al., 2011), however, Japanese people are more familiar with umami foods, such as *dashi* (Satoh-Kuriwada et al., 2014). Nevertheless, these interpretations of cultural differences particularly in sound, smell, and taste are only tentative and must be verified in the future. Furthermore, the Psi-Q has already been translated into Spanish (Pérez-Fabello and Campos, 2020) and Dutch (Woelk et al., 2022), and comparative studies with more diverse cultures should be conducted.

4.4. Limitation and future direction

Although we were able to develop a Japanese version of the Psi-Q and validate its validity and reliability in this study, several limitations must be mentioned. First, as with the original and translated versions (Andrade et al., 2014; Pérez-Fabello and Campos, 2020; Woelk et al., 2022), the short version was created by reducing items from the long version, and some concerns have been noted with this method of creation (c.f., Aquino et al., 2018), as the same item may be regarded differently in the long and short versions. Second, although we made cultural comparisons, these are exploratory results, and based on the results obtained in this study, hypotheses need to be developed and factors that cause cultural differences in multisensory mental imagery abilities need to be further explored.

5. Conclusion

In this study, we developed a Japanese version of the Psi-Q, a scale measuring multisensory imagery ability with high validity and reliability and were able to produce a sufficient measure. Additionally, correlations with various scales, including the Big Five and life satisfaction scales, which had not been examined before, were clarified, and the adequate construct validity was demonstrated. In the Japanese-British comparison, cultural differences in multisensory imagery ability between the Japanese and British cultures could be noted for the first time. In the future, comparing the Psi-Q with other scales and using it in combination with behavioral experiments and functional brain imaging studies can increase our understanding of multisensory imagery ability. Future research should also conduct multicultural comparisons, not only between Japanese and British samples. The number of studies dealing with multiple modality senses simultaneously remains small, and it is expected that the Psi-Q will enable us to advance this field.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Kyoto University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JH: conceptualization, methodology, software, formal analysis, writing – original draft, and visualization. MN: conceptualization, methodology, writing – review and editing, and supervision. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by Grants-In-Aid for Scientific Research (JSPS KAKENHI; Grant Number: 20K20863 and 22H01103).

Acknowledgments

We are grateful to Jon May (University of Plymouth) for allowing us to translate the scale into Japanese and providing feedback. We also thank Editage for professional English proofreading.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1166543/full#supplementary-material

References

Andrade, J., May, J., Deeprose, C., Baugh, S. J., and Ganis, G. (2014). Assessing vividness of mental imagery: the Plymouth sensory imagery questionnaire. *Br. J. Psychol.* 105, 547–563. doi: 10.1111/bjop.12050

Aquino, A., Picconi, L., and Alparone, F. R. (2018). Validation of the Italian version of the need for cognition scale-short version. BPA-Appl. *Psychol. Bull.* 66, 18–29. doi: 10.1037/t74529-000

Ataka, K. (2022). Konokuni deha fact ya ronri yori kuuki nohouga omoi. [Between Neuroscience and Markenting]. Available at: https://kaz-ataka.hatenablog.com/entry/2022/10/30/150350 (Accessed February 8, 2023).

Betts, G. H. (1909). *The distribution and functions of mental imagery*. New York, NY: Columbia University.

Budnik-Przybylska, D., Syty, P., Kaźmierczak, M., Łabuda, M., Doliński, Ł., Kastrau, A., et al. (2023). Exploring the influence of personal factors on physiological responses to mental imagery in sport. *Sci. Rep.* 13:2628. doi: 10.1038/s41598-023-29811-6

Chiao, J. Y., and Blizinsky, K. D. (2010). Culture–gene coevolution of individualism–collectivism and the serotonin transporter gene. *Proc. Royal Soc. B-Biol. Sci.* 277, 529–537. doi: 10.1098/rspb.2009.1650

Clark, I. A., Monk, A. M., Hotchin, V., Pizzamiglio, G., Liefgreen, A., Callaghan, M. F., et al. (2020). Does hippocampal volume explain performance differences on hippocampal-dependant tasks? *NeuroImage* 221:117211. doi: 10.1016/j.neuroimage.2020.117211

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., and Eagleman, D. M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vis. Res.* 47, 474–478. doi: 10.1016/j.visres.2006.11.013

Del Re, A. C. (2013). compute.es: Compute Effect Sizes. R package version 0.2-2. Available at: http://cran.r-project.org/web/packages/compute.es (Accessed May 28, 2023)

Diener, E., Emmons, R. A., Larsen, R. J., and Griffin, S. (1985). The satisfaction with life scale. *J. Pers. Assess.* 49, 71–75. doi: 10.1207/s15327752jpa4901_13

Gilbert, A. N., Crouch, M., and Kemp, S. E. (1998). Olfactory and visual mental imagery. *J. Ment. Imag.* 22, 137–146.

Grebot, E. (2003). Validation with a French sample of the four scales of Switras's survey of mental imagery. *Percept. Mot. Skills* 97, 763–769. doi: 10.2466/pms.2003.97.3.763

Hatakeyama, T. (2019). Individual differences in imagery ability and cognition: a review of researches. *Jpn. J. Ment. Imag.* 16, 1–37. doi: 10.32145/jia.16.1_1

Hishitani, S. (2005). "Image nouryoku no sokutei" in *Kokoro Wo Hakaru*. eds. S. Hishitani and Y. Hakoda (Tokyo, Japan: Yachiyo Publishing), 140–141.

Hishitani, S. (2009). Auditory imagery questionnaire: its factorial structure, reliability, and validity. *J. Ment. Imag.* 33, 63–80.

Hishitani, S. (2011). "Image nouryoku no kojinsa" in *Gendai No Ninchishinrigaku 7*. ed. Y. Hakoda (Kyoto, Japan: Kitaohjishobou), 52–75.

Hitsuwari, J., and Nomura, M. (2021). How individual states and traits predict aesthetic appreciation of haiku poetry. *Empir. Stud. Arts* 40, 81–99. doi: 10.1177/0276237420986420

Holmes, E. A., Blackwell, S. E., Burnett Heyes, S., Renner, F., and Raes, F. (2016). Mental imagery in depression: phenomenology, potential mechanisms, and treatment implications. *Annu. Rev. Clin. Psychol.* 12, 249–280. doi: 10.1146/annurev-clinpsy-021815-092925

Kharlas, D. A., and Frewen, P. (2016). Trait mindfulness correlates with individual differences in multisensory imagery vividness. *Pers. Individ. Differ.* 93, 44–50. doi: 10.1016/j.paid.2015.09.027

Koivisto, M., and Grassini, S. (2022). Mental imagery of nature induces positive psychological effects. *Curr. Psychol.*, 1–16. doi: 10.1007/s12144-022-04088-6

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *Br. J. Psychol.* 64, 17–24. doi: 10.1111/j.2044-8295.1973.tb01322.x

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

McAvinue, L. P., and Robertson, I. H. (2007). Measuring visual imagery ability: a review. *Imagin. Cogn. Pers.* 26, 191–211. doi: 10.2190/3515-8169-24J8-7157

McDougall, S., and Pfeifer, G. (2012). Personality differences in mental imagery and the effects on verbal memory. *Br. J. Psychol.* 103, 556–573. doi: 10.1111/j.2044-8295.2011.02094.x

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0

Mueller, C. A., Pintscher, K., and Renner, B. (2011). Clinical test of gustatory function including umami taste. *Ann. Otol. Rhinol. Laryngol.* 120, 358–362. doi: 10.1177/000348941112000602

Namikawa, T., Tani, I., Wakita, T., Kumagai, R., Nakane, A., and Noguchi, H. (2012). Development of a short form of the Japanese big-five scale, and a test of its reliability and validity. *Jpn. J. Psychol.* 83, 91–99. doi: 10.4992/jjpsy.83.91

Olofsson, J. K., and Pierzchajlo, S. (2021). Olfactory language: context is everything. *Trends Cogn. Sci.* 25, 419–420. doi: 10.1016/j.tics.2021.02.004

Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nat. Rev. Neurosci.* 20, 624–634. doi: 10.1038/s41583-019-0202-9

Pearson, J., Naselaris, T., Holmes, E. A., and Kosslyn, S. M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends Cogn. Sci.* 19, 590–602. doi: 10.1016/j.tics.2015.08.003

Pérez-Fabello, M. J., and Campos, A. (2020). Spanish version of the Plymouth sensory imagery questionnaire. *Front. Psychol.* 11:916. doi: 10.3389/fpsyg.2020.00916

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychol. Bull.* 80, 1–24. doi: 10.1037/h0034650

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: https://www.R-project.org/ (Accessed May 28, 2023)

Reisberg, D., Pearson, D. G., and Kosslyn, S. M. (2003). Intuitions and introspections about imagery: the role of imagery experience in shaping an investigator's theoretical views. *Appl. Cognit. Psychol* 17, 147–160. doi: 10.1002/acp.858

Revelle, W. R. (2022). Psych: Procedures for personality and psychological research. Available at: https://CRAN.R-project.org/package=psych (Accessed May 28, 2023)

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Soft.* 48, 1–36. doi: 10.18637/jss.v048.i02

Satoh-Kuriwada, S., Kawai, M., Iikubo, M., Sekine-Hayakawa, Y., Shoji, N., Uneyama, H., et al. (2014). Development of an umami taste sensitivity test and its clinical use. *PLoS One* 9:e95177. doi: 10.1371/journal.pone.0095177

Sekiyama, K. (1994). Differences in auditory-visual speech perception between Japanese and Americans: McGurk effect as a function of incompatibility. *J. Acoust. Soc. Jp.* 15, 143–158. doi: 10.1250/ast.15.143

Shaw, G. (2008). The multisensory image and emotion in poetry. *Psychol. Aesthet. Creat. Arts* 2, 175–178. doi: 10.1037/1931-3896.2.3.175

Sheehan, P. W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *J. Clin. Psychol.* 23, 386–389. doi: 10.1002/1097-4679(196707)23:3<386::AID-JCLP2270230328>3.0.CO;2-S

Sumino, Z. (1994). An attempt to create a Japanese version of the satisfaction with life scale (SWLS). *Annu. Conf. Jpn. Soc. Psychol.* 58:192. doi: 10.20587/pamjaep.36.0_192

Switras, J. E. (1978). An alternate-form instrument to assess vividness and controllability of mental imagery in seven modalities. *Percept. Mot. Skills* 46, 379–384. doi: 10.2466/pms.1978.46.2.379

Takahashi, T., Saito, J., Fujino, M., Sato, M., and Kumano, H. (2022). The validity and reliability of the short form of the five facet mindfulness questionnaire in Japan. *Front. Psychol.* 13, 1–15. doi: 10.3389/fpsyg.2022.833381

Talamini, F., Vigl, J., Doerr, E., Grassi, M., and Carretti, B. (2022). Auditory and visual mental imagery in musicians and non-musicians. *Music. Sci.* 27, 428–441. doi: 10.1177/10298649211062724

Uno, K., and Yokosawa, K. (2022). Pitch-elevation and pitch-size cross-modal correspondences do not affect temporal ventriloquism. *Atten. Percept. Psychophys.* 84, 1052–1063. doi: 10.3758/s13414-022-02455-w

Wilson, A. C., Schwannauer, M., McLaughlin, A., Ashworth, F., and Chan, S. W. (2018). Vividness of positive mental imagery predicts positive emotional response to visually presented project soothe pictures. *Br. J. Psychol.* 109, 259–276. doi: 10.1111/bjop.12267

Woelk, M., Hagenaars, M. A., and Krans, J. (2022). Validation of the Dutch version of the Plymouth sensory imagery questionnaire. *Eur. J. Psychol. Assess. Adv*. doi: 10.1027/1015-5759/a000729

Yamamoto, K., Inomata, K., Susami, K., and Ayabe, S. (2018). Development of the Japanese version of the vividness of odor imagery questionnaire. *Jpn. J. Pers.* 27, 87–89. doi: 10.2132/personality.27.1.10

Yong, A. G., and Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutor. Quant. Methods Psychol.* 9, 79–94. doi: 10.20982/tqmp.09.2.p079

Zaleskiewicz, T., Bernady, A., and Traczyk, J. (2020). Entrepreneurial risk taking is related to mental imagery: a fresh look at the old issue of entrepreneurship and risk. *Appl. Psychol.* 69, 1438–1469. doi: 10.1111/apps.12226

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 31 July 2023
DOI 10.3389/fpsyg.2023.1190364

Check for updates

OPEN ACCESS

EDITED BY
Varun Dutt,
Indian Institute of Technology Mandi, India

REVIEWED BY
Carmen Adams,
University of Hasselt, Belgium
Bruno Mesz,
National University of Tres de Febrero, Argentina

*CORRESPONDENCE
Anders Crichton-Fock
✉ anders.crichton-fock@oru.se

RECEIVED 20 March 2023
ACCEPTED 06 July 2023
PUBLISHED 31 July 2023

CITATION
Crichton-Fock A, Spence C and Pettersson N (2023) Using crossmodal correspondences as a tool in wine communication.
Front. Psychol. 14:1190364.
doi: 10.3389/fpsyg.2023.1190364

Using crossmodal correspondences as a tool in wine communication

Anders Crichton-Fock[1*], Charles Spence[2] and Nicklas Pettersson[3]

[1]School of Hospitality, Culinary Arts & Meal Science, Örebro University, Örebro, Sweden, [2]Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, United Kingdom, [3]Örebro University School of Business, Örebro, Sweden

Introduction: This research investigates consumer acceptance of alternative methods for communicating information about wine, focusing on the alignment between sensory attributes and consumer expectations.

Methods: A survey was administered to wine enthusiasts to assess their attitudes toward crossmodal communication.

Results: The findings reveal significant associations between consumer behaviors and acceptance of alternative communication methods, highlighting the emerging field of crossmodal correspondences.

Discussion: These results suggest that leveraging crossmodal correspondences can enhance the match between a product's sensory qualities and consumer expectations, potentially reducing wine wastage resulting from unmet consumer preferences. These findings have implications for improving communication strategies in the wine industry and enhancing consumer experiences.

KEYWORDS

crossmodal correspondence, multisensory experience, crossmodal communication, product matching, wine communication

1. Introduction

Multisensory research is important if one is to understand how to optimize communication, since it involves much more than just spoken or written language. As humans, we use multiple senses such as sight, hearing, touch, taste, and smell, to gather information and communicate with others. Research on the crossmodal correspondences can provide an efficient tool for communication by leveraging the connections between the senses, not least when it comes to wine (see Spence, 2023). This is accomplished by involving a combination of different senses to increase the potential success of product communication, depending on the product, potential consumer, and purpose. For example, this could include different kinds of visual cues, such as shapes and colors (Spence, 2023), to support the communication of complex multisensory products, which on a single-sense basis can be hard to fully communicate/understand for the regular consumer. It could also be the use of mental pictures, metaphors and analogies relating to common recognizable human characteristics (Herdenstam et al., 2009, 2018, 2020). In this context it could be argued that consumers in general are well-experienced using crossmodal descriptions and sensory metaphors in their daily speech. For example: *warm welcoming* and *bubbly personality* (touch), *bright idea* and *glowing review* (vision), *end on sour note* and *such a sweet personality* (taste), *love stinks* and *sweet smell of success* (odor), *music to my ears* and *the world is listening* (sound).

In this context, a question arises as to whether this could be a tool not only to help understand the correspondence between different senses, but also as a tool to meet future challenges in regard to food communication and ultimately food waste (Uiterkamp and Vlek, 2007; Rosen, 2009; Adams, 2010; Hopton et al., 2010; Stock and Burton, 2011; De Soete, 2016; Rinaldi, 2017; Dondi et al., 2020; Martini et al., 2021). As mentioned already, there is a need to implement new strategies to decrease the climate footprint, both in the wine industry as well as beyond (Christ and Burritt, 2013; Galbreath et al., 2020). Strong communication and learning from the sensory perspective of crossmodal correspondence can potentially be used to enhance sensory training by creating more immersive and interactive learning experiences as a tool to help individuals better retain information about a product (Ahn, 2011; Ghosh et al., 2016), especially one with a complex flavor profile such as wine. This is a critical subject when it comes to the communication of olfactory experiences, both with regard to limitations in verbally and linguistically grasping the message as well as in terms of understanding sensory complex food products such as wine (Paradis and Eeg-Olofsson, 2013).

In the realm of sensory marketing, the utilization of multisensory stimulation has traditionally served as a valuable tool for understanding consumer responses to products in relation to the fulfillment of their expectations (Elder and Krishna, 2010; Varela and Ares, 2012; Spence et al., 2013; Krishna and Schwarz, 2014; Croijmans and Wang, 2021; Spence, 2022a). This line of research focuses, in part, on consumer acceptance while acknowledging the efficacy of imagery and text in marketing, as well as the interaction between visual and linguistic elements in such contexts (Bolognesi and Strik Lievers, 2018). Nonetheless, it is crucial to explore how various visual cues (e.g., Lick et al., 2017; Baptista et al., 2022; Nguyen and Durner, 2023; Spence, 2023) and haptic sensations (Piqueras-Fiszman and Spence, 2012; Wang and Spence, 2018) may influence consumer preferences and correlate with taste and texture. Within this context, research has indicated that imagery depicting wine plays a significant role in the recollection and communication of sensory experiences (Herdenstam et al., 2009, 2018; Croijmans et al., 2020) and that certain training and expertise can enhance imagery abilities among wine-interested consumers (Croijmans et al., 2020; Herdenstam et al., 2020). In a recent study, consumers with varying levels of self-reported imagery vividness were examined (Croijmans and Wang, 2021). The findings suggested that the vividness of mental imagery might be a crucial factor to consider, as multisensory wine descriptions can help stimulate purchase intentions among consumers with lower imagery vividness, particularly in terms of their desire to drink the wine (Croijmans and Wang, 2021). Conversely, consumers with higher vividness reported a greater desire to drink the wine even in the absence of multisensory descriptions in imagery (Wang et al., 2022).

Overall, the understanding of mental imagery vividness, is an important factor to consider when it comes to finding better tools for the communication of wine (Spence, 2017; Spence, 2022c). This approach might be applied for understanding how other complex food products communicate, as they contain layers of volatile odors, flavors, and tactile sensations, and this tool might be critical in understanding linguistic communication and different consumer group's attraction to a product (Paradis and Eeg-Olofsson, 2013). In this context, investigating the multisensory environment and its impact on consumer acceptance has also shown how important it is to consider the sound and hearing aspects when tasting wine (Spence and Wang, 2015a,b,c).

Recent research studies have shifted their focus towards comprehending the comprehensive impact of multisensory environments, surpassing earlier studies that examined sensory experiences on an individual basis. These studies aim to gain insights into consumer experiences (Maziriri et al., 2021; Spence and Van Doorn, 2022; Wörfel et al., 2022; Spence, 2022a). This emphasis is particularly significant in unraveling the reasons behind consumers' product choices (Spence, 2020, 2022a,b).

In the past, a large body of wine research has demonstrated a wide variety of influences affecting the consumer's esthetic and hedonic relation to sensory experience (Spence et al., 2014). One possible reason is that traditional and cultural aspects of the wine industry are reflected in idyllic images with beautiful landscapes, which has led to a perception that wine is considered as an environmentally friendly product (Ruggieri et al., 2009; Christ and Burritt, 2013). On the contrary, the United Nations has made it clear that there is an inevitable need to increase resources used in the food industry through environmental, sustainable, and cost-effective solutions (Roy et al., 2009; Ghosh et al., 2016; Merli et al., 2018). According to the above-mentioned studies, it can be argued that finding alternative ways to communicate about wine and other complex food products, novel or otherwise, could be helpful to both the consumer experience and the environment (Ahn, 2011; Ghosh et al., 2016). This is partially due to the demonstrated limits of language in describing sensory experiences that derive from the olfactory domain when experiencing wine (Paradis and Eeg-Olofsson, 2013).

This questionnaire-based study aimed to explore the attitudes of wine consumers towards crossmodal communication and assess the effectiveness of crossmodal correspondence as a potential communication tool for wine and other complex food products. The primary motivation behind this study was to explore innovative approaches to enhance the alignment between consumer expectations and sensory experiences, particularly by identifying novel and effective tools to address challenges associated with food production and waste.

2. Materials and methods

2.1. Ethics statement

All of the participants were over 20 years of age, and informed consent was obtained from all of those taking part in the study. All of the data and analysis files were kept in accordance with legislated and regulated data handling practices.

2.2. Participants

The participants consisted of 329 students from different sections of the 7.5-credit, 15-week distance course 'Beverage knowledge' offered through Örebro University (Sweden). The students received training in which they learned a common methodology to analyze wine. One important factor in the selection of these group of participants was their common experience of partly being exposed of effects of crossmodal correspondence during their training sessions, learning the basics of wine tasting. Both theoretically, by taking part of research during the course, and also practically, through the tasting exercises, experiencing crossmodal correspondences affects when moving from one sense modality to another. They were instructed first to look at a wine's appearance qualities in order to make visual judgments concerning its color, intensity, maturity, age, freshness, acidity, and concentration. They were then encouraged to validate these first impressions on the nose, and subsequently on the palate, observing the effect of crossmodal correspondence and their notable impact during the professional wine tasting procedures, assessing first visual impressions, then the bouquet with layers of aromas, and lastly taste and oral-somatosensation while during this multisensory experience also being intervened by the retronasal effect of the odors in the palate. The majority of the participants were female (59%) and lived in the city (76%). Almost all had previously studied at university (94%), and most of them had received a bachelor's degree, or higher (74%). As a group, they considered that they had better than average wine knowledge (76%). Most of the participants consumed wine on a weekly basis (87%), which they typically purchased at Systembolaget (Sweden's nationally regulated liquor monopoly; 81%) and then consumed at home (80%; see Appendix A).

The participants shared the following traits:

i. They had all tasted the same wines and other beverages and had therefore shared a variety of sensory experiences within their training program.
ii. They had all learned a common approach and methodology when it came to the sensory analysis of wine. It can therefore be presumed that, as a group, they had an awareness of the importance of (all) the senses in the analysis process, including vision, olfaction, taste, touch, and sound.
iii. They had all experienced crossmodal correspondence during the tasting methodology practice. In other words, they were aware that the senses could not be totally isolated from one other during the process of analyzing the wine and, subsequently, communicating about the multisensory experience.

2.3. The questionnaire

The questionnaire used in this study comprised four sections. This section focuses on the analysis of the first section, which included inquiries about demographics, as well as single-choice and multiple-choice (check-all-that-apply; CATA) questions pertaining to purchase and consumption behaviors, communication practices, and sensory experiences related to wine. The remaining sections, which investigated visual aspects, the use of specific symbols, and alternative communication methods, were analyzed and presented separately based on the results of the statistical analysis. The questionnaire was distributed to participants upon completion of the course. In addition to collecting demographic information, it encompassed single-choice and multi-choice (CATA) questions concerning purchase behaviors, consumption habits, communication preferences, and sensory experiences (see Appendix B).

2.3.1. Examples of single-choice questions

1) Where do you primarily purchase wine? 2) Where do you primarily consume wine? 3) To what extent would you consider buying a wine based only on sensory information – i.e., if no other conventional information were available?

2.3.2. Examples multi-choice questions (CATA)

1) When purchasing wine from your selected choice in the question above, what factors influence your choice? 2) When consuming your wine at your selected choice in the question above, which factors primarily influence "your experience"? 3) If you could freely choose an alternative to regular communication on a wine label, what would you prefer?

2.4. Data analysis

EyeQuestion version 5 (Logic 8, Elst, Netherlands), a software program for sensory and consumer testing, was used to collect the participants' responses. The software package R (R Core Team, 2021) was used to analyze the data.

3. Results

We analyze the influential factors when purchasing and consuming wine (in section 3.1) and reported preferences towards alternative wine communication (in section 3.2) including their associations to demographics and influential factors (in section 3.2.1–7). The relations between preferences are then analyzed (in section 3.3) and followed up by investigating the most frequent preference combinations and their associations to demographics and influential factors (in section 3.4).

3.1. Reported factors influencing purchasing and consumption of wine

The most influential factors reported by the participants when purchasing wine were (in descending order of importance): type of grape (77%), prior experience of the wine (72%), country of origin (72%), external recommendations (68%), style (65%), and price (62%). At the same time, the sensory descriptors communicated on the wine label or on the shelf were also reported to be an important factor for just under half of the participants.

The dominant factors influencing the choices reported by this group of engaged wine consumers would appear to be of a more general character. On the other hand, more specific aspects such as: sensory indicators of the experience, name of the producer, alcohol or vintage information, overall product communication in the store, shelf information and producer information (on the bottle design) were considered to be less important (see Table 1).

The most influential sensory factors reported by the group of wine enthusiasts when consuming wine were as follows: taste (94%), tasting/dinner setting (66%), and smell of the wine (65%). Far fewer of the participants reported being influenced by the other senses during their consumption: vision (26%), touch (12%), sound (8%). Here, a discrepancy was noted, with the reported factors influencing purchasing

TABLE 1 Reported factors influencing wine purchasing behavior (CATA).

Influence factor	*n*	Frequency % (*n* = 329)
The grape	253	76.90%
Previous experience (having tried the wine before)	237	72.04%
Country of origin	237	72.04%
External recommendations (professional, friends, others)	224	68.09%
Style of wine	215	65.35%
Price	204	62.01%
Sensory indicators (concerning the taste, aroma profile, and/or mouthfeel)	163	49.54%
The wine producer	112	34.04%
Climate impact (How eco-friendly is the wine)	98	29.79%
Vintage	65	19.76%
Illustrations on the label	61	18.54%
The front label	61	18.54%
Bottle design (bag in box/bottle)	55	16.72%
The back label	12	3.65%
Other	9	2.74%

TABLE 2 Reported factors influencing experience of consuming wine (CATA).

Influence factor	*n*	Frequency % (*n* = 329)
Taste	309	93.92%
Dinner/tasting setting	216	65.65%
Smell/odor sensations of the wine	215	65.35%
"Meal" companions/other guests	164	49.85%
The visual impression (vision)	84	25.53%
Overall room environment	75	22.80%
Tactile (touch) sensations	40	12.16%
Hosts/Professionals/Staff	31	9.42%
Sound environment (sounds)	25	7.60%
Others	3	0.91%

being dominated by more general information rather than sensory indicators, and the sensory indicator of taste was reported as the single most important factor when consuming the wine (see Table 2).

3.2. Reported attitudes towards alternative wine communication

When the participants were asked whether they might consider buying wine based only on sensory information and no conventional label information, most of them had positive reactions to buying a bottle just so long as the sensory descriptors matched their own personal preferences (91%). Another question asked if they would also consider buying a wine with no specific origin, for example, a blend of different wines from different origins. Here, the participants reacted positively to the thought of buying a bottle if the descriptors happened to be equivalent to their sensory profile (95%). In terms of sustainability, 76% of them answered that they, to varying extents, would consider climate change into when making a decision about what wine to buy (see Appendix C).

TABLE 3 Reported alternative communication to regular wine label (CATA).

[Article section] Alternative communication	*n*	Frequency % (*n* = 329)
[3.2.1] Shapes/Colors (other visual symbols)	178	54.10%
[3.2.2] Flavors (tastes)	148	44.98%
[3.2.3] Odors/aromas (smells)	125	37.99%
[3.2.4] Touch (tactile)	49	14.89%
[3.2.5] Speech (hearing)	30	9.12%
[3.2.6] Music (hearing)	26	7.90%
[3.2.7] Sounds (hearing)	20	6.08%

In the questionnaire, the participants were introduced to alternative ways of communicating about the sensory experiences of wine and were asked to freely choose from examples of different modalities (see Appendix B). Visual communication using shapes and colors were the most frequently requested (54%), while flavor (45%) and odor/aroma (38%) were also highly ranked alternative means of communication. The least popular alternatives were those communicated auditorily, such as speech (9%), music (8%), and sounds (6%), see Table 3.

In sections 3.2.1–3.2.7 we explore each of the alternative ways of communicating in relation to reported demographics (see also Appendix D) and purchase choice (Appendix E) and consumption experience (Appendix F), using two-sided hypothesis tests of, association (Kendall's Tau rank correlation coefficient, Fisher's exact test and Chi-square test), and of equal location among groups (Mann–Whitney U test and for significance to categories Kruskall-Wallis test, Wilcoxon rank sum test and t-test), see Appendix D. When studying the relation to purchase choice and consumption experience, as to rule out the influence of the reported locale, we also use binary logistic regression to regress the preferred alternative way of communication on the locale and the influential factor under study. Logged odds ratios (logOR) are used to represent 2*2 categorical associations, as exemplified at the start of section 3.2.1.

3.2.1. Shapes/colors (other visual symbols)

There were 62 males and 115 females who preferred shapes/colors as a means of communication, while 115 males and 80 females did not. A positive (negative) logged odds ratio indicates larger (smaller) odds amongst males who preferred shapes/colors as communication, relative to the odds among females, while a value of zero indicates no difference in odds. Since the odds amongst females (115/80 = 1.44) is higher than amongst males (62/70 = 0.89), the positive logged odds ratio, logOR = log(1.44/0.89) = 0.48, indicates that females were more likely than males to prefer shapes/colors as a means of communicating about wine. Using Fisher's exact test, under the null hypothesis that gender and preference for shapes/colors are independent, the probability of an outcome at least as extreme as the observed (i.e., the *p*-value) is 0.042, which is less than 0.05. Thus, females (59%) are significantly more open to communications involving vision (using shapes/colors) to describe the multisensory experience of wine, compared to males (41%), see Tables AD.1, AD.2.

Regarding reported influencing purchasing factors (Table AE.1), the odds of preferring visual communication were significantly smaller for respondents influenced (relative to those not influenced) by vintage (logOR = −0.63), and significantly larger for respondents influenced by the climate impact (0.66), the bottle design (0.76), the illustrations on the label (1.15) or the front label (0.67). The sensory indicators (0.48) and the price (0.50) were also significantly larger, but not after controlling for the locale of purchase (Table AE.2).

Among the factors influencing the experience (Table AF.1), the odds of preferring visual communication were significantly larger for respondents influenced by taste sensations (logOR = 1.34) or hosts/professionals/staff (0.55), also after controlling for the locale of consumption (Table AF.2).

3.2.2. Flavors (tastes)

Neither any demographical factors (Appendix D) nor any factors influencing wine purchasing (Appendix E) were significantly related to preferring flavors (tastes) as means of communication. The only association found to be statistically significant (Table AF.1), which also held after controlling for the locale of consumption (Table AF.2), was the smaller odds of preferring flavors (logOR = −0.55) for respondents whose experience was affected by the overall room environment.

3.2.3. Odors/aromas (smells)

The median income was significantly lower amongst those respondents who were open to (35,000 SEK) vs. those who were not open to (40,000 SEK) communication through odor/aromas (Table AD.4 and Figure AD.2).

Among the factors influencing the purchase (Table AE.1), the odds of preferring communication via odors/aromas were significantly larger for respondents influenced by the country of origin (logOR = 0.74), the type of grape (0.70), the style (0.54), the sensory indicators (0.63) or the bottle design (0.72) respectively, also after controlling for the locale of purchase (Table AE.2).

Regarding influence on consumption experience (Table AF.1), only one association was found to be statistically significant, where the odds of preferring odor/aromas was (intuitively) larger for those whose experiencing was influenced by smell/odor (logOR = 0.86), also after controlling for the locale of consumption (Table AF.2).

3.2.4. Touch (tactile)

Consumers open for communication involving touch were on average significantly younger (35.2 years) than those who were not open (42.4 years) to this form of communication (Table AD.3 and Figure AD.1). Their median income was also significantly lower (34.3' SEK) compared to those who were not open to communication through touch (39' SEK), see Table AD.4 and Figure AD.2. The respondents open to vs. not open to communication through touch differed significantly in their distribution of wine consumption, and those more open to this modality were less likely to consume wine frequently (33% more than once weekly) compared to respondents that were not open to this modality (55% more than once weekly). Also, they were significantly more likely to primarily consume wine in restaurants (12%) and not with friends or family/in a wine tasting group (8%) compared to others (4 and 15% respectively), see Table AD.5.

Regarding statistically significant associations between reported factors influencing purchase of wine (Table AE.1) and the preference for tactile alternative communication the odds was (intuitively) larger for respondents who preferred communication via touch for the illustrations on the label (logOR = 1.61) but as well for the front label (1.16), and the bottle design (0.84), also after controlling for the locale of purchase (Table AE.2).

The only association found to be statistically significant was the (intuitively) larger logged odds (logOR = 0.91) of preferring touch for consumers whose experience were influenced by tactile sensations (Table AF.1), which also held after controlling for the locale of consumption (Table AF.2).

3.2.5. Speech (hearing)

Neither any demographical factors (Appendix D) nor any factors influencing the consumption (Appendix F) were significantly related to preferring speech (hearing) as a means of communication. The only association found to be statistically significant (Table AE.1) was the larger logged odds of preferring speech as alternative communication for respondents whose wine purchasing were influenced by the country of origin (logOR = 1.34), also after controlling for the locale of purchase (Table AE.2).

3.2.6. Music (hearing)

Respondents open to alternative communication via music were on average significantly younger (36.4 years) compared to others (41.8), and lived in more densely populated areas (96% vs. 74% in capital or city), see Table AD.3 and Figure AD.1.

For influence on purchase and consumption of wine, only a single factor each was significantly associated to preferring music as alternative communication, also after controlling for the locale. Regarding purchase, the odds of preferring music were higher for the bottle design (logOR = 1.27), see Appendix E, and for consumption the odds of preferring music were higher for the sound environment (logOR = 1.75), see Appendix F.

3.2.7. Sounds (hearing)

Respondents open to alternative communication via sounds were on average significantly younger (34.8 years) compared to others (41.7), and lived in more densely populated areas (95% vs. 74% in capital or city), see Table AD.3 and Figure AD.1.

None of the factors influencing consumption experience (Appendix F) were significantly associated with the preference for sounds as alternative communication. Regarding influence on purchase of wine a single factor was statistically significant (Table AE.1) with a larger odds of preferring sounds in relation to sensory indicators (logOR = 1.18), also after controlling for the locale of consumption (Table AE.2).

3.3. Associations between reported alternatives to regular wine communication

We study the pairwise associations between the seven preferences for alternative wine communication using logOR and correlations, and then look for higher dimensional structures. While most consumers reported only one alternative (45.9%), very few reported

TABLE 4 Total number of reported alternatives (Shapes/Colors; Sounds; Music; Speech; Touch; Odors/aromas; Flavors) to regular wine communication among respondents.

Number of reported alternatives	0	1	2	3	4	5	6	7	Total
Number of respondents	20	151	90	45	13	6	0	4	329
Percentage of total respondents (%)	6.1	45.9	27.4	13.7	4.0	1.8	0	1.2	100

TABLE 5 Pairwise logged odds ratios (upper right) and correlations (lower left) of reported alternatives to regular wine communication.

	Shapes/colors	Sounds	Music	Speech	Touch	Odors/aromas	Flavors
Shapes/colors		0.26	−0.35	−0.78*	1.11***	−0.66***	−0.90***
Sounds	0.03		2.89***	1.33**	1.47***	0.74	0.64
Music	−0.05	0.40***		1.49***	0.83	0.70	0.39
Speech	−0.11*	0.14*	0.18***		1.02**	0.69	−0.22
Touch	0.18**	0.18**	0.10	0.13*		1.02***	0.09
Odors/aromas	−0.16**	0.09	0.10	0.10	0.18***		1.32***
Flavors	−0.22***	0.08	0.05	−0.03	0.02	0.31***	

Log odds ratio with Fisher's exact test (Haldane-Anscombe), and Pearson correlation coefficient with t-test. Two-sided p-value: 0.05 > * > 0.01 > ** > 0.001 > ***.

no alternative (6.1%), so almost half (49.0%) checked at least two of these alternatives (see Table 4).

When pairs of reported alternatives were considered, about half of them had a significant relationship (see Table 5). Since all significant logORs/correlations were positive, except for shapes/colors, where only touch was positive and the others significantly negative, the occurrence of a reported alternative generally had an increased probability of also having other alternatives reported. Particularly, the logOR/correlation between sounds and music were large (2.89/0.40), although both alternatives were quite infrequent in total (6.1 and 7.0%). The logOR/correlation was also fairly high between Odors/aromas and Flavor (1.32/0.31). The other logORs involving Speech, Music, Sounds and Touch (except for Music and Touch) were about as high (1.02–1.49) although the correlations were somewhat lower (0.10–0.18). For preferred alternatives Shape/colors and Flavors, the logOR was not as high (0.90) but the correlation was somewhat higher (0.22).

Among the respondents there were 51 different combinations of preferences for alternative wine communication, including the combination with no preferences at all. After applying logistic principal component analysis for dimensionality reduction on the preferences, we kept the first three extracted principal components (PCs) since they accounted for as much as 79% of the total deviance among the preferences (44, 19 and 16% respectively), involved all preferences with (very) high loadings and resembled the correlational structure well, see Table 6. The first PC had high loadings (absolute value larger than 0.3) of Sounds, Music, Speech and Touch, the second had a very high loading (absolute value larger than 0.6) of Shapes/colors but also high of Odors/aromas and Flavors, while the third PC had a very high loading of Shape/colors, but also high of Odors/aromas and Touch.

After plotting the respondents scores on the three extracted PCs, see Figure 1, we identified three groups (plus two deviant single combination groups) of observations which (by construction) differed significantly in terms of their preferences of alternative wine communication, see Table 6.

The number of preferences also differed significantly between the groups, the bottom group (in red, named G1) with fewest (0–3), the middle group (green, G2) with slightly more (1–4) and the top group (blue, G3) with most preferences (2–5), see Figure 1. and Table 7. The left single combination group (black, G4) had 80 respondents preferring only Shapes/colors, while the utmost right (black, G5) had four respondents preferring all the seven ways of alternative wine communication.

Relatively many in group G1 preferred hearing and tactile, but fewer in G2 and G3. While none in group G1 preferred smells and taste, about half in G2 and everyone in G3 did. The groups also differed significantly in terms of age (G5 youngest, G3 and G4 oldest), see Table 7, and to the degree which they selected wine with regard to sustainability (Table AG.1).

In studying the relation to purchase choice and consumption experience we use multinomial regression and regress the group (but to enable reliable estimation we exclude group G5) on the locale (as to rule it its effect) and the influential factor under study, see Table AG.2. In addition to previously found significant factors (see sections 3.2.1–7) influencing purchase (country of origin, grape, illustrations on the label, vintage), the wine producer was also found to be significant. Regarding consumption experience (Table AG.3) the situation was similar (with smell/odor and tactile sensations, overall room environment and hosts/staffs from before) adding visual impression.

We summarize the description of the groups in Table 8.

3.4. Most frequent combinations of reported alternatives to regular communication of wine

Given the frequencies in Table 3 it is as expected that the most frequent combinations (63.2%) of reported alternatives (see Table 9) only involved the univariately most frequent alternatives: shapes/colors (54.1%), flavors (38.0%) and odors/aromas (45.0%). Almost half of those who reported any three alternatives (13.7%) specifically reported these three (6.4%). However, no one reported the specific

TABLE 6 Loading matrix from logistic principal component analysis on preferences for alternative wine communication.

Principal component (% explained deviance)	Shapes/colors	Sounds	Music	Speech	Touch	Odors/ aromas	Flavors
First PC (44%)	0.090	**−0.496**	**−0.488**	**−0.468**	**−0.464**	−0.242	−0.119
Second PC (19%)	**−0.462**	−0.092	−0.034	−0.096	−0.272	**0.449**	***0.702***
Third PC (16%)	***0.655***	−0.106	−0.180	−0.265	**0.373**	**0.537**	0.173

Loadings with absolute value larger than 0.3 (0.6) are in bold (bold italic).

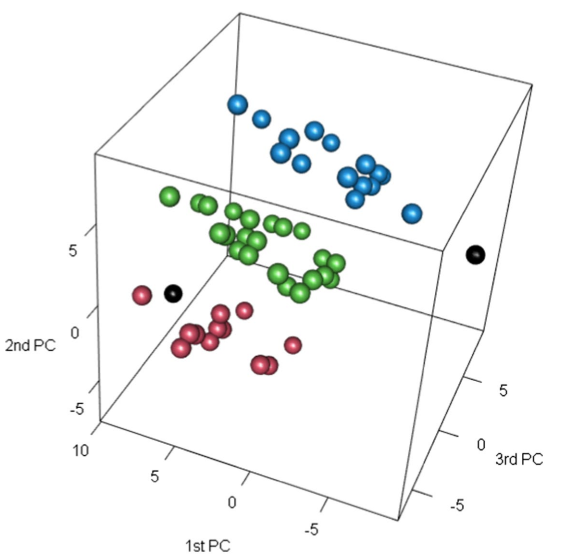

FIGURE 1
Respondents scores (with groupings colored and single group combinations in black) on three extracted PCs (from preferences of alternative wine communication).

combination of shapes/colors and odors/aromas. The most common combination was the group G4 (see section 3.3) with 80 respondents only preferring shapes/colors as alternative communication.

It is hard to detect significant differences between the combinations since they are relatively small (except for G4). Still, the age of respondents was significantly higher among the combination odors/ aromas and flavors (49.7 years), relative to those only reporting the alternative odors/aromas (37.9 years), but not among those only preferring flavors (45.4 years), see Figure AH.1. A few influential factors on the choice of purchase (illustrations on the label, sensory indicators), see Table AH.1, and on the consumption experience (taste sensations), see Table AH.2, differed significantly between the combinations. Due to scarcity of observations, we did not control for the locale.

4. Discussion

The findings of this study revealed that taste (94%), tasting/dinner setting (66%), and the smell of the wine (65%) were the most influential self-reported sensory factors during wine consumption. In contrast, a smaller percentage of participants reported being influenced by other senses, such as vision (26%), touch (12%), and sound (8%). These results suggest that there is relatively low awareness of crossmodal correspondence and the effects of the multisensory environment among this group of engaged wine consumers. This highlights the potential for more serious attempts to implement the findings of this research in the food industry, considering the valuable role that alternative multisensory communication tools have played in sensory marketing and increasing sales (Elder and Krishna, 2010; Varela and Ares, 2012; Spence et al., 2013; Krishna and Schwarz, 2014; Croijmans and Wang, 2021; Spence, 2022a).

Despite the extensive research demonstrating the influence of senses like vision and touch on crossmodal correspondence (Lick et al., 2017; Baptista et al., 2022; Nguyen and Durner, 2023; Spence, 2023) and the role of haptic sensations (Piqueras-Fiszman and Spence, 2012; Wang and Spence, 2018), the participants in this study reported these factors as being influential for only a few of them. However, when asked if they would consider purchasing wine based solely on sensory information, the majority of participants reacted positively, particularly if the sensory descriptors aligned with their personal preferences (91%). Similarly, when asked about purchasing a wine with no specific origin but blended from different wines, participants responded positively (95%) if the descriptors matched their sensory profile.

Furthermore, when it came to the primary choice of alternative communication for wine, visual cues were rated the highest (54%). This supports the idea that this group of participants is open to buying wine based solely on sensory information, without considering origin or blending. By analyzing the respondents' scores on the three extracted principal components, three groups (along with two deviant single combination groups) were identified. These groups exhibited significant differences in their preferences for alternative wine communication. This result suggests the potential for a strategic communication approach that employs different forms of communication tools to target different consumer groups through alternative communication methods. It also highlights the importance of utilizing such tools to support specific consumer groups in need of special assistance (Spence, 2022c).

As mentioned earlier, one of the motivations behind this research is to explore how novel sustainable products that are resource-efficient and have a reduced carbon footprint can be effectively communicated and validated for consumer acceptance (Lévy et al., 2006; Peschel et al., 2019; Herdenstam et al., 2022). This research aims to enhance researchers' understanding of how consumers perceive a product in relation to meeting their expectations, thereby contributing to the communication and acceptance of such products.

5. Conclusion

These results provide insights into the factors that influence wine purchasing and consumption, as well as the preferences and attitudes towards alternative wine communication. They highlight the importance of sensory information, particularly taste, and the potential for using visual and other alternative means to communicate

10.3389/fpsyg.2023.1190364

TABLE 7 General characteristics and preferences for wine communication by group (G1-G5).

General characteristics					Preference for altern. wine communication (%)						
Group	Nr of respondents***	Average nr of alternatives for communication**	Average age (years)*	Median income (1,000 SEK)	Shapes/ colors (%)***	Sounds (%)***	Music (%)***	Speech (%)***	Touch (%)***	Odors/ aromas (%)***	Flavors (%)***
G1	57	1.04	37.6	37.4	33	9	14	21	26	0	0
G2	111	1.72	40.5	37.0	37	6	7	7	14	40	60
G3	77	2.83	43.5	38.0	44	5	8	8	18	100	100
G4	80	1	43.5	41.0	100	0	0	0	0	0	0
G5	4	7	33.0	24.5	100	100	100	100	100	100	100

Test of equal size: Nr of respondents (Chi-square). Test of association: Preference for alternative (Chi-square). Test of equal location: Age (Anova F-test); Nr of alternatives and Income (Kruskal-Wallis). p-value: 0.05 > * > 0.01 > ** > 0.001 > ***.

TABLE 8 Salient features among groups compared to the least salient (reference) group.

Group (size in %)	Alt. wine communication	Demographics and other characteristics	Factors influencing purchase choice	Factors influencing consumption experience
G1 (17%)	Hearing, tactile	Younger, consume in bars/restaurants, willing to try blended wine	Country of origin, illustrations on the label, wine producer	Tactile sensations
G2 (34%)		Consume with friends/family/wine tasting group, regard sustainability	Grape	Overall room environment, hosts/professionals/ staff
G3 (23%)	Smell, taste	Older, consume at home, consider buying with only sensory information	Country of origin, wine producer, vintage	Visual impression, smell/odor sensation
G4 (24%)	Visual	Older, consume with friends/family/wine tasting group, regard sustainability		Hosts/professionals/staff

TABLE 9 Observed combinations of the most frequently reported alternatives to regular wine communication (shapes/colors, flavors and odors/aromas).

Shapes/ Colors (other visual symbols)	Flavors (tastes)	Odors/ aromas (smells)	Sound; Music; Speech; Touch	*n*	%
YES	NO	NO	NO	80	24.3
NO	YES	NO	NO	37	11.2
NO	YES	YES	NO	35	10.6
YES	YES	YES	NO	21	6.4
YES	YES	NO	NO	18	5.5
NO	NO	YES	NO	17	5.2
YES	NO	YES	NO	0	0
Sum of observations among the combinations in the table				208	63.2
All observations				329	100

the sensory experience of wine. Understanding these preferences and associations can assist in developing effective strategies for wine marketing and communication, addressing better resource use. Overall, these findings suggest that wine consumers consider various factors when purchasing and consuming wine, including sensory indicators, personal preferences, and influential factors. There is openness to alternative means of wine communication, particularly visual communication using shapes and colors. The associations between preferences and demographics/influential factors highlight the individual differences in wine communication preferences. It's important to note that these conclusions are based on the information provided in the given sections. Further analysis and research may be necessary to validate and expand upon these findings.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent was taken from the participants.

Author contributions

AC-F contributed to conception and design of the study and wrote the first draft of the manuscript. CS wrote sections of the manuscript and approved the submitted version. AC-F and CS contributed to manuscript revision and read. NP performed the statistical analysis and contributed to manuscript revision. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to especially acknowledge Markus Ekman, School of Hospitality, Culinary Arts & Meal Science at Örebro, for collecting data. This study is the second pilot within the multidisciplinary research project Rewine The World, which has a mission to rescue wine and other complex food products.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1190364/full#supplementary-material

References

Adams, C. A. (2010). Sustainability research in need of a multi-disciplinary approach and a practice and policy focus? *Sustain. Account. Manag. Policy J.* 1. doi: 10.1108/sampj.2010.46801aaa.001

Ahn, Y. H. (2011). Building for the future: United Nations approach for their buildings. *J. Green Build.* 6, 1–6. doi: 10.3992/jgb.6.4.1

Baptista, I., Spence, C., Shimizu, R., Ferreira, E., and Behrens, J. (2022). Color is to flavor as shape is to texture: a choice-based conjoint study of visual cues on chocolate packaging. *J. Sens. Stud.* 38:e12793. doi: 10.1111/joss.12793

Bolognesi, M., and Strik Lievers, F. (2018). How language and image construct synaesthetic metaphors in print advertising. *Vis. Commun.* 19, 431–457. doi: 10.1177/1470357218782001

Christ, K. L., and Burritt, R. L. (2013). Critical environmental concerns in wine production: an integrative review. *J. Clean. Prod.* 53, 232–242. doi: 10.1016/j.jclepro.2013.04.007

Croijmans, I., Hendrickx, I., Lefever, E., Majid, A., and Van Den Bosch, A. (2020). Uncovering the language of wine experts. *Nat. Lang. Eng.* 26, 511–530. doi: 10.1017/S1351324919000500

Croijmans, I., Speed, L. J., Arshamian, A., and Majid, A. (2020). Expertise shapes multimodal imagery for wine. *Cogn. Sci.* 44:e12842. doi: 10.1111/cogs.12842

Croijmans, I., and Wang, Q. J. (2021). Do you want a description with that wine? The role of wine mental imagery in consumer's desire to drink using the revised vividness of wine imagery questionnaire (VWIQ-II). *J. Sens. Stud.* 37:e12712. doi: 10.1111/joss.12712

De Soete, W. (2016). Towards a multidisciplinary approach on creating value: sustainability through the supply chain and ERP systems. *Systems* 4:16. doi: 10.3390/systems4010016

Dondi, L., Ripamonti, F., Ugolini, M., and Varvaro, S. (2020). A multidisciplinary approach as an assumption for design sustainability in developing countries. IOP conference series. *Mater. Sci. Eng.* 960:032102. doi: 10.1088/1757-899X/960/3/032102

Elder, R. S., and Krishna, A. (2010). The effects of advertising copy on sensory thoughts and perceived taste. *J. Consum. Res.* 36, 748–756. doi: 10.1086/605327

Galbreath, J., Tisch, D., Quaddus, M., and Rabbanee, F. (2020). The impact of climate change on firm adaptation: the case of the wine industry. *Int. J. Wine Bus. Res.* 32, 373–389. doi: 10.1108/IJWBR-07-2019-0045

Ghosh, P. R., Fawcett, D., Sharma, S. B., and Poinern, G. E. J. (2016). Progress towards sustainable utilisation and management of food wastes in the global economy. *Int. J. Food Sci.* 2016:3563478. doi: 10.1155/2016/3563478

Herdenstam, A., Hammarén, M., Ahlström, R., and Wiktorsson, P.-A. (2009). The Professional language of wine: Perception, training and dialogue. *Journal of Wine Research* 20, 53–84. doi: 10.1080/09571260902978543

Herdenstam, A. P. F., Kurtser, P., Swahn, J., and Arunachalam, A. (2022). Nature versus machine: a pilot study using a semi-trained culinary panel to perform sensory evaluation of robot-cultivated basil affected by mechanically induced stress. *Int. J. Gastron. Food Sci.* 29:100578. doi: 10.1016/j.ijgfs.2022.100578

Herdenstam, A., Nilsen, A. N., and Öström, Å. (2020). Breaking the silence: a pilot study investigating communication skills of sommeliers and chefs after analogical training. *Int. J. Gastron. Food Sci.* 20:100210. doi: 10.1016/j.ijgfs.2020.100210

Herdenstam, A., Nilsen, A., Öström, Å., and Harrington, R. J. (2018). Sommelier training – dialogue seminars and repertory grid method in combination as a pedagogical tool. *Int. J. Gastron. Food Sci.* 13, 78–89. doi: 10.1016/j.ijgfs.2018.07.002

Hopton, M. E., Cabezas, H., Campbell, D., Eason, T., Garmestani, A. S., Heberling, M. T., et al. (2010). Development of a multidisciplinary approach to assess regional sustainability. *Int J Sust Dev World* 17, 48–56. doi: 10.1080/13504500903488297

Krishna, A., and Schwarz, N. (2014). Sensory marketing, embodiment, and grounded cognition: a review and introduction. *J. Consum. Psychol.* 24, 159–168. doi: 10.1016/j.jcps.2013.12.006

Lévy, C. M., Macrae, A., and Köster, E. P. (2006). Perceived stimulus complexity and food preference development. *Acta Psychol.* 123, 394–413. doi: 10.1016/j.actpsy.2006.06.006

Lick, E., König, B., Kpossa, M. R., and Buller, V. (2017). Sensory expectations generated by colors of red wine labels. *J. Retail. Consum. Serv.* 37, 146–158. doi: 10.1016/j.jretconser.2016.07.005

Martini, D., Ragone, G., Cazzini, F., Cheli, F., Formici, G., la Porta, C. A. M., et al. (2021). The need for a multidisciplinary approach to face challenges related to food, health, and sustainability: the contribution of CRC I-WE. *Sustainability* 13:13720. doi: 10.3390/su132413720

Maziriri, E. T., Rukuni, T. F., and Chuchu, T. (2021). Factors influencing food consumption satisfaction and purchase decisions of restaurant consumers. *Cogent Bus. Manag.* 8. doi: 10.1080/23311975.2021.1968731

Merli, R., Preziosi, M., and Acampora, A. (2018). Sustainability experiences in the wine sector: toward the development of an international indicators system. *J. Clean. Prod.* 172, 3791–3805. doi: 10.1016/j.jclepro.2017.06.129

Nguyen, T. H., and Durner, D. (2023). Sensory evaluation of wine aroma: Should color-driven descriptors be used? *Food Qual. Prefer.* 107:104844. doi: 10.1016/j.foodqual.2023.104844

Paradis, C., and Eeg-Olofsson, M. (2013). Describing sensory experience: the genre of wine reviews. *Metaphor. Symb.* 28, 22–40. doi: 10.1080/10926488.2013.742838

Peschel, A. O., Kazemi, S., Liebichová, M., Sarraf, S. C. M., and Aschemann-Witzel, J. (2019). Consumers' associative networks of plant-based food product communications. *Food Qual. Prefer.* 75, 145–156. doi: 10.1016/j.foodqual.2019.02.015

Piqueras-Fiszman, B., and Spence, C. (2012). The weight of the bottle as a possible extrinsic cue with which to estimate the price (and quality) of the wine? Observed correlations. *Food Qual. Prefer.* 25, 41–45. doi: 10.1016/j.foodqual.2012.01.001\

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/

Rinaldi, C. (2017). Food and gastronomy for sustainable place development: a multidisciplinary analysis of different theoretical approaches. *Sustainability* 9:1748. doi: 10.3390/su9101748

Rosen, M. A. (2009). Sustainability: a crucial quest for humanity – welcome to a new open access journal for a growing multidisciplinary community. *Sustainability* 1, 1–4. doi: 10.3390/su1010001

Roy, P., Nei, D., Orikasa, T., Xu, Q., Okadome, H., Nakamura, N., et al. (2009). A review of life cycle assessment (LCA) on some food products. *J. Food Eng.* 90, 1–10. doi: 10.1016/j.jfoodeng.2008.06.016

Ruggieri, L., Cadena, E., Martínez-Blanco, J., Gasol, C. M., Rieradevall, J., Gabarrell, X., et al. (2009). Recovery of organic wastes in the Spanish wine industry. Technical, economic and environmental analyses of the composting process. *J. Clean. Prod.* 17, 830–838. doi: 10.1016/j.jclepro.2008.12.005

Spence, C. (2017). *Gastrophysics: The new science of eating*. New York, NY: Viking.

Spence, C. (2020). Senses of place: architectural design for the multisensory mind. *Cogn. Res. Princ. Implic.* 5:46. doi: 10.1186/s41235-020-00243-4

Spence, C. (2022a). Experimental atmospherics: a multi-sensory perspective. *Qual. Mark. Res. Int. J.* 25, 662–673. doi: 10.1108/QMR-04-2022-0070

Spence, C. (2022b). The form of taste: on the origins, implications, and applications of shape-taste crossmodal correspondences. *UOU Scient. J.* 4, 150–163. doi: 10.14198/UOU.2022.4.15

Spence, C. (2022c). Exploring group differences in the crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079

Spence, C. (2023). Explaining visual shape-taste crossmodal correspondences. *Multisens. Res.* 36, 313–345. doi: 10.1163/22134808-bja10096

Spence, C., Hobkinson, C., Gallace, A., and Piqueras Fiszman, B. (2013). A touch of gastronomy. *Flavour* 2, 1–15. doi: 10.1186/2044-7248-2-14

Spence, C., and Van Doorn, G. (2022). Visual communication via the design of food and beverage packaging. *Cogn. Res. Princ. Implic.* 7:42. doi: 10.1186/s41235-022-00391-9

Spence, C., Velasco, C., and Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour* 3:8. doi: 10.1186/2044-7248-3-8

Spence, C., and Wang, Q. (2015a). Wine and music (I): on the crossmodal matching of wine and music. *Flavour* 4:34. doi: 10.1186/s13411-015-0045-x

Spence, C., and Wang, Q. J. (2015b). Wine and music (II): can you taste the music? Modulating the experience of wine through music and sound. *Flavour* 4:33. doi: 10.1186/s13411-015-0043-z

Spence, C., and Wang, Q. J. (2015c). Wine and music (III): so what if music influences the taste of the wine? *Flavour* 4:36. doi: 10.1186/s13411-015-0046-9

Stock, P., and Burton, R. J. F. (2011). Defining terms for integrated (multi-inter-trans-disciplinary) sustainability research. *Sustainability* 3, 1090–1113. doi: 10.3390/su3081090

Uiterkamp, A. J. M. S., and Vlek, C. (2007). Practice and outcomes of multidisciplinary research for environmental sustainability. *J. Soc. Issues* 63, 175–197. doi: 10.1111/j.1540-4560.2007.00502.x

Varela, P., and Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Res. Int.* 48, 893–908. doi: 10.1016/j.foodres.2012.06.037

Wang, H., Feng, X., Suo, H., Yuan, X., Zhou, S., Ren, H., et al. (2022). Comparison of the performance of the same panel with different training levels: flash profile versus descriptive analysis. *Food Qual. Prefer.* 99:104582. doi: 10.1016/j.foodqual.2022.104582

Wang, Q. J., and Spence, C. (2018). A smooth wine? Haptic influences on wine evaluation. *Int. J. Gastron. Food. Sci.* 14, 9–13. doi: 10.1016/j.ijgfs.2018.08.002

Wörfel, P., Frentz, F., and Tautu, C. (2022). Marketing comes to its senses: a bibliometric review and integrated framework of sensory experience in marketing. *Eur. J. Mark.* 56, 704–737. doi: 10.1108/EJM-07-2020-0510

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 14 August 2023
DOI 10.3389/fpsyg.2023.1226661

Check for updates

OPEN ACCESS

EDITED BY
Carlos Cornejo,
Pontificia Universidad Católica de Chile, Chile

REVIEWED BY
Nicola Di Stefano,
National Research Council (CNR), Italy
Vivian Ciaramitaro,
University of Massachusetts Boston,
United States

*CORRESPONDENCE
Xianwei Meng
✉ meng@i.nagoya-u.ac.jp
Na Chen
✉ imminana7@gmail.com

RECEIVED 22 May 2023
ACCEPTED 01 August 2023
PUBLISHED 14 August 2023

CITATION
Meng X, Chen N, Ishida J, Watanabe K and Murakami T (2023) Crossmodal correspondences between visual features and tastes in preschoolers: an exploratory study. *Front. Psychol.* 14:1226661.
doi: 10.3389/fpsyg.2023.1226661

Crossmodal correspondences between visual features and tastes in preschoolers: an exploratory study

Xianwei Meng[1]*, Na Chen[2]*, Junya Ishida[3], Katsumi Watanabe[4,5] and Taro Murakami[3]

[1]Graduate School of Informatics, Nagoya University, Nagoya, Japan, [2]The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel, [3]Faculty of Education and Care of Early Childhood, Tokoha University, Shizuoka, Japan, [4]Faculty of Science and Engineering, Waseda University, Tokyo, Japan, [5]Department of Psychology, University of New South Wales, Sydney, NSW, Australia

Introduction: Adults possess a natural inclination to associate sensory cues derived from distinct modalities, such as the pairing of sweet with pink. However, studies exploring crossmodal correspondences in children, particularly in the sensory pairing of visual features and tastes, are scant, leaving unanswered questions regarding the developmental trajectory of crossmodal correspondences. The present study investigates whether Japanese preschool children demonstrate specific biases in shape–color, shape–taste, and color–taste associations.

Methods: In a series of in-person experiments, 92 children between 3 to 6 years of age completed matching tasks utilizing paper stimuli.

Results: Children exhibit crossmodal correspondences in shape-color (circle-red and asymmetrical star-yellow), shape–taste (triangle-salty and circle-sweet), and color–taste (yellow-sour, black-bitter, and pink-sweet) associations. Moreover, children's choices are not influenced by their individual preferences.

Discussion: The crossmodal correspondences observed in this study have been observed in previous research on adults from the same (Japanese) culture, although adults showed more crossmodal correspondences than the children in this study (e.g., pink-circle, triangle-sour, and green-bitter). Thus, while some crossmodal correspondences emerge during childhood, others may require additional time to develop, thereby highlighting the importance of understanding the cognitive mechanisms underlying crossmodal correspondences from an ontogenic perspective.

KEYWORDS

crossmodal correspondence, shape, color, taste, child development, sensory pairing, association

1. Introduction

The cognitive ability to integrate information derived from various features and stimuli dimensions is fundamental in everyday life (Spence, 2011; Stein, 2012). For example, people use the association between size and sound to make various practical decisions: people often link the size of an animal to the sound it produces. Larger animals, like elephants or whales, are associated with deep, booming sounds, while smaller animals like birds are linked to higher-pitched chirping. People also often use the association to assess the size of objects when hearing them drop on the floor. When an object falls and makes a sound upon impact, the sound produced can give us some clues about the object's size and weight: larger and heavier objects tend to produce louder and lower-pitched sounds when they hit the floor.

Cognitive scientists have investigated the factors or circumstances producing or facilitating such integration to understand how individuals integrate information from different features and dimensions. Traditional research has demonstrated that stimuli are more likely to be associated if presented closely in time and space (Stein et al., 1988). Additionally, individuals tend to associate information when their meanings are related rather than unrelated (e.g., pairing a woofing sound with a static image of a dog) (Molholm et al., 2004; Hein et al., 2007). Recently, associations have been identified among fundamental stimulus characteristics across various modalities, such as color, shape, taste, pitch, size, and brightness. These are shared by many individuals, and some of these seem to be culture-independent (e.g., black color-bitter taste; Wan et al., 2014) or have deep developmental origins [e.g., the low pitch-big size association observed in 3-month-old infants; (Pietraszewski et al., 2017)]. These mappings are commonly referred to as "crossmodal correspondences" (Spence, 2011; Parise and Spence, 2014).

Crossmodal correspondence has been explained through four distinct categories of accounts: emotional, semantic, statistical, and structural. The first posits that correspondence arises because unimodal stimuli elicit similar emotional experiences, providing a plausible explanation for certain color–taste correspondences, such as pink-sweet and black-bitter, mediated by positive and negative emotions, respectively (Velasco et al., 2016; Spence, 2020a; Di Stefano et al., 2021; Spence and Di Stefano, 2022). By contrast, the second asserts that correspondence arises from the shared meaning of stimuli. For example, in some languages, pitch is described using the spatial terms "thin and thick," and these labels can promote crossmodal mappings (Shayan et al., 2014). This account may also elucidate the way we associate animal appearances with voices (Molholm et al., 2004; Hein et al., 2007). According to the third, correspondence results from statistical learning through exposure to co-occurring stimuli in an environment. For example, the crossmodal correspondence between auditory pitch and spatial elevation (e.g., high or low sounds) has been considered to be extracted from experience; when analyzing natural sounds from the environment, researchers found a clear mapping between frequency and elevation (i.e., higher-pitched sounds occur more frequently in higher space; Parise et al., 2014). This account also enables us to understand the reason why some crossmodal correspondences (e.g., correspondences between certain colors/shapes and tastes) vary between cultures (Wan et al., 2014). Finally, the fourth posits that correspondence arises because of unimodal stimuli sharing the same or interconnected neural basis, which may be innate (Wagner and Dobkins, 2011; Deroy and Spence, 2013; Ludwig and Simner, 2013; Pietraszewski et al., 2017). Wagner and Dobkins (2011) demonstrated that certain shapes influence color preferences in typical 2- and 3-month-olds, but not in 8-month-olds or adults, suggesting that exuberant neural connectivity facilitates synesthetic associations during infancy. These categories of crossmodal correspondences are not mutually exclusive and often overlap, making it challenging to explain them using a single hypothesis.

Although a multitude of empirical studies have investigated crossmodal correspondence, the majority have solely focused on adult participants (Wan et al., 2014; Chen et al., 2021). Consequently, there exists a gap in the literature regarding the developmental trajectory of crossmodal correspondences, particularly related to the sensory pairing of visual features (color/shape) and tastes (for developmental studies of other correspondences, see Simpson et al., 1956; Lewkowicz and Turkewitz, 1980; Ludwig and Simner, 2013; Pejovic and Molnar, 2017; Thomas et al., 2017; Chow et al., 2021; Speed et al., 2021). Previous research has found adults exhibiting specific color–taste/shape–taste crossmodal correspondences, such as yellow–sour, red–sweet, angular–sour/bitter, and round–sweet associations (Spence and Gallace, 2011; Spence and Ngo, 2012; Spence, 2012, 2019; Bremner et al., 2013; Wan et al., 2014; Spence et al., 2015; Velasco et al., 2015; Saluja and Stevenson, 2018; Spence and Levitan, 2021). However, literature on younger participants, such as preschool children, is lacking. Studying preschoolers can illuminate whether these crossmodal associations occur independent of the written language and understanding of metaphorical uses of language, and limited life experiences, as compared to adults (Spector and Maurer, 2009, 2011; Meng et al., 2022). Moreover, testing toddlers over infants is advantageous because they understand simple verbal instructions and can be tested using methods yielding easily interpreted data from more test points (Spector and Maurer, 2009).

Previous studies suggest that learning that involves acquisition of associations between environmental stimuli, may account for a certain crossmodal correspondence. For instance, infants as young as 4 months of age have demonstrated the ability to learn arbitrary bimodal correspondences between color and taste, as evidenced by their consistent selection of the color that had been paired with sweetness (Reardon and Bushnell, 1988). Moreover, cross-cultural variations have been observed in the correspondence between specific colors, shapes, and taste terms, indicating that adults' associations may stem from their daily experiences (Wan et al., 2014). However, no study has provided robust evidence on whether children exhibit adult-like color–taste/shape–taste crossmodal correspondences. To date, only Fateminia et al. (2020) have investigated the association between color and taste in preschool children. The study involved presenting a group of Iranian preschoolers, aged 2–6 years, with eleven colors categorized as primary and secondary, and asking them to select a taste corresponding to each color from a set of options, including sour, salty, sweet, bitter, pungent, astringent, and tasteless. Although the authors interpreted the results as revealing "meaningful and clear relationships" between colors and tastes, the data analysis relied solely on descriptive statistics, failing to investigate the likelihood that the outcomes were due to chance or actual effects, and whether they could be generalized to a wider population (Fateminia et al., 2020).

To bridge this gap, the present study investigates the crossmodal correspondences between visual features and tastes in preschool-aged children. The selection of this age group has enabled the utilization of more sophisticated experimental methods such as manual or verbal responses, typically employed in adult studies, while ensuring that the children possessed the cognitive capacity to comprehend and complete the tasks. To ensure comparability with previous research conducted on adult populations, we modeled our study design after the work of Chen et al. (2021), who investigated shape–color, shape–taste, and color–taste correspondences in Japanese adults. The tasks were modified to be more age-appropriate for children (Chen et al., 2021).

In a series of in-person experiments, children engaged in matching tasks using paper stimuli. The stimuli involved the following elements. (1) Five shapes: a triangle, circle, square, blob (bouba), and an asymmetric star (kiki) (Köhler, 1947; Chow and Ciaramitaro, 2019); (2) nine colors: black, blue, green, orange, pink, purple, red, white, and yellow; and (3) four basic taste words: sweet, sour, bitter, and salty (Spence and Levitan, 2021). This is the first study to investigate crossmodal associations between visual features and taste in preschool-aged children by examining three different types of sensory pairings. We conducted this research with an exploratory approach because we did not formulate any specific hypotheses regarding the potential correspondence patterns that might have emerged from the participating children.

2. Method

2.1. Participants

The final sample included 92 3- to 6-year-old children living in Kitakyushu, Japan (50 girls and 42 boys; mean age = 62.21 months, SD = ±9.83). We determined our sample size following Chen et al. (2021) who investigated crossmodal correspondences in adults using similar tasks. One 3-year-old child did not provide answers to some of the questions (regarding the taste of purple, potato, strawberry, and green pepper); the child's data from the provided answers was nonetheless included in the analysis. Prior written consent was obtained from the children's caregivers. This study was approved by the Ethics Committee of Tokoha University (No. 22-20) and conducted in accordance with the ethical standards of the American Psychological Association (2017) and the Declaration of Helsinki (2001).

2.2. Set-up

The experiment was conducted in a quiet room at a kindergarten. Three researchers individually tested each child. Each pair of children and the experimenter sat face-to-face across a child's table. The experimenter presented the paper stimuli on the table and asked the children questions. All experimental sessions were video recorded.

2.3. Materials

All stimuli were created using PowerPoint (Microsoft) and printed onto white hard paper. Nine color paper patches and five geometric shapes were used as "target" stimuli (Chen et al., 2021; Figure 1). The patches were circular with a diameter of 5.5 cm and were black, blue, green, orange, pink, purple, red, white, and yellow in color. The geometric shapes included circle, square, triangle, asymmetrical star, and blob, and the sizes were similar to the color patch size.

Four basic taste terms and six colors were used as "selection" stimuli. The taste items were presented in black Japanese text as "にがい (bitter)," "しょっぱい (salty)," "すっぱい (sour)," and "あまい (sweet)." The colors selected from the target stimuli were blue, green, orange, purple, red, and yellow. All colors were printed on a piece of paper. The order of the taste items and colors from left to right was created in two patterns, and each participant was randomly assigned to either pattern.

2.4. Procedure

The experiment consisted of three phases: warm-up, crossmodal correspondence, and follow-up. Prior to the experiment, the researcher explained to each child that they would participate in an interesting quiz game. All participants categorically affirmed that they wanted to play the game.

The warm-up phase familiarized the children with the experimental circumstances and materials. The researcher showed the children the target color patches and asked them to indicate the names of the colors. Then, they were presented with two black-and-white pictures: a girl and a boy cartoon characters. For each character, children were asked to choose a color from the color selections for their t-shirts. The pictures were presented randomly.

The crossmodal correspondence phase consisted of two sessions in a fixed order: the basic shape and bouba–kiki, in which circles, squares, and triangles and asymmetrical stars and blobs, respectively, were used as the target shapes. Each session consisted of a shape–color matching task followed by a shape–taste matching task. In the first task, the target shapes were presented individually at the center of the table, and the participants were asked to choose a color from the color selections that best matched each shape. Whereas, in the second task, they had to choose a taste from the taste selections that best matched the target shapes. To ensure that children understood the words for the tastes presented as selections, the researcher read each word (e.g., "This word is *sweet*") and confirmed that children could match each word on the paper with what they read. Additionally, a color–taste matching task was conducted between the two sessions, in which participants were shown the target color patches and asked to choose a taste from the taste selection. In each matching task, the target stimuli were presented in random orders.

The follow-up phase had two aims. The first was to test whether the participants understood the concept of taste. Therefore, in the taste-comprehension task, they were shown pictures of French fries, green peppers, lemons, and strawberries and asked to choose a taste from the taste selections that best matched each food. These foods were selected assuming that they are typical foods recognized by children as salty, bitter, sour, and sweet. The second aim was to test whether the children's color choices were influenced by their preferences. Therefore, a color-preference task was conducted where participants were asked to choose one color that they liked best from the color selection.

2.5. Analysis plan

Chi-square tests were used to determine whether the children showed crossmodal correspondences. For each matching test, we examined whether the choices varied by category. Furthermore, an adjusted residual analysis was used to determine which choice was significantly associated with a certain target (i.e., color–taste/shape–taste/shape–color associations). The analysis shows which selection is chosen more frequently than expected by chance for each target. Positive and negative residual values indicate more and less frequently chosen than expected,

10.3389/fpsyg.2023.1226661

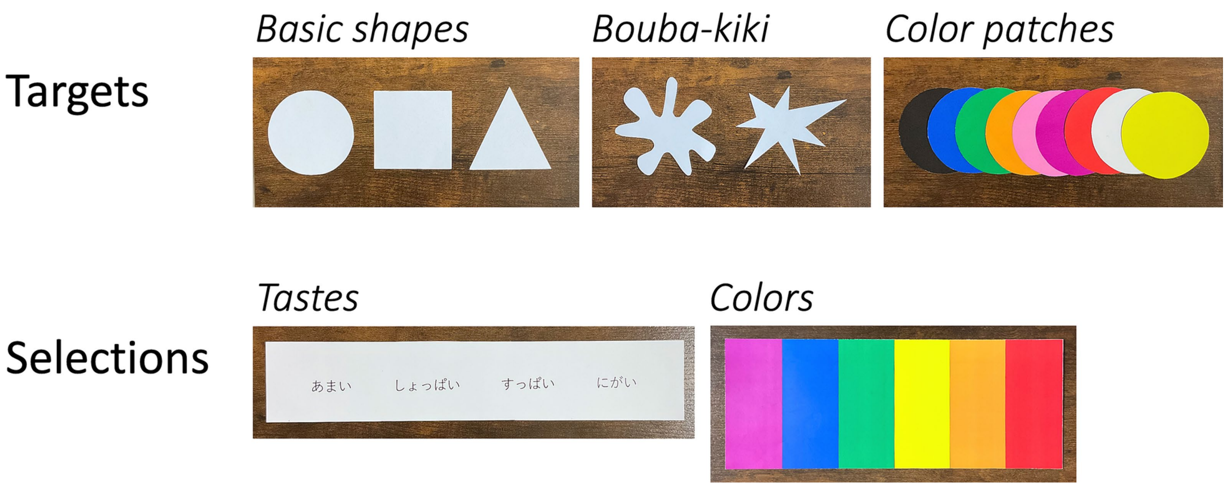

FIGURE 1
Materials of the experiment.

respectively; here, only positive residuals were used (Agresti, 2012; Chen et al., 2015a). A similar analysis was conducted for the taste-comprehension task. Furthermore, the correlation coefficient between the children's judgment in the crossmodal correspondence phase and the color preference task was evaluated.

3. Results

3.1. Basic shapes session (shape–color and shape–taste matching tasks)

The results of the Chi-square test show that the color choices vary by the shape categories ($\chi^2=24.65$, $df=10$, $p=0.006$, Cramer's $V=0.211$; Figure 2). The adjusted residual results show the association of a circle with red color (z = 4.13, adjusted $p=0.001$, p-values are corrected using Holm's method for multiple comparisons). Furthermore, the taste choices vary by shape categories ($\chi^2=27.17$, $df=6$, $p<0.001$, Cramer's $V=0.222$). The adjusted residual results indicate the association of the triangle with salty (z = 3.34, adjusted $p=0.010$), and circle with sweet taste ($z=3.34$, adjusted $p=0.010$), respectively.

3.2. Color–taste matching task

The results of the Chi-square test show that taste choices vary by color categories ($\chi^2=72.33$, $df=24$, $p<0.001$, Cramer's $V=0.171$; Figure 3). Specifically, yellow, black, and pink are associated with sour ($z=3.82$, adjusted $p=0.005$), bitter ($z=4.13$, adjusted $p=0.001$), and sweet taste ($z=3.34$, adjusted $p=0.029$), respectively.

3.3. Bouba–kiki session (shape–color and shape–taste matching tasks)

The results of the chi-square test show that color choices vary according to the shape categories of the bouba–kiki session ($\chi^2=32.18$, $df=5$, $p<0.001$, Cramer's $V=0.418$; Figure 4); however, there is no significant association between taste and shape ($\chi 2=2.23$, df = 3, $p=0.526$, Cramer's $V=0.110$). The adjusted residual results show that a kiki is associated with yellow (z = 5.57, adjusted $p=< 0.001$).

3.4. Follow-up

For the taste-comprehension task, the results of the Chi-square test show that taste choices vary by foods ($\chi^2=395.38$, $df=9$, $p<0.001$, Cramer's $V=0.573$; Figure 5). The adjusted residual results show that children associate French fries with salty (z = 5.52, adjusted $p<0.001$) and sweet taste ($z=3.62$, adjusted $p=0.001$). Furthermore, they associate green pepper with bitter ($z=12.77$, adjusted $p=< 0.001$), lemons with sour ($z=12.19$, adjusted $p=< 0.001$), and strawberries with sweet taste ($z=9.71$, adjusted $p=< 0.001$), respectively.

Correlation statistics indicate that children's color choices in crossmodal correspondence tasks inadequately agree with their color preferences (McHugh, 2012). The kappa coefficients of the paired data between color preference and color choices for triangles, circles, squares, blobs, and asymmetrical stars were 0.032, 0.041, 0.023, 0.172, and −0.020, respectively.

4. Discussion

Previous studies indicate that adults consistently associate features of different sensory modalities, such as sweet taste with pink colors. To better understand the mechanisms underlying the acquisition of crossmodal correspondence, this study explores crossmodal associations between visual features and taste in preschoolers using paper-based stimuli. The participants were asked to match color to shape, taste to shape, and taste to color. The results indicate that children associate red with circles and yellow with asymmetric stars, salty with triangles and sweet with circles, and sour with yellow, bitter with black, and sweet with pink in shape–color, shape–taste, and color–taste tasks, respectively.

Several shape–color correspondences observed in the current study sample of children differed from those found in previous studies

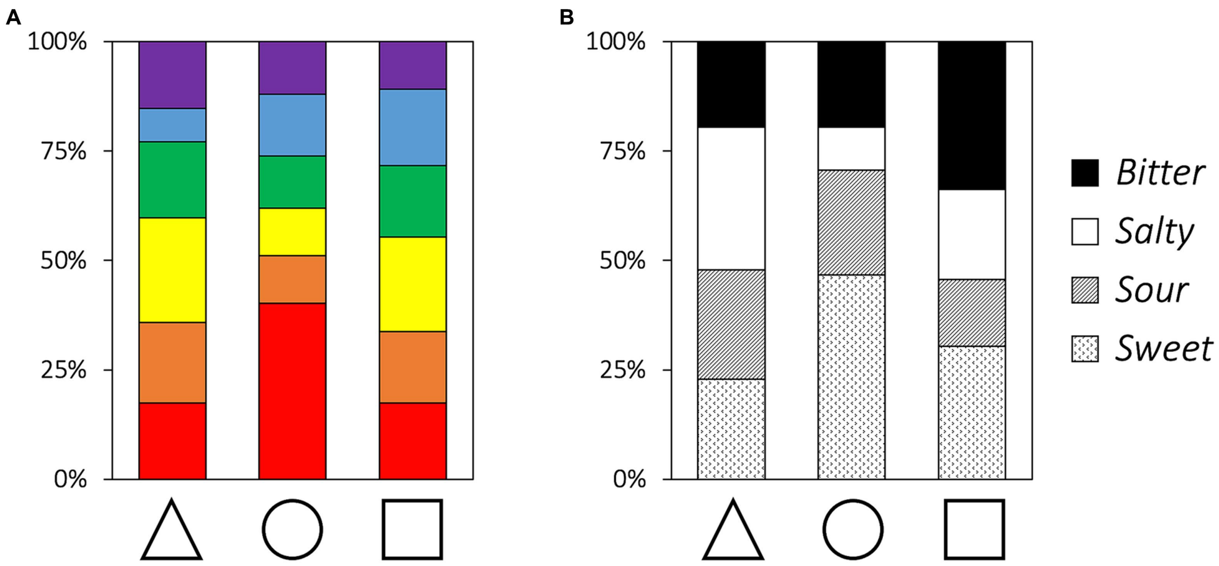

FIGURE 2
Proportions of color **(A)** and taste **(B)** choices for each basic shape.

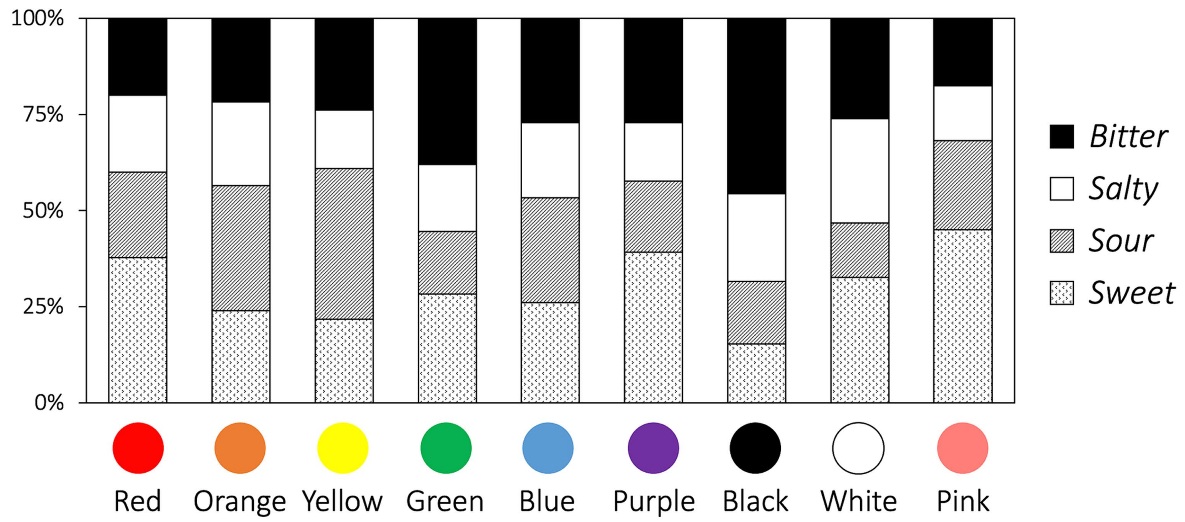

FIGURE 3
Proportion of taste choices for each color.

involving adult participants who shared the same cultural background (Chen et al., 2021). According to Chen et al. (2021), adults tend to associate red and pink with circles, yellow with triangles and asymmetric stars, blue with squares, and green with blobs. The pink–circle, yellow–triangle, blue–square, and green–blob associations were not found in the current sample. Note that the red–circle and yellow–asymmetrical star associations might be the result of exposure to frequent combination occurrences in the environment, such as the Japanese flag and night sky stars (Chen et al., 2015a,b). Therefore, even though some color–shape associations might be observed from early infancy (Wagner and Dobkins, 2011), the current findings may provide evidence supporting the hypothesis that shape–color correspondences are the result of statistical learning from environmental co-occurrences (Parise et al., 2014). One plausible explanation for the absence of a yellow–triangle association is that the yellow–asymmetrical star association is acquired earlier, and then generalized to a yellow–triangle association during development, given that both shapes have sharp angles. In line with the learning hypothesis, studies have demonstrated that even the most basic shapes used in this study, such as circles, squares, and triangles, exhibit cultural variations in their shape–color correspondences; for instance, German participants showed red–triangle, blue–square and yellow–circle association correspondences, whereas Japanese participants showed yellow–triangle, blue–square and red–circle association correspondences (Jacobsen, 2002; Chen et al., 2015a; Dreksler and Spence, 2019). In another stream of research, Spector and Maurer

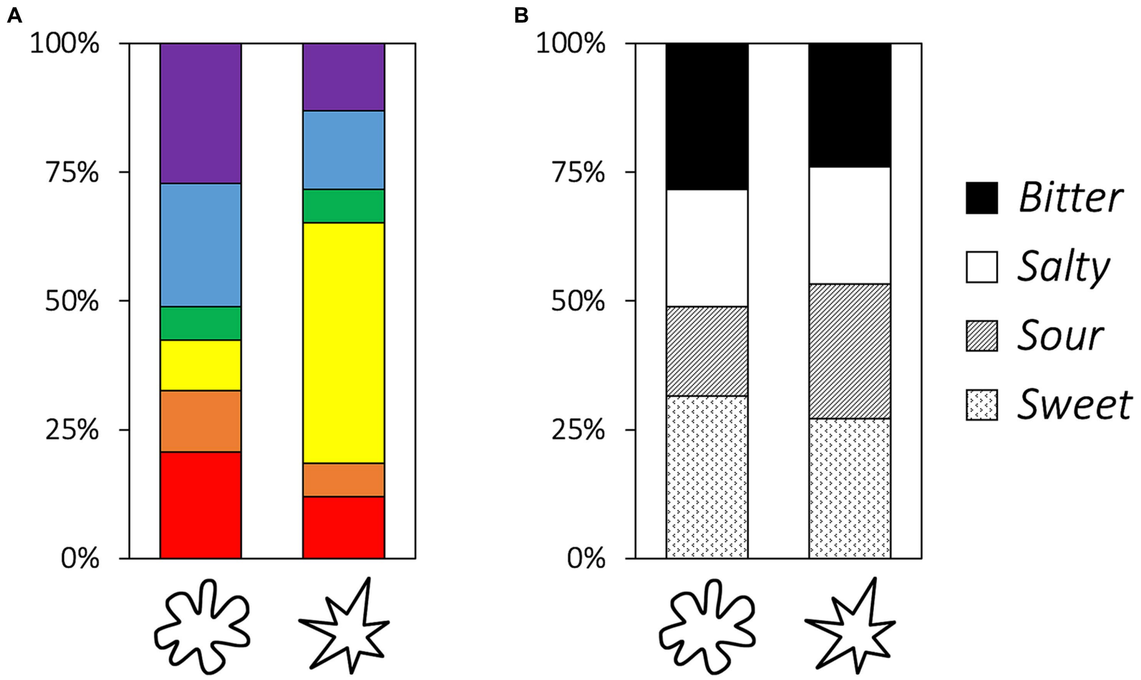

FIGURE 4
Proportion of color **(A)** and taste **(B)** choices for blob and asymmetrical star.

FIGURE 5
Proportion of taste choices for French fries, green pepper, lemon, and strawberry.

(2008, 2009, 2011) explored the associations between alphabets and colors in pre-literate and literate children and adults in Canada. Their findings revealed that the participants in all groups associate specific letters with colors (e.g., O with white and X with black), presumably based on their shapes but not their phonetic or morphological properties (e.g., B with blue). Furthermore, preliterate children did not exhibit the same associations observed in literate children and adults (e.g., A with red and G with green). These results, combined with the ones of the present study, suggest that even pre-literate children's shape–color correspondence is influenced by cultural factors (e.g., circle is associated with red in Japanese children and with white in Canadian children), consistent with the learning hypothesis.

The participating children also exhibited several shape–taste correspondences: associating salty with triangles and sweet with circles. This finding is consistent with that of previous research on adults in Japanese culture, where sharp-angled (e.g., triangles, asymmetrical stars), circular (e.g., circles), right-angled (e.g., squares), and blob shapes are associated with sour, sweet, salty, and bitter taste, respectively (Chen et al., 2021). Notably, our follow-up taste comprehension task confirmed that participating children could accurately differentiate between salty and sour tastes, thus supporting the significance of the observed salty-triangle association. A developmental shift in shape–taste correspondence may occur from preschoolers to adults, whereby the association between salty taste and triangles shifts to sour taste and triangles. Several mechanisms have been proposed to account for the development of shape–taste correspondences, including shape–taste experiences [e.g., sweet candies being typically round in shape; Spence, 2023; see also Speed et al. (2021)'s notion on the perspective of odor–shape association] and emotional mediation (e.g., the association between sweet taste and roundness owing to their positive valence; Blazhenkova and Kumar, 2018). Future research should investigate the potential role of these mechanisms in shaping shape–taste correspondences in children.

The present findings on color–taste correspondence in children are consistent with those of previous studies conducted on adults across various cultures (O'mahony, 1983; Koch and Koch, 2003; Tomasik-Krótki and Strojny, 2008; Wan et al., 2014; Spence et al., 2015; Woods et al., 2016; Woods and Spence, 2016; Spence, 2019; Chen et al., 2021). For adult participants sharing the same cultural background as that of the children in our study, Chen et al. (2021) observed additional color–taste associations, specifically orange with sour, blue and white with salty, and green and purple with bitter. Based on our follow-up taste-comprehension task results, which indicated that the participants were able to correctly identify sour, salty, and bitter tastes, the observed color–taste correspondences may reflect a developmental change (Chen et al., 2021). A green–bitter association has been documented in adults across different cultures (Saluja and Stevenson, 2018; Chen et al., 2021). Interestingly, although the children also associated bitter taste with green peppers, they did not exhibit a significant bias toward linking bitter taste with green. Adult participants also often associate green with sourness (O'mahony, 1983; Koch and Koch, 2003; Tomasik-Krótki and Strojny, 2008; Wan et al., 2014; Woods et al., 2016; Woods and Spence, 2016), which may be because of their experience consuming sour green fruits. Similar associative learning mechanisms may account for the development of the white–salty association, which may be established through exposure to salt. Our results indicate that children judged French fries as not only salty but also sweet. This may be because children are prevented from consuming excessive amounts of salt in their daily diet, including that coming from French fries, for health-related reasons. Another reason might be that the presence of added sugar in certain industrial fries could contribute to the perception of sweetness in the context of French fries (Atsumi et al., 2023).

Regarding previous findings on color–taste correspondences in children having different cultural backgrounds, the descriptive statistics of Fateminia et al. (2020) showed that 2- to 6-year-old Iranian children tend to associate blue with sour, red with sweet and sour, yellow with sweet and sour, purple with bitter, orange with sour, green with sour, black with bitter, white with sweet, pink with sweet (Fateminia et al., 2020). Some of these correspondences (i.e., yellow–sour, black–bitter, and pink–sweet), but not others, are shared by the current study sample of Japanese children. These imply that the acquisition of color–taste correspondences may be both culture-dependent and culture-independent, depending on the nature of certain correspondences.

Nevertheless, the current experimental evidence supports the notion that experience-based explanations may underlie color–taste associations. Higgins and Hayes (2019) demonstrated using a repeated brief exposure paradigm that color–taste associations may be learned through repeated exposure, even when the exposure period is relatively brief (e.g., several days).

Contrastingly, even though the participating children associated green pepper with bitter taste, their association with it was black, rather than green. Moreover, young children may have limited opportunities to be exposed to a bitter taste. These findings and considerations suggest that alternative mechanisms for the acquisition of color–taste correspondence in children must be considered. One such mechanism may be emotional or hedonic associations, as previous research has shown that bitter taste is generally disliked or associated with danger (Higgins and Hayes, 2019; Spence, 2023). Additionally, black is the least favorite color of children and elicits negative emotions (Boyatzis and Varghese, 1994; Koleoso et al., 2014). Therefore, it is plausible that the association between bitter taste and black color in children is based on shared negative values (Salgado-Montejo et al., 2015; Velasco et al., 2015; Turoman et al., 2018; Spence, 2023). Future research should use more detailed measures, such as taste and color evaluation, to directly test this hypothesis.

This study has several limitations which raise important questions that can form the basis for future research. First, to make the tasks more child-friendly, we limited the shapes to basic geometric shapes and the Bouba–Kiki pair (Chow and Ciaramitaro, 2019). Additionally, we limited the tastes to four basic tastes (sweet, sour, bitter, and salty), although many food scientists suggest that there are more than 20 basic tastes (Stuckey, 2012). To explore other crossmodal correspondences, future studies should include a broader range of stimuli, while ensuring that children make reliable choices.

Second, children were asked to select colors for shapes, tastes for shapes, and tastes for colors. These selections are unidirectional, making it unclear whether children link items bidirectionally (Wan et al., 2014; Chen et al., 2021). For instance, when asked to choose a taste for yellow color, children may first imagine yellow food, such as lemons, and select sour for yellow because lemons have a sour taste. However, if the procedure was reversed and children were asked to choose a color for sour taste, whether a yellow–sour correspondence would still occur cannot be said (Fateminia et al., 2020). Some crossmodal correspondences have been thought to be bidirectional; for instance, larger objects are matched with lower-pitched sounds, and lower-pitched sounds are just as strongly associated with larger objects (Deroy and Spence, 2013; Spence, 2020b). However, this does not mean that all crossmodal correspondences are bidirectional. Future studies need to investigate the directionality of cross-modal correspondences in children, as it contributes to our understanding of the precise nature of the relationships between associated stimuli (Posner et al., 1976; Spence et al., 2001; Spence, 2019). Collaterally, future studies could also investigate whether individuals' crossmodal correspondences show consistency across different sensory modalities (e.g., a yellow–triangle–sour correspondence), and whether and how such tendencies vary between individuals. Chen et al. (2021) found that Japanese adult participants show certain color–taste and shape–color associations and that such associations were observed more frequently in individuals with lower autistic quotient scores. This was in line with the notion that individuals with autism tend to have a decreased ability to perform statistical learning with co-occurrences and regularities in the environment, leading to difficulties in associating and integrating sensory information across modalities. Therefore, one hypothesis might be that one crossmodal correspondence (e.g., color–taste association) could be predicted based on other crossmodal correspondences (e.g., taste–shape and color–shape associations), especially for individuals who have a higher ability to perform statistical learning from the environment.

Third, the current study had a limited sample size and included a wide range of ages, which precluded testing for potential developmental changes in the observed crossmodal correspondences in the participating children. While the preschoolers in this study can be classified as preliterate, their experiences with daily visual features and tastes, as well as their cognitive abilities to complete tasks, undergo significant developmental changes during the preschool period. Moreover, because the current study was the first attempt to test crossmodal associations between visual features and taste in Japanese preschoolers, there was no information on effect size that could be used for the sample size calculation. Therefore, we followed the sample size applied in the Japanese adult study (Chen et al., 2021). As the current findings revealed relatively small effect sizes of the crossmodal correspondences, future studies need larger sample sizes to enhance the power of the tests in children. Taken together, larger-scale studies are needed to investigate the developmental trajectory of crossmodal correspondence in children.

The present study reveals that children exhibit fewer and more broadly distributed crossmodal correspondences compared to those observed in adults coming from different cultural backgrounds, supporting the notion that the cross-modal associations between visual features and tastes are learned from environmental co-occurrences (Spector and Maurer, 2008, 2009, 2011; Deroy and Spence, 2016; Higgins and Hayes, 2019; Spence and Levitan, 2021). Nonetheless, certain correspondences, such as the black–bitter association, may be better explained by other basic mechanisms. These results represent a new wave of developmental research aimed at elucidating how individuals map and integrate information based on various features and dimensions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Tokoha University (no. 22-20). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

XM, NC, JI, KW, and TM designed the study. XM created the stimuli. XM, JI, and TM conducted the experiments. XM analyzed the data and drafted the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the JSPS grants-in-aid for Scientific Research (20K20156, 20K14176, 23K02873, 23H03702, and 22H00090).

Acknowledgments

The authors would like to thank the staff at Seiwa Kindergarten for their advice and support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agresti, A. (2012). *Categorical data analysis*. New Jersey: John Wiley & Sons.

Atsumi, N., Yasumatsu, K., Takashina, Y., Ito, C., Yasui, N., Margolskee, R. F., et al. (2023). Chloride ions evoke taste sensations by binding to the extracellular ligand-binding domain of sweet/umami taste receptors. *Elife* 12, e84291. doi: 10.7554/eLife.84291

Blazhenkova, O., and Kumar, M. M. (2018). Angular versus curved shapes: correspondences and emotional processing. *Perception* 47, 67–89. doi: 10.1177/0301006617731048

Boyatzis, C. J., and Varghese, R. (1994). Children's emotional associations with colors. *J. Genet. Psychol.* 155, 77–85. doi: 10.1080/00221325.1994.9914760

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., and Spence, C. (2013). "Bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to westerners. *Cognition* 126, 165–172. doi: 10.1016/j.cognition.2012.09.007

Chen, N., Tanaka, K., Matsuyoshi, D., and Watanabe, K. (2015a). Associations between color and shape in Japanese observers. *Psychol. Aesthet. Creat. Arts* 9, 101–110. doi: 10.1037/a0038056

Chen, N., Tanaka, K., and Watanabe, K. (2015b). Color-shape associations revealed with implicit association tests. *PLoS One* 10, 1–15. doi: 10.1371/journal.pone.0116954

Chen, N., Watanabe, K., and Wada, M. (2021). People with high autistic traits show fewer consensual crossmodal correspondences between visual features and tastes. *Front. Pyschol.* 12:714277. doi: 10.3389/fpsyg.2021.714277

Chow, H. M., and Ciaramitaro, V. (2019). What makes a shape "baba"? The shape features prioritized in sound–shape correspondence change with development. *J. Exp. Child Psycho.* 179, 73–89. doi: 10.1016/j.jecp.2018.10.005

Chow, H. M., Harris, D. A., Eid, S., and Ciaramitaro, V. M. (2021). The feeling of "kiki": comparing developmental changes in sound-shape correspondence for audio-visual and audio-tactile stimuli. *J. Exp. Child Psycho.* 209:105167. doi: 10.1016/j.jecp.2021.105167

Deroy, O., and Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neurosci. Biobehav. Rev.* 37, 1240–1253. doi: 10.1016/j.neubiorev.2013.04.001

Deroy, O., and Spence, C. (2016). Crossmodal correspondences: four challenges. *Multisens. Res.* 29, 29–48. doi: 10.1163/22134808-00002488

Di Stefano, N., Murari, M., and Spence, C. (2021). "Crossmodal correspondences in art and science: Odours, poetry, and music" in *Olfaction: an interdisciplinary perspective from philosophy to life sciences* . Eds. N. D. Stefano and M. T. Russo (Cham: Springer), 155–189.

Dreksler, N., and Spence, C. (2019). A critical analysis of colour–shape correspondences: examining the replicability of colour–shape associations. *Iperception* 10:2041669519834042. doi: 10.1177/2041669519834042

Fateminia, M., Ghotbabadi, T. D., and Azad, K. M. (2020). Perceptions of the taste of colors in children and adults. *Color. Res. Appl.* 45, 743–753. doi: 10.1002/col.22503

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., and Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *J. Neurosci.* 27, 7881–7887. doi: 10.1523/JNEUROSCI.1740-07.2007

Higgins, M. J., and Hayes, J. E. (2019). Learned color taste associations in a repeated brief exposure paradigm. *Food Qual. Prefer.* 71, 354–365. doi: 10.1016/j.foodqual.2018.08.013

Jacobsen, T. (2002). Kandinsky's questionnaire revisited: fundamental correspondence of basic colors and forms? *Percept. Mot. Skills* 95, 903–913. doi: 10.2466/pms.2002.95.3.903

Koch, C., and Koch, E. C. (2003). Preconceptions of taste based on color. *J. Psychol.* 137, 233–242. doi: 10.1080/00223980309600611

Köhler, W. (1947). *Gestalt psychology* (*2nd ed.*). New York: Liveright.

Koleoso, O. N., Ehigie, B. O., and Akhigbe, K. O. (2014). Colour preference among children in a Nigerian Montessori school. *Mediterr. J. Soc. Sci.* 5, 325–332. doi: 10.5901/mjss.2014.v5n1p325

Lewkowicz, D. J., and Turkewitz, G. (1980). Cross-modal equivalence in early infancy: auditory-visual intensity matching. *Dev. Psychol.* 16, 597–607. doi: 10.1037//0012-1649.16.6.597

Ludwig, V. U., and Simner, J. (2013). What colour does that feel? Tactile-visual mapping and the development of cross-modality. *Cortex* 49, 1089–1099. doi: 10.1016/j.cortex.2012.04.004

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem. Med.* 22, 276–282. doi: 10.11613/BM.2012.031

Meng, X., Wang, J. J., Yoshikawa, Y., Ishiguro, H., and Itakura, S. (2022). A cross-cultural investigation of people's intuitive beliefs about the origins of cognition. *Front. Pyschol.* 13, 1–14. doi: 10.3389/fpsyg.2022.974434

Molholm, S., Ritter, W., Javitt, D. C., and Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb. Cortex* 14, 452–465. doi: 10.1093/cercor/bhh007

O'mahony, M. (1983). Gustatory responses to nongustatory stimuli. *Perception* 12, 627–633. doi: 10.1068/p120627

Parise, C. V., Knorre, K., and Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proc. Natl. Acad. Sci. U. S. A.* 111, 6104–6108. doi: 10.1073/pnas.1322705111

Parise, C., and Spence, C. (2014). "Audiovisual cross-modal correspondences in the general popul ation" in *The Oxford handbook of synesthesia*. eds. J. Simner and E. Hubbard (Oxford: Oxford University Press), 790–815.

Pejovic, J., and Molnar, M. (2017). The development of spontaneous sound-shape matching in monolingual and bilingual infants during the first year. *Dev. Psychol.* 53, 581–586. doi: 10.1037/dev0000237

Pietraszewski, D., Wertz, A. E., Bryant, G. A., and Wynn, K. (2017). Three-month-old human infants use vocal cues of body size. *Proc. Royal Soc. B.* 284:20170656. doi: 10.1098/rspb.2017.0656

Posner, M. I., Nissen, M. J., and Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychol. Rev.* 83, 157–171. doi: 10.1037/0033-295X.83.2.157

Reardon, P., and Bushnell, E. W. (1988). Infants' sensitivity to arbitrary pairings of color and taste. *Infant Behav. Dev.* 11, 245–250. doi: 10.1016/S0163-6383(88)80010-9

Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., and Spence, C. (2015). The sweetest thing: the influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Front. Pyschol.* 6:1382. doi: 10.3389/fpsyg.2015.01382

Saluja, S., and Stevenson, R. J. (2018). Cross-modal associations between real tastes and colors. *Chem. Senses* 43, 475–480. doi: 10.1093/chemse/bjy033

Shayan, S., Ozturk, O., Bowerman, M., and Majid, A. (2014). Spatial metaphor in language can promote the development of cross-modal mappings in children. *Dev. Sci.* 17, 636–643. doi: 10.1111/desc.12157

Simpson, R. H., Quinn, M., and Ausubel, D. P. (1956). Synaesthesia in children: association of colors with pure tone frequencies. *J. Genet. Psychol.* 89, 95–103. doi: 10.1080/00221325.1956.10532990

Spector, F., and Maurer, D. (2008). The colour of Os: naturally biased associations between shape and colour. *Perception* 37, 841–847. doi: 10.1068/p5830

Spector, F., and Maurer, D. (2009). Synesthesia: a new approach to understanding the development of perception. *Dev. Psychol.* 45, 175–189. doi: 10.1037/a0014171

Spector, F., and Maurer, D. (2011). The colors of the alphabet: naturally-biased associations between shape and color. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 484–495. doi: 10.1037/a0021437

Speed, L. J., Croijmans, I., Dolscheid, S., and Majid, A. (2021). Crossmodal associations with olfactory, auditory, and tactile stimuli in children and adults. *Iperception* 12:204166952110485. doi: 10.1177/20416695211048513

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2012). Managing sensory expectations concerning products and brands: capitalizing on the potential of sound and shape symbolism. *J. Consum. Psychol.* 22, 37–54. doi: 10.1016/j.jcps.2011.09.004

Spence, C. (2019). On the relationship(s) between color and taste / flavor. *Exp. Psychol.* 66, 99–111. doi: 10.1027/1618-3169/a000439

Spence, C. (2020a). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisens. Res.* 33, 1–29. doi: 10.1163/22134808-20191469

Spence, C. (2020b). Simple and complex crossmodal correspondences involving audition. *Acoust. Sci. Technol.* 41, 6–12. doi: 10.1250/ast.41.6

Spence, C. (2023). Explaining visual shape-taste crossmodal correspondences. *Multisens. Res.* 82, 1–33. doi: 10.1163/22134808-bja10096

Spence, C., and Di Stefano, N. (2022). Crossmodal harmony: looking for the meaning of harmony beyond hearing. *Iperception* 13:20416695211073817. doi: 10.1177/20416695211073817

Spence, C., and Gallace, A. (2011). Tasting shapes and words. *Food Qual. Prefer.* 22, 290–295. doi: 10.1016/j.foodqual.2010.11.005

Spence, C., and Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and Tatstes. *Iperception* 12:20416695211018223. doi: 10.1177/20416695211018223

Spence, C., and Ngo, M. K. (2012). Assessing the shape symbolism of the taste, flavour, and texture of foods and beverages. *Flavour* 1, 1–13. doi: 10.1186/2044-7248-1-12

Spence, C., Shore, D. I., and Klein, R. M. (2001). Multisensory prior entry. *J. Exp. Psychol. Gen.* 130, 799–832. doi: 10.1037//0096-3445.130.4.799

Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., et al. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour* 4, 1–17. doi: 10.1186/s13411-015-0033-1

Stein, B. E. (2012). *The new handbook of multisensory processing*. Cambridge, Massachusetts: Mit Press

Stein, B. E., Scott Huneycutt, W., and Alex Meredith, M. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Res.* 448, 355–358. doi: 10.1016/0006-8993(88)91276-0

Stuckey, B. (2012). *Taste what you're missing: The passionate eater's guide to why good food tastes good*. New York: Simon and Schuster.

Thomas, R. L., Nardini, M., and Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *J. Exp. Child Psychol.* 162, 72–88. doi: 10.1016/j.jecp.2017.04.020

Tomasik-Krótki, J., and Strojny, J. (2008). Scaling of sensory impressions. *J. Sens. Stud.* 23, 251–266. doi: 10.1111/j.1745-459X.2008.00153.x

Turoman, N., Velasco, C., Chen, Y. C., Huang, P. C., and Spence, C. (2018). Symmetry and its role in the crossmodal correspondence between shape and taste. *Atten. Percept. Psychophys.* 80, 738–751. doi: 10.3758/s13414-017-1463-x

Velasco, C., Woods, A. T., Deroy, O., and Spence, C. (2015). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Qual. Prefer.* 41, 151–158. doi: 10.1016/j.foodqual.2014.11.010

Velasco, C., Woods, A. T., Marks, L. E., Cheok, A. D., and Spence, C. (2016). The semantic basis of taste-shape associations. *PeerJ* 4:e1644. doi: 10.7717/peerj.1644

Wagner, K., and Dobkins, K. R. (2011). Synaesthetic associations decrease during infancy. *Psychol. Sci.* 22, 1067–1072. doi: 10.1177/0956797611416250

Wan, X., Woods, A. T., Van Den Bosch, J. J. F., McKenzie, K. J., Velasco, C., and Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Front. Pyschol.* 5, 1–13. doi: 10.3389/fpsyg.2014.01365

Woods, A. T., Marmolejo-Ramos, F., Velasco, C., and Spence, C. (2016). Using single colors and color pairs to communicate basic tastes II: foreground-background color combinations. *Iperception* 7, 1–20. doi: 10.1177/2041669516663750

Woods, A. T., and Spence, C. (2016). Using single colors and color pairs to communicate basic tastes. *Iperception* 7, 1–15. doi: 10.1177/2041669516658817

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 31 August 2023
DOI 10.3389/fpsyg.2023.1168258

Check for updates

OPEN ACCESS

EDITED BY
Charles Spence,
University of Oxford, United Kingdom

REVIEWED BY
Lieve Doucé,
University of Hasselt, Belgium
Ryan J. Ward,
University of Liverpool, United Kingdom

*CORRESPONDENCE
Mats B. Küssner
✉ mats.kuessner@hu-berlin.de

RECEIVED 17 February 2023
ACCEPTED 08 August 2023
PUBLISHED 31 August 2023

CITATION
Mesz B, Tedesco S, Reinoso-Carvalho F, Ter Horst E, Molina G, Gunn LH and Küssner MB (2023) Marble melancholy: using crossmodal correspondences of shapes, materials, and music to predict music-induced emotions.
Front. Psychol. 14:1168258.
doi: 10.3389/fpsyg.2023.1168258

Marble melancholy: using crossmodal correspondences of shapes, materials, and music to predict music-induced emotions

Bruno Mesz[1,2], Sebastián Tedesco[1], Felipe Reinoso-Carvalho[3], Enrique Ter Horst[3], German Molina[4], Laura H. Gunn[5,6,7] and Mats B. Küssner[8]*

[1]Instituto de Investigación en Arte y Cultura, Universidad Nacional de Tres de Febrero, Sáenz Peña, Argentina, [2]Programa de Investigación STSEAS, EUdA, UNQ, Bernal, Argentina, [3]Universidad de los Andes School of Management, Bogotá, Colombia, [4]Bayesian Solutions LLC, Charlotte, NC, United States, [5]Department of Public Health Sciences, University of North Carolina at Charlotte, Charlotte, NC, United States, [6]School of Data Science, University of North Carolina at Charlotte, Charlotte, NC, United States, [7]Faculty of Medicine, Department of Primary Care and Public Health, Imperial College London, London, United Kingdom, [8]Department of Musicology and Media Studies, Humboldt-Universität zu Berlin, Berlin, Germany

Introduction: Music is known to elicit strong emotions in listeners, and, if primed appropriately, can give rise to specific and observable crossmodal correspondences. This study aimed to assess two primary objectives: (1) identifying crossmodal correspondences emerging from music-induced emotions, and (2) examining the predictability of music-induced emotions based on the association of music with visual shapes and materials.

Methods: To achieve this, 176 participants were asked to associate visual shapes and materials with the emotion classes of the Geneva Music-Induced Affect Checklist scale (GEMIAC) elicited by a set of musical excerpts in an online experiment.

Results: Our findings reveal that music-induced emotions and their underlying core affect (i.e., valence and arousal) can be accurately predicted by the joint information of musical excerpt and features of visual shapes and materials associated with these music-induced emotions. Interestingly, valence and arousal induced by music have higher predictability than discrete GEMIAC emotions.

Discussion: These results demonstrate the relevance of crossmodal correspondences in studying music-induced emotions. The potential applications of these findings in the fields of sensory interactions design, multisensory experiences and art, as well as digital and sensory marketing are briefly discussed.

KEYWORDS

crossmodal correspondences, music-induced emotions, shapes, materials, machine learning, random forests, sensory interactions

1. Introduction

Crossmodal correspondences have been defined as the ability to map or associate features across different sensory modalities (Spence, 2011; Spence and Parise, 2012). In the auditory domain, crossmodal correspondences between pitch and visual or spatial features have been a recurrent topic of studies. For instance, most people match high-pitched sounds with small, bright objects located high up in space (Spence, 2011). However, there is also evidence for stable mappings between pitch and other sensory modalities such as taste and smell (Ward et al., 2022). In a study by Crisinel and Spence (2012), high-pitched sounds were associated with sweet and sour tastes, while low-pitched sounds were preferably matched with umami and bitter tastes. Belkin et al. (1997) demonstrated matches between certain auditory pitches with specific odorants according to their odor quality; just as specific odors are matched to certain types of instruments (Crisinel and Spence, 2012), or basic tastes to colors and shapes (Turoman et al., 2018; Lee and Spence, 2022). Turning to musical stimuli, Palmer et al. (2013) found music-color correspondences using classical orchestral pieces. In a recent study, Albertazzi et al. (2020) even found robust crossmodal associations between paintings by Kandinsky and music by Schönberg (for a narrative historical review of crossmodal correspondences between color and sound, see Spence and Di Stefano, 2022).

Spence (2011) proposed three different types of crossmodal correspondences: structural, statistical, and semantic. Structural correspondences are based on common neural encoding across the senses. Statistical correspondences come from the statistical regularities of the multisensory environment such as, for example, the physical correlation between pitch and size. Semantic correspondences are based on a common vocabulary describing stimuli in different sensory modalities, as in the use of "sweet" to describe music and taste (Mesz et al., 2011). Spence (2011) further points out that crossmodal correspondences between features of stimuli can also be established based on the emotional effects these stimuli have on an observer. Such emotionally-mediated correspondences are thought to be based on the matching of similar emotions or hedonic valence related to each of the associated stimuli (Palmer et al., 2013).

Stimuli that often give rise to (strong) emotional responses such as music are therefore likely to give rise to emotionally-mediated crossmodal correspondences. In fact, music-color and music-painting correspondences can often be predicted by the emotional ratings of the stimuli involved (Spence, 2020). And color has been shown to influence music-induced emotions. For instance, in a study by Hauck et al. (2022), judgments on the emotional impact of musical pieces changed in accordance to the emotions attributed to colored lighting. This mechanism of "emotion transfer" has been found also between other senses. For example, the experience of drinking coffee while listening to music was largely determined by the emotional effect of the music that was playing at that moment (Galmarini et al., 2021). Another study showed that pleasant sounds enhanced odor pleasantness (Seo and Hummel, 2011). Below, we review previous findings on musical emotions, the emotional responses elicited by visual shapes and materials and their crossmodal correspondences with music features, and introduce the research hypotheses that shaped the rationale of the present work.

1.1. Theoretical framework

Our aim was to investigate crossmodal correspondences between (a) music-induced emotions and freely drawn visual shapes representing those emotions, and between (b) music-induced emotions and materials associated with them (such as those employed, for instance, in concert halls, furniture, or present in natural environments). A further aim was to predict music-induced emotions from musical excerpts, visual shapes, and materials matched with those emotions.

1.1.1. Music-induced emotions

Music can arouse a wide range of powerful emotions in a listener (Juslin, 2019), and two well-known scales have been developed in order to discretely model music-induced emotions. GEMS – The Geneva Emotional Musical Scale - is a model and instrument specifically designed to capture emotions evoked by music (Trost et al., 2012). GEMS comprises nine categories of musical emotions (wonder, transcendence, tenderness, nostalgia, peacefulness, energy, joyful activation, tension, and sadness). More recently, Coutinho and Scherer (2017) introduced the GEneva Music-Induced Affect Checklist (GEMIAC), as a brief instrument for analysis of music-induced emotions. GEMIAC was designed to extend and complement the GEMS, across intensity and affective responses to music.

Several researchers also propose a more parsimonious model for music-induced emotions, suggesting that music essentially communicates two dimensions of core affect: valence and arousal (Eerola and Vuoskoski, 2011; Flaig and Large, 2014; Cespedes-Guevara and Eerola, 2018). Both of these dimensions have emerged during experiments involving crossmodal correspondences between music and visual features (Palmer et al., 2013), as well as correspondences of music with other sensory modalities, for example between music and taste/flavor (Reinoso-Carvalho et al., 2019, 2020; Motoki et al., 2022). Importantly, valence and arousal (or activity) also represent two core dimensions of Osgood's semantic differential technique. Osgood et al. (1957) showed that most variation in individuals' connotative meanings of aesthetic stimuli can be explained by three dimensions: valence, activity, and potency. However, crossmodal correspondences between music and other (sensory) stimuli (e.g., colors or paintings) can often be accounted for by emotional mediation of specific emotions associated with both music and visuals (for a more detailed discussion, see Spence, 2020). The question then remains whether an emotional mediation account (using the GEMIAC instrument) has more explanatory power than the semantic differential technique (with its core dimensions of valence and arousal) when induced or felt (as opposed to perceived or associated) emotions are concerned.

1.1.2. Emotions evoked by visual shapes

Reliable associations between perceived emotions and visual shapes have been documented in a number of studies, where shapes were either selected from a repertory or freely drawn (Poffenberger and Barrows, 1924; Karwoski et al., 1942; Lyman, 1979; Collier, 1996). Three of the most studied qualitative properties of visual shape, in relation with their emotional effects, are curvature, symmetry, and complexity. Visual shape curvature or roundness has been extensively assessed in connection with the "curvature effect," the preference for curved over sharp-angled contours (for a review, see Corradi and Munar, 2020). Vartanian et al. (2013) combine behavioral and neural evidence to show that this effect is probably driven by pleasantness. With respect to emotional arousal induced by visual contours, Blazhenkova and Kumar (2018) found curved and angular shapes to be associated with relieved and excited/surprised emotions, respectively.

Visual symmetry is a property traditionally associated with beauty and aesthetic preference (Weyl, 2015; Weichselbaum et al., 2018; also cf. Leder et al., 2019). Bertamini et al. (2013) reported that aesthetic responses to visual patterns with reflectional symmetry involved both positive valence and high arousal. Salgado-Montejo et al. (2015) also demonstrated preferential matching of symmetry (asymmetry) of visual shapes with the word "pleasant" ("unpleasant"). Pleasantness has also been shown to mediate associations between the (a)symmetry of visual shapes and taste (Turoman et al., 2018).

Ratings of visual shape complexity have been shown to depend on diverse sources as well, such as the number of sides or turns in the shape, asymmetry, compactness, degree of self-similarity, or shape skeletons based on local axes of symmetry (Sun and Firestone, 2021). Physiological arousal has been shown to increase with complexity, while pleasantness has been found to approximate a Wundt curve or inverted "U" shape, that is, to increase with complexity up to a certain point but then decrease for highly complex shapes (Berlyne, 1970; Madan et al., 2018).

1.1.3. Emotions evoked by materials

Compared to music and other visual features (e.g., shapes), materials seem to have a weaker capacity to evoke emotions (Crippa et al., 2012), with a tendency to neutrality with respect to valence (Marschallek et al., 2021). Nevertheless, some studies on aesthetics of materials have considered their capacity to evoke emotions, as well as analyzing the associated set of elicited emotions. Crippa et al. (2012), for instance, asked individuals to assess the emotions elicited by 9 different materials. They found that, in general, emotions evoked by materials were rather weak, with the most frequent emotions reported being satisfaction, joy, fascination, surprise, dissatisfaction and boredom. For a study investigating the emotions associated with a range of different materials (cotton, satin, tinfoil, sandpaper, and abrasive sponge), see Etzi et al. (2016).

1.1.4. Crossmodal correspondences between sound/music and visual shapes/materials

When it comes to crossmodal correspondences between sound/music and visual shapes/materials, previous studies have shown that pitch height is consistently associated with visual shape (for a summary, see Küssner, 2017). Lower pitch tends to be congruently framed with curvy shapes, while higher pitch is matched with sharper angular shapes (Melara and O'Brien, 1987). Parise and Spence (2012) also found that tones with a sinusoidal waveform were associated with a curvy shape, while tones with a square waveform were more associated with a jagged one (i.e., a kind of bouba-kiki effect; Reinoso Carvalho et al., 2017). In Fernay et al. (2012), a synesthete and 10 control participants were asked to draw a shape for different vowel sounds and to define two colors, together with a vertical and horizontal position for it. Control participants showed crossmodal correspondences agreeing with the synesthete's perceptions, such that vertical position and color brightness increased as pitch increased (see also Salgado-Montejo et al., 2016 and Küssner et al., 2014, for pitch-space associations during free hand movement tasks in a two- and three-dimensional space, respectively). Interestingly, the speaker's gender influenced the way participants set the size of the shape as well (i.e., larger shapes for a male voice). The influence of individuals' background and training on forming crossmodal correspondences between music/sound and visual shapes is exemplified in Küssner (2013) and Küssner and Leech-Wilkinson (2014). For instance, it was shown that musically untrained participants produce more diverse visual shapes than musically trained participants when asked to draw visual representations of a series of sine tones and short musical excerpts. Notably, the most complex and asymmetrical visual shapes associated with the auditory stimuli were produced by a dancer without musical background (Küssner, 2013). Adeli et al. (2014) reported that softer musical timbres were associated with blue, green, or light gray rounded shapes, while harsher timbres were matched with red, yellow, or dark gray sharper angular shapes. In their study, timbres involving elements of both softness and harshness were associated with a mixture of the two kinds of visual shapes being analyzed.

Crossmodal correspondences between music and materials have been much less studied, which is surprising in view of the importance of materiality for instrumental timbre. In a recent study on the timbre semantics of Western orchestral instruments, Wallmark (2019) reported that the domain of musical timbre is often conceptualized by terms that are applied also to properties of materials (i.e., soft, dry, hard, metallic, smooth), as well as to visual shapes (i.e., smooth, round, open, sharp). In Murari et al. (2015), non-verbal sensory scales for qualitative music description are proposed. Among other scales, the authors used wood, polystyrene, and sandpaper samples of different roughness to represent the qualities of hard vs. soft and smooth vs. rough, finding that their participants were consistent in their ratings of musical excerpts with respect to smoothness/roughness (but not to softness/hardness).

1.2. Hypotheses

Based on the body of research reviewed above, we formulated the following hypotheses focusing on specific properties of crossmodal correspondences, as well as on the predictability[1] of the GEMIAC emotions and core affect from the associated musical excerpts, visual shape features, and materials.

1. Common emotional associations with stimuli in different senses have often been shown to underlie crossmodal correspondences, particularly those involving music (emotionally-mediated correspondences: Spence, 2020). In other words, when music is associated with a non-auditory object or feature, both are often perceived to convey the same emotion, enabling one to infer musical emotions from those elicited by non-musical stimuli. While musical emotions have been shown to be predicted by perceptual musical features alone, such as pitch, tempo, mode, or dynamics (Lange and Frieler, 2018), we hypothesize that music-induced emotions may be recovered more efficiently from joint information on the musical excerpts, visual shapes, and materials that have been associated with them:

1 Predictability relates to the ability to use the information contained in the explanatory variables in order to infer what the corresponding emotions would be (future experiments) or have been (past experiments where the individual may not have disclosed the emotion). The hypothesis is that these associations between the covariates and the emotions are sufficiently consistent within and between individuals so that emotions are predictable if only the covariates are observed. However, such associations can potentially be substantially complex and non-linear, where covariates combine with each other and with themselves in complex ways, so a complex machine learning algorithm is needed to mimic how the brain operates linking those covariates in complex forms with the associated emotions.

> *H1*: Music-induced emotions can be predicted by the joint information of musical excerpt and features of visual shape and materials arising in crossmodal correspondences with the emotions induced by this excerpt.

2. Core affect is a universal and ubiquitous basic aspect of subjective emotional experience, capable of describing the affective connotations of percepts in different sensory modalities (Collier, 1996; Russell, 2003). However, the GEMIAC emotion pairs have been selected because of their specific relevance to describe musical emotions; in consequence, some of them, such as "joyful, wanting to dance" and "enchanted, in awe" might not be easily applicable to the visual shapes and usual design materials considered in this study. Moreover, the possibility of choosing between several similar GEMIAC emotions may diminish the predictive relevance of each individual one. Due to its universality, we hypothesize that we will obtain more accurate predictability of core affect than of discrete music-specific emotions such as those solely relying on GEMIAC:

> *H2*: Valence and arousal of core affect induced by music will be predicted more accurately by the associated musical excerpts, visual shapes, and materials, compared to discrete GEMIAC emotions.

3. Materials seem to have a weaker capacity for evoking emotions, when compared to music, other visual features, and even other sensory modalities, such as taste and olfaction (i.e., Crippa et al., 2012). Emotions evoked by materials seem to be rather weak, with a tendency toward neutrality with respect to valence (Marschallek et al., 2021). Nevertheless, some studies on aesthetics of materials have considered their capacity to evoke emotions and to elicit crossmodal associations (Barbosa Escobar et al., 2023), as well as analyzing the associated set of elicited emotions. Some of these results suggest that materials may play a role for the emotional experience. However, and based on the above, we hypothesize that visual shapes are still better predictors of music-induced emotion than any specific material. We also hypothesize that a musical excerpt is an even better predictor of music-induced emotions, compared to visual shapes and materials crossmodally associated to these emotions:

> *H3*: Materials by themselves will be poorer predictors of music-induced emotions and their valence and arousal than visual shapes. In turn, musical excerpts will be better predictors of music-induced emotions than visual shapes and materials.

2. Materials and methods

2.1. Participants

A total of 176 participants completed the experiment, 23 in English (12 female, 11 male) and 153 in Spanish (75 female, 77 male, 1 other), with a mean age of 33.20 years (SD = 11.02) and age range of 19–56 years. Participants resided in 10 countries (Germany, Argentina, Colombia, Australia, Italy, United Kingdom, United States, India, France, and Belgium) and were recruited by means of convenience sampling among the authors' networks. None of the participants reported auditory or visual limitations. In order to assess the participants' level of musical training/sophistication, we applied the single item measure described by Zhang and Schubert (2019). The resulting self-evaluation percentages were: nonmusicians, 19%; music-loving nonmusicians, 45%; amateur musicians, 18%; serious amateur musicians, 7%; semiprofessional musicians, 5%; and professional musicians, 6%. Thus, participants had a wide range of musical competence, with a majority of non-musicians. A participant information and consent form were built into the first page of the study, and all participants gave their informed consent before proceeding to the study itself.

2.2. Musical stimuli

Eight excerpts from the Eerola and Vuoskoski dataset of emotional film music (Eerola and Vuoskoski, 2011) were used as stimulus material. These excerpts have been consistently rated in valence and arousal and classified across discrete emotions *expressed* by the music and were shown to elicit a variety of induced emotions when modeled by the GEMS (Vuoskoski and Eerola, 2011), of which GEMIAC is an extension. Importantly, cluster analysis showed that low-level clusters of ratings of emotional excerpts were the same for the GEMS and discrete models (namely, four clusters corresponding respectively, in the discrete model, to Scary, Happy, Sad and Tender emotions; Vuoskoski and Eerola, 2011). Specifically, for our study, we selected two *Scary*, two *Happy*, two *Sad* and two *Tender* musical excerpts from those used in the aforementioned study: "Oliver Twist" and "Dances with Wolves" (named here *Happy 1* and *Happy 2*, respectively), "The English Patient" and "Running Scared" (*Sad 1* and *Sad 2*, respectively), "The Portrait of a Lady" and "Shine" (*Tender 1* and *Tender 2*, respectively), and "The Alien Trilogy" and "Batman Returns" (*Scary 1* and *Scary 2*, respectively). The duration of these musical excerpts varied between 46 and 72 s.

2.3. GEMIAC

The GEneva Music-Induced Affect Checklist (GEMIAC) comprises 14 classes of musical emotions, denoted by term pairs, such as "*melancholic, sad*" and "*moved, touched*" (Coutinho and Scherer, 2017). In the usage proposed by its authors, it is required to rate the experienced intensity of each emotion class after listening to a piece of music. Instead, here we asked participants to choose the better matching emotion class in response to each musical excerpt. We did so in order to later obtain a single representation of the (most characteristic) music-induced emotion for each excerpt, both as visual shape and material from a list (see Procedure). For the Spanish translation of the GEMIAC, we followed the methodology proposed by Vallerand (1989).

2.4. Procedure

The experiment was designed in Qualtrics and conducted online, either in English or in Spanish depending on the respondent's language preferences. Informed consent was a prerequisite for taking part in the study. Participants were asked to use headphones at all times.

First, participants listened to a sample audio (not included in the set of stimuli) to adjust the sound volume to a comfortable level. Second, as the main task, the eight excerpts were presented in a randomized order. After listening to each excerpt, participants were asked to choose a single term pair from the GEMIAC scale to describe their induced emotion. The precise instruction for the participant was as follows: "*Please indicate which of the following term pairs best describes the emotion you experienced when listening to this audio. DO NOT DESCRIBE the music* (*Example: "this music is melancholic, sad"*) *or what the music seemed to express* (*Example: "this music expresses joy"*). *Describe YOUR OWN EMOTION while listening to the music* (i.e., *"I feel melancholic/sad while listening to this music"*). *If you consider that your emotion does not correspond to any of the term pairs, choose from the list the term pair closest to the emotion you experienced*."

Having selected a GEMIAC term, they were further asked to draw their emotion ("*Please draw a CLOSED SHAPE that represents THE EMOTION you experienced while listening to the music. You can erase the drawing as many times as you like*"). To record these shapes, we used the Signature feature provided by Qualtrics, which allows drawing using the mouse or touchpad of the computer.

As a final step, participants were asked to select a material that they thought would match the corresponding induced emotion. The list of materials, included in Table 1, was taken from Keyshot, a 3D rendering software used by designers (Jo, 2012). The instruction was: "*Please select from the given list of materials the material that you think best corresponds with the previously chosen emotion*."

2.5. Shape features

The visual shapes reported by participants mentioned in the paragraph above were evaluated independently by two raters who did not participate in the study or its design. Using graphic sliders with a 0–100 range, they rated the three shape features described in Section 1.1.2. in scales for symmetry, either reflectional or rotational (from very asymmetrical to very symmetrical), roundness (from no roundness to high roundness), and complexity (from very simple to very complex). The raters had been previously instructed about 2D symmetry and the notion of curvature. Complexity was left as an intuitive notion.

2.6. Data

A total of $N = 1{,}408$ responses were gathered from 176 participants. Covariates include the musical excerpt, shape characteristics, and materials, though the interest resides in the latter two upon adjusting for the former. Shape characteristics were assessed based on symmetry, roundness, and complexity and mapped to a 0–100 scale, while the musical excerpt and materials were kept as categorical explanatory variables.

2.7. Statistical analysis

The software used for statistical analysis was R, version 4.2.1. Initially, we conducted an analysis of the predictive quality of the covariates separately for each GEMIAC emotion class. Covariate relevance was ranked using the Decreased Gini score metric (Breiman, 2001). Predictive quality was explored by calculating the area under the curve (AUC) for each GEMIAC class independently, and 95% confidence intervals for the AUC were also reported. Values above 0.80 are oftentimes considered to represent an excellent classifier (Hosmer et al., 2013), with an upper bound of 1 representing perfect classification. Several emotion classes stood out with high AUC: "tense, uneasy," "powerful, strong" "energetic, lively" and "melancholic, sad." However, other emotion pairs had relatively low AUC predictability.

Therefore, emotions were grouped according to five relatively homogeneous valence-arousal pairs, as described in Figure 1, which form the categories within the primary analysis. These pairs represent the spectrum of emotions anticipated to be experienced by respondents, so that intra-group emotions share similarities (and represent common traits in the underlying emotions), but inter-group differences in emotion type are more clearly identifiable by the respondent (and represent more relevant differences in the underlying emotions).

Summary statistics were produced for the covariates as well as both the ungrouped and grouped emotions. A derivation cohort was defined using 75% of the data, and the remainder of the sample was used as a validation (out-of-sample) cohort. Since the associations between the covariates (musical excerpt, shapes, and materials) and the responses (pairs of valence-arousal emotions) were expected to

TABLE 1 Summary statistics of covariates.

Covariates		Code	N/mean	%/sd
Materials	Ceramic	X1	62	4.40%
	Glossy plastic	X2	9	0.64%
	Grained wood	X3	99	7.03%
	Granite	X4	89	6.32%
	Leather	X5	6	3.27%
	Fine grain wood	X6	55	3.91%
	Marble	X7	103	7.32%
	Nylon	X8	57	4.05%
	Opaque glass	X9	52	3.69%
	Opaque liquid	X10	68	4.83%
	Opaque metal	X11	97	6.89%
	Opaque plastic	X12	34	2.41%
	Porcelain	X13	87	6.18%
	Rubber	X14	46	3.27%
	Shiny metal	X15	119	8.45%
	Shiny plastic	X16	43	3.05%
	Transparent glass	X17	80	5.68%
	Transparent liquid	X18	89	6.32%
	Transparent plastic	X19	21	1.49%
	Transparent stone	X20	45	3.20%
	Velvet	X21	107	7.60%
Shapes	Complexity	C	20.01	11.99
	Roundness	R	40.83	12.00
	Symmetry	S	38.67	22.41

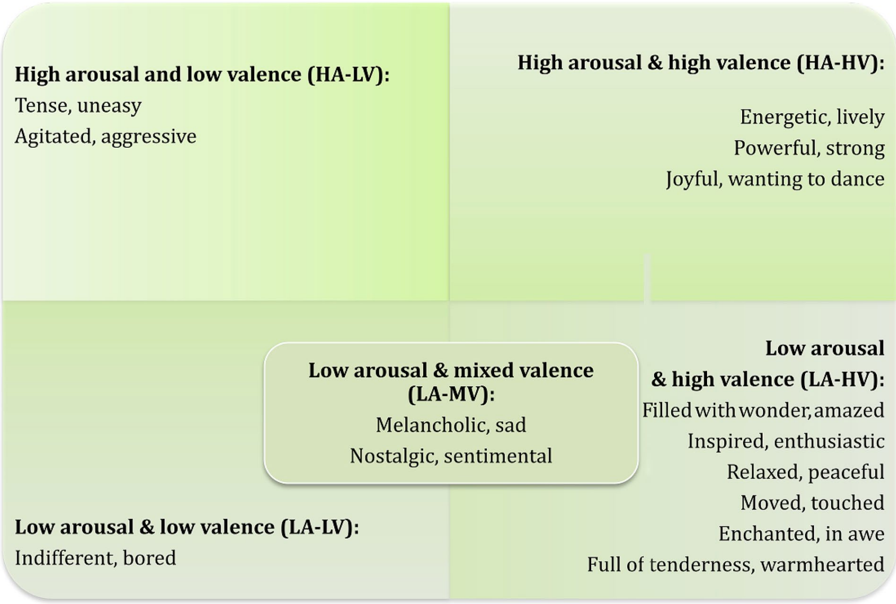

FIGURE 1
GEMIAC emotion terms grouped by arousal category (rows) and valence category (columns).

have a complex and non-linear form, in line with the complexity associated with human information processing and decision-making, a machine learning-based method, random forest, was fit for the derivation cohort.[2] The fitted random forest was subsequently used to demonstrate the joint out-of-sample predictive power of the covariates using solely the information in the validation cohort. In order to maintain full out-of-sample validity of the study, a participant effect was not included, though in practice the approach can be enhanced by including any known participant characteristics or prior responses if those are available prior to the experiment.

Predictive quality was explored by calculating the AUC for each valence-arousal pair independently, and 95% confidence intervals for the AUC were also reported. In order to understand whether core emotion features (arousal or valence) were independently associated with the covariates, two sensitivity analyses were performed: (1) collapsing valence across arousal categories, resulting in two broader categories for emotions (low & high arousal); and (2) collapsing arousal across valence categories, resulting in three broader categories for emotions (low, mixed, and high valence).

3. Results

Table 1 contains the summary statistics for the covariates, and Table 2 contains both the ungrouped and grouped emotions, where grouping of emotions is depicted in Figure 1. The majority of respondents selected low arousal (61.51%) and high valence (59.66%) emotions, with the combination of both categories representing the majority of responses (39.77%). The category with least responses corresponds to the combination of low valence and low arousal, which was selected by respondents in 4.29% of the experiments. Among materials, the most common ones were shiny metal (8.45%) and velvet (7.60%), while glossy plastic was the least common (0.64%). The metrics representing complexity and symmetry are right-skewed, while the distribution of roundness is more symmetric, as depicted in Figure 2. Complexity and symmetry were negatively correlated ($r = -0.21$; 95% CI –0.26, −0.16), while complexity and roundness were not found to be associated ($r = 0.02$; 95% CI –0.03, 0.07). Conversely, roundness and symmetry were found to be positively associated ($r = 0.23$; 95% CI 0.18, 0.28).

Figure 3 includes a graphical representation of the random forest by mean decreased Gini scores in the derivation cohort for the overall analysis (by arousal-valence pair) as well as the two sensitivity analyses (grouping valence by arousal and arousal by valence). Variables are ranked by contribution to node homogeneity. Upon controlling for the musical excerpt, shape variables provided a higher contribution to enhanced node homogeneity than materials, which aligns with hypothesis H3.

Figure 4 portrays the out-of-sample predictive/classification power (AUC) for each emotion category across the primary and sensitivity analyses for the validation cohort. Our results demonstrate a high level of predictive power across all emotions except for the low arousal-low valence combination, as seen in the left panel of Figure 4. This category could represent a basket choice of very heterogeneous emotions, such as true low arousal-low valence emotions or simply exhaustion and lack of interest in the experiment from participants. Also, this category contains the fewest number of data points (59 responses across both validation and derivation datasets, or 4.19%), which affects the ability of the random forest algorithm to learn

2 For another innovative analysis approach using non-linear models (a Gaussian process regression model with a linear plus squared-exponential covariance function) to study crossmodal associations between music/sound and visual shapes, see Noyce et al. (2013).

TABLE 2 Summary statistics of emotions: both ungrouped and grouped by valence and arousal category and further collapsed across valence and arousal categories.

Responses	*N*	%
Agitated, aggressive	50	3.55%
Enchanted, in awe	73	5.18%
Energetic, lively	107	7.60%
Filled with wonder, amazed	77	5.47%
Full of tenderness, warmhearted	46	3.27%
Indifferent, bored	59	4.19%
Inspired, enthusiastic	154	10.94%
Joyful, wanting to dance	49	3.48%
Melancholic, sad	122	8.66%
Moved, touched	96	6.82%
Nostalgic, sentimental	125	8.88%
Powerful, strong	124	8.81%
Relaxed, peaceful	114	8.10%
Tense, uneasy	212	15.06%
Low arousal, low valence (LA-LV)	59	4.19%
Low arousal, mixed valence (LA-MV)	247	17.54%
Low arousal, high valence (LA-HV)	560	39.73%
High arousal, low valence (HA-LV)	262	18.61%
High arousal, high valence (HA-HV)	280	19.89%
Low valence (LV)	321	22.80%
Mixed valence (MV)	247	17.54%
High valence (HV)	840	59.66%
Low arousal (LA)	866	61.51%
High arousal (HA)	542	38.49%

efficiently about this category. For all other categories in the primary analysis, the approach demonstrates exceptional out-of-sample predictive power within the validation cohort, with AUC values above 0.85. When collapsing by arousal or valence, the predictive power across all categories remains excellent, with all AUC point estimates between 0.89 and 0.95, as demonstrated in the middle and right panels of Figure 4. This indicates that levels of both arousal and valence can be identified independently and jointly.

In Figure 5, the associations between the covariates and the outcome variable from the primary analysis are (univariately) descriptively visualized. For example, opaque glass (Material 9 as per the ordering in Table 1) was associated with a lower observed frequency of high arousal, high valence emotions than materials such as glossy plastic (Material 2) or shiny metal (Material 15). Similarly, low arousal and low valence were negligible for shiny metal (Material 15), while they were observed in higher proportion for transparent plastic (Material 19). When grouping observations across valence (with HA-HV plus HA-LV constituting the high arousal broader category), we observed that high arousal is largest among observations of opaque metal and leather (Materials 11 and 5, respectively), while low arousal categories prevail among observations of velvet and porcelain (Materials 21 and 13, respectively). Figure 6A demonstrates, for example, how low arousal categories are associated with smaller levels of complexity than high arousal categories.

4. Discussion

The observed crossmodal correspondences between music-induced emotions and visual shapes/materials demonstrate some level of heterogeneity in emotional response univariately by covariate (Figures 5, 6). Consistently with previous findings, we observed an increase of arousal with shape complexity (Figure 6A), whereas our results on associations between materials and emotions appear to be new. The strong explanatory power demonstrated in Figure 4 comes from more complex, non-linear multivariate associations between the covariates, such as those extracted by the random forest approach. This indicates that emotional responses cannot be simply explained by compounding one-dimensional associations, and a more complex structure is needed, in line with how individuals process information. This predictive power is in agreement with our hypothesis H1.

We obtained further evidence for hypothesis H1 in the predictive analysis of the full 14 GEMIAC emotion pairs. Some of these emotions showed high AUC: "tense, uneasy" (AUC = 0.85), "powerful, strong" (AUC = 0.74), "energetic, lively" (AUC = 0.78) and "melancholic, sad" (AUC = 0.79). As such, we obtained good predictability for some highly specific musical emotion classes (Palmer et al., 2016). It could be that the names of these classes have more multimodal applicability, for instance, "sad" can be applied to shapes (Sievers et al., 2019), music, and materials (Zuo et al., 2001).

However, overall, the mixed performance of the random forest algorithm for music-specific emotions is in agreement with hypothesis H2, that is, there was less effectiveness in predicting specific GEMIAC emotion classes compared to core affect. This may have been due in part to the similarity of some of those emotion classes, such as "enchanted, in awe" (AUC = 0.56) and "filled with wonder, amazed" (AUC = 0.61), or to their low representativity for our excerpts, such as "full of tenderness, warmhearted," amounting to only 3.27% of the responses (AUC = 0.59). These results can be compared to those of Meuleman and Scherer (2013), where predictive accuracy decreased as the number of clusters of emotions increased from two (positive vs. negative emotions) to twelve. Moreover, our results can be linked to the traditional Osgood semantic differential technique (SDT) which aims to measure connotative meanings of (aesthetic) objects, concepts or events (Osgood et al., 1957). While empirical studies of music in the tradition of Osgood's SDT approach deal with associative, connotative meanings of music (for overviews, see Schubert and Fabian, 2006 and Spence, 2020), here we show that emotions *felt* by the listener can be predicted better when they are mapped onto core dimensions of SDT (i.e., valence and arousal) compared to the specific GEMIAC emotion classes.

Interestingly, in order to obtain this level of prediction efficiency for music-induced core affect, it was important to categorize the terms "melancholic, sad" and "nostalgic, sentimental"—which would seem to be low valence, low arousal (LV-LA) emotions—in a different valence-arousal group. These emotions—sometimes framed as "negative"—have been considered, nonetheless, as emotions that make a positive contribution to aesthetic liking. People may enjoy feelings of nostalgia and melancholy when listening to sad music,

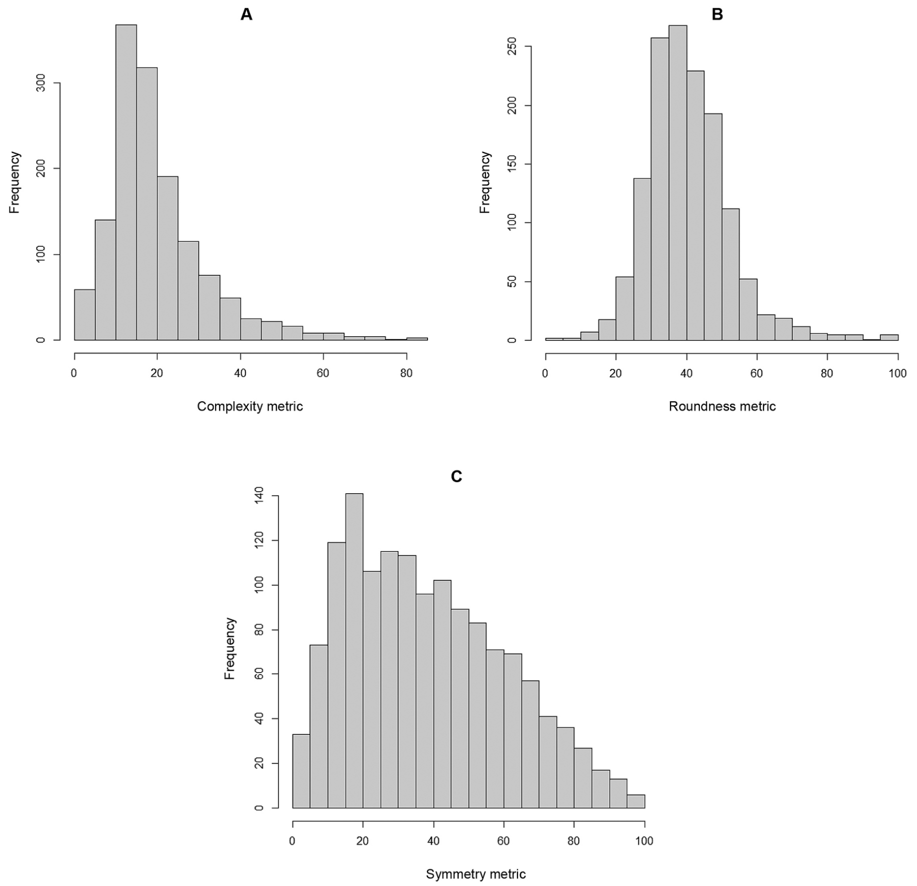

FIGURE 2
Distributions of the shape metrics **(A)** complexity, **(B)** roundness, and **(C)** symmetry across musical excerpts and participants.

which can evoke not only sadness, but also a wide range of complex and partially positive emotions, such as peacefulness, tenderness, transcendence, and wonder (Taruffi and Koelsch, 2014). In particular, nostalgia has been conceptualized variously as a negative, ambivalent, or positive emotion (Sedikides et al., 2004). Consequently, we have categorized these terms in a separate mixed valence, low arousal (MV-LA) group. In contrast with its excellent performance when we take this group into account, running the random forest algorithm instead with "melancholic, sad" and "nostalgic, sentimental" grouped as LV-LA (so considering 4 emotion clusters instead of 5), we obtained that these emotions were predicted worse upon grouping them when compared to predicting each individual one. In fact, with this 4-class grouping, LV-LA emotions had AUC = 0.56, while in the full GEMIAC analysis "melancholic, sad" had AUC = 0.79 and "nostalgic, sentimental" had AUC = 0.67. We noted also that HV-HA, HV-LA and LV-HA emotions were predicted worse in the 4-class than in the 5-class grouping, having AUC = 0.62, 0.62, and 0.78, respectively, in the former case.

In accordance with our hypothesis H3, materials by themselves were less relevant than visual shapes for predicting core affect (Figure 3), and the musical excerpts were the most relevant predictors. This lower predictive power of materials can be partially attributed to the diversity of options presented. We also expected this on the basis of the relative emotional neutrality of materials (Marschallek et al., 2021), which would prevent them from capturing emotional connotations. However, despite this supposed neutrality, we observed distinctive, nonrandom distributions of core affect associated with each material (Figure 5). Interestingly, in a study comparing prediction of discrete music-induced emotions using perceptual musical features (such as pitch, tempo, mode, dynamics) vs. crossmodal associations (warm, cold, rough, smooth, dark, bright), models based on crossmodal features (tactile and visual) performed better than those based on perceptual features in four out of six emotions, suggesting that music-induced emotions may be captured more clearly and directly by extra-musical characteristics than by music-specific dimensions (Lange and Frieler, 2018).

Our results aligned with previous similar findings using random forests classifiers. In fact, these classifiers have been shown to improve predictive accuracy of emotions, with respect to other linear and nonlinear methods, in different contexts, involving both music-specific and more general emotions. For instance, Meuleman and Scherer (2013) found random forests to have the best performance, among 14 linear and nonlinear models, for predicting event-related emotions from appraisal criteria of events such as relevance, consequences, causes, or coping. Vempala and Russo (2018) studied predictability of musical emotion judgments from audio features and physiological signals with machine learning methods, finding that linear and nonlinear methods achieved similar prediction performance from audio features, while more flexible nonlinear models such

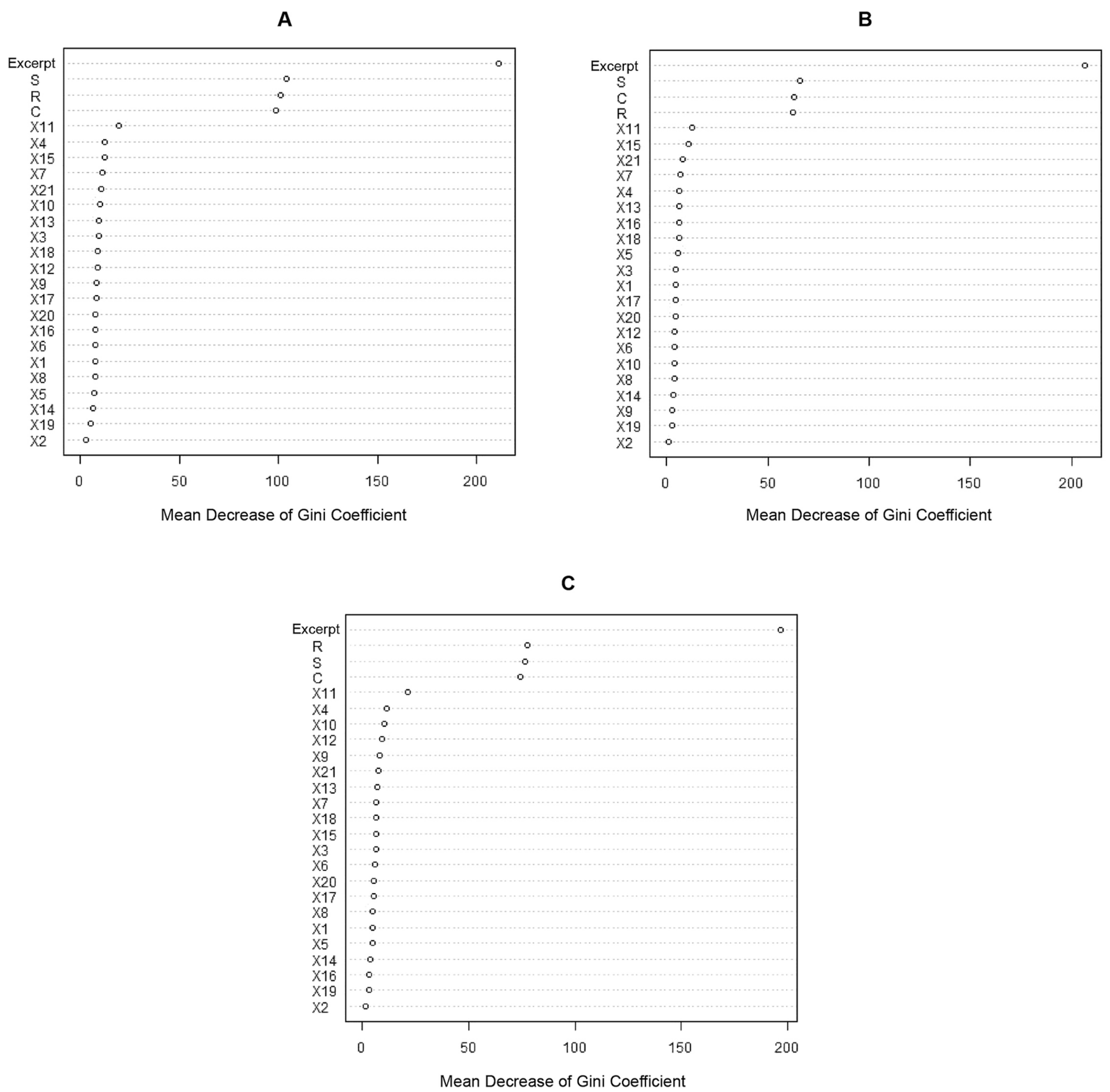

FIGURE 3
Mean decrease of Gini coefficient by covariate for the primary analysis **(A)**, and the two secondary analyses: **(B)** arousal grouped by valence and **(C)** valence grouped by arousal. C, complexity; R, roundness; S, symmetry. For all other codes (X1–X21), see Table 1.

as neural networks or random forests were needed to capture the predictive capacity of physiology features.

5. Limitations, applications, and future work

Some limitations need to be considered regarding the methodology and generalizability of our findings. For instance, we explicitly asked our participants to wear headphones at all times (see Section 2.4.). Since this was an online survey, we could not ensure they complied with this instruction. We also decided to allow our participants to start developing their associations related to the emotions induced by the music, first, via the selection of a GEMIAC term and second, via a drawing, while leaving the material association at the very end. We adopted such an order because materials are poor predictors of music-induced emotions (see H3). Nevertheless, a balanced order of these tasks could be explored in future research. Also, to obtain a single response in terms of materials and visual shapes, we asked participants to select the emotion they experienced by choosing a single item from the GEMIAC, while it is possible that they may have felt several ones, in different intensities. Moreover, the graphical interface to draw the visual shapes is designed for signatures and does not seem to allow a flexibility and ease of input equivalent to hand drawing. However, none of the participants complained or reported any issues using it.

There has been little research on the impact of the visual environment and its materiality on musical emotions. The rare

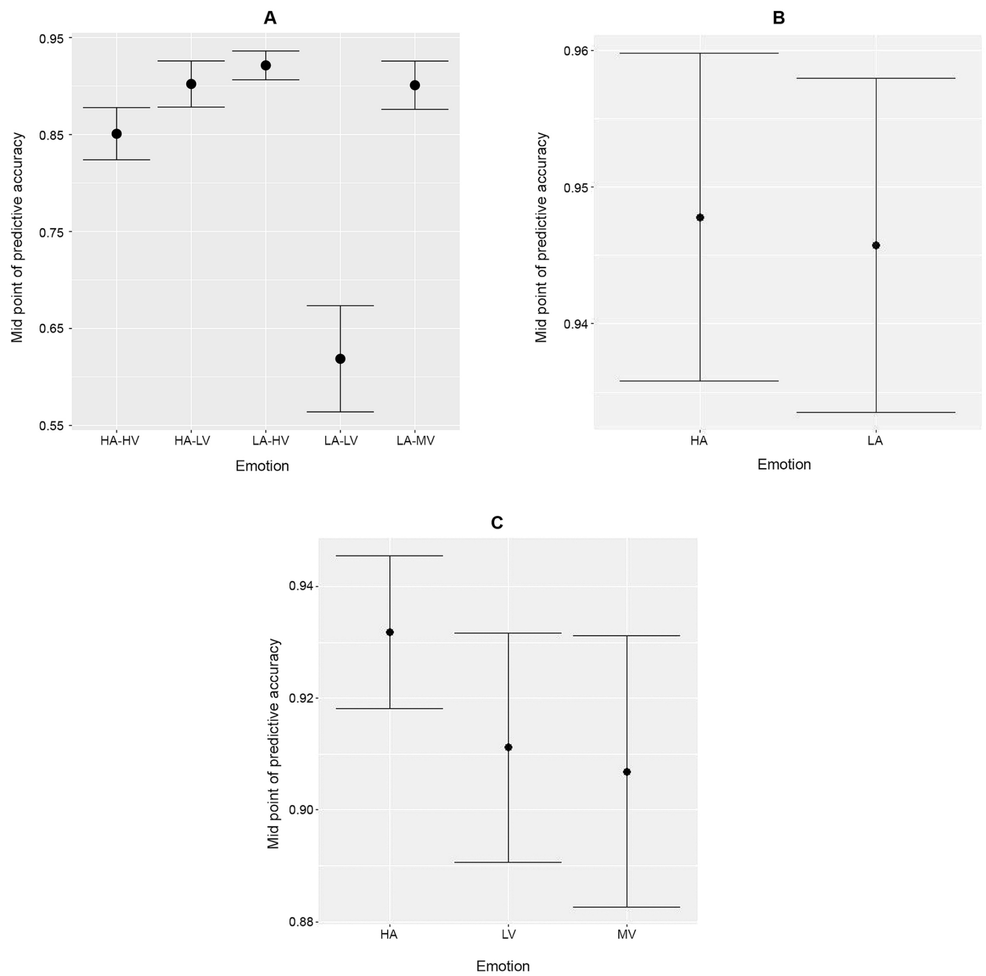

FIGURE 4
Out-of-sample AUC midpoint estimates and 95% confidence intervals for the validation cohort under the response categorizations in the primary analysis **(A)**, and the two secondary analyses: **(B)** arousal grouped by valence and **(C)** valence grouped by arousal. HA, high arousal; HV, high valence; LA, low arousal; LV, low valence; MV, mixed valence.

existing work has shown effects of videos on the emotional appraisal of music, as well as on the perception of musical features such as tempo and loudness (Boltz et al., 2009; Boltz, 2013). Another recent example is the study of Hauck et al. (2022) in which musical emotion ratings have been shown to be shifted by lighting conditions such as hue, brightness, and saturation. For example, a musical piece paired with red light was rated as more powerful than the same musical piece paired with green or blue light. In the same vein, given the crossmodal associations and predictability between visual shapes, materials, and musical emotions shown in our study, we would predict that emotions induced by a given musical piece are moderated by different environments, e.g., shown as projected images or in virtual reality, exhibiting various characteristics of materiality and visual shape design.

Thus, our findings suggest that visual shapes and materiality may be important factors to consider in sensory interactions where emotional synergy is sought between environment and music, for instance in hospitals, shopping centers, theaters and opera houses, art installations, and music therapy environments. In the field of art and aesthetics, our research may be extended to the analysis of works of visual music such as those of Oskar Fischinger and Norman McLaren (Gibson, 2023), as well as to the design of systems for visualizing music-induced emotions.

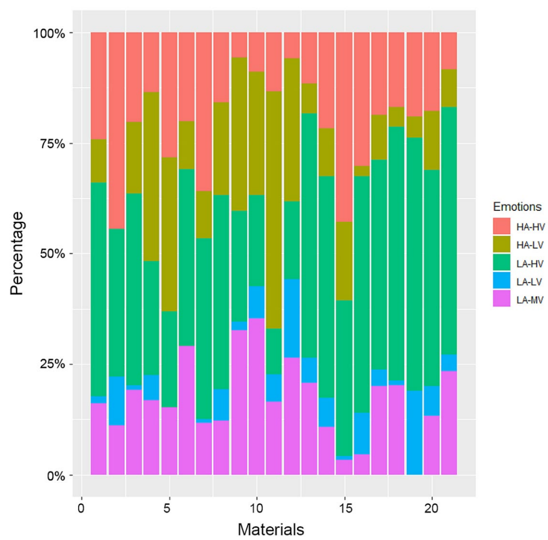

FIGURE 5
Observed proportions of emotion by material type (categorized as per Table 1). HA, high arousal; HV, high valence; LA, low arousal; LV, low valence; MV, mixed valence.

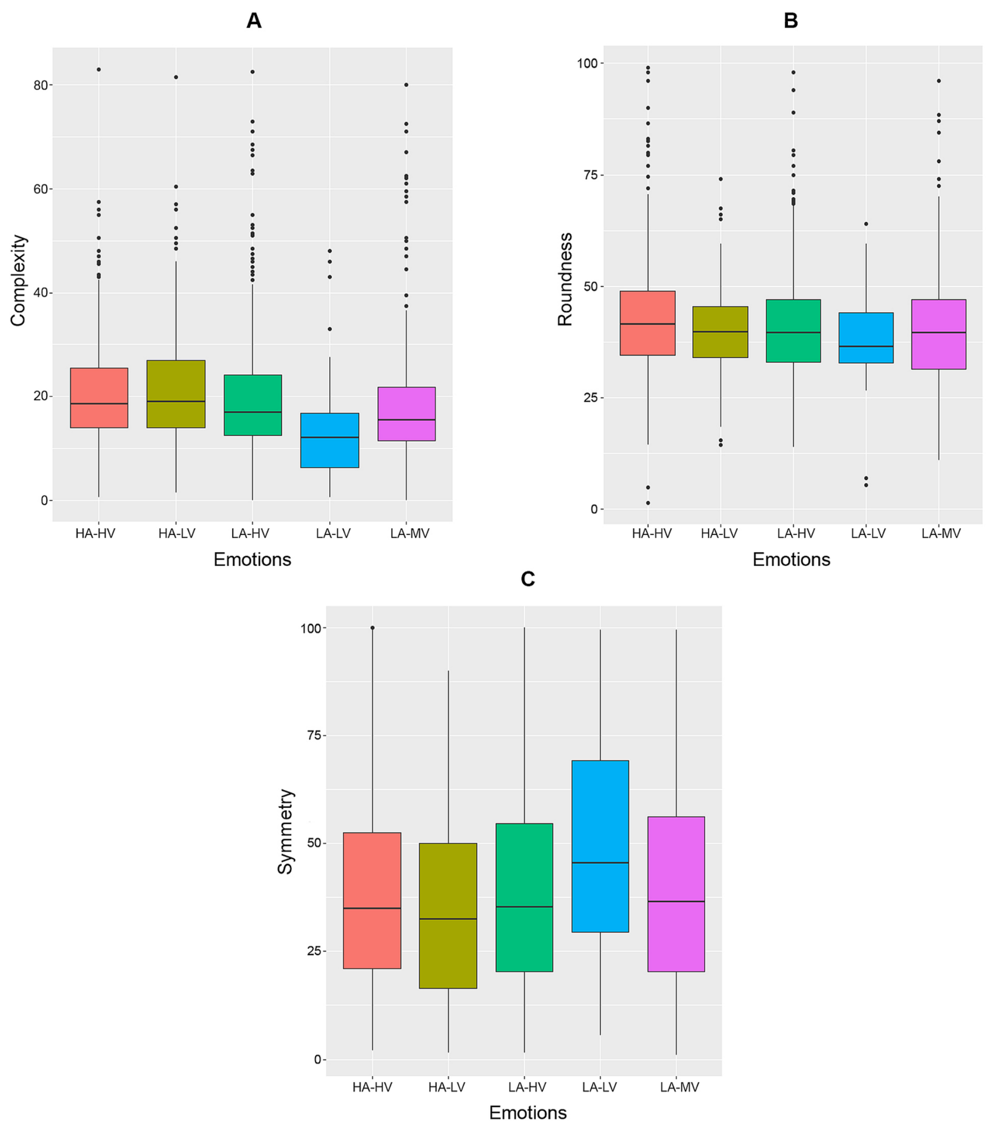

FIGURE 6
Boxplots of **(A)** complexity, **(B)** roundness, and **(C)** symmetry by emotion. HA, high arousal; HV, high valence; LA, low arousal; LV, low valence; MV, mixed valence.

Other applications may arise in sensory interactions design and multisensory experiences. For instance, in 2020, a scented visual installation entitled Emotional Plateware (Mesz and Tedesco, 2021) introduced digital conceptual designs of tableware intended to augment gastronomic experiences from a multisensory perspective. This installation was conceived from the results of a study on crossmodal correspondences of visual shapes, colors, smells, and materials with four emotions induced by music, which inspired both the shapes and materiality of the plateware and the visuals displayed in digital tablets embedded in it. Further applications may arise in the field of sensory marketing as well. For instance, it is well-known that store atmospherics affect consumer behavior (Spence et al., 2014). In the constant look for novelty, as well as more engagement from consumers, brands may rely on correspondences such as the ones assessed in this study (i.e., music, shapes, and materials) in order to augment the user's experience in retail from a multisensory perspective, either physically or digitally (Petit et al., 2019). Think of store atmospherics designed to evoke certain sensations and/or emotions congruently across music, visual shapes, and wall materials. We can also think of specific industries, such as fashion, where visual shapes and materials can be congruently associated with certain music in order to evoke certain emotions during the experience of consumers. Likewise, when thinking about the customer journey (e.g., Lemon and Verhoef, 2016), offline/online touch-point interactions may be reframed inspired by correspondences such as the ones being analyzed in this, and similar, work.

In conclusion, our results show the emergence of crossmodal correspondences between music-induced emotions and visual shapes and design materials, complementing previous findings on crossmodal correspondences between sound/music and their extra-musical elements (Küssner, 2013, 2017; Murari et al., 2015). We also show that visual shapes and materials capture and predict music-induced emotions, suggesting the possibility of designing congruent multimodal objects, environments, or messages oriented toward specific emotions, in spite of the interindividual diversity of "translations" across the senses (Spence, 2022). Future work should explore multisensory emotional design, combining music, visuals and materials.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Research Ethics Committee of Universidad Nacional de Tres de Febrero. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

BM, ST, and MK conceptualized the study. BM, ST, MK, and FR-C developed the methodology and experimental protocol. BM and FR-C supervised the data collection. EH, GM, and LG analyzed the data and provided a report of the results. BM wrote the first draft of the manuscript. All authors revised and agreed on the final version of the manuscript.

Funding

The article processing charge was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 491192747 and the Open Access Publication Fund of Humboldt-Universität zu Berlin. We also gratefully acknowledge the financial support provided by Universidad Nacional de Tres de Febrero.

Acknowledgments

We would like to express our deep gratitude for the help of our research students from the Universidad de Tres de Febrero: Camilo Alvarez, Leonardo Potenza, Victoria Balay, and Tomas Levita.

Conflict of interest

Author GM was employed by company Bayesian Solutions LLC, Charlotte, NC, United States.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adeli, M., Rouat, J., and Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Front. Hum. Neurosci.* 8:352. doi: 10.3389/fnhum.2014.00352

Albertazzi, L., Canal, L., Micciolo, R., and Hachen, I. (2020). Cross-modal perceptual organization in works of art. *i-Perception* 11, 1–22. doi: 10.1177/2041669520950750

Barbosa Escobar, F., Velasco, C., Byrne, D. V., and Wang, Q. J. (2023). Crossmodal associations between visual textures and temperature concepts. *Q. J. Exp. Psychol.* 76, 731–761. doi: 10.1177/17470218221096452

Belkin, K., Martin, R., Kemp, S. E., and Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological Science* 8, 340–342. doi: 10.1111/j.1467-9280.1997.tb00450.x

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Percept. Psychophys.* 8, 279–286. doi: 10.3758/BF03212593

Bertamini, M., Makin, A., and Rampone, G. (2013). Implicit association of symmetry with positive valence, high arousal and simplicity. *i-Perception* 4, 317–327. doi: 10.1068/i0601jw

Blazhenkova, O., and Kumar, M. M. (2018). Angular versus curved shapes: correspondences and emotional processing. *Perception* 47, 67–89. doi: 10.1177/0301006617731048

Boltz, M. (2013). "Music videos and visual influences on music perception and appreciation: should you want your MTV" in *The psychology of music in multimedia*. eds. S.-L. Tan, A. J. Cohen, S. D. Lipscomb and R. A. Kendall (New York: Oxford University Press), 217–234.

Boltz, M. G., Ebendorf, B., and Field, B. (2009). Audiovisual interactions: the impact of visual information on music perception and memory. *Music Percept.* 27, 43–59. doi: 10.1525/mp.2009.27.1.43

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Cespedes-Guevara, J., and Eerola, T. (2018). Music communicates affects, not basic emotions–a constructionist account of attribution of emotional meanings to music. *Front. Psychol.* 9:215. doi: 10.3389/fpsyg.2018.00215

Collier, G. L. (1996). Affective synesthesia: extracting emotion space from simple perceptual stimuli. *Motiv. Emot.* 20, 1–32. doi: 10.1007/BF02251005

Corradi, G., and Munar, E. (2020). "The curvature effect" in *The Oxford handbook of empirical aesthetics*. eds. M. Nadal and O. Vartanian (Oxford: Oxford University Press), 35–52.

Coutinho, E., and Scherer, K. R. (2017). Introducing the GEneva music-induced affect checklist (GEMIAC): a brief instrument for the rapid assessment of musically induced emotions. *Music Percept* 34, 371–386. doi: 10.1525/mp.2017.34.4.371

Crippa, G., Rognoli, V., and Levi, M. (2012). "Materials and emotions, a study on the relations between materials and emotions in industrial products" in *In 8th international conference on design & emotion: out of control* (London, England: Central Saint Martin's College of Arts & Design), 1–9.

Crisinel, A.-S., and Spence, C. (2012). A fruity note: crossmodal associations between odors and musical notes. *Chem. Senses* 37, 151–158. doi: 10.1093/chemse/bjr085

Eerola, T., and Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychol. Music* 39, 18–49. doi: 10.1177/0305735610362821

Etzi, R., Spence, C., Zampini, M., and Gallace, A. (2016). When sandpaper is 'kiki' and satin is 'bouba': an exploration of the associations between words, emotional states, and the tactile attributes of everyday materials. *Multisens. Res.* 29, –155. doi: 10.1163/22134808-00002497

Fernay, L., Reby, D., and Ward, J. (2012). Visualized voices: a case study of audio-visual synesthesia. *Neurocase* 18, 50–56. doi: 10.1080/13554794.2010.547863

Flaig, N. K., and Large, E. W. (2014). Dynamic musical communication of core affect. *Front. Psychol.* 5:72. doi: 10.3389/fpsyg.2014.00072

Galmarini, M. V., Paz, R. S., Choquehuanca, D. E., Zamora, M. C., and Mesz, B. (2021). Impact of music on the dynamic perception of coffee and evoked emotions evaluated by temporal dominance of sensations (TDS) and emotions (TDE). *Food Res. Int.* 150:110795. doi: 10.1016/j.foodres.2021.110795

Gibson, S. (2023). "Moving towards the performed image (colour organs, synesthesia and visual music): early modernism (1900–1955)" in *Live visuals: history, theory, practice*. eds. S. Gibson, S. Arisona, D. Leishman and A. Tanaka (London: Routledge), 41–61.

Hauck, P., Castell, C. von, and Hecht, H. (2022). Crossmodal correspondence between music and ambient color is mediated by emotion. *Multisens. Res.* 35, 407–446. doi: 10.1163/22134808-bja10077

Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013) *Applied logistic regression* (*3rd Edn.*). Hoboken, NJ: John Wiley and Sons.

Jo, J. L. (2012). *KeyShot 3D Rendering*. Birmingham: Packt Publishing Ltd.

Juslin, P. N. (2019). *Musical emotions explained: unlocking the secrets of musical affect*. New York: Oxford University Press.

Karwoski, T. F., Odbert, H. S., and Osgood, C. E. (1942). Studies in synesthetic thinking: II. The role of form in visual responses to music. *J. Gen. Psychol.* 26, 199–222. doi: 10.1080/00221309.1942.10545166

Küssner, M. B. (2013). Music and shape. *Lit. Linguist. Comput.* 28, 472–479. doi: 10.1093/llc/fqs071

Küssner, M. B. (2017). "Shape, drawing and gesture: empirical studies of cross-modality" in *Music and shape*. eds. D. Leech-Wilkinson and H. M. Prior (New York: Oxford University Press), 33–56.

Küssner, M. B., and Leech-Wilkinson, D. (2014). Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing paradigm. *Psychol. Music* 42, 448–469. doi: 10.1177/0305735613482022

Küssner, M. B., Tidhar, D., Prior, H. M., and Leech-Wilkinson, D. (2014). Musicians are more consistent: gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Front. Psychol.* 5:789. doi: 10.3389/fpsyg.2014.00789

Lange, E. B., and Frieler, K. (2018). Challenges and opportunities of predicting musical emotions with perceptual and automatized features. *Music Percept.* 36, 217–242. doi: 10.1525/mp.2018.36.2.217

Leder, H., Tinio, P. P., Brieber, D., Kröner, T., Jacobsen, T., and Rosenberg, R. (2019). Symmetry is not a universal law of beauty. *Empir. Stud. Arts* 37, 104–114. doi: 10.1177/0276237418777941

Lee, B. P., and Spence, C. (2022). Crossmodal correspondences between basic tastes and visual design features: a narrative historical review. *i-Perception* 13, 1–27. doi: 10.1177/20416695221127325

Lemon, K. N., and Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *J. Mark.* 80, 69–96. doi: 10.1509/jm.15.0420

Lyman, B. (1979). Representation of complex emotional and abstract meanings by simple forms. *Percept. Mot. Skills* 49, 839–842. doi: 10.2466/pms.1979.49.3.839

Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., and Sommer, T. (2018). Visual complexity and affect: ratings reflect more than meets the eye. *Front. Psychol.* 8:2368. doi: 10.3389/fpsyg.2017.02368

Marschallek, B. E., Wagner, V., and Jacobsen, T. (2021). Smooth as glass and hard as stone? On the conceptual structure of the aesthetics of materials. *Psychology of Aesthetics, Creativity, and the Arts*. doi: 10.1037/aca0000437

Melara, R. D., and O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *J. Exp. Psychol. Gen.* 116, 323–336. doi: 10.1037/0096-3445.116.4.323

Mesz, B., and Tedesco, S. (2021) Bruno multisensory gastrosonic and osmotactile installations, Vimeo. STT21 workshop / ACM CHI 2021 on human factors in computing systems - human computer interaction. Available at: https://vimeo.com/561472087.

Mesz, B., Trevisan, M. A., and Sigman, M. (2011). The taste of music. *Perception* 40, 209–219. doi: 10.1068/p6801

Meuleman, B., and Scherer, K. R. (2013). Nonlinear appraisal modeling: an application of machine learning to the study of emotion production. *IEEE Trans. Affect. Comput.* 4, 398–411. doi: 10.1109/T-AFFC.2013.25

Motoki, K., Takahashi, N., Velasco, C., and Spence, C. (2022). Is classical music sweeter than jazz? Crossmodal influences of background music and taste/flavour on healthy and indulgent food preferences. *Food Qual. Prefer.* 96:104380. doi: 10.1016/j.foodqual.2021.104380

Murari, M., Rodà, A., Canazza, S., De Poli, G., and Da Pos, O. (2015). Is Vivaldi smooth and takete? Non-verbal sensory scales for describing music qualities. *J. New Music Res.* 44, 359–372. doi: 10.1080/09298215.2015.1101475

Noyce, G. L., Küssner, M. B., and Sollich, P. (2013). Quantifying shapes: mathematical techniques for analysing visual representations of sound and music. *Empiri. Musicol. Rev.* 8, 128–154. doi: 10.18061/emr.v8i2.3932

Osgood, C., Suci, G., and Tannenbaum, P. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Palmer, S. E., Langlois, T. A., and Schloss, K. B. (2016). Music-to-color associations of single-line piano melodies in non-synesthetes. *Multisens. Res.*, 29(1–3),:157–.

Palmer, S. E., Schloss, K. B., Xu, Z., and Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proc. Natl. Acad. Sci.* 110, 8836–8841. doi: 10.1073/pnas.1212562110

Parise, C. V., and Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Exp. Brain Res.* 220, 319–333. doi: 10.1007/s00221-012-3140-6

Petit, O., Velasco, C., and Spence, C. (2019). Digital sensory marketing: integrating new technologies into multisensory online experience. *J. Interact. Mark.* 45, 42–61. doi: 10.1016/j.intmar.2018.07.004

Poffenberger, A. T., and Barrows, B. E. (1924). The feeling value of lines. *J. Appl. Psychol.* 8, 187–205. doi: 10.1037/h0073513

Reinoso Carvalho, F., Wang, Q. J., Van Ee, R., Persoone, D., and Spence, C. (2017). "Smooth operator": music modulates the perceived creaminess, sweetness, and bitterness of chocolate. *Appetite* 108, 383–390. doi: 10.1016/j.appet.2016.10.026

Reinoso-Carvalho, F., Dakduk, S., Wagemans, J., and Spence, C. (2019). Not just another pint! The role of emotion induced by music on the consumer's tasting experience. *Multisens. Res.* 32, 367–400. doi: 10.1163/22134808-20191374

Reinoso-Carvalho, F., Gunn, L., Molina, G., Narumi, T., Spence, C., Suzuki, Y., et al. (2020). A sprinkle of emotions vs a pinch of crossmodality: towards globally meaningful sonic seasoning strategies for enhanced multisensory tasting experiences. *J. Bus. Res.* 117, 389–399. doi: 10.1016/j.jbusres.2020.04.055

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., and Spence, C. (2015). The sweetest thing: the influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Front. Psychol.* 6:1382. doi: 10.3389/fpsyg.2015.01382

Salgado-Montejo, A., Marmolejo-Ramos, F., Alvarado, J. A., Arboleda, J. C., Suarez, D. R., and Spence, C. (2016). Drawing sounds: representing tones and chords spatially. *Exp. Brain Res.* 234, 3509–3522. doi: 10.1007/s00221-016-4747-9

Schubert, E., and Fabian, D. (2006). The dimensions of baroque music performance: a semantic differential study. *Psychol. Music* 34, 573–587. doi: 10.1177/0305735606068105

Sedikides, C., Wildschut, T., and Baden, D. (2004). "Nostalgia: conceptual issues and existential functions" in *Handbook of experimental existential psychology*. eds. J. Greenberg, S. L. Koole and T. Pyszczynski (New York: The Guilford Press), 200–214.

Seo, H. S., and Hummel, T. (2011). Auditory–olfactory integration: congruent or pleasant sounds amplify odor pleasantness. *Chem. Senses* 36, 301–309. doi: 10.1093/chemse/bjq129

Sievers, B., Lee, C., Haslett, W., and Wheatley, T. (2019). A multi-sensory code for emotional arousal. *Proc. R. Soc. B* 286:20190513. doi: 10.1098/rspb.2019.0513

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2020). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisens. Res.* 33, 1–29. doi: 10.1163/22134808-20191469

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079

Spence, C., and Di Stefano, N. (2022). Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and pitch. *i-Perception* 13, 1–42. doi: 10.1177/20416695221092802

Spence, C., and Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i-Perception* 3, 410–412. doi: 10.1068/i0540ic

Spence, C., Puccinelli, N. M., Grewal, D., and Roggeveen, A. L. (2014). Store atmospherics: a multisensory perspective. *Psychol. Mark.* 31, 472–488. doi: 10.1002/mar.20709

Sun, Z., and Firestone, C. (2021). Curious objects: how visual complexity guides attention and engagement. *Cogn. Sci.* 45:e12933. doi: 10.1111/cogs.12933

Taruffi, L., and Koelsch, S. (2014). The paradox of music-evoked sadness: an online survey. *PLoS One* 9:e110490. doi: 10.1371/journal.pone.0110490

Trost, W., Ethofer, T., Zentner, M., and Vuilleumier, P. (2012). Mapping aesthetic musical emotions in the brain. *Cereb. Cortex* 22, 2769–2783. doi: 10.1093/cercor/bhr353

Turoman, N., Velasco, C., Chen, Y.-C., Huang, P.-C., and Spence, C. (2018). Symmetry and its role in the crossmodal correspondence between shape and taste. *Atten. Percept. Psychophys.* 80, 738–751. doi: 10.3758/s13414-017-1463-x

Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: implications pour la recherche en langue française. *Can. Psychol.* 30, 662–680. doi: 10.1037/h0079856

Vartanian, O., Navarrete, G., Chatterjee, A., Fich, L. B., Leder, H., Modroño, C., et al. (2013). Impact of contour on aesthetic judgments and approach-avoidance decisions in architecture. *Proc. Natl. Acad. Sci.* 110, 10446–10453. doi: 10.1073/pnas.13012271

Vempala, N. N., and Russo, F. A. (2018). Modeling music emotion judgments using machine learning methods. *Front. Psychol.* 8:2239. doi: 10.3389/fpsyg.2017.02239

Vuoskoski, J. K., and Eerola, T. (2011). Measuring music-induced emotion: a comparison of emotion models, personality biases, and intensity of experiences. *Music. Sci.* 15, 159–173. doi: 10.1177/1029864911403367

Wallmark, Z. (2019). A corpus analysis of timbre semantics in orchestration treatises. *Psychol. Music* 47, 585–605. doi: 10.1177/0305735618768102

Ward, R. J., Wuerger, S., and Marshall, A. (2022). Smelling sensations: olfactory crossmodal correspondences. *J. Percept. Imaging* 5, 000402-1–000402-12. doi: 10.2352/J.Percept.Imaging.2022.5.000402

Weichselbaum, H., Leder, H., and Ansorge, U. (2018). Implicit and explicit evaluation of visual symmetry as a function of art expertise. *i-Perception* 9, 1–24. doi: 10.1177/2041669518761464

Weyl, H. (2015). *Symmetry*. Princeton, NY: Princeton University Press.

Zhang, J. D., and Schubert, E. (2019). A single item measure for identifying musician and nonmusician categories based on measures of musical sophistication. *Music Percept.* 36, 457–467. doi: 10.1525/mp.2019.36.5.457

Zuo, H., Hope, T., Castle, P., and Jones, M. (2001). "An investigation into the sensory properties of materials," in *Proceedings of the Second International Conference on Affective Human Factors Design* (pp. 500–507). London, UK: Asean Academic Press.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 07 September 2023
DOI 10.3389/fpsyg.2023.1234417

OPEN ACCESS

EDITED BY
Adrian Von Mühlenen,
University of Warwick, United Kingdom

REVIEWED BY
Marcelo Fernandes Costa,
University of São Paulo, Brazil
Christine Mohr,
Université de Lausanne, Switzerland

*CORRESPONDENCE
Na Chen
✉ imminana7@gmail.com

RECEIVED 04 June 2023
ACCEPTED 21 August 2023
PUBLISHED 07 September 2023

CITATION
Chen N, Nakamura K and Watanabe K (2023) Red biases sex categorization of human bodies. *Front. Psychol.* 14:1234417.
doi: 10.3389/fpsyg.2023.1234417

Red biases sex categorization of human bodies

Na Chen[1]*, Koyo Nakamura[2,3] and Katsumi Watanabe[4]

[1]The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel, [2]Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria, [3]Japan Society for the Promotion of Science, Tokyo, Japan, [4]Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Color is associated with gender information (e.g., red-female). However, little has been known on the effect of color on sex recognition of human bodies. This study aimed to investigate whether the color red could influence the categorization of human bodies by sex, and the effect of contextual information. Visual stimuli were created using body silhouettes varying along the waist-to-hip ratio from female to male shapes. These stimuli were presented in conjunction with red, green, and gray colors, which were used either as body color (Experiment 1) or background color (Experiment 2). Participants were instructed to categorize the sex of the body stimuli as either male or female by pressing labeled keys. The results showed that when red was used as a body color, it induced a bias toward feminine body perception, while when used as a background color, it induced a bias toward masculine body perception. Thus, the color red influenced the sex categorization of human bodies, which being modulated by contextual information. These findings provided novel insights into the effect of contextual color cues in sex recognition of human bodies.

KEYWORDS

red, sex categorization, color-gender association, body, background color

1. Introduction

Recognizing the sex of an individual from a distance, particularly when facial or body features are not easily distinguishable, often relies on various cues. Colors are sometimes used as a means of recognition. For instance, the color red could be stereotypically associated with women based on their clothing colors. In visual media, positioning a character against a red background, especially in intense scenes like walking away from an explosion, can evoke perceptions of heroism or villainy, as seen in Hollywood action movies like Wolverine and the X-Men (2011), which are often associated with male characters. Those differing effects of perceiving red highlight the importance of context in sex recognition. In addition to physical information from the visual body itself, color serves as the fastest and most direct visual cue for sex recognition, playing a crucial role in quickly assessing potential rivals or mates (Nestor and Tarr, 2008). Despite its potential significance, to the best of our knowledge, no previous study has examined the effect of color on the categorization of sex in human bodies. Specifically, the effect of contextual information, when applied directly to the body or present in the background, on this sex process remains unexplored.

Color carries gender information in both Western and Eastern countries. In Western societies, pink has been stereotypically associated with girls and blue with boys since the early 20th century (Pomerleau et al., 1990; Cunningham and Macrae, 2011). There is a prevalent tendency for individuals to have different color preferences for their children based on gender. They tend to choose pink colors for their female children, including pink toys, clothing, and room decorations, while choosing blue or green colors for male children (Del Giudice, 2017; Endendijk, 2022). These gender-stereotyped color perceptions in early life may lead to color-gender associations. In Japan, gender-stereotyped colors are frequently used in social and public service systems. For instance, female toilet icons are typically colored in red or pink, while male toilet icons are colored in blue, green, or black. Previous studies have shown that Japanese people associate reddish colors with females and bluish/black colors with males (Kitagami et al., 2009, 2010; Ishii et al., 2019; Mochizuki and Ota, 2022; Chen et al., 2023a,b). It is common for reddish colors to be widely used in clothing, cosmetics, and activities for females, while blue/green colors are more frequently used in male clothing and behaviors/interests in modern daily life. The mere exposure to the co-occurrence of color and gender information during social interactions may contribute to the establishment of color-gender associations (e.g., red-female). This color-gender association plays a significant role in shaping people's gender-related cognition and behavior. In a recent study by Chen et al. (2023b), the red-female association could be automatically activated and influenced both conceptual gender categorization and perceptual font color discrimination through Stroop-word categorization tasks. It is possible that when the color red was applied to human bodies, it may evoke the learned red-female associations based on the mere exposure to co-occurrences of red clothing and female body perception.

Meanwhile, it should be noted that red has also been associated with masculine traits, such as dominance, aggression, danger, anger, and testosterone levels (Stephen et al., 2012; Young et al., 2013; Wiedemann et al., 2015; Jonauskaite et al., 2021a; Chen et al., under review). People can use the presence and intensity of red facial color information when categorizing the sex of faces, as male faces are typically redder and darker than female faces (Van den Berghe and Frost, 1986; Yip and Sinha, 2002; Frost, 2006; Nestor and Tarr, 2008; Carrito and Semin, 2019). Thus, in certain contexts, red can also be associated with males. For instance, when the color red is present in the background of an ambiguous human body, it may evoke impressions of dominance, aggression, and danger, potentially leading to the perception of a male body. According to the color-in-context theory proposed by Elliot and colleagues, colors acquire specific meanings through environmental and social interactions, while also being influenced by evolutionary factors (Elliot and Maier, 2014). Importantly, the influence of color is contingent on psychological contexts. For instance, the effects of the color red on approach or avoidance behaviors can vary depending on the context. In a romance-related context, red may elicit approach behavior, whereas in a competition-related context, it may lead to avoidance behavior (Meier et al., 2012).

The categorization of human bodies based on sex is strongly influenced by their visual appearance, including features such as the waist-to-hip ratio (WHR) and shoulder-to-hip ratio (Johnson et al., 2012; Gandolfo and Downing, 2020). Individuals with a lower WHR are more likely to be identified as females rather than males (Pazhoohi and Liddle, 2012). Additionally, social contextual information also plays a role in sex categorization. Studies have shown that a fear condition can enhance a bias toward perceiving bodies as male (Johnson et al., 2012), while the presence of a child next to a target can lead to a higher likelihood of attributing female sex (Brielmann et al., 2015). It has been proposed that sex categorization involves a multimodal process, governed by both perception and high-level cognition processes, including stereotypes (Freeman and Ambady, 2011; Clifford et al., 2015). In other words, the categorization process is shaped by both low-level processes that involve visual features and higher-level cognitions formed through prior experiences. Given these findings, it is plausible that the sex categorization of human bodies could be influenced by various factors, including color-gender associations.

Based on the previous findings, the current study aimed to investigate the effect of the color red on the categorization of human body by sex, in combination with contextual effects related to both body color and background color. Through two experiments, we manipulated the perceived masculinity and femininity of body shapes, incorporating the contextual information of red color. This information was used both when the color was applied as the actual body color and when it served as a background color. In the first experiment, we examined whether a red body color could lead to a bias in the sex categorization of human bodies, comparing it to green and gray body colors. In the subsequent experiment, we investigated how contextual color cues might modulate the impact of color on body-sex perception. This was achieved by presenting the body stimuli against different background colors, namely red, green, and gray. The body stimuli were adapted from Johnson et al. (2012) and encompassed a range of waist-to-hip ratios (WHR) representing variations in perceptions of feminine and masculine bodies (see Figure 1). The colors were selected based on their gender associations. Green, for instance, is commonly linked with male, and it's situated opposite to red in many well-established color models (Fehrman and Fehrman, 2004). We did not select blue as a choice due to its prevalence as a more preferred color among both males and females (Jonauskaite et al., 2019), as well as its lack of gender bias in previous research (Li et al., 2021; Jonauskaite et al., 2021a). Gray, an achromatic color, was chosen as a contrasting color that could be standardized for saturation and lightness (Elliot and Niesta, 2008; Young et al., 2013). Participants were required to categorize the gender of body stimuli presented with the three different body colors in the first experiment or against the three different background colors in the second experiment. We hypothesized that a red body color would lead to a bias toward perceiving the body as female, while a red background would instead induce a bias toward perceiving the body as male.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Twenty-nine Japanese undergraduate students (17 males, mean age = 20.2 years, $SD = 1.8$) from Waseda University participated in the experiment. All participants had normal or corrected-to-normal visual acuity and normal color vision, and were naïve regarding the purpose of the experiment. The sample size was set *a priori* at 23 participants, based on a target of 0.8 power with a medium effect size by power analysis (Cohen's $d = 0.6$). We collected more data in case of some participants' data could not fit the psychometric functions (e.g., Deviation >100). This experiment, as well as the subsequent experiment, was approved by the institutional review board (IRB) of Waseda University (2015-033), and conducted in accordance with the

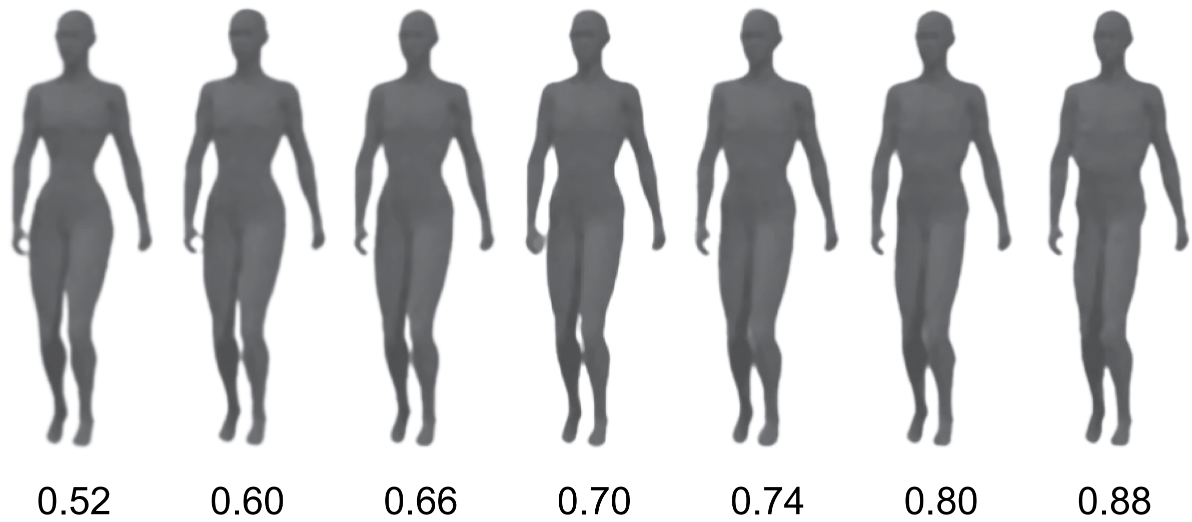

FIGURE 1
Body stimuli varying in seven levels of waist-to-hip ratio (WHR).

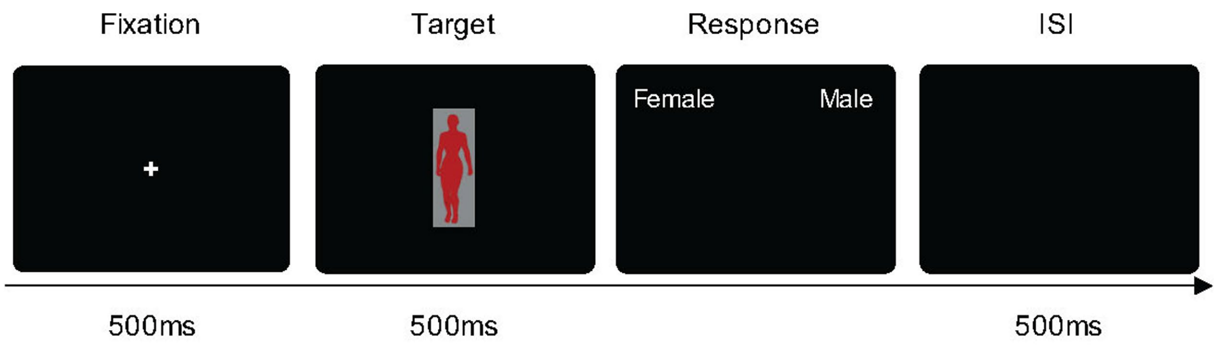

FIGURE 2
The paradigm of Experiment 1. Each trial began with a fixation cross for 500 ms, followed by a colored-body target for 500 ms, and participants pressed the key to report the sex judgment of the body.

ethical standards of the 1964 Declaration of Helsinki. Written informed consent was obtained from all participants in advance.

2.1.2. Apparatus and stimuli

E-Prime 2.0 (Psychology Software Tools, Inc.) was used to present the stimuli and collect the data. Stimuli were displayed on a 24-inch LCD monitor (EIZO FG2421, EIZO corp, Hakusan, Japan), with a 1920 × 1080-pixel resolution and a refresh rate of 100 Hz. Participants viewed the monitor binocularly at a distance of approximately 60 cm.

Seven computer-generated bodies that varied in a sexually dimorphic cue [WHR = 0.52, 0.6, 0.66, 0.7, 0.74, 0.8, 0.88 from the body stimuli in Johnson et al. (2012) and Figure 1] were used. The body stimuli were fitted in a frame of 135 × 380 pixels with gray background color (with 57.37/1.37/255.96 in the CIE LCh color space), presented in three different body colors created by Photoshop CS2 (Adobe Systems Inc., San Jose, CA, United States). The body stimuli were presented against a black screen background throughout the experiment (see Figure 2). The three colors used in the present study were measured by PR-655 (Photo Research, Chatsworth, CA, USA). The color information was as follows: Red: $L^* = 56.66$, $a^* = 80.59$, $b^* = 57.60$; Green: $L^* = 60.03$, $a^* = -64.56$, $b^* = 41.64$; Gray: $L^* = 57.59$, $a^* = -2.46$, $b^* = -30.15$.

2.1.3. Procedure

The experiment was conducted in a laboratory under dim lighting conditions. At the beginning of the experiment, participants were instructed to denote the sex of a centrally presented body shape stimulus by pressing a key on the computer keyboard (i.e., pressing '*z*' for a female body using the left index finger or '*m*' for a male body using the right index finger). Each trial began with a fixation for 500 ms, followed by a body stimulus in body colors (i.e., red, blue, or gray) for 500 ms, after which a black screen was presented until the participants responded (see Figure 2). They were requested to identify the sex of the body stimulus (male or female) as quickly as possible. Trials were separated by an inter-stimulus interval (ISI) of 500 ms. Each body stimulus was presented 15 times in a random order, resulting in 315 trials (7 bodies × 3 body colors × 15 repetitions = 315 trials). The experiment was preceded by ten practice trials. At the end of every 35 trials, participants took a self-determined break. The entire experiment took approximately 20 min to finish.

2.2. Analysis

All statistical analyses were performed using R (R Core Team, 2020). For each participant and condition of body color (i.e., red, green, and gray), we determined a psychometric function based on the proportion of trials in which the body was judged as female. That is, each visual stimulus was repeated 15 times, the proportion of responses with judging as female was calculated. The sex categorization rate from each participant was fitted with a psychometric function using a generalized linear model, with a binomial distribution. By fitting a psychometric function, the level of stimulus variable at which perceptual performance transitions from one perceptual category to the other [i.e., the point of subjective equality (PSE)], and the precision with which participant is able to perceptually differentiate stimuli along the dimension [i.e., the just noticeable difference (JND)] can be revealed (Klein, 2001). To examine the effect of body color on body-sex categorization, the PSE and JND obtained from the fitted curves were used as indicators of perceived differences in sex judgments and discrimination sensitivity, respectively. The PSE was taken from the x value at which the fitted curve had a y value of 0.5, i.e., the WHR of the body stimuli being judged switching from "female" to "male" (where the body looked equally male and female). Thus, the PSE is the WHR value denoting the maximum uncertainty of sex information. The JND was calculated from one-half of the difference between the x values at which the fitted curve had y values of 0.50 and 0.75. This is the amount of change necessary in a body stimulus for the change to be detected 50% of the time. When the JND value is small, a participant is better able to discriminate the sex information. The PSEs and JNDs of the three body colors from the fitted curves of each participant were compared using Repeated measure ANOVA and paired-sample t-test with Bonferroni's correction for multiple comparisons ($\alpha = 0.05/3 = 0.017$). The Bayes Factors (BF10) were further referred to determine whether or not there was support in favor of the alternative (H1) or null (H0) hypotheses (Morey and Rouder, 2018). A value of 1 means that null and alternative are equally likely, larger values suggest that the data are in favor of the alternative hypothesis, and smaller values indicate that the data are in favor of the null hypothesis.

To verify the consistency and robustness of the results, the categorical responses were also fitted by means of Generalized Linear Mixed Model (GLMM) with the *glmer* function (Jaeger, 2008; Bates et al., 2015). We included all fixed and random effects as a first instance and then to eliminate those factors that reduce the model's overall goodness of fit (Bolker, 2008). The Akaike information criterion (AIC) was also used to compare the goodness of fit of the alternative models (a lower AIC value indicating a higher quality model; Mazerolle, 2019). After the model selection, analysis of variance (Type II Wald x^2 test) was performed to identify significant effects and interactions, and Tukey *post hoc* pairwise comparisons were also performed. All data and code are available online.[1]

1 https://osf.io/djygf/

2.3. Results

The data of five participants which failed to fit the psychometric function were removed from the data analysis (the deviances of the fitted models of the excluded participants were all above 120, and the mean deviance of the included participants was 80.11; see the individual plots-Exp1 in the online repository (see footnote 1). Data from 24 participants were used for the data analysis.

For the total amount of male and female responses, irrespective of the WHR manipulation, participants made male responses more frequently than female responses ($t(23) = 3.32$, $p = 0.003$, Cohen's $d = 0.68$). The mean proportion of female responses to each body stimulus and their psychometric curves for the three body colors are shown in Figure 3. The data points represent the mean proportion of trials in which participants responded with female judgment for each body stimulus in the three body colors. The psychometric function showed a horizontal left shift for all the three body colors, indicating that participants' sex judgment of the body stimuli was biased toward a male perception. Repeated measure ANOVA showed that there are significant difference on PSEs between the three body colors, $F(2, 46) = 13.72$, $p < 0.01$, $\eta_p^2 = 0.08$. The PSEs in red body color (Mean = 0.69, $SD = 0.06$) were significantly larger than those in gray body color (Mean = 0.66, $SD = 0.05$), $t(23) = 3.86$, Bonferroni corrected $p = 0.002$, Cohen's $d = 0.79$, $BF_{10} = 42.56$, and in green body color (Mean = 0.66, $SD = 0.05$), $t(23) = 4.76$, Bonferroni corrected $p = 0.0003$, Cohen's $d = 0.97$, $BF_{10} = 316.73$. No difference was observed in PSEs between gray and green background colors, $t(23) = 0.15$, Bonferroni corrected $p = 1$, Cohen's $d = 0.03$, $BF_{10} = 0.22$. Thus, red body color biased the sex categorization of bodies toward a female body perception, reduced the general shift toward a male body bias. Repeated measure ANOVA analysis of JNDs showed that there was no significant difference between the three body colors, $F(2, 46) = 0.13$, $p = 0.88$, $\eta_p^2 = 0.01$. Thus, there was little effect of body color on the sensitivity of sex perception of human bodies.

As a further test of population categorical responses, these responses were fitted using a GLMM. We anticipated that the sex judgment of body would depend on those two factors with the body color and the WHRs. The generalized linear mixed-effect model analysis showed that there was no significant interaction effect between the body color and WHRs, $x(12)^2 = 8.85$, $p = 0.72$. After removing the interaction effect from model analysis, a final model including the body color and WHRs as fixed factors (AIC = 6143.93) showed a significant main effect of body color, $x(2)^2 = 62.31$, $p < 0.001$, and WHRs, $x(6)^2 = 1790.08$, $p < 0.001$. Post-hoc multiple comparison showed a significant difference on the logit scale between the red and gray body colors (Tukey's HSD, $p = 0.00001$), and between the red and green body colors (Tukey's HSD, $p = 0.00001$), and no difference between the gray and green body colors (Tukey's HSD, $p = 0.92$). Thus, these analyses performed on the categorical responses led to the same conclusion: the red body color was more likely to bias body stimulus toward a feminine body than green and gray body colors.

In summary, this experiment demonstrated that the presence of a red body color led to a bias in sex categorization toward a female body perception, compared with green and gray body colors. These

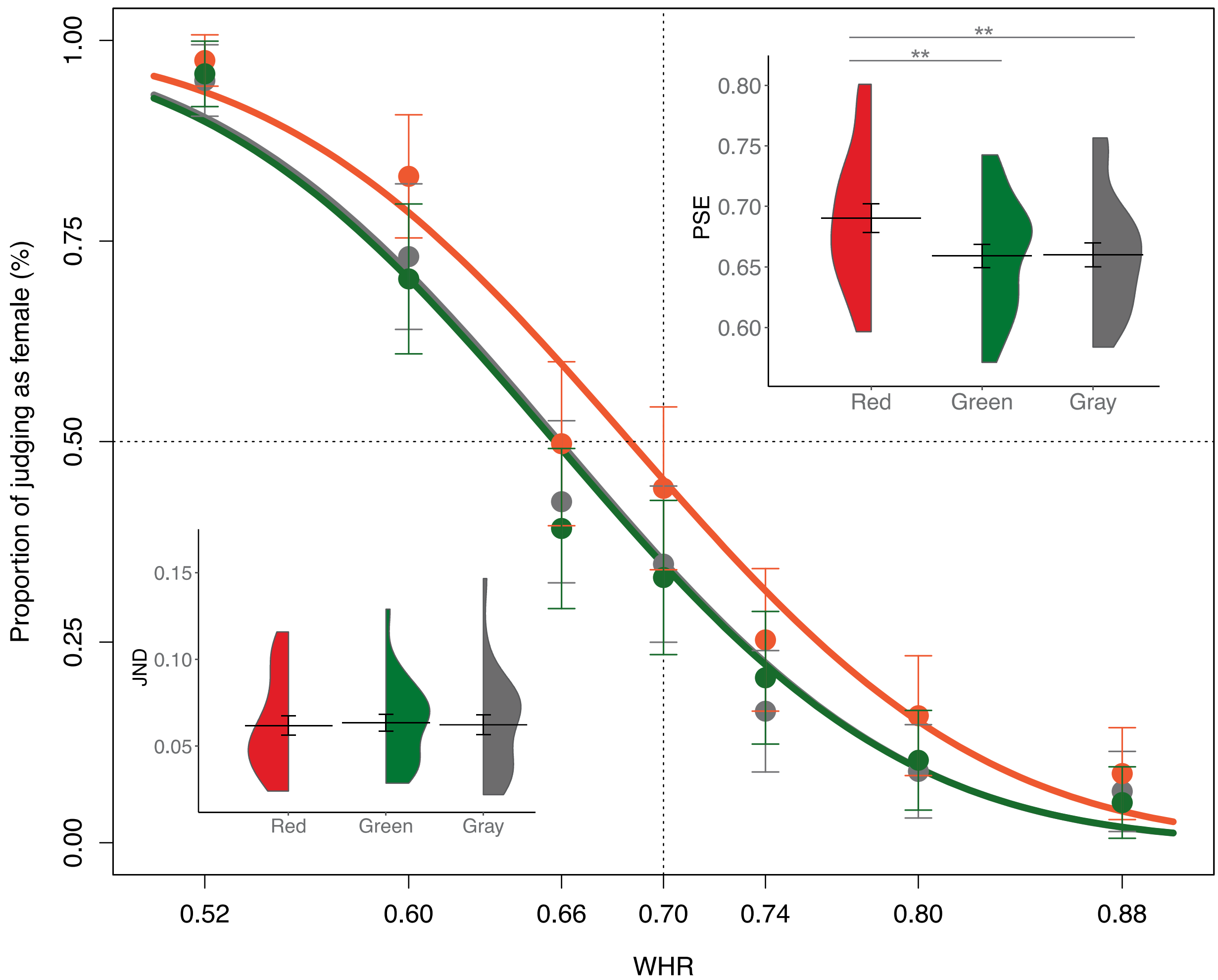

FIGURE 3
The female proportion for each body stimulus. The mean PSEs and JNDs for the three body color conditions are presented. Error bars represent the standard mean error. Asterisks indicate significant differences (**$p < 0.01$, Bonferroni corrected).

findings suggest that the red body color might trigger an automatic red–female association, influencing the sex categorization process. In the next experiment, we investigated whether the contextual information of color cues could alter the observed effect of red-female association on body sex categorization. This was accomplished by comparing the sex categorization responses for bodies presented against three background colors: red, green, and gray.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Twenty-five newly recruited Japanese undergraduate students (11 males, mean age = 20.3 years, *SD* = 1.1) from Waseda University participated in Experiment 2. All participants had normal or corrected-to-normal visual acuity and normal color vision and were naïve regarding the purpose of the experiment.

3.1.2. Stimuli and procedure

The body silhouettes were identical to those used in Experiment 1. The visual stimuli were body silhouettes in gray color [set as gray (57.37/1.37/255.96) in the CIE LCh color space] and presented in the three background colors (i.e., the same color value of red, green, and gray as used in Exp. 1; Figure 4). The experimental setting and procedure were identical to those of Experiment 1.

3.2. Analysis

The data analysis was identical to that in Experiment 1.

3.3. Results

One participant whose response data failed to fit the psychometric function were removed from the data analysis (the deviance of the fitted models of the excluded participant was above 100, and the mean

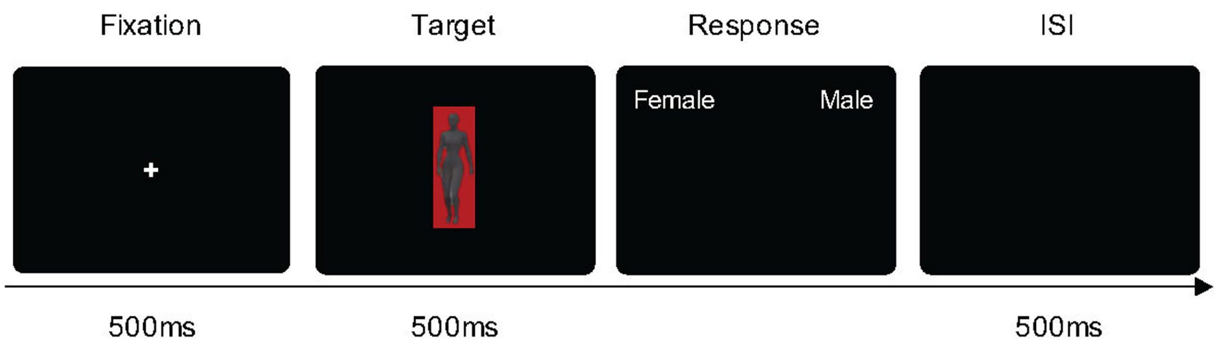

FIGURE 4
The paradigm of Experiment 2. Each trial began with a fixation cross for 500 ms, followed by a target body in a colored-background for 500 ms, and then participants pressed the key to report the sex judgment of the body.

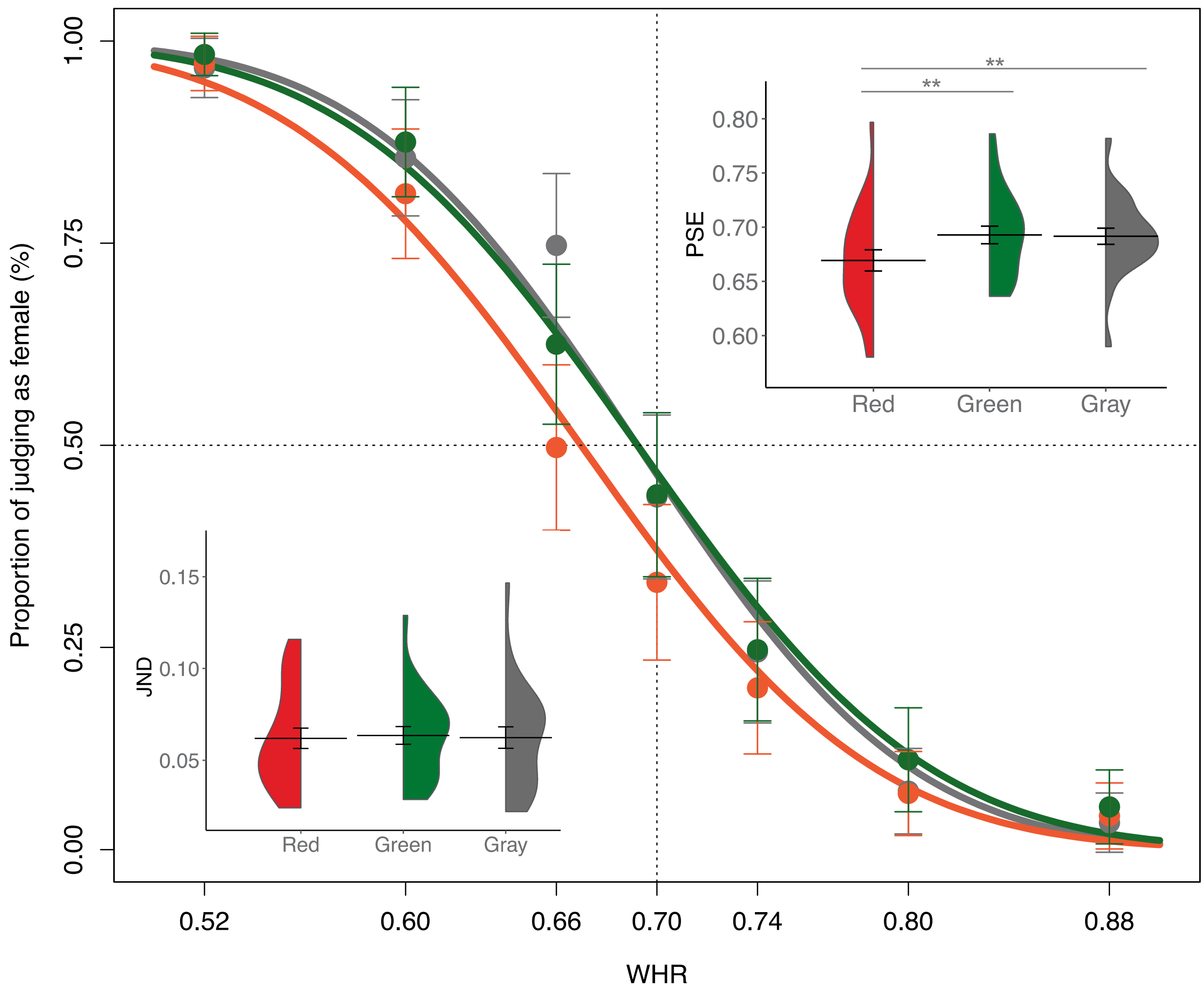

FIGURE 5
The female proportion for each body stimulus. The mean PSEs and JNDs for the three background color conditions are presented. Error bars represent the standard mean error. Asterisks indicate significant differences (**$p < 0.01$, Bonferroni corrected).

deviance of the fitted models of the included participants was 75.6; see the individual plots-Exp2 in the online repository (see footnote 1).

For the total amount of male and female responses, participants made male responses more frequently than female responses ($t(23) = 2.18$, $p = 0.04$, Cohen's $d = 0.45$). The proportion of female response to each body stimulus and their psychometric curves in the three background colors are shown in Figure 5. Repeated measure ANOVA showed that there are significant difference on

PSEs between the three background colors, $F(2, 46) = 9.83$, $p < 0.01$, $\eta_p^2 = 0.07$. The PSEs in red background color (Mean = 0.67, $SD = 0.05$) were significantly smaller than in gray background color (Mean = 0.69, $SD = 0.04$; $t(23) = 3.78$, Bonferroni corrected $p = 0.003$, Cohen's $d = 0.77$, $BF_{10} = 36.40$), and in green background color (Mean = 0.69, $SD = 0.04$; $t(23) = 3.38$, Bonferroni corrected $p = 0.008$, Cohen's $d = 0.69$, $BF_{10} = 15.28$). No difference was observed in PSEs between the gray and green background colors ($t(23) = 0.22$, Bonferroni corrected $p = 1$, Cohen's $d = 0.05$, $BF_{10} = 0.22$). Thus, a red background color could bias the body-sex toward a male body perception, compared with gray and green background colors. Repeated measure ANOVA analysis of JNDs showed no significant difference between the three background colors, $F(2, 46) = 0.87$, $p = 0.43$, $\eta_p^2 = 0.01$, suggesting that the sensitivity of body-sex perception may not be influenced by the background colors.

A final model including the interaction and main fixed factors (AIC = 5780.36) showed a significant interaction effect between the background color and WHRs, $x(12)^2 = 33.41$, $p = 0.0008$, a significant main effect of background color, $x(2)^2 = 49.48$, $p < 0.001$, and a significant main effect of WHRs, $x(6)^2 = 1955.58$, $p < 0.001$. Post-hoc multiple comparison showed a significant difference on the logit scale between the red and gray background colors (Tukey's HSD, $p = 0.026$), and between the red and green background colors (Tukey's HSD, $p = 0.0001$), and no difference between the gray and green background colors (Tukey's HSD, $p = 0.23$). Thus, categorical responses also showed that the red background color was more likely to bias the body stimulus toward a masculine body than green and gray background colors.

This experiment showed that a red background color could bias sex categorization of bodies toward a male body perception. As background color is a task-irrelevant visual cue, it may influence body-sex categorization in different levels of visual and cognition processing, compared with the body color that on main body figures (Minami et al., 2018; Peromaa and Olkkonen, 2019). Thus, the contextual information also plays a role in red-sex bias on human bodies.

4. Discussion

This study is the first to provide evidence that red color can serve as an effective cue for the sex categorization of human bodies. Specifically, our findings revealed that the contextual information influenced the effect of red on body-sex categorization, as the presence of red cues had distinct effects when presented as either the body color or the background color. Through two experiments, we demonstrated that an ambiguous body depicted in red color induced a bias toward perceiving it as female, whereas when the body was present against a red background, it induced a bias toward perceiving it as male. These results contribute to filling the research gap by revealing the effect of color as a contextual information on the sex categorization of human bodies.

One intriguing aspect of the effect of red on body-sex categorization is the opposite effect between the two conditions, where red used as a body color induces a female body bias, while red used as a background color induces a male body bias. One possible explanation for this result may be that the visual processing of sex for human bodies involves multimodal interactions, including interactions between bottom-up perceptual cues and top-down stereotype effects. Previous research has highlighted the dynamic interactive processes in this context, emphasizing the interplay between sensory information, prior knowledge, and attentional factors (Freeman et al., 2008; Freeman and Ambady, 2011; Clifford et al., 2015; Jonauskaite et al., 2021b). According to the dynamic interactive model, bottom-up processing and top-down attention jointly constrain the activation of sex categories through an ongoing interactive process. During the processing of the sex of human body, the sensory information of body features (e.g., WHR) and color cues interact with stored knowledge, represented as the prior probability distribution, leading to a decision of sex categorization.

When red was used as the body color, the interaction between the stored knowledge of red-female association and the sexually dimorphic visual cue of the waist-to-hip ratio appeared to contribute to the formation of a cohesive perception of body sex. This interaction might have activated top-down effects, such as stereotypes related to the association between red body and feminine characteristics, potentially being automatically triggered. It is worth noting that, in many cultural contexts, females tend to wear red clothing more frequently compared to males (Frank, 1990). Consequently, the activation of the red-female association, combined with the impact of expectations and memories, could have influenced the processing of visual cues and led to a bias toward perceiving the bodies as feminine. These findings align with previous research demonstrating the influence of red-female associations on perceptions and categorizations (Cunningham and Macrae, 2011; Chen et al., 2023a,b). For instance, Cunningham and Macrae (2011) showed that a pink *t*-shirt increased perception of femininity, while a blue *t*-shirt increased perception of masculinity.

When red was used as a background color, it induced a bias toward perceiving the bodies as masculine, contradicting the feminine bias observed with red body color. This suggests that the contextual information plays a significant role in shaping the impact of red on the processing of body-sex. The background color, as a contextual task-irrelevant visual cues, interacting with body silhouettes, contributed to the visual configuration in body-sex processing. Plenty of research has found that the color red is associated with visual cues related to testosterone, dominance, aggression, threat, anger, and danger, all of which are associated with masculine characteristics (Elliot and Niesta, 2008; Elliot and Maier, 2012, 2014; Fetterman et al., 2012; Stephen et al., 2012; Young et al., 2013; Pravossoudovitch et al., 2014; Wiedemann et al., 2015; Mentzel et al., 2017; Ruba et al., 2021). In many non-human animals, the presence of red on a conspecific is interpreted as a threat cue, signaling dominance, aggression, and fighting ability (Pryke, 2009; Khan et al., 2011). In humans, testosterone surges during aggressive/dominance encounters can cause visible reddening of the face (Montoya et al., 2005; Stephen et al., 2012), and the experience of anger also leads to reddening of the skin (Changizi et al., 2006). Consequently, red colors carry significant biological signals, upon which impressions of masculinity may be built. In a previous study, Chen et al. (under review) observed that red background color enhanced the perception of dominance in human faces. Facial dominance and masculinity are positively correlated (i.e., the more masculine a face appears, the more dominant it is perceived; Hill et al., 2013). Thus, a red background color might evoke associated impressions of masculinity (e.g., dominance/aggression/anger/danger) through a specific cognitive processing pathway, leading to a bias toward perceiving the bodies as male (Schneider, 2005; Stephen et al., 2012; Young et al., 2013; Wiedemann et al., 2015).

The contextual information provided by color cues influences the recognition of the sex of human bodies, with body color and background color leading to different patterns of sex recognition. A red background color may trigger a biological masculine association, whereas red body color may induce learned red-female associations in a top-down manner. When red is applied to the body interacting with features such as waist-to-hip ratio (WHR), it may reduce the masculine bias and potentially trigger red-love-feminine associations. Numerous studies have shown that red is associated with love, romance, and sexual attractiveness (Elliot et al., 2013; Jonauskaite et al., 2020a; Pazda et al., 2021). For instance, red roses and hearts symbolize love on Valentine's Day, and red lingerie signifies romance in literature. Those results suggest that color-gender interactions are context-dependent and highly dynamic, processed at different levels of cognition.

One limitation of the current study is that we only used three colors (red, green, and gray) to examine the effect of red on sex categorization. Gray was used as a control color; however, it might also be associated with masculinity to some extent. Defining a gender-neutral color can be challenging. Future research may explore the gender associations linked with each individual color, such as natural skin colors (e.g., yellow, brown, black, white), to reveal the strongest and least associated colors for female and male. Another limitation is that each visual stimulus was repeated 15 times, which may have resulted in a response frequency bias, such as participants made male responses more frequently than female responses. Future studies should use an even number of repetitions to avoid response bias. Moreover, future studies are needed to explore the mechanism underlying the effect of color on body-sex perception by examining the effect of perceptual color dimensions (e.g., hue/lightness/chroma; Dael et al., 2016), top-down processing (e.g., color–sex associations/ color semantic/color emotion; Jonauskaite et al., 2020b), and their interactional effects.

In conclusion, our findings highlight the role of red as contextual information in the sex categorization of human bodies. We demonstrated that participants encoded specific patterns of color information related to bodies, with red color playing a role at different levels of body-sex processing. Specifically, when used as a body color, red induced a bias toward female body, whereas when used as a background color, red induced a bias toward male body. These findings provide novel insights into the effect of color as contextual information on the processing of sex-related information in human bodies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Review Board (IRB) of Waseda University (2015-033). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NC coded the experimental paradigm, collected and analyzed the data, and drafted the manuscript. KN and KW provided critical revisions. All authors contributed to the study design and approved the final version of the manuscript for submission.

Funding

This work was supported by Grant-in-Aid for Scientific Research (grand numbers: 19K20387, 20K22296, 21K13759, 22K13801, and 22H00090) from Japan Society for the Promotion of Science.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Bates, D., Kliegl, R., Vasishth, S., and Baayen, H. (2015). Parsimonious mixed models. *arXiv*. doi: 10.48550/arXiv.1506.04967

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton: Princeton University Press.

Brielmann, A. A., Gaetano, J., and Stolarova, M. (2015). Man, you might look like a woman—if a child is next to you. *Adv. Cogn. Psychol.* 11, 84–96. doi: 10.5709/acp-0174-y

Carrito, M. L., and Semin, G. R. (2019). When we don't know what we know–sex and skin color. *Cognition* 191:103972. doi: 10.1016/j.cognition.2019.05.009

Changizi, M. A., Zhang, Q., and Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biol. Lett.* 2, 217–221. doi: 10.1098/rsbl.2006.0440

Chen, N., Nakamura, K., and Watanabe, K. (2023a). Red background color biases gender categorization of human faces. *Sci. Rep.* 13:7792. doi: 10.1038/s41598-023-34644-4

Chen, N., Nakamura, K., and Watanabe, K. (2023b). An automatic red-female association tested by the stroop task. *Acta Psychol.* 238:103982. doi: 10.1016/j.actpsy.2023.103982

Chen, N., Yang, Y., Kobayashi, M., Nakamura, K., and Watanabe, K. (Under review). Red background color enhances dominance perception of faces and shapes. doi: 10.31234/osf.io/dq53a

Clifford, C. W. G., Mareschal, I., Otsuka, Y., and Watson, T. L. (2015). A Bayesian approach to person perception. *Conscious. Cogn.* 36, 406–413. doi: 10.1016/j.concog.2015.03.015

Cunningham, S. J., and Macrae, C. N. (2011). The colour of gender stereotyping. *Br. J. Psychol.* 102, 598–614. doi: 10.1111/j.2044-8295.2011.02023.x

Dael, N., Perseguers, M. N., Marchand, C., Antonietti, J. P., and Mohr, C. (2016). Put on that colour, it fits your emotion: colour appropriateness as a function of expressed emotion. *Q. J. Exp. Psychol.* 69, 1619–1630. doi: 10.1080/17470218.2015.1090462

Del Giudice, M. (2017). Pink, blue, and gender: An update. *Arch. Sex. Behav.* 46, 1555–1563. doi: 10.1007/s10508-017-1024-3

Elliot, A. J., and Maier, M. A. (2012). "Color-in-context theory" in *Advances in experimental social psychology*. eds. P. Devine and A. Plant, vol. *45* (Academic Press), 61–125.

Elliot, A. J., and Maier, M. A. (2014). Color psychology: effects of perceiving color on psychological functioning in humans. *Annu. Rev. Psychol.* 65, 95–120. doi: 10.1146/annurev-psych-010213-115035

Elliot, A. J., and Niesta, D. (2008). Romantic red: red enhances men's attraction to women. *J. Pers. Soc. Psychol.* 95, 1150–1164. doi: 10.1037/0022-3514.95.5.1150

Elliot, A. J., Tracy, J. L., Pazda, A. D., and Beall, A. T. (2013). Red enhances women's attractiveness to men: first evidence suggesting universality. *J. Exp. Soc. Psychol.* 49, 165–168. doi: 10.1016/j.jesp.2012.07.017

Endendijk, J. J. (2022). Welcome to a pink and blue world! An analysis of gender-typed content in birth announcement cards from 1940–2019 in the Netherlands. *Sex Roles* 86, 1–13. doi: 10.1007/s11199-021-01249-y

Fehrman, K. R., and Fehrman, C. (2004). *Color: the secret influence*, *2nd* (Upper Saddle River, NJ: Pearson Prentice Hall).

Fetterman, A. K., Robinson, M. D., and Meier, B. P. (2012). Anger as "seeing red": evidence for a perceptual association. *Cognit. Emot.* 26, 1445–1458. doi: 10.1080/02699931.2012.673477

Frank, J. (1990). Gender differences in color naming: direct mail order advertisements. *Am. Speech* 65, 114–126. doi: 10.2307/455532

Freeman, J. B., and Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychol. Rev.* 118, 247–279. doi: 10.1037/a0022327

Freeman, J. B., Ambady, N., Rule, N. O., and Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *J. Exp. Psychol. Gen.* 137, 673–690. doi: 10.1037/a0013875

Frost, P. (2006). European hair and eye color: a case of frequency-dependent sexual selection? *Evol. Hum. Behav.* 27, 85–103. doi: 10.1016/j.evolhumbehav.2005.07.002

Gandolfo, M., and Downing, P. E. (2020). Asymmetric visual representation of sex from human body shape. *Cognition* 205:104436. doi: 10.1016/j.cognition.2020.104436

Hill, A. K., Hunt, J., Welling, L. L., Cardenas, R. A., Rotella, M. A., Wheatley, J. R., et al. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evol. Hum. Behav.* 34, 334–341. doi: 10.1016/j.evolhumbehav.2013.05.004

Ishii, K., Numazaki, M., and Tado'oka, Y. (2019). The effect of pink/blue clothing on implicit and explicit gender-related self-cognition and attitudes among men. *Jpn. Psychol. Res.* 61, 123–132. doi: 10.1111/jpr.12241

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446. doi: 10.1016/j.jml.2007.11.007

Johnson, K. L., Iida, M., and Tassinary, L. G. (2012). Person (mis) perception: functionally biased sex categorization of bodies. *Proc. R. Soc. B Biol. Sci.* 279, 4982–4989. doi: 10.1098/rspb.2012.2060

Jonauskaite, D., Abu-Akel, A., Dael, N., Oberfeld, D., Abdel-Khalek, A. M., Al-Rasheed, A. S., et al. (2020a). Universal patterns in color-emotion associations are further shaped by linguistic and geographic proximity. *Psychol. Sci.* 31, 1245–1260. doi: 10.1177/0956797620948810

Jonauskaite, D., Camenzind, L., Parraga, C. A., Diouf, C. N., Ducommun, M. M., Müller, L., et al. (2021b). Colour-emotion associations in individuals with red-green colour blindness. *PeerJ* 9:e11180. doi: 10.7717/peerj.11180

Jonauskaite, D., Dael, N., Chèvre, L., Althaus, B., Tremea, A., Charalambides, L., et al. (2019). Pink for girls, red for boys, and blue for both genders: colour preferences in children and adults. *Sex Roles* 80, 630–642. doi: 10.1007/s11199-018-0955-z

Jonauskaite, D., Parraga, C. A., Quiblier, M., and Mohr, C. (2020b). Feeling blue or seeing red? Similar patterns of emotion associations with colour patches and colour terms. *i-Perception* 11:2041669520902484. doi: 10.1177/2041669520902484

Jonauskaite, D., Sutton, A., Cristianini, N., and Mohr, C. (2021a). English colour terms carry gender and valence biases: a corpus study using word embeddings. *PLoS One* 16:e0251559. doi: 10.1371/journal.pone.0251559

Khan, S. A., Levine, W. J., Dobson, S. D., and Kralik, J. D. (2011). Red signals dominance in male rhesus macaques. *Psychol. Sci.* 22, 1001–1003. doi: 10.1177/0956797611415543

Kitagami, S., Suga, S., Kim, H., Komeda, H., and Miyamoto, Y. (2009). Is color important in toilet sign? A cross-cultural study of Stroop-like effect on cognition of toilet sign. *Proceeding of the 7th Conference of the Japanese Society for Cognitive Psychology*, Saitama, Japan, pp. 19. (In Japanese)

Kitagami, S., Suga, S., Kim, H., Komeda, H., and Miyamoto, Y. (2010). "Is color important in toilet sign? (2) A cross-cultural study of Stroop-like effect on cognition of toilet sign," in *Proceeding of the 8th Conference of the Japanese Society for Cognitive Psychology*. 28. (In Japanese).

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: a commentary. *Percept. Psychophys.* 63, 1421–1455. doi: 10.3758/BF03194552

Li, Y., Du, J., Song, Q., Wu, S., and Liu, L. (2021). An ERP study of the temporal course of gender-color stroop effect. *Front. Psychol.* 11:613196. doi: 10.3389/fpsyg.2020.613196

Mazerolle, M. J. (2019). Model selection and multimodel inference based on (Q) AIC (c). R package version 2.2-2.

Meier, B. P., D'agostino, P. R., Elliot, A. J., Maier, M. A., and Wilkowski, B. M. (2012). Color in context: psychological context moderates the influence of red on approach-and avoidance-motivated behavior. *PLoS One* 7:e40333. doi: 10.1371/journal.pone.0040333

Mentzel, S. V., Schücker, L., Hagemann, N., and Strauss, B. (2017). Emotionality of colors: an implicit link between red and dominance. *Front. Psychol.* 8:317. doi: 10.3389/fpsyg.2017.00317

Minami, T., Nakajima, K., and Nakauchi, S. (2018). Effects of face and background color on facial expression perception. *Front. Psychol.* 9:1012. doi: 10.3389/fpsyg.2018.01012

Mochizuki, M., and Ota, N. (2022). A matter of brightness, hue, or both? Gender–color associations in Japanese participants. *Curr. Psychol.* 1–10. doi: 10.1007/s12144-022-04179-4

Montoya, P., Campos, J. J., and Schandry, R. (2005). See red? Turn pale? Unveiling emotions through cardiovascular and hemodynamic changes. *Span. J. Psychol.* 8, 79–85. doi: 10.1017/S1138741600004984

Morey, R. D., and Rouder, J. N. (2018). BayseFactor: computation of Bayes factors for common designs. R package version 0.9.12-4.2.

Nestor, A., and Tarr, M. J. (2008). Gender recognition of human faces using color. *Psychol. Sci.* 19, 1242–1246. doi: 10.1111/j.1467-9280.2008.02232.x

Pazda, A. D., Thorstenson, C. A., and Elliot, A. J. (2021). The effect of red on attractiveness for highly attractive women. *Curr. Psychol.* 1–8. doi: 10.1007/s12144-021-02045-3

Pazhoohi, F., and Liddle, J. R. (2012). Identifying feminine and masculine ranges for waist-to-hip ratio. *J. Soc. Evol. Cult. Psychol.* 6, 227–232. doi: 10.1037/h0099212

Peromaa, T., and Olkkonen, M. (2019). Red color facilitates the detection of facial anger—but how much? *PLoS One* 14:e0215610. doi: 10.1371/journal.pone.0215610

Pomerleau, A., Bolduc, D., Malcuit, G., and Cossette, L. (1990). Pink or blue: environmental gender stereotypes in the first two years of life. *Sex Roles* 22, 359–367. doi: 10.1007/BF00288339

Pravossoudovitch, K., Cury, F., Young, S. G., and Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red–danger association. *Ergonomics* 57, 503–510. doi: 10.1080/00140139.2014.889220

Pryke, S. R. (2009). Is red an innate or learned signal of aggression and intimidation? *Anim. Behav.* 78, 393–398. doi: 10.1016/j.anbehav.2009.05.013

R Core Team. (2020). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ruba, A. L., Thorstenson, C. A., and Repacholi, B. M. (2021). Red enhances the processing of anger facial configurations as a function of target gender. *Soc. Cogn.* 39, 396–407. doi: 10.1521/soco.2021.39.3.396

Schneider, D. J. (2005). *The psychology of stereotyping*. New York: Guilford Press.

Stephen, I. D., Oldham, F. H., Perrett, D. I., and Barton, R. A. (2012). Redness enhances perceived aggression, dominance and attractiveness in men's faces. *Evol. Psychol.* 10, 562–572. doi: 10.1177/147470491201000312

Van den Berghe, P. L., and Frost, P. (1986). Skin color preference, sexual dimorphism and sexual selection: a case of gene culture coevolution? *Ethn. Racial Stud.* 9, 87–113. doi: 10.1080/01419870.1986.9993516

Wiedemann, D., Burt, D. M., Hill, R. A., and Barton, R. A. (2015). Red clothing increases perceived dominance, aggression and anger. *Biol. Lett.* 11:20150166. doi: 10.1098/rsbl.2015.0166

Yip, A. W., and Sinha, P. (2002). Contribution of color to face recognition. *Perception* 31, 995–1003. doi: 10.1068/p3376

Young, S. G., Elliot, A. J., Feltman, R., and Ambady, N. (2013). Red enhances the processing of facial expressions of anger. *Emotion* 13, 380–384. doi: 10.1037/a0032471

TYPE Original Research
PUBLISHED 25 September 2023
DOI 10.3389/fpsyg.2023.1176794

Enhancing the design of wine labels

Anders Crichton-Fock[1]*, Charles Spence[2], Maria Mora[3,4] and Nicklas Pettersson[5]

[1]School of Hospitality, Culinary Arts & Meal Science, Örebro University, Örebro, Sweden, [2]Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, United Kingdom, [3]BCC Innovation, Technology Center in Gastronomy, Basque Culinary Center, Donostia-San Sebastián, Spain, [4]Basque Culinary Center, Faculty of Gastronomic Sciences, Mondragon Unibersitatea, Donostia-San Sebastián, Spain, [5]School of Business, Örebro University, Örebro, Sweden

EDITED BY
Eddy J. Davelaar,
Birkbeck, University of London,
United Kingdom

REVIEWED BY
Erhard Lick,
School of Business and Management, France
Vincenzo Russo,
Università IULM, Italy

*CORRESPONDENCE
Anders Crichton-Fock
anders.crichton-fock@oru.se

RECEIVED 28 February 2023
ACCEPTED 08 September 2023
PUBLISHED 25 September 2023

CITATION
Crichton-Fock A, Spence C, Mora M and Pettersson N (2023) Enhancing the design of wine labels.
Front. Psychol. 14:1176794.
doi: 10.3389/fpsyg.2023.1176794

Introduction: The knowledge accrued through research in the domain of crossmodal correspondences has had a significant influence on a diverse array of disciplines, including behavioral studies, neuroscience, computational modeling, and notably, marketing, with the objective of aligning sensory experiences to help shape patterns of consumer behavior. A study is reported that explores the extension of these principles to the communication of products having a notably complex sensory profile, specifically within the context of wine. The central aim of the project is to explore the feasibility of using crossmodal communication as a strategic tool to augment the congruence between the consumers' multisensory expectations and their sensory experiences. For consumers venturing into the realm of wine selection without the advantage of prior tasting experience, it is of paramount importance to possess a robust understanding of the mandated information. This encompasses critical elements such as the wine's origin, grape varietal(s) used, geographical indications, producer qualifications, and the potential implications of these factors on the final wine product. This level of comprehension stands as a necessary prerequisite, enabling these consumers to make informed choices that align with their preferences, even in the absence of previous sensory encounters. Nonetheless, semiotic investigations underscore the significance attributed to symbolic components such as signs, logos, colors, gestures, and linguistic cues. Research from the field performing multisensory studies, presents a counterpoint to prevailing communication paradigms, advocating for a heightened incorporation of metaphors, analogies, symbols, metonymies, and allegories. This alternative approach aims to enhance the efficacy of communication strategies, offering a more profound and evocative means of conveying intricate messages on a more holistic level.

Methods: A questionnaire was sent to a specific group of engaged wine consumers ($n = 329$). Besides questions regarding demographics, purchase behavior, and consumption behavior, the questionnaire included examples of multisensory communication through a selection of symbols, as well as alternative wine information.

Results: The results showed significant correlations between demographics, consumption behavior, and attitudes toward the tested multisensory symbols and alternative information, thus helping to gain a better understanding of the sensory properties that should be communicated on wine labels.

Discussion: The findings reported here highlight the effectiveness of visual crossmodal communication as a promising pathway capable of skillfully capturing consumer attributes, conveying multisensory experiences, and portraying the comprehensive timeline of taste evolution. As a result, it assumes a pivotal role as a communicative tool for intricate consumables, like wine, functioning at the crossroads of visual and sensory dimensions.

KEYWORDS

crossmodal communication, correspondence, multisensory, vision, wine assessment, complex foods, resource efficiency, wine labels

1. Introduction

In the realm of sensory marketing, the utilization of crossmodal correspondences has emerged as a strategic approach to harness the manifold multisensory effects elucidated through an expanding array of controlled laboratory investigations. This primary strategy involves leveraging insights into the intricacies of consumer behavior and contentment, thereby serving as a mechanism to augment sales and fortify market standing within the competitive market environment (Goldkuhl and Styvén, 2007; Krishna, 2010; Krishna and Schwarz, 2014). While research on the crossmodal correspondences that has been published to date has primarily focused on understanding cognitive stimulation and intermodal connections, the application in sensory marketing has targeted various sensory experiences that might benefit from it. By collectively influencing the manner in which consumers perceive and engage with products, particularly in the context of communication and alignment with consumers' expectations of the products (Weil, 2007). Vision is widely considered the dominant sense to use in this context (Hutmacher, 2019).

In essence, sensory marketing seeks to investigate how sensory cues influence the consumer's encounters from a commercial perspective (Wang and Li, 2022). Notably, within this field, vision tends to take precedence. Consequently, it becomes essential to establish a seamless alignment between consumer preferences and the attributes that a potential product can offer and by so doing create a more harmonious multisensory experience (Elder and Krishna, 2010; Varela and Ares, 2012; Paradis and Eeg-Olofsson, 2013; Krishna and Schwarz, 2014; Croijmans and Wang, 2021).

Consumer research includes many possible approaches and multiple cultural and genetic factors to consider (Bartoshuk et al., 1996; Bartoshuk, 2000; Reed and Knaapila, 2010; Pagliarini et al., 2021). For instance, at the sensory level, researchers have explored the genetic impact of taste sensations in regard to consumers' perceptual sensitivity to, and preference for, certain key attributes, such as sweetness (Gent and Bartoshuk, 1983), sourness (Breslin, 1996; Pagliarini et al., 2021), bitterness (Bartoshuk et al., 1988), saltiness (Breslin and Beauchamp, 1997; Bartoshuk et al., 1998), and umami (Keast and Breslin, 2003; Kim et al., 2015; Linscott and Lim, 2016). Beyond crossmodal interactions, integrating genetic factors becomes pertinent in the pursuit of enhancing communication by targeting pivotal sensory attributes that influence consumer perception and acceptance of specific food products amongst particular groups of consumers. This becomes particularly relevant when examining the divergent reactions of various groups of consumers to a given product, even though genetic research might exhibit certain limitations in pinpointing such responses. This challenge is notably intricate when addressing multifaceted flavor profiles and the aromas of certain food products, with wine serving as a prime example of such stimulus complexity (Shepherd, 2006; Parr, 2015; Spence and Wang, 2018). Adding to the communication challenge, the production of wine involves multiple stages of refinement, spanning from cultivation to bottling which, in turn, contributes to a notable climate impact as this intricate process unfolds (Christ and Burritt, 2013; Iannone et al., 2016).

In the domain of packaging, despite the transient ebb and flow of diverse trends involving motifs such as critters and idiomatic expressions, wines continue to be characterized by labels that can be classed as conventional. These labels, primarily affixed to the front and back of wine bottles, predominantly serve as conduits for obligatory and regulated content, as dictated by prevailing legislative frameworks. The conventional labeling conventionally encompasses details pertaining to the wine's provenance, country of origin, grape varietal, alcohol concentration, and vintage year. However, it is worth noting that these traditional designs may not inherently convey the intricate nuances of a wine's sensory properties to the discerning consumer. In light of the genetic influences and crossmodal factors elucidated by prior research, it becomes relevant to explore the feasibility of incorporating these conceptual frameworks into the design of wine labels in order to investigate their potential use in consumer communication.

The present study aimed to investigate consumer attitudes toward crossmodal and multisensory approaches to wine communication, using both visual and non-verbal cues to help communicate various multisensory information on the label. The second aim was to investigate critical attributes and other information requested by consumers in order to examine the paradigm of conventional wine labeling. More effective communication can thus better cater to specific target groups while optimizing the use of natural resources in terms of satisfying the consumer.

1.1. Literature review

Studies in semiotics, exploring symbolic communication and understanding, propose that various forms of meaning such as signs, logos, gestures, illustrations, linguistic and non-linguistic communication can serve as essential tools when it comes to engaging different groups of consumers (König and Lick, 2014; van Tonder and Mulder, 2015; Lick et al., 2017; Celhay and Remaud, 2018; Pelet et al., 2020). Beyond semiotics, researchers have also advocated for the use of rhetorical figures such as metaphors, analogies, symbols, metonymies, and allegories to enhance communication effectiveness (Moreno Lara, 2014; Alousque, 2015; Kelley et al., 2015; Costello et al., 2018; Herdenstam et al., 2020). Furthermore, exploration into modern and innovative communication strategies has considered sensory descriptors to evoke olfactory mental imagery (Shepherd, 2006; Tomiczek and Stevenson, 2009), as well as the use of scene or country descriptions or origin to help construct a mental sensory experience to promote sale (Tomiczek and Stevenson, 2009; Williamson et al., 2016; Croijmans and Wang, 2021).

Furthermore, researchers have ventured into novel sensory strategies, including the integration of multisensory or crossmodal stimuli, in order to align consumer expectations with the tasting experience. This involves communicating the impact of specific food combinations (Harrington, 2005, 2008; Koone et al., 2014; Herdenstam et al., 2018; Spence, 2020b) and identifying attributes in these combinations that might impact consumer acceptance (Harrington, 2005, 2007; Harrington and Hammond, 2006, 2007, 2009; Harrington et al., 2010; Koone et al., 2014; Harrington and Seo, 2015).

In the realm of crossmodal correspondences, associations between the stimuli presented (or merely imagined) in one sensory modality affecting responses in another modality have been studied (Spence, 2011). Notably, within sensory analysis, crossmodal interactions have demonstrated varied impacts on consumer perceptions within different dining and food contexts. Research in this domain has explored influences ranging from frequency of sound and music (Spence et al., 2014; Hagtvedt and Brasel, 2016; De Luca et al., 2019), lighting and colors (Spence et al., 2014; Biswas et al., 2017; Heatherly et al., 2019; Maziriri et al., 2021), visually-presented shapes (Hanson-Vaux et al., 2013; Liu et al., 2018; Spence, 2020a,b), touch and tactile sensations (Gallace and Spence, 2010; Spence et al., 2013; Etzi et al., 2014; Olzak and Craig, 2014; Wang and Spence, 2018a,b), and, not least, the significant influence of odors and scents (Goldkuhl and Styvén, 2007; Krishna, 2010; Crisinel and Spence, 2012; Deroy et al., 2013; Reid et al., 2015; Ward et al., 2022; Spence, 2022b).

Furthermore, other researchers have explored contemporary and innovative methods for communicating essential and desirable product attributes. These approaches include using sensory descriptors to help conjure up olfactory mental imagery (González et al., 2006; Shepherd, 2006; Tomiczek and Stevenson, 2009), as well as investigating the effects of describing scenes or countries to craft a detailed mental image of a particular sensory encounter (Tomiczek and Stevenson, 2009; Williamson et al., 2016; Croijmans and Wang, 2021).

In summary, the phenomenon of crossmodal correspondence has been extensively studied in various consumer contexts, revealing its substantial influence on consumer satisfaction. However, to the best of the authors' knowledge, there appears to be a research gap regarding the potential application of crossmodal communication to enhance the consumer's comprehension of the expected intricate sensory attributes in a complex tasting experience, such as offered by a quality wine. These products, characterized by layers of volatile odors, flavors, and oral-somatosensory sensations on a multisensory level, could hold critical importance for achieving consumer approval (Wang and Spence, 2018a,b). If effectively harnessed to cater to specific target audiences, crossmodal communication might thus not only help to bolster marketing strategies and consumer contentment but also serve as a tool for optimizing the use of resources within complex food products, potentially contributing to the broader goal of reducing food waste (Galbreath et al., 2020).

2. Materials and methods

2.1. Ethics statement

The questionnaire was performed in accordance with the Declaration of Helsinki and the European Code of Conduct for Research Integrity. All of the respondents were over 20 years of age, and informed consent was obtained from all respondents. All data and analysis files were kept in accordance with legislated and regulated data handling practices.

2.2. Respondents

The sample consisted of 329 students from different sections of the 7.5-credits, 15-week distance course 'Beverage knowledge' offered at Örebro University (Sweden). These students underwent comprehensive training and gained collective proficiency in wine analysis, along with experience in crossmodal correspondence. This experience was particularly evident during training and tasting sessions, wherein respondents engaged in meticulous evaluations while transitioning between senses, leading to significant preconceived notions. The assessment process involved a sequence, starting with visual evaluations encompassing aspects like color, intensity, maturity, age, freshness, acidity, and concentration. These preliminary impressions were subsequently corroborated through olfactory assessments and later confirmed on the palate. This approach provided respondents with firsthand encounters of crossmodal influences and their noteworthy impact during professional wine tasting procedures. This impact was exemplified when respondents engaged in diverse tasting exercises. For instance, they initially perceived fragrance notes of ripe pineapple and sweet mango, thus forming initial impressions concerning the wine's perceived level of sweetness. However, upon tasting, they realized that the wine was, in fact, completely dry. This experiential interplay of senses distinctly highlighted the intricate interrelationship between sensory modalities and their potential to substantially influence the overall perceptual experience.

The majority of the respondents were female (59%) and lived in the city (76%). Almost all had previously studied at the university (94%), and most of them had received a bachelor's degree or higher (74%). Many considered themselves to have better wine knowledge than the population at large (76%). Most of them consumed wine on a weekly base (87%), which they typically purchased at Systembolaget (Sweden's nationally regulated liquor monopoly) (81%) and consumed at home (80%) (see Supplementary Appendix A).

The respondents shared the following traits:

i. They had all tasted the same wines and other beverages, and therefore shared a variety of sensory experiences (see Supplementary Appendix F).
ii. They had all learned a common approach and methodology for analyzing wine. It can thereby be presumed that, on a group level, they had an awareness of the importance of all sensory modalities in the analysis process, including vision, smell, taste, touch, and sound.
iii. They had all been exposed to crossmodal correspondence during the course. This by performing the large number of

tasting exercises involved in the course. Especially when moving from one sense to another during the tasting process, and, subsequently, communicating it, while each sensory modality is not separately (see Supplementary Appendix F).

2.3. Questionnaire

The questionnaire consisted of four sections. The first section included questions relating to *demographics*, as well as single-choice and multiple-choice (check-all-that-apply; CATA) questions relating to purchase and consumption behaviors, communication, and sensory experiences (see Supplementary Appendix B). The second section of the questionnaire aimed to test different *design and symbols* developed in dialogue with wine experts and researchers. The symbols and illustrations used in this section attempted to assess the perception of several of the multisensory factors that have been shown to affect crossmodal experiences. The illustrations were developed by the art and food designer Elin Aronsen Beis, who also specializes in food packaging. For each design question, a short background was given to briefly illustrate the communicative purpose of each symbol, item, or other piece of information. After exposure to the different designs, the respondents were asked to indicate on a 7-point Likert scale (1 = "Not very helpful"; 4 = "Neutral"; 7 = "Very helpful") how helpful each design was in interpreting the potential sensory characteristics of the product.

Here follows an example of a creative design question included the questionnaire and respective background information given to the respondents before answering each question:

On the group level, the research shows that consumers have different sensitivities to bitterness that affect our preference for wine depending on consumers' taste type (tolerant, sensitive, very sensitive, and hyper-sensitive). To what extent do you think it would be helpful to communicate the optimal consumer taste group on the label (see example)? (see Figure 1A). The idea being both to present research in this field as well as how it could be implemented in consumer communication. One aspect being that the consumer already has awareness of their own "optimal consumer taste group," another whether this information would be helpful if added to a wine label. The other design questions included other symbols and illustrations (see Figures 1B–D) as well as background information to stimulate the creative process and understand the context of use for each symbol.

In the third section, the respondents were asked about their preferred *textual sensory descriptions and assessments* to be included in label of wines of different origins (see Figure 2).

Within this segment, various other inquiries concerning labels were also presented. These included evaluations conducted by experts, encompassing factors like the readiness of the wine for consumption and judgments on quality. Moreover, the third section of the questionnaire directed respondents' attention toward the potential inclusion of insights from professional tasters. This section aimed to gauge whether communicating common faults and defects typically associated with a specific type of wine would be beneficial. The underlying rationale behind these inquiries shifted from the prior questions, which had focused on more personalized engagement and self-awareness. For instance, respondents were asked about their awareness of their "optimal consumer taste group." The intention behind incorporating these queries involving wine experts, as opposed to the previous questions concerning personal knowledge, was to facilitate a comparison between diverse communication strategies. This comparison aimed to shed light on the efficacy of different approaches in conveying information to consumers. In the fourth and final section, the respondents were asked questions related to alternative sensory communication in general as well as attitudes toward buying blended wines, wines made from already existing wines, and sustainability.

2.4. Data analysis

EyeQuestion version 5 (Logic 8, Elst, The Netherlands), a software program for sensory and consumer testing, was used to collect the respondents' responses. Statistical analysis was undertaken using the software R (R Core Team, 2021).

3. Results

3.1. Attitudes toward tested wine label design and symbols (questionnaire section 2)

Regarding attitudes toward alternative communication using the tested symbols, the respondents showed positive responses (above neutral) toward symbols illustrating dominant *flavor intensity* (76%) and *taste timeline/flavor development* (65%). Symbols for *non-existing qualities* (52%) and *genetics and taste sensitivity* profiling (49%) received a positive response from approximately half of the respondents (see Table 1). For more details, see Supplementary Appendix D.

3.2. Attitudes toward tested textual sensory descriptions and assessments by wine experts (questionnaire section 3)

For the tested text information, respondents showed a positive response toward *level of readiness* (89%) and *quality assessment* by a wine expert (71%). Just over half of respondents (53%) responded positively toward highlighting *potential faults* (see Table 2).

3.3. Attitudes toward alternative sensory communication, blending, and sustainability (questionnaire section 4)

A majority of the respondents (64%) said that they would be open to at least try a bottle based on sensory information alone, while approximately 9% answered that they would never consider it. As for buying wine that had been made through a blend of other existing wines, most respondents (74%) reported that they would try a bottle, while about 5% reported that they would never consider it. Regarding attitudes toward sustainability, a majority of the respondents (76%) reported that they take this into consideration at least to some degree when purchasing wine. By contrast, about 6% of respondents answered that they would never take this into account when buying wine (see Table 3). For more information, see Supplementary Appendix D.

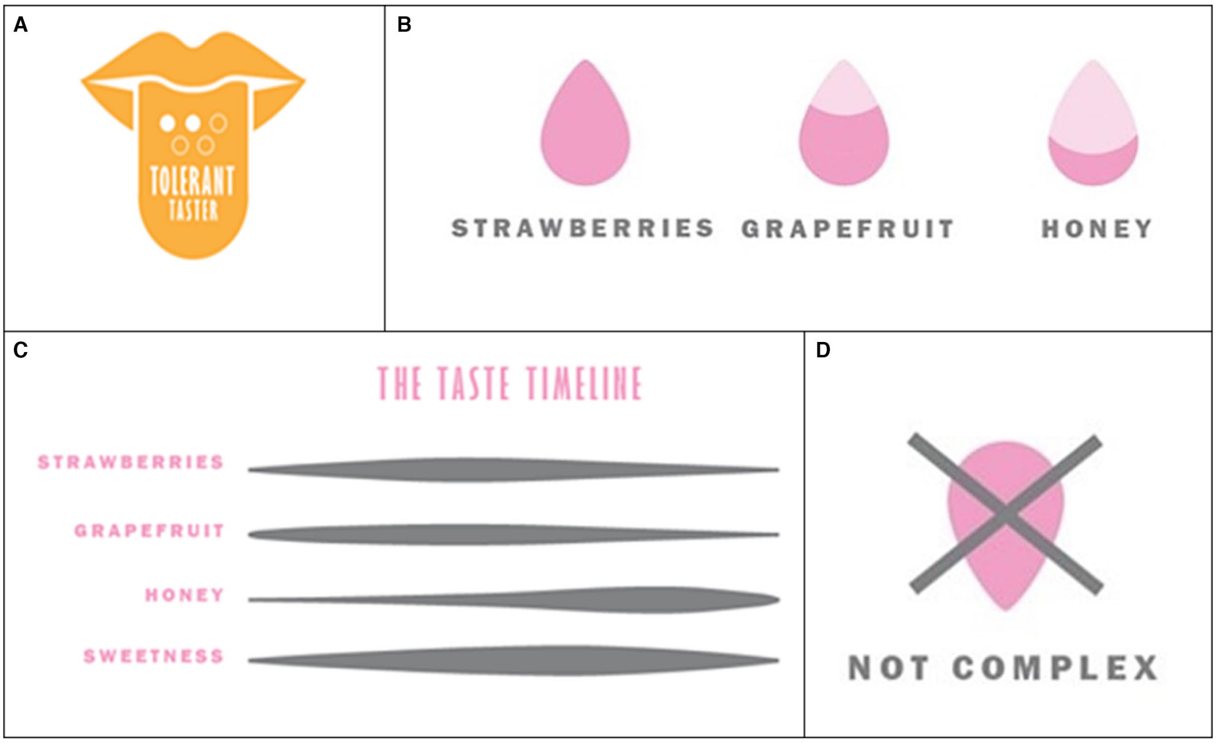

FIGURE 1
Symbol in the questionnaire used to illustrate an alternative communication approach based on; **(A)** genetics and sensitivity; **(B)** dominant flavors in the wine; **(C)** the temporal dynamic change during the tasting experience – from when the wine hits the nose and mouth (attack), development on the palate (mid-palate), and the duration of flavors in the end (finish); **(D)** of non-existent qualities in the wine – such as the lack of complexity.

3.4. Bivariate analysis of consumer attitudes toward alternative wine label communication related to reported demographics and purchase and consumption behaviors

The bivariate relationships between attitudes, demographics, and behaviors were analyzed *via* the Kendall's Tau rank correlation coefficient, except for the non-binary nominal demographical factors, where the Kruskal–Wallis test of equality of average rank among groups was used.

3.4.1. Results – associations between demographics and the perceived helpfulness of alternative label communication

As shown in Table 4 and Supplementary Appendix D, no significant associations were found between any of the demographic traits and the helpfulness of certain communication—neither the optimal consumer taste group based on genetics, nor the taste timeline in order to match a certain preference, nor the intensity of the dominating flavors. Likewise, the helpfulness of alternative label communication was not significantly associated with where the respondents happened to live, their knowledge about wine, where they primarily consume their wine, or the degree to which they select wine based on sustainability (decreasing the climate footprint). There were, however, certain significant associations:

- The helpfulness of being informed of properties that the wines do not have was positively correlated with how often respondents consume wine.
- The helpfulness of professionals to communicate the quality was negatively correlated with being born outside of Sweden and positively correlated with age, the level of highest completed education, and being a Swedish resident.
- The helpfulness of a professional assessment of the consumption readiness of the wine was positively correlated with being female, the level of highest completed education, and the level of income.
- The helpfulness of communicating potential faults was positively correlated with age (q2) and where respondents primarily purchased wine.

For more details, see Supplementary Appendix D.

3.4.2. Results – correlations between purchasing behavior and perceived helpfulness of alternative label communication

As shown in Table 4, no signification correlations were found between the helpfulness of being informed about properties that the wines did not have and purchasing behavior. Similarly, the helpfulness of alternative label communication was not significantly correlated with the factors influencing the choice of wine such as price, grape, country of origin, climate impact, style, vintage, front label, back label, and label illustrations. The following significant correlations were found with factors influencing purchasing behavior:

FIGURE 2
An image that was included in the questionnaire, intended to illustrate an innovative method of label communication. This approach involves using relabeled wine bottles featuring solely sensory descriptors. The design of this concept was created by Elin Aronsen Beis.

TABLE 1 Attitudes toward tested symbols.

Frequency (*n* = 329)	Genetics (%)	Flavor intensity (%)	Taste timeline (%)	Non-existing qualities (%)*
7. Very helpful	6.6	18.0	14.2	11.4
6.	9.8	29.3	20.2	17.0
5.	32.2	29.0	30.9	23.7
4. Neutral	32.2	13.3	20.2	21.5
3.	7.6	6.6	7.3	7.9
2.	6.6	1.9	3.8	8.8
1. Not very helpful	5.1	1.9	3.5	9.8

* Communication of absent sensory qualities that might affect acceptance on individual level.

TABLE 2 Attitudes toward other assessments/scale.

Frequency (n = 329)	Quality assessed by professional wine expert/ Poor – Outstanding. (%)	Readiness (to drink) assessed by professional wine expert/ Too young – Too old. (%)	High-risk faults assessed by professional wine expert/ Frequent wine faults. (%)
7. Very helpful	19.2	40.1	12.0
6.	28.4	27.4	18.3
5.	23.0	21.5	23.0
4. Neutral	15.1	6.6	23.7
3.	6.0	1.9	11.7
2.	5.0	0.3	5.0
1. Not very helpful	3.2	2.2	6.3

Only text and no symbol presented.

- The helpfulness of the optimal consumer taste group based on genetics was negatively correlated with previous experience.
- The helpfulness of the taste timeline in order to match a certain preference was positively correlated with sensory indicators.
- The helpfulness of the intensity of the dominating flavors was positively correlated with sensory indicators and external recommendations.
- The helpfulness of professionals to communicate the quality was negatively correlated with the wine producer and bottle design.
- The helpfulness of professional assessment of wines readiness was positively correlated with external recommendations.
- The helpfulness of communicating potential faults was positively correlated with previous experience.

3.4.3. Results – correlations between perceived helpfulness of alternative label communication and consumption behavior

Regarding the question concerning influencing factors during the consumption of wine and the sensory experience, no significant correlations could be found between influencing factors and a preference for certain label communication. One possible reason for this might be the complexity of the question, which presupposes an understanding of the influencing factors. The lack of understanding of these influencing factors and their impact in the consumption context may be potential areas of further exploration to improve consumer communication (see Supplementary Appendix E).

3.5. Regression analysis of perceived helpfulness of alternative wine label communication in relation to reported demographics, purchase behaviors, and consumption behaviors

To compare the rating of the various attitudes toward the helpfulness of different types of alternative communication as well as

TABLE 3 Other attitudes regarding sensory labeling, blending wines and sustainability.

Choice (*n* = 329)	Sensory information (%)	Blending (%)	Choice (*n* = 329)	Sustainability (%)
This would suit me	1.6	3.5	Always	2.2
This sounds like something that I would like	15.9	14.9	Most of the time	22.9
I would be willing to try a bottle	46.0	55.2	Sometimes	50.8
I might if the descriptors suited my palate	27.9	21.0	Rarely	18.4
Never	8.6	5.4	Never	5.7

TABLE 4 Correlations (Kendall's tau) between demographics and perceived helpfulness of alternative label communication.

	Perceived helpfulness of alternative label communication						
Demographical variable	Genetics	Flavor intensity	Taste timeline	Non-existing qualities (x)	Quality by pro	Readiness (to drink)	High-risk faults
Age	0.029	−0.012	−0.004	0.038	0.083 *	0.080	0.144 ***
Highest completed education	−0.027	−0.013	−0.003	0.083	0.112 *	0.131 **	0.068
How often do you consume wine?	0.025	−0.004	0.026	0.105 *	0.050	0.068	0.074
How much do you know about wine?	−0.065	−0.062	−0.018	0.008	0.008	−0.037	−0.035
What is your monthly income? (Before taxes.)	−0.014	−0.033	0.01	0.078	0.078	0.132 **	0.048
To what degree do you select wine in regard to sustainability - decreasing the climate footprint?	0.061	0.040	0.012	0.001	−0.026	0.001	−0.061

Two-tailed value of p: 0.05 > * > 0.01 > ** > 0.001 > ***. (x) Communication of absent sensory qualities that might affect acceptance on individual level.

their relationship to demographics, purchase behaviors, and consumption behaviors, a cumulative linked mixed model was fitted. The within-participant rating correlation was modeled by participant random intercepts. First, the ratings were regressed on the; variable categories; demographics; purchase behaviors; and consumption behaviors. Then, the non-significant variables were eliminated until all the remaining estimated effects significantly differed from zero, see Table 5.

The magnitude of the helpfulness questions (value of $p = 0.000$) resembled the univariate results in Sections 3.1 and 3.2, where the level of readiness had the highest and genetics and taste sensitivity profiling and non-existing qualities had the lowest helpfulness rating, while the other categories fell in between. The demographical and behavioral variables found to be significant in the regression provide indications of general patterns among attitudes toward the helpfulness of alternative wine label communication. All of them also had some significant bivariate relationships to the helpfulness questions (see Section 3.4). Being older (value of $p = 0.014$) and being female (value of $p = 0.035$) both had positive estimated effects, corresponding to higher expected helpfulness ratings. Furthermore, wine purchasing behavior influenced by sensory indicators had a positive effect (value of $p = 0.009$) while primarily consuming wine in restaurants and bars (value of $p = 0.005$) had a negative effect (Table 6).

4. Discussion

In general, the results of the questionnaire revealed an overall positive response for all tested symbols. A broader question is whether this is an indicator, or signal, of the need to introduce alternative ways of communicating about wine (and hence, by extension, other sensorially-complex products). It also raises the question of whether knowledge within the field of crossmodal correspondence could further transfer into the field of consumer communication to optimize consumer-product matching. This study thus aligns with earlier research aiming to improve matching between consumer groups and potential products (Elder and Krishna, 2010; Varela and Ares, 2012; Krishna and Schwarz, 2014; Croijmans and Wang, 2021), but with a different motivation for implementing these strategies.

The respondents showed an overall positive response to the symbols used in the study, especially those focusing on the multisensory experience and indicators focusing on dominant flavors and their dynamic change during the expected palate experience by

TABLE 5 Correlations (Kendall's tau) between purchasing behavior and perceived helpfulness of alternative label communication.

	Perceived helpfulness of alternative label communication						
When purchasing wine from your selected choice in the above question, what factors influence your choice?	Genetics	Flavor intensity	Taste timeline	Non-existing qualities (x)	Quality by pro	Readiness (to drink)	High-risk faults
The country of origin	−0.023	−0.003	−0.03	0.020	0.029	0.015	−0.003
The grape	0.026	0.051	−0.013	0.085	0.034	0.037	0.041
The style	−0.003	0.039	0.030	0.022	0.060	−0.001	−0.017
Illustrations on the label	−0.012	0.065	−0.028	0.059	−0.020	−0.033	−0.030
The wine producer	−0.006	−0.008	0.014	0.068	−0.133 **	0.053	0.040
Sensory indicators	0.039	0.153 **	0.103 *	0.037	0.038	0.078	0.095
External recommendations	−0.075	−0.051	0.121 *	0.028	0.004	0.120 *	0.017
Previous experience	−0.102 *	−0.080	0.000	0.018	0.045	−0.012	0.144 **
Climate impact	−0.014	0.041	0.002	−0.031	0.005	0.049	−0.059
Price	−0.020	0.029	0.009	−0.010	−0.056	−0.026	−0.027
Bottle design	−0.069	−0.036	−0.023	0.034	−0.117 *	−0.014	−0.055
Front label	−0.032	0.029	0.002	0.079	−0.07	−0.015	−0.064
Back label	−0.085	−0.050	−0.042	0.082	0.002	−0.087	−0.096
Vintage	−0.025	−0.094	−0.065	0.037	0.029	−0.003	0.035

Two-sided value of p: 0.05 > * > 0.01 > ** > 0.001 > ***. (x) Communication of absent sensory qualities in the product that might affect acceptance on individual level.

TABLE 6 Significant effects from the regression of attitudinal ratings toward the perceived helpfulness of alternative communication on demographical, consumption and purchasing behavioral variables (cumulative linked mixed model).

	Perceived helpfulness of alternative label communication variable						
Category	Genetics	Flavor intensity	Taste timeline	Non-existing qualities (x)	Quality by pro	Readiness (to drink)	High-risk faults
Estimated effect (standard error)	0	0.83 (0.14) ***	1.33 (0.15) ***	0.13 (0.14)	1.18 (0.15) ***	2.46 (0.16) ***	0.37 (0.14) **

	Demographic, consumption and purchasing behavioral variables						
Variable and category	Age	Gender		Consume in		Influential on choice	
	10-year effect	Male	Female	Other	Bar or restaurant	Sensory indicator No	Sensory indicator Yes
Estimated effect (standard error)	0.14 (0.06) *	0	0.34 (0.16) *	0	−0.93 (0.35) **	0	0.44 (0.16) **

Two-sided value of p: 0.05 > * > 0.01 > ** > 0.001 > ***. (x) Communication of absent sensory qualities that might affect acceptance on individual level.

the consumer. This indicates that symbols may be a consumer-friendly tool in communicating both multisensory changes and more holistic sensory profiles of wines, conveying the overall expected sensory experience of appearance, odor, taste, and tactile sensations. This finding thus answers the call of semiotic studies to further investigate the use of symbols as a potential tool when it comes to communicating sensory attributes and flavor profiles (König and Lick, 2014; van Tonder and Mulder, 2015; Celhay and Remaud, 2018; Pelet et al., 2020). Furthermore, since there was no significant correlation between demographics and flavor intensity, timeline, or the most popular symbols, this type of visual symbolic approach might be a useful crossmodal tool in addressing a broader population by not being a tool for communication to a specific demographic group (Hutmacher, 2019).

A significant majority of the respondents also demonstrated a positive inclination toward seeking evaluations from wine experts to assess aspects like the wine's readiness, quality, and potential flaws. This inclination underscores the inherent sensory intricacies associated with wine as a product, as well as the persistent desire to enhance communication strategies aimed at bridging the gap between sensory expectations and the ensuing sensory experience. This need for expert assessment sheds light on a possible explanation for the proliferation of websites and mobile applications that prioritize assisting consumers in making informed wine purchases, examples of which include vivino.com, wine-searcher.com, and cellertracker.com.

These online platforms typically offer a diverse array of communication tools designed to support consumers in selecting wine. These tools encompass quality ratings, geographical descriptions, insights into vinification and viticultural practices, purchasing and maturation guidance, and vintage charts. Importantly, they often transcend conventional information parameters such as wine origin, grape variety, producer details, vintage year, and legal specifications. The prevalence and diversity of these online resources may signify a growing recognition of the limitations inherent in traditional approaches to communicating about complex food products such as wine. It's plausible that the sheer number of these websites and the multifaceted communication tools they provide serve as a testament to the evolving landscape of wine communication, one that seeks to address the nuanced and multifaceted aspects of this sensory-rich domain.

In the present study, respondents' positive response toward symbols and other visual tools to be applied for crossmodal communication also supports earlier findings, which suggest that linguistic symbolic tools, like metaphors, analogies, metonymies, and allegories, may complement crossmodal communication (Paradis and Eeg-Olofsson, 2013; Moreno Lara, 2014; Alousque, 2015; Kelley et al., 2015; Costello et al., 2018; Herdenstam et al., 2020). Other aspects to consider when developing visual tools for crossmodal communication are which crucial sensory descriptors to select when trying to create olfactory mental images (Tomiczek and Stevenson, 2009) or a more general mental image of the overall sensory experience (Williamson et al., 2016; Croijmans and Wang, 2021; Spence and Van Doorn, 2022; Spence, 2022a).

Other attitudes regarding sensory labeling, blending wines, and sustainability indicate a positive response toward wine with alternative labeling using symbols and strict sensory information. Respondents' positive response toward testing wine that had been made by blending existing wines, see Wang and Spence (2018a,b), combined with their willingness to make sustainable choices when purchasing, indicate the potential for future wine rescuing programs,. Such programs could use different batches of wines that are left over due to overproduction and/or changes in sensory profile.

Within the respondent group under examination in this study, it was observed that individuals who primarily relied on their prior wine experiences during the purchasing process exhibited reduced interest in communication that pertained to consumer taste group classifications based on genetics. This finding suggests that once consumers have identified a particular style or type of wine that aligns with their preferences, it becomes a potent determinant for their future wine purchases. This influence seems to outweigh the significance of genetic classifications, which can often be challenging to relate to. An alternative explanation could be rooted in the substantial body of research focusing on genetics and preference, which, due to its complexity, may be challenging for consumers to grasp. This complexity arises from the multitude of variables beyond genetics, including environmental factors and cultural influences, which, to a certain extent, necessitate self-awareness, a foundational understanding of genetics, and knowledge of how this genetic information corresponds to their individual sensory experiences (Bartoshuk et al., 1996; Bartoshuk, 2000; Keast and Breslin, 2003; Reed and Knaapila, 2010; Kim et al., 2015; Linscott and Lim, 2016; Herdenstam et al., 2018; Pagliarini et al., 2021).

Conversely, the positive correlations identified between purchasing behaviors and the examined multisensory symbols—comprising the intensity and composition of dominant flavors, as well as the temporal development of pivotal sensory flavors—suggest that these symbols possess potential as crossmodal tools. These tools use visual cues to communicate not only taste, aroma, and tactile sensations, but also the anticipated progression of taste experiences on the palate. This holistic approach aids consumers in grasping the sensory encounter comprehensively. Furthermore, the outcomes of this study could have implications for the context in which the wine is consumed and potential recommendations for certain food pairings that help to enhance the attributes of the wine in the context of the wine-food matching (Harrington, 2005, 2008; Koone et al., 2014; Herdenstam et al., 2018). This study's focus on vision as a crossmodal tool for communication highlights one part of the multisensory reality that the consumer faces, whether it is the purchasing or the consuming situation or both. The environment in which individuals interact with wine labels is a multisensory, atmospheric, and crossmodal experience on many levels, as has been proposed by earlier studies investigating the multisensory environment (Spence et al., 2014; Spence, 2020b, 2022a).

5. Conclusion

Applying alternative labeling approaches with sensory indicators and symbols may better communicate the expected sensory experience in relation to different consumers and their actual preferences. Taken together, accomplishing better communication for food products—in this case wine, which has been refined at many levels, from cultivation, production, maturation, and storage to final distribution to end consumer—also contributes to improved use of natural resources, thus decreasing the climate footprint. Based on the results of the present study, visual crossmodal communication may potentially both grasp critical consumer attributes and convey multisensory experiences, as well as the holistic timeline of taste development. This form of communication may thus be a useful tool in communicating wine and other complex food products.

6. Limitations and further research

While the study provides valuable insights into the potential of alternative labeling approaches and visual crossmodal communication for wine and other complex food products, it is important to acknowledge some limitations. The study might have benefited from a larger and more diverse sample. The respondents' demographics and wine preferences could have been more varied to obtain a broader perspective on the effectiveness of the symbols and visual tools across different consumer groups. The study primarily focused on the visual aspect of crossmodal communication and did not extensively consider other contextual factors that influence wine perception, such as the environment, social context, or individual differences in sensory sensitivity. Future research could explore the interaction between visual symbols and these contextual factors to gain a more comprehensive understanding of crossmodal communication. Investigating the cultural and individual differences in symbol interpretation and understanding would provide valuable insights for effective crossmodal communication.

Future research could address these limitations by conducting larger-scale studies with diverse samples, considering contextual factors, investigating symbol interpretation and design optimization, examining long-term effects, exploring practical implementation challenges, and extending the scope to other sensory-complex products. By addressing these areas, researchers can further advance the understanding and application of crossmodal communication strategies in consumer product matching.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

AC-F: manuscript and direction. CS: manuscript and direction. MM: philosophical input. NP: statistics and general analysis. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would especially like to acknowledge Markus Ekman, School of Hospitality, Culinary Arts & Meal Science at Örebro University, for collecting the data. This study is the second pilot within the multidisciplinary research project *Rewine The World*®, which has a mission to rescue wine and other complex food products by combining knowledge from different research areas and fields of expertise to increase the use of global resources.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1176794/full#supplementary-material

References

Alousque, I. N. (2015). Visual wine metaphor and metonymy in ads. *Procedia Soc. Behav. Sci.* 173, 125–131. doi: 10.1016/j.sbspro.2015.02.041

Bartoshuk, L. M. (2000). Comparing sensory experiences across individuals: recent psychophysical advances illuminate genetic variation in taste perception. *Chem. Senses* 25, 447–460. doi: 10.1093/chemse/25.4.447

Bartoshuk, L. M., Duffy, V. B., Lucchina, L. A., Prutkin, J., and Fast, K. (1998). PROP (6-n-propylthiouracil) supertasters and the saltiness of NaCl. *Ann. N. Y. Acad. Sci.* 855, 793–796. doi: 10.1111/j.1749-6632.1998.tb10660.x

Bartoshuk, L. M., Duffy, V. B., Reed, D., and Williams, A. (1996). Supertasting, earaches and head injury: genetics and pathology alter our taste worlds. *Neurosci. Biobehav. Rev.* 20, 79–87. doi: 10.1016/0149-7634(95)00042-D

Bartoshuk, L. M., Rifkin, B., Marks, L. E., and Hooper, J. E. (1988). Bitterness of KCl and benzoate: related to genetic status for sensitivity to PTC/PROP. *Chem. Senses* 13, 517–528. doi: 10.1093/chemse/13.4.517

Biswas, D., Szocs, C., Chacko, R., and Wansink, B. (2017). Shining light on atmospherics: how ambient light influences food choices. *J. Mark. Res.* 54, 111–123. doi: 10.1509/jmr.14.0115

Breslin, P. A. S. (1996). Interactions among salty, sour and bitter compounds. *Trends Food Sci. Technol.* 7, 390–399. doi: 10.1016/S0924-2244(96)10039-X

Breslin, P. A. S., and Beauchamp, G. K. (1997). Salt enhances flavour by suppressing bitterness. *Nature* 387:563. doi: 10.1038/42388

Celhay, F., and Remaud, H. (2018). What does your wine label mean to consumers? A semiotic investigation of Bordeaux wine visual codes. *Food Qual. Prefer.* 65, 129–145. doi: 10.1016/j.foodqual.2017.10.020

Christ, K. L., and Burritt, R. L. (2013). Critical environmental concerns in wine production: an integrative review. *J. Clean. Prod.* 53, 232–242. doi: 10.1016/j.jclepro.2013.04.007

Costello, M., Smith, R., and Lane, L. (2018). Australian wine labels: terroir without terror. *Gastronomica* 18, 54–65. doi: 10.1525/gfc.2018.18.3.54

Crisinel, A.-S., and Spence, C. (2012). A fruity note: Crossmodal associations between odors and musical notes. *Chem. Senses* 37, 151–158. doi: 10.1093/chemse/bjr085

Croijmans, I., and Wang, Q. J. (2021). Do you want a description with that wine? The role of wine mental imagery in consumer's desire to drink using the revised vividness of wine imagery questionnaire (VWIQ-II). *J. Sens. Stud.* 37:e12712. doi: 10.1111/joss.12712

De Luca, M., Campo, R., and Lee, R. (2019). Mozart or pop music? Effects of background music on wine consumers. *Int. J. Wine Bus. Res.* 31, 406–418. doi: 10.1108/IJWBR-01-2018-0001

Deroy, O., Crisinel, A.-S., and Spence, C. (2013). Crossmodal correspondences between odors and contingent features: odors, musical notes, and geometrical shapes. *Psychon. Bull. Rev.* 20, 878–896. doi: 10.3758/s13423-013-0397-0

Elder, R. S., and Krishna, A. (2010). The effects of advertising copy on sensory thoughts and perceived taste. *J. Consum. Res.* 36, 748–756. doi: 10.1086/605327

Etzi, R., Spence, C., and Gallace, A. (2014). Textures that we like to touch: an experimental study of aesthetic preferences for tactile stimuli. *Conscious. Cogn.* 29, 178–188. doi: 10.1016/j.concog.2014.08.011

Galbreath, J., Tisch, D., Quaddus, M., and Rabbanee, F. (2020). The impact of climate change on firm adaptation: the case of the wine industry. *Int. J. Wine Bus. Res.* 32, 373–389. doi: 10.1108/IJWBR-07-2019-0045

Gallace, A., and Spence, C. (2010). The science of interpersonal touch: an overview. *Neurosci. Biobehav. Rev.* 34, 246–259. doi: 10.1016/j.neubiorev.2008.10.004

Gent, J. F., and Bartoshuk, L. M. (1983). Sweetness of sucrose, neohesperidin dihydrochalcone, and saccharin is related to genetic ability to taste the bitter substance 6-n-propylthiouracil. *Chem. Senses* 7, 265–272. doi: 10.1093/chemse/7.3-4.265

Goldkuhl, L., and Styvén, M. (2007). Sensing the scent of service success. *Eur. J. Mark.* 41, 1297–1305. doi: 10.1108/03090560710821189

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., et al. (2006). Reading cinnamon activates olfactory brain regions. *NeuroImage* 32, 906–912. doi: 10.1016/j.neuroimage.2006.03.037

Hagtvedt, H., and Brasel, S. A. (2016). Cross-modal communication: sound frequency influences consumer responses to color lightness. *J. Mark. Res.* 53, 551–562. doi: 10.1509/jmr.14.0414

Hanson-Vaux, G., Crisinel, A.-S., and Spence, C. (2013). Smelling shapes: Crossmodal correspondences between odors and shapes. *Chem. Senses* 38, 161–166. doi: 10.1093/chemse/bjs087

Harrington, R. J. (2005). The wine and food pairing process. *J. Culinary Sci. Technol.* 4, 101–112. doi: 10.1300/J385v04n01_11

Harrington, R. J. (2007). The impact of oak: implications for food and wine pairing. *J. Culinary Sci. Technol.* 5, 97–102. doi: 10.1300/J385v05n04_08

Harrington, R. J. (2008). *Food & wine pairing: a sensory experience*. Hoboken, NJ: John Wiley.

Harrington, R. J., and Hammond, R. (2006). Body deviation-from-match. *J. Culinary Sci. Technol.* 5, 51–69. doi: 10.1300/J385v05n01_06

Harrington, R. J., and Hammond, R. (2009). The impact of wine effervescence levels on perceived palatability with salty and bitter foods. *J. Foodserv. Bus. Res.* 12, 234–246. doi: 10.1080/15378020903158509

Harrington, R. J., McCarthy, M., and Gozzi, M. (2010). Perceived match of wine and cheese and the impact of additional food elements: a preliminary study. *J. Foodserv. Bus. Res.* 13, 311–330. doi: 10.1080/15378020.2010.524541

Harrington, R. J., and Seo, H.-S. (2015). The impact of liking of wine and food items on perceptions of wine–food pairing. *J. Foodserv. Bus. Res.* 18, 489–501. doi: 10.1080/15378020.2015.1093455

Heatherly, M., Dein, M., Munafo, J. P., and Luckett, C. R. (2019). Crossmodal correspondence between color, shapes, and wine odors. *Food Qual. Prefer.* 71, 395–405. doi: 10.1016/j.foodqual.2018.08.019

Herdenstam, A., Nilsen, A. N., and Öström, Å. (2020). Breaking the silence: a pilot study investigating communication skills of sommeliers and chefs after analogical training. *Int. J. Gastron. Food Sci.* 20:100210. doi: 10.1016/j.ijgfs.2020.100210

Herdenstam, A., Nilsen, A., Öström, Å., and Harrington, R. J. (2018). Sommelier training – dialogue seminars and repertory grid method in combination as a pedagogical tool. *Int. J. Gastron. Food Sci.* 13, 78–89. doi: 10.1016/j.ijgfs.2018.07.002

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Front. Psychol.* 10. doi: 10.3389/fpsyg.2019.02246

Iannone, R., Miranda, S., Riemma, S., and De Marco, I. (2016). Improving environmental performances in wine production by a life cycle assessment analysis. *J. Clean. Prod.* 111, 172–180. doi: 10.1016/j.jclepro.2015.04.006

Keast, R. S. J., and Breslin, P. A. S. (2003). An overview of binary taste–taste interactions. *Food Qual. Prefer.* 14, 111–124. doi: 10.1016/S0950-3293(02)00110-6

Kelley, K., Hyde, J., and Bruwer, J. (2015). U.S. wine consumer preferences for bottle characteristics, back label extrinsic cues and wine composition: a conjoint analysis. *Asia Pac. J. Mark. Logist.* 27, 516–534. doi: 10.1108/APJML-09-2014-0140

Kim, M. J., Son, H. J., Kim, Y., Misaka, T., and Rhyu, M.-R. (2015). Umami–bitter interactions: the suppression of bitterness by umami peptides via human bitter taste receptor. *Biochem. Biophys. Res. Commun.* 456, 586–590. doi: 10.1016/j.bbrc.2014.11.114

König, B., and Lick, E. (2014). Wine labels in Austrian food retail stores: a semiotic analysis of multimodal red wine labels. *Semiotica* 2014, 313–334. doi: 10.1515/sem-2014-0014

Koone, R., Harrington, R. J., Gozzi, M., and McCarthy, M. (2014). The role of acidity, sweetness, tannin and consumer knowledge on wine and food match perceptions. *J. Wine Res.* 25, 158–174. doi: 10.1080/09571264.2014.899491

Krishna, A. (2010). *Sensory marketing: Research on the sensuality of products*. New York, NY: Routledge.

Krishna, A., and Schwarz, N. (2014). Sensory marketing, embodiment, and grounded cognition: a review and introduction. *J. Consum. Psychol.* 24, 159–168. doi: 10.1016/j.jcps.2013.12.006

Lick, E., König, B., Kpossa, M. R., and Buller, V. (2017). Sensory expectations generated by colours of red wine labels. *J. Retail. Consum. Serv.* 37, 146–158. doi: 10.1016/j.jretconser.2016.07.005

Linscott, T. D., and Lim, J. (2016). Retronasal odor enhancement by salty and umami tastes. *Food Qual. Prefer.* 48, 1–10. doi: 10.1016/j.foodqual.2015.08.004

Liu, S. Q., Bogicevic, V., and Mattila, A. S. (2018). Circular vs. angular servicescape: "shaping" customer response to a fast service encounter pace. *J. Bus. Res.* 89, 47–56. doi: 10.1016/j.jbusres.2018.04.007

Maziriri, E. T., Rukuni, T. F., and Chuchu, T. (2021). Factors influencing food consumption satisfaction and purchase decisions of restaurant consumers. *Cogent Business Manage* 8. doi: 10.1080/23311975.2021.1968731

Moreno Lara, M. A. (2014). Multimodal representations of metaphor and metonymy on the labels of wine from the denominacion de origen of Rioja. *Revist. Española Lingüíst. Aplic.* 27, 454–468. doi: 10.1075/resla.27.2.10mor

Olzak, L. A., and Craig, J. C. (2014). Review: visual psychophysics: from laboratory to theory, in touch with the future: the sense of touch from cognitive neuroscience to virtual reality. *Perception* 43, 711–714. doi: 10.1068/p4307rvw

Pagliarini, E., Proserpio, C., Spinelli, S., Lavelli, V., Laureati, M., Arena, E., et al. (2021). The role of sour and bitter perception in liking, familiarity and choice for phenol-rich plant-based foods. *Food Qual. Prefer.* 93:104250. doi: 10.1016/j.foodqual.2021.104250

Paradis, C., and Eeg-Olofsson, M. (2013). Describing sensory experience: the genre of wine reviews. *Metaphor. Symb.* 28, 22–40. doi: 10.1080/10926488.2013.742838

Parr, W. V. (2015). Unraveling the nature of perceived complexity in wine. *Practical Winery & Vineyard* 5–8.

Pelet, J.-É., Durrieu, F., and Lick, E. (2020). Label design of wines sold online: effects of perceived authenticity on purchase intentions. *J. Retail. Consum. Serv.* 55:102087. doi: 10.1016/j.jretconser.2020.102087

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org

Reed, D. R., and Knaapila, A. (2010). "Chapter 8- genetics of taste and smell: poisons and pleasures" in *Progress in molecular biology and translational science*. ed. C. Bouchard, vol. *94* (Netherlands: Academic Press), 213–240.

Reid, C. A., Green, J. D., Wildschut, T., and Sedikides, C. (2015). Scent-evoked nostalgia. *Memory* 23, 157–166. doi: 10.1080/09658211.2013.876048

Shepherd, G. M. (2006). Smell images and the flavour system in the human brain. *Nature* 444, 316–321. doi: 10.1038/nature05405

Spence, C. (2011). Crystal clear or gobbletigook? *World Fine Wine* 33, 96–101.

Spence, C. (2020a). Designing for the multisensory mind. *Archit. Des.* 90, 42–49. doi: 10.1002/ad.2630

Spence, C. (2020b). Senses of place: architectural design for the multisensory mind. *Cognit. Res.* 5:46. doi: 10.1186/s41235-020-00243-4

Spence, C. (2022a). Experimental atmospherics: a multi-sensory perspective. *Qual. Mark. Res. Int. J.* 25, 662–673. doi: 10.1108/QMR-04-2022-0070

Spence, C. (2022b). On the use of ambient odours to influence the multisensory experience of dining. *Int. J. Gastron. Food Sci.* 27:100444. doi: 10.1016/j.ijgfs.2021.100444

Spence, C., Hobkinson, C., Gallace, A., and Piqueras Fiszman, B. (2013). A touch of gastronomy. *Flavour* 2, 1–15. doi: 10.1186/2044-7248-2-14

Spence, C., and Van Doorn, G. (2022). Visual communication via the design of food and beverage packaging. *Cognit. Res.* 7:42. doi: 10.1186/s41235-022-00391-9

Spence, C., Velasco, C., and Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour* 3:8. doi: 10.1186/2044-7248-3-8

Spence, C., and Wang, Q. J. (2018). What does the term 'complexity' mean in the world of wine? *Int. J. Gastron. Food Sci.* 14, 45–54. doi: 10.1016/j.ijgfs.2018.10.002

Tomiczek, C., and Stevenson, R. J. (2009). Olfactory imagery and repetition priming: the effect of odor naming and imagery ability. *Exp. Psychol.* 56, 397–408. doi: 10.1027/1618-3169.56.6.397

van Tonder, E. M., and Mulder, D. (2015). Marketing communication for organic wine: semiotic guidelines for wine bottle front labels. *Communication* 41, 131–151. doi: 10.1080/02500167.2015.1011179

Varela, P., and Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Res. Int.* 48, 893–908. doi: 10.1016/j.foodres.2012.06.037

Wang, M., and Li, D. (2022). Visualise the tastes from the label: a study on the taste-colour crossmodal association of crisp and dry. *Front. Psychol.* 13. doi: 10.3389/fpsyg.2022.980049

Wang, Q. J., and Spence, C. (2018a). A smooth wine? Haptic influences on wine evaluation. *Int. J. Gastron. Food Sci.* 14, 9–13. doi: 10.1016/j.ijgfs.2018.08.002

Wang, Q. J., and Spence, C. (2018b). Wine complexity: an empirical investigation. *Food Qual. Prefer.* 68, 238–244. doi: 10.1016/j.foodqual.2018.03.011

Ward, R. J., Rahman, S., Wuerger, S., and Marshall, A. (2022). Predicting the crossmodal correspondences of odors using an electronic nose. *Heliyon* 8:e09284. doi: 10.1016/j.heliyon.2022.e09284

Weil, R. L. (2007). Debunking critics' wine words: can amateurs distinguish the smell of asphalt from the taste of cherries? *J Wine Econ.* 2, 136–144. doi: 10.1017/S1931436100000390

Williamson, P. O., Lockshin, L., Francis, I. L., and Mueller Loose, S. (2016). Influencing consumer choice: short and medium term effect of country of origin information on wine choice. *Food Qual. Prefer.* 51, 89–99. doi: 10.1016/j.foodqual.2016.02.018

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 October 2023
DOI 10.3389/fpsyg.2023.1175703

Check for updates

OPEN ACCESS

EDITED BY
Thomas Alrik Sørensen,
Aalborg University, Denmark

REVIEWED BY
Patrick Bruns,
University of Hamburg, Germany
Nicola Di Stefano,
National Research Council (CNR), Italy

*CORRESPONDENCE
Ryan J. Ward
r.j.ward@ljmu.ac.uk

RECEIVED 28 February 2023
ACCEPTED 28 August 2023
PUBLISHED 06 October 2023

CITATION
Ward RJ, Ashraf M, Wuerger S and Marshall A (2023) Odors modulate color appearance.
Front. Psychol. 14:1175703.
doi: 10.3389/fpsyg.2023.1175703

Odors modulate color appearance

Ryan J. Ward[1,2]*, Maliha Ashraf[3], Sophie Wuerger[4] and Alan Marshall[2]

[1]Department of Computer Science and Mathematics, Liverpool John Moores University, Liverpool, United Kingdom, [2]Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, United Kingdom, [3]Department of Psychology, University of Liverpool, Liverpool, United Kingdom, [4]Department of Computer Science and Technology, Cambridge University, Cambridge, United Kingdom

Our brain constantly combines multisensory information from our surrounding environment. Odors for instance are often perceived with visual cues; these sensations interact to form our own subjective experience. This integration process can have a profound impact on the resulting experience and can alter our subjective reality. Crossmodal correspondences are the consistent associations between stimulus features in different sensory modalities. These correspondences are presumed to be bidirectional in nature and have been shown to influence our perception in a variety of different sensory modalities. Vision is dominant in our multisensory perception and can influence how we perceive information in our other senses, including olfaction. We explored the effect that different odors have on human color perception by presenting olfactory stimuli while asking observers to adjust a color patch to be devoid of hue (neutral gray task). We found a shift in the perceived neutral gray point to be biased toward warmer colors. Four out of five of our odors also trend toward their expected crossmodal correspondences. For instance, when asking observers to perform the neutral gray task while presenting the smell of cherry, the perceptually achromatic stimulus was biased toward a red-brown. Using an achromatic adjustment task, we were able to demonstrate a small but systematic effect of the presence of odors on human color perception.

KEYWORDS

crossmodal correspondences, crossmodal associations, odors, olfaction, colors, neutral gray, color perception, olfactory perception

1. Introduction

In our everyday life, we are simultaneously bombarded with information from different sensory modalities. Our brain combines this information to better understand our surrounding environment (Calvert et al., 2005); this integration process has been shown to influence our perception in different senses, for example (Morrot et al., 2001). Crossmodal correspondences is the tendency for a sensory attribute to be associated with a stimulus feature in a different sensory modality (Spence, 2011; Spence and Parise, 2012). For instance, people have consistent correspondences between odors and a variety of different sensory modalities, including but not limited to the angularity of shapes, the smoothness of texture, pitch, colors, and musical dimensions (Ward R. J. et al., 2021). These correspondences still occur outside of olfaction, including, but not limited to, sound-taste (Knöferle and Spence, 2012), temperature-color (Motoki et al., 2019), and pitch-vertical position (Bernstein and Edelstein, 1971). The nature and origin of these correspondences have diverse characterization in the literature with hedonics (Crisinel et al., 2013; Hanson-Vaux et al., 2013), semantics (Demattè et al., 2006; Jacquot et al., 2016), and natural co-occurrence (Spence, 2011; Spence and Deroy, 2013) being frequently deducted.

Colors have a profound effect on the perception of odors (Zellner and Kautz, 1990; Zellner and Whitten, 1999). For instance, observers may perceive an orange-flavored drink if it is colored orange when in fact, it is cherry-flavored (DuBose et al., 1980). A later study by Morrot et al. (2001) has shown that this effect can bias the judgment of expert wine tasters. That is when a glass of white wine is artificially colored red with an odorless dye; a panel described it as a red wine. Crossmodal correspondences are presumed to be bidirectional in nature (Evans and Treisman, 2010), meaning if odor-color correspondences are strong enough, one could expect the presence of odors to slightly influence our color perception. A study by Hansen et al. (2006) found that when asking observers to adjust the color of fruit objects to appear achromatic, participants overcorrected the color stimuli to the opponent color of the presented visual aid. That is, when asking their observers to adjust the color of a banana to a neutral gray, the observers adjusted the banana to a "blueish hue," which is the opponent color of yellow. Hansen et al. (2006) state that this is due to a visual memory modulation effect, meaning that knowledge of the color that corresponds with the identity of visual cue (e.g., banana being associated with the color yellow) would modulate the color appearance.

The existence of stable and strong crossmodal correspondences between odors and colors may reflect interactions taking place at a perceptual level (Demattè et al., 2009). For instance, olfactory crossmodal correspondences have recently been shown to be predictable based on their physical and chemical characteristics (Ward R. et al., 2021; Ward et al., 2022). Kemp and Gilbert (1997) asked participants to select color chips that best matched the presented odor. They found that the weaker the perceived intensity, the lighter the selected colors. Kemp and Gilbert (1997) state that this finding could be attributed to the interaction taking place at a perceptual level rather than a decisional (see also Gilbert et al., 1996). However, it is important to note the importance of semantic information and language in the nature and origin of odor-color correspondences (see de Valk et al., 2017; Kaeppler, 2018; Ward R. J. et al., 2021). In terms of olfactory discrimination and identification, color (Zellner et al., 1991; Narumi et al., 2010) and other visual information play an important role (Demattè et al., 2009).

In our prior work (Ward R. J. et al., 2021), we uncovered the crossmodal correspondences between odors and the angularity of shapes, smoothness of texture, pitch, colors, perceived pleasantness, and emotional/musical dimensions. Additionally, we probed the nature and origin of these correspondences. We found that the knowledge of an odor's identity (semantics) plays a role when judging the emotional and musical dimensions as well as color correspondences. Hedonics was also found to be the main mediator for the angularity of shapes, smoothness of texture, pleasantness, and pitch correspondences. Of particular interest for this manuscript is color correspondences where we found robust crossmodal correspondences between odors and colors, see Ward R. J. et al., (2021; Figure 3B). We used a subset of olfactory stimuli used in Ward R. J. et al. (2021) that induced the most consistent odor-color correspondences (caramel, cherry, coffee, lemon, and peppermint); in addition to these odors, we added a control (plain unscented water). Here we hypothesize that the presence of odors would influence the perception of color. Based on the findings of Hansen et al. (2006), we anticipated that when asking the observers to create a color patch that is achromatic (a neutral gray), the findings would reveal a shift to a color that is opposite to the one associated with the particular odor. That is, if the presented odor is primarily associated with the color yellow, we anticipate that the perceived neutral gray point would be shifted more toward blue.

2. Methodology and materials

2.1. Participants

Twenty-four people participated in the experiment (eleven males and thirteen females). The mean age of the participants was 29.12. The youngest participant was 20 years of age, and the oldest was 57 years of age. No participants reported any impairment affecting their sense of smell (i.e., cold, flu, COVID) or color vision (i.e., color blindness). Participants were briefed before the experiment started on potential allergens and the task. Participants were instructed not to wear any scented deodorants, or perfumes on the day of the experiment. Color vision was not tested before the observer's participation and was assumed based on the observer's report. The experiment was approved by the University of Liverpool and conducted in accordance with the declaration of Helsinki's standard for medical research involving human subjects. Observers were recruited via mass e-mail. The sample size for this study was considered from a variety of resources. For the circular statistics, a sample size of at least 20 is required (NCSS, 2010). To estimate the sample size needed for the rest of the statistics, *a priori* statistical power analysis was performed in G*Power. For a one-tailed t-test with a power of 0.9, $\alpha = 0.05$ and a computed effect size of 0.66, a sample size of at least 22 is required.

2.2. Experimental conditions

The windows in the room were blacked out using blackout window film and the lights were turned off during the experiment, to ensure consistent illumination. The observers were given 5 min to adjust to the dim light before starting the experiment. The task was performed in full-screen mode, to stop the observers from using a color located elsewhere on the screen to aid in their decision. This was to stop observers from using colors elsewhere on the screen (i.e., white text) in an endeavor to make a neutral gray. The task was conducted in four parts. The first part consisted of odor removal using an IQAir HealthPro 250 Air Purifier that lasted 4 min and was done even before the first odor was presented to the observer. In the second part, ambient odor was introduced into the experiment room using ultrasonic diffusers for 5 min. In the third part, the observer was asked to adjust a color patch, so that it was devoid of any hue (neutral gray task). In the fourth part, observers were asked to identify the odor that was being presented (classification task). This process was repeated until the observer had completed all four parts for each one of the four odors and the control. The timings for these parts were determined by a small test of five participants. This test involved running through a reduced version of the experiment, where only the odors were presented to the participants and involved noting down the timings for when either they could smell the odor that was being released or when they felt they could no longer smell the last presented odor. The largest timing across odors was used as the measurement for how long to present each odor and the largest timing for the smell evacuation process was used. Additionally, a minute was added to these timings. These timings are the timings used in the main experiment. The odors were presented in a random order to the participants in accordance with a number displayed on the screen. This number was an identifier, so the experimenter knew which odor to present and when. The color patch was adjusted five times for each condition. The diffusers were cleaned, and new smells were added each time the experiment was conducted. Diffusers were used for this experiment as a COVID safety measure. The experiment was programmed in MATLAB R2021b. See Figure 1 for a diagram of the experimental setup.

2.3. Neutral gray task

The neutral gray task involved asking the observers, "Please adjust the color patch so that it is neither red nor green and neither yellow nor blue" (Chauhan et al., 2014). Five repetitions were taken for each condition resulting in 25 recorded neutral gray adjustments for each observer. The observer's final neutral gray selections were saved in RGB format, then corrected and converted back into the CIELAB color space before any analyses took place. This color space was chosen due to its approximate perceptual uniformity; see (Luo, 2016) for more information. The colors were corrected to better reflect the color that was presented on the screen during the experiment, as the actual and requested colors may differ depending on the monitor's screen settings (e.g., brightness). When the observers were adjusting the color patch, they had two sliders available to them, one that changes the green to red (a*) value and another that changes the blue to yellow (b*) value. The RGB values of the observer's final neutral gray selections were saved. Before any analyses took place, the monitor's native RGB color space was analyzed using a spectroradiometer (SpectraScan® Spectroradiometer PR-670). The RGB values of the neutral gray selections of the participants were then converted into the CIELAB color space using a transformation from the monitor's calibration. This was done so that any color values reported in this section better reflect the colors displayed on the computer screen when the observers were making their judgments. See Figure 2 for a screenshot of the neutral gray task. See Section "Color stimuli" for information on the color stimuli.

2.4. Classification task

After the neutral gray selection task was completed for each condition, the observer was asked to identify the odor that was presented. This task involved selecting the odor the observer thought was presented from a pre-compiled list of 26 different options; one option was "Can not identify," and the remaining 25 elements were different odors, presented in alphabetical order. These odors are basil, camphor, caramel, cherry, chocolate, cinnamon, coffee, eucalyptus, freshly cut grass, ginger, grapefruit, lavender, lemon, lime, onion, orange, pepper, peppermint, pine, rose, sage, vanilla, wood, and ylang ylang.

2.5. Olfactory stimuli

Five different odors were used in the experiment; two from Miaroma™ (lemon and peppermint) and three from Mystic Moments™ (coffee, caramel, and cherry). These odors were chosen as they induced the most robust odor-color correspondences in our prior work, see Ward R. J. et al., (2021; Figure 3B). In addition to these odors, plain, unscented water was used as a control. The odors were presented to the users using ultrasonic diffusers. The diffusers were marked with a random number between one and six; this aligned with the number presented on a computer screen during the experiment. This number was for the experimenter to indicate what odor to present to the observer. The experimenter operated the diffusers which were hidden from the observer behind a divider ≈ 1 meter away to the observer's left. To create the olfactory stimuli five drops of the respective essential oil were added to 500 mL of tap water approximately 20 min before the experiment started. To avoid

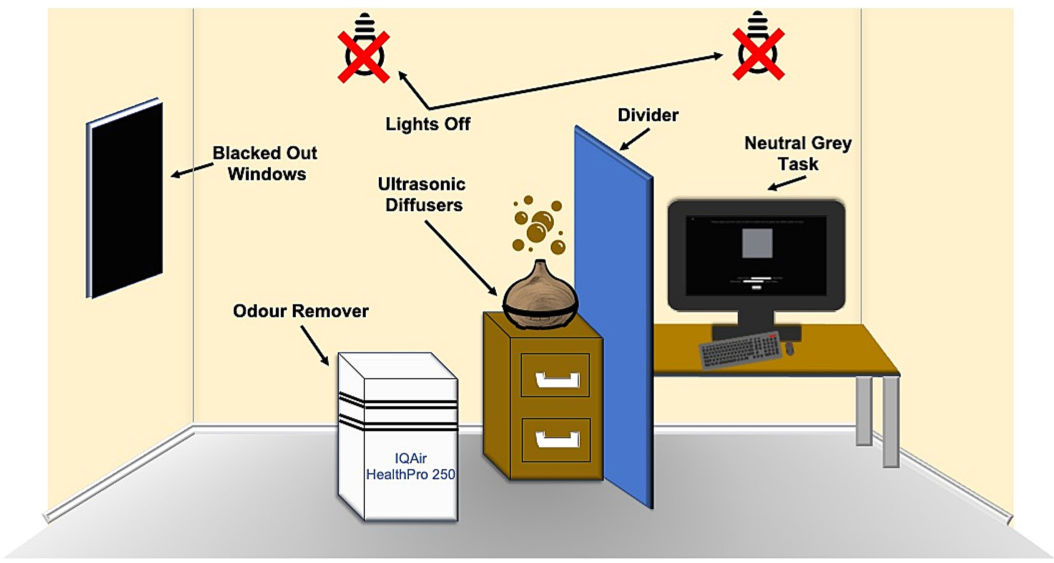

FIGURE 1
Diagram of the experimental setup.

FIGURE 2
Screenshot of the neutral gray task. This figure has been cropped to improve readability.

olfactory contamination each diffuser was only used for a singular odor, and the hole in the diffuser was taped over to minimize any odor leakage.

2.6. Color stimuli

The monitor used for the experiments was a calibrated 28" BenQ GL2450-B. The lightness of the presented stimuli inside of MATLAB was fixed at L* = 50 and could not be adjusted by the observer. The color patch was presented in the center of the screen horizontally and consisted of a 500 × 500 pixel patch and was presented at a resolution of 1080p. All text on the screen was adjusted to be the same color as the color patch; this included the instructions, the number corresponding to an odor, the slider labels, and the "Done" button. The application's background color was black; ($X = 0.1496$, $y = 0.1360$, $z = 0.2763$) after correction. The sliders were visually offset on the x-plane to stop the participants from simply centering the sliders together rather than aligning them based on the color patch. The starting color was set to be randomly between −5 and + 5 for the a* and b* channels separately. The starting position of the sliders was also adjusted to compensate. The range of −5 and + 5 was chosen to give the observers finer control over their final neutral gray selections, as a larger range would have resulted in larger jumps between colors when adjusting the slides. Larger jumps in the slides would make a truly achromatic stimuli harder to obtain as they might miss the values required for the task. As all external illumination was blacked out D65 was used as the white point as the observers were already dark-adapted. D65 is a CIE standard device-independent illuminate that simulates noon daylight and is commonly used in the relevant literature (Maric and Jacquot, 2013; Chauhan et al., 2014; Nehmé et al., 2016; Cuskley et al., 2019) as it makes it easy for the results to be reproducible.

2.7. Data analysis

The data analysis, figure preparation, and the programming of the experiment were performed in MATLAB 2021b. The CircStat toolbox (Berens, 2009a) was used for the circular statistics.

3. Results

First, the mean value for each of the observer's a* and b* color ratings across repetitions was calculated to reduce the signal-to-noise ratio of the reported neutral gray selections. Hue angles were then calculated using the a* and b* values collapsed across repetitions. Rayleigh tests were conducted (Bonferroni–Holm corrected) on these hue angles for each condition (all of the odors and the control). Circular data involves measurements that wrap around itself circularly and therefore, requires appropriate statistical tests that consider the circular nature of underlying data. The Rayleigh test (Fisher, 1995; Berens, 2009b) checks for the non-uniformity of circular data. In other words, it tests to see if the data points are evenly distributed around a circle. The null hypothesis is that the population is uniformly distributed around a circle, indicating a random distribution (chance selection). The alternate hypothesis is that the population is not uniformly distributed around a circle, indicating a non-random distribution (not attributed to chance). The Rayleigh test is often denoted as Z or R, which refers to the vector sum of the numbers representing the data points on the unit circle. Rayleigh tests indicated that all of our conditions induced a non-random distribution; caramel

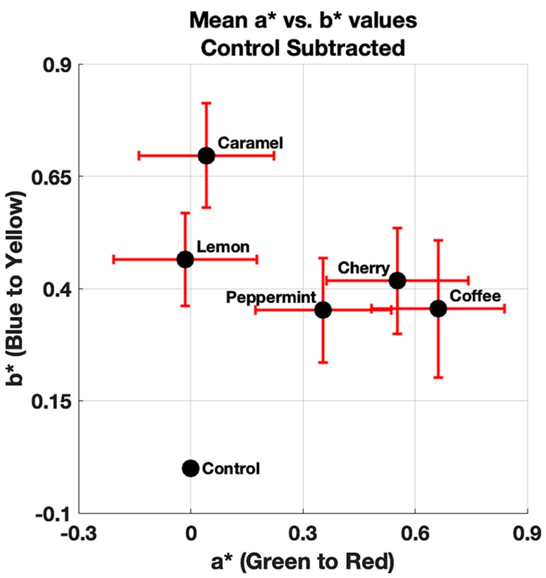

FIGURE 3
Mean a* and b* values for each of the odors and the control. Error bars denote standard mean error for the a* (*x*-axis) and b* (*y*-axis) dimensions.

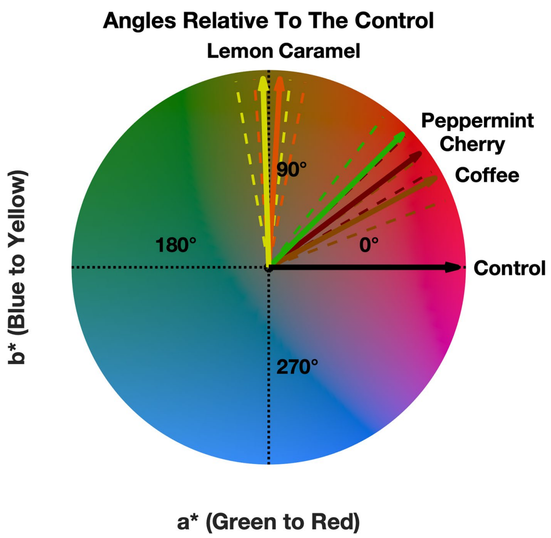

FIGURE 4
Angle for each of the odors relative to the control. Error bars (dashed lines) denote the standard mean error.

($Z = 22.37$, $p < 0.05$), cherry ($Z = 22.05$, $p < 0.05$), coffee ($Z = 22.20$, $p < 0.05$), peppermint ($Z = 22.26$, $p < 0.05$), lemon ($Z = 21.68$, $p < 0.05$), and the control ($Z = 22.78$, $p < 0.05$). In other words, the neutral gray selections for all of our conditions differed from chance selection. Next, we proceeded to correct for the control by subtracting the raw a* and b* values for each of the odors from the observer's rating for the control. The mean a* and b* values after correcting for the control for each condition is shown in Figure 3.

From Figure 3, a pattern is observed that suggests a small effect on the presence of odors influencing the observer's neutral gray selections. Figure 3 also shows that this shift in the CIELAB color space is generally small (less than one unit). Next, we decided to compare the direction of the mean a* and b* values with respect to the control in the CIELAB color space. This was done to visually see if there was a systematic effect. That is, does the presence of the odors shift the perceived neutral gray point toward their anticipated color correspondences? The angles relative to the control for the mean a* and b* are shown in Figure 4.

Figure 4 suggests that the angles all point toward warm colors, indicating that the presence of any of our odors induced a general shift to warmer colors, which is independent of the presented odor. Figure 4 shows that four (lemon, caramel, cherry, and coffee) out of five odors roughly trend toward their anticipated crossmodal correspondences except for peppermint, which we initially expected to be trending toward a green-blue color. See the discussion for more information about Figure 4. Next, we proceeded to statistically validate these findings.

We first tested to see if these results could be attributed to a spatial response bias due to the x-axis offset of the slider's visual location on the screen. A two-sample t-test was conducted on the control corrected a* and b* values collapsed across odors. This revealed no significant difference between the a* and b* values $t(119) = 0.5296$, $p = 0.69$, suggesting that if a spatial response bias was present, it has been rectified when correcting for the control. Now, it is known that there is no bias in the underlying data; hue angles were then calculated using the a* and b* values that have been corrected for the control (control shifted angle matrix). To determine if the hue angles of all of our odors were not attributed to chance after correcting for the control, Rayleigh tests (Bonferroni–Holm corrected) were then conducted on the control shifted angle matrix, revealing that the odors cherry ($Z = 5.15$, $p < 0.05$), coffee ($Z = 5.23$, $p < 0.05$), and peppermint ($Z = 5.14$, $p < 0.05$) induced a non-random distribution. The odors caramel ($Z = 1.18$, $p > 0.05$) and lemon ($Z = 0.36$, $p > 0.05$) were randomly distributed. In other words, the neutral gray selections for cherry, coffee, and peppermint demonstrated a systematic effect and did not originate from chance selection; the inverse could be said about caramel and lemon. However, as the odors may have multiple color correspondences, for instance, lemon is consistently associated with various shades of yellow, green, and pink (Ward R. J. et al., 2021), we decided to take this finding lightly. Pair-wise *post-hoc* V-tests were conducted to determine if the odors shared the same mean direction as the other odors and the control. The V-test is similar to the Rayleigh test with a difference in the alternative hypothesis, which is assumed to have a known mean direction (Zar, 1999; Berens, 2009b). The null hypothesis is that the data points are uniformly distributed around a circle. The alternate hypothesis is that the data points are not uniformly distributed around a circle and have a pre-specified mean angle. The V-test has more power than the Rayleigh test if there is a reason to believe there is a specific mean direction. The V test is often denoted as V, which refers to the multiplication of the Rayleigh test with the cosine of the difference between the actual and anticipated mean direction. *Post-hoc* pair-wise V-tests (Bonferroni–Holm corrected) were conducted between all odors and the control. Most importantly, this revealed that caramel ($V = 7.23$, $p > 0.05$), cherry ($V = 19.55$, $p > 0.05$), peppermint ($V = 20.10$, $p > 0.05$), and lemon ($V = 6.52$, $p > 0.05$) do not share the same mean direction as the control. However, the mean direction for coffee ($V = 23.26$, $p < 0.05$) did overlap with the control. Suggesting that the presence of some odors, but potentially not all, may modulate color appearance. See

Supplementary Table S1 for all of the results for the pair-wise comparisons.

To further validate this finding, pair-wise *t*-tests (Bonferroni–Holm corrected) were conducted on distance matrices. These distance matrices are the Euclidean distance from each condition to every other condition. The results indicate that all odors are significantly different from the control, and all conditions are significantly different from each other (all *p*-values less than 0.05). Zero was used as the hypothesized mean, as zero indicates both odors occupy the same space. Therefore, providing further evidence that the presence of odors could modulate color appearance. See Supplementary Table S2 for all of the results for this pair-wise comparison.

In terms of correct identification of the presented odors, caramel was correctly identified 33% of the time, with cherry being 33%, coffee at 12%, peppermint at 33%, and lemon at 12%, indicating the observers had trouble correctly labeling the odors. However, odor naming, even for common odors, has been shown to be a difficult task but is typically accompanied by a strong feeling of knowing (Jönsson et al., 2005).

In summary, these results suggest that the presence of different odors influenced the observer's neutral gray selections and that the angles of the odors trend toward warmer colors. Based on Figure 4, four (lemon, caramel, cherry, and coffee) out of five of the odors shift the perceived neutral gray point toward the color that is associated with the respective odor (Ward R. J. et al., 2021) except for peppermint.

4. Discussion

Our hypothesis was that the presence of odors would influence the perception of color, with observers overcorrecting their neutral gray selections to counteract their odor-color correspondences. The second half of this hypothesis did not hold; however, we did find that the presence of any of our odors induced a shift toward warmer colors, supporting the first half of our hypothesis. For the second half of our hypothesis, four (lemon, caramel, coffee, and caramel) out of five odors shifted the neutral gray point toward the anticipated crossmodal correspondences for the respective odor, indicating that our expectation that the observers would overcorrect to counteract this effect proved false. These findings demonstrate the effect that odors have on our color perception, specifically inducing a shift in a neutral gray setting.

Based on the crossmodal odor-color profiles from Ward R. J. et al. (2021; Figure 3B), caramel was typically associated with dark brown-yellow, coffee with dark brown and red, cherry with pink, red, and purple, peppermint with green and blue, and lemon with yellow, green, and pink. Here we found that caramel induced a shift in the neutral gray point toward a yellow-brown, cherry and coffee toward a red-brown, peppermint toward a brown-red, and lemon toward yellow–green. With the exception of peppermint, we have shown that the presence of odors induced a shift roughly toward the expected crossmodal correspondences for each of the odors. The most popular explanation for the existence of odor-color correspondences is in terms of semantics (Spence, 2020b). It may be the case that our findings stem partly from a semantic basis, but at a perceptual level. That is, as crossmodal odor-color correspondences are strong and stable (Demattè et al., 2009) semantic knowledge of odor could have changed the observer's perception of what a neutral gray is and due to the nature of the task presented to the observers, this lead us to conclude that this happened at a perceptual level, as they were asked to make a color patch achromatic, and there was no reference or language referring to the odors during the briefing or task. Therefore, suggesting that this effect does not occur at a decisional level.

In terms of cognitive penetrability, the idea that our expectations, thoughts, and beliefs can impact our perception of the world around us (Deroy, 2013) could play a role in our findings. That is, there might be top-down influences of higher cognitive processes that modify our perception, see (Delk and Fillenbaum, 1965; Macpherson, 2012; Deroy, 2013). Earlier research by Hansen et al. (2006) suggests that visual memory can modulate color appearance. Therefore, it could be the case that knowledge of the source of the odor penetrates/alters the perception of color. It is important to note that this knowledge of the identity of the odor could be induced by either a strong feeling of knowing (Jönsson et al., 2005) or by directly being able to name the odor. In turn, suggesting that odor-color correspondences are strong enough to influence color appearance. Stevenson et al. (2012) noted that when an odor is neither familiar nor nameable, odor-color correspondences may be modulated by intensity, hedonics, and/or irritancy. For odor-color correspondences, semantic attributes (identifiability and familiarity) are associated with more saturated colors and perceptual attributes (irritancy, intense, and unpleasant) are associated with brighter colors (Stevenson et al., 2012). It would stand to reason that, at least for some of our participants and potentially some of our odors, the semantic account might fail. In turn, suggesting that both semantic and perceptual attributes of the odors could also have played a role in explaining our findings. Crossmodal correspondences, in some cases, have been shown to be effected by emotions (Spence, 2020a; Di Stefano et al., 2021). Both unfamiliar odors (Schiffman, 1974; Benderly, 1988) and colors (Palmer et al., 2013; Gilbert et al., 2016) have been shown to be linked to emotions. Therefore, in the case of lack of knowledge of the identity of the odor the neutral gray selections could have been mediated by emotions.

Another, albeit more controversial, way of thinking about these findings, is in terms of crossmodal harmony. Crossmodal harmony refers to the way our senses complement and collaborate with other sensory modalities, such as audition (Spence and Di Stefano, 2022) and color perception (Judd and Wyszenk, 1975). In terms of olfaction, crossmodal harmony occurs when our sense of smell interacts and harmonizes with the other senses. Crossmodal harmony implies a balanced relationship between the component parts (Spence and Di Stefano, 2022). The experimental task could be conceived as a matching task between visual stimuli and ambient odors. It may be the case the perceived neutral gray selections differed due to crossmodal harmony, where the perceived neutral gray point shifted due to a harmony between the ambient odor and the visual stimuli. That is, the neutral gray selections have been modulated or matched in terms of pleasantness, or combinations of stimuli that were in a state of perceptual harmony.

A few limitations of this study should be mentioned. Firstly, the observers were restricted in adjusting the L* channel when asked to make an achromatic stimulus; this was done to make the experiment as simple as possible for the observer. However, including this additional dimension may supply more information to uncover why this effect occurs. For instance, it has been shown that when an odor cannot be named or is unfamiliar, the observer's color matches may be modulated by the irritancy, intensity, and/or hedonics instead (Stevenson et al., 2012; Spence, 2020b). Considering that odor intensity and color lightness are correlated sensory dimensions (Kemp and Gilbert, 1997), the inclusion of this additional dimension may help uncover if there is any involvement from the intensity of the presented odor. Another limitation of the current study is the number of tested odors. We only tested five odors, and although we found that four out of five of the odors had a mean direction significantly different from the control, it would be beneficial to uncover the extent in which this bias is present and to uncover if the trend toward the expected crossmodal correspondences for a given odor is only true if it also happens to be associated with a warm color.

Future work could include investigating to see how the familiarity of the presented odors affects the observer's ratings during the neutral gray task. That is, were the observers more biased if the presented odor was nameable or familiar compared to the observers where the presented odor was neither nameable nor familiar. To investigate this a larger sample size of observers would be required.

5. Conclusion

The results in this study suggest that the presence of an odor can bias an observer's decision of what their perceived version of a neutral gray is. That is the hue angles relative to the control stimuli for all odors point toward warm colors. It was also found that the hue angles for four out of five of our odors had a mean direction significantly different from the control and all of the odors are not located in the same place in the CIELAB color gamut as each other. Indicating that the presence of some odors' induces a shift of the perceived neutral gray point toward warmer colors and that four out of five of the odors are shifted toward their respective odor-color correspondences, further suggesting a small but systematic effect of the presence of odors on human color perception.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Department of Psychology, University of Liverpool. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RW conceived and designed the experiments, performed the experiments, analyzed and interpreted the data, and wrote the paper. MA performed the experiments and wrote the paper. SW conceived and designed the experiments, interpreted the data, wrote the paper, provided supervisory support, and provided materials and reagents. AM wrote the paper and provided supervisory support. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1175703/full#supplementary-material

References

Benderly, B. L. (1988). Aroma driven: on the trail of our most emotional sense. *Health (Irvine. Calif).* 20, 62–65.

Berens, P. (2009a). CircStat: a MATLAB toolbox for circular statistics. *Can. Vet. J.* 31, 1–21.

Berens, P. (2009b). CircStat: a MATLAB toolbox for circular statistics. *J. Stat. Softw.* 31, 1–21.

Bernstein, I. H., and Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *J. Exp. Psychol.* 87:241. doi: 10.1037/h0030524

Calvert, G., Spence, C., and Stein, B. E. (2005). *The handbook of multisensory processes.* Cambridge: MIT press.

Chauhan, T., Perales, E., Xiao, K., Hird, E., Karatzas, D., and Wuerger, S. (2014). The achromatic locus: effect of navigation direction in color space. *J. Vis.* 14, 1–11. doi: 10.1167/14.1.25

Crisinel, A.-S., Jacquier, C., Deroy, O., and Spence, C. (2013). Composing with cross-modal correspondences: music and odors in concert. *Chemosens. Percept.* 6, 45–52. doi: 10.1007/s12078-012-9138-4

Cuskley, C., Dingemanse, M., Kirby, S., and van Leeuwen, T. M. (2019). Cross-modal associations and synesthesia: categorical perception and structure in vowel–color mappings in a large online sample. *Behav. Res. Methods* 51, 1651–1675. doi: 10.3758/s13428-019-01203-7

de Valk, J. M., Wnuk, E., Huisman, J. L. A., and Majid, A. (2017). Odor–color associations differ with verbal descriptors for odors: a comparison of three linguistically diverse groups. *Psychon. Bull. Rev.* 24, 1171–1179. doi: 10.3758/s13423-016-1179-2

Delk, J. L., and Fillenbaum, S. (1965). Differences in perceived color as a function of characteristic color. *Am. J. Psychol.* 78, 290–293. doi: 10.2307/1420503

Demattè, M. L., Sanabria, D., and Spence, C. (2006). Cross-modal associations between odors and colors. *Chem. Senses* 31, 531–538. doi: 10.1093/chemse/bjj057

Demattè, M. L., Sanabria, D., and Spence, C. (2009). Olfactory discrimination: when vision matters? Chem. *Senses* 34, 103–109. doi: 10.1093/chemse/bjn055

Deroy, O. (2013). Object-sensitivity versus cognitive penetrability of perception. *Philos. Stud.* 162, 87–107. doi: 10.1007/s11098-012-9989-1

Di Stefano, N., Murari, M., and Spence, C. (2021). "Crossmodal correspondences in art and science: odours, poetry, and music" in *Olfaction: An interdisciplinary perspective from philosophy to life sciences*. eds. N. D. Stefano and M. T. Russo (Singapore: Springer).

DuBose, C. N., Cardello, A. V., and Maller, O. (1980). Effects of colorants and Flavorants on identification, perceived flavor, and hedonic quality of fruit flavored beverages and cake. *J. Food Sci.* 45, 1393–1399. doi: 10.1111/j.1365-2621.1980.tb06562.x

Evans, K. K., and Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *J. Vis.* 10, 1–12. doi: 10.1167/10.1.6

Fisher, N. I. (1995). *Statistical analysis of circular data*. Cambridge: Cambridge University Press.

Gilbert, A. N., Fridlund, A. J., and Lucchina, L. A. (2016). The color of emotion: a metric for implicit color associations. *Food Qual. Prefer.* 52, 203–210. doi: 10.1016/j.foodqual.2016.04.007

Gilbert, A. N., Martin, R., Kemp, S. E., The, S., Journal, A., and Autumn, N. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *Am. J. Psychol.* 109, 335–351. doi: 10.2307/1423010

Hansen, T., Olkkonen, M., Walter, S., and Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nat. Neurosci.* 9, 1367–1368. doi: 10.1038/nn1794

Hanson-Vaux, G., Crisinel, A.-S., and Spence, C. (2013). Smelling shapes: Crossmodal correspondences between odors and shapes. *Chem. Senses* 38, 161–166. doi: 10.1093/chemse/bjs087

Jacquot, M., Noel, F., Velasco, C., and Spence, C. (2016). On the colours of odours. *Chemosens. Percept.* 9, 79–93. doi: 10.1007/s12078-016-9209-z

Jönsson, F. U., Tchekhova, A., Lönner, P., and Olsson, M. J. (2005). A metamemory perspective on odor naming and identification. *Chem. Senses* 30, 353–365. doi: 10.1093/chemse/bji030

Judd, D. B., and Wyszenk, G. (1975). *Color in business, science and industry*. *3rd* New York: John Wiley and Sons.

Kaeppler, K. (2018). Crossmodal associations between olfaction and vision: color and shape visualizations of odors. *Chemosens. Percept.* 11, 95–111. doi: 10.1007/s12078-018-9245-y

Kemp, S. E., and Gilbert, A. N. (1997). Odor intensity and color lightness are correlated sensory dimensions. *Am. J. Psychol.* 110, 35–46. doi: 10.2307/1423699

Knöferle, K., and Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychon. Bull. Rev.* 19, 1–15. doi: 10.3758/s13423-012-0321-z

Luo, R. (2016). *Encyclopedia of color science and technology*. Berlin: Springer Reference.

Macpherson, F. (2012). Cognitive penetration of colour experience: rethinking the issue in light of an indirect mechanism. *Philos. Phenomenol. Res.* 84, 24–62. doi: 10.1111/j.1933-1592.2010.00481.x

Maric, Y., and Jacquot, M. (2013). Contribution to understanding odour-colour associations. *Food Qual. Prefer.* 27, 191–195. doi: 10.1016/j.foodqual.2012.05.001

Morrot, G., Brochet, F., and Dubourdieu, D. (2001). The color of odors. *Brain Lang.* 79, 309–320. doi: 10.1006/brln.2001.2493

Motoki, K., Saito, T., Nouchi, R., Kawashima, R., and Sugiura, M. (2019). Light colors and comfortable warmth: Crossmodal correspondences between thermal sensations and color lightness influence consumer behavior. *Food Qual. Prefer.* 72, 45–55. doi: 10.1016/j.foodqual.2018.09.004

Narumi, T., Sato, M., Tanikawa, T., and Hirose, M. (2010). Evaluating cross-sensory perception of superimposing virtual color onto real drink. Proc. 1st augment. *Hum. Int. Conf.* 10, 1–6. doi: 10.1145/1785455.1785473

NCSS (2010). Circular data analysis, 1–28. doi: 10.1038/npp.2010.217

Nehmé, L., Barbar, R., Maric, Y., and Jacquot, M. (2016). Influence of odor function and color symbolism in odor-color associations: a French-Lebanese-Taiwanese cross-cultural study. *Food Qual. Prefer.* 49, 33–41. doi: 10.1016/j.foodqual.2015.11.002

Palmer, S. E., Schloss, K. B., Xu, Z., and Prado-León, L. R. (2013). Music-color associations are mediated by emotion. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8836–8841. doi: 10.1073/pnas.1212562110

Schiffman, S. S. (1974). Physicochemical correlates of olfactory quality: a series of physicochemical variables are weighted mathematically to predict olfactory quality. *Science* 80, 112–117. doi: 10.1126/science.185.4146.112

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2020a). Assessing the role of emotional mediation in explaining Crossmodal correspondences involving musical stimuli. *Multisens. Res.* 33, 1–29. doi: 10.1163/22134808-20191469

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

Spence, C., and Deroy, O. (2013). How automatic are crossmodal correspondences? *Conscious. Cogn.* 22, 245–260. doi: 10.1016/j.concog.2012.12.006

Spence, C., and Di Stefano, N. (2022). Crossmodal harmony: looking for the meaning of harmony beyond hearing. *Iperception* 13, 1–40. doi: 10.1177/20416695211073817

Spence, C., and Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *Iperception* 3, 410–412. doi: 10.1068/i0540ic

Stevenson, R. J., Rich, A., and Russell, A. (2012). The nature and origin of cross-modal associations to odours. *Perception* 41, 606–619. doi: 10.1068/p7223

Ward, R., Rahman, S., Wuerger, S., and Marshall, A. (2021). Predicting the colour associated with odours using an electronic nose. 1st workshop on multisensory experiences-SensoryX'21.

Ward, R. J., Rahman, S., Wuerger, S., and Marshall, A. (2022). Predicting the crossmodal correspondences of odors using an electronic nose. *Heliyon* 8:e09284. doi: 10.1016/j.heliyon.2022.e09284

Ward, R. J., Wuerger, S. M., and Marshall, A. (2021). Smelling sensations: olfactory crossmodal correspondences. *J. Percept. Imaging* 4, 1–12. doi: 10.2352/j.percept.imaging.2021.4.2.020402

Zar, J. H. (1999). *Biostatistical analysis*. Noida: Pearson Education India.

Zellner, D. A., Bartoli, A. M., and Eckard, R. (1991). Influence of color on odor identification and liking ratings. 547–561.

Zellner, D. A., and Kautz, M. A. (1990). Color affects perceived odor intensity. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 391–397. doi: 10.1037/0096-1523.16.2.391

Zellner, D. A., and Whitten, L. A. (1999). The effect of color intensity and appropriateness on color-induced odor enhancement. *Am. J. Psychol.* 112, 585–604. doi: 10.2307/1423652

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 January 2024
DOI 10.3389/fpsyg.2024.1250781

Check for updates

OPEN ACCESS

EDITED BY
Karl Schweizer,
Goethe University Frankfurt, Germany

REVIEWED BY
Olivier Penacchio,
University of St Andrews, United Kingdom
Givago Silva Souza,
Federal University of Pará, Brazil

*CORRESPONDENCE
Na Chen
✉ imminana7@gmail.com

RECEIVED 30 June 2023
ACCEPTED 11 January 2024
PUBLISHED 24 January 2024

CITATION
Yang Y, Chen N, Kobayashi M and Watanabe K (2024) Color-taste correspondence tested by the Stroop task. *Front. Psychol.* 15:1250781.
doi: 10.3389/fpsyg.2024.1250781

Color-taste correspondence tested by the Stroop task

Yidie Yang[1], Na Chen[2]*, Maiko Kobayashi[1] and Katsumi Watanabe[1]

[1]Faculty of Science and Engineering, Waseda University, Tokyo, Japan, [2]The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

People consistently associate colors with tastes (e.g., pink-sweet, yellow-sour). However, little has been known on the strength of those color-taste correspondences. The current study examined the congruency effect of color-taste correspondence using two Stroop word categorization tasks. The visual stimuli consisted of food names associated with sweet and sour tastes, presented in different shades of pink and yellow font colors. Participants were instructed to categorize the taste (sweet or sour) of the words in the Stroop word-taste categorization task and to discriminate the font color (pink or yellow) of the words in the Stroop word-color discrimination task. Results showed that participants responded faster in congruent conditions (sweet-pink and sour-yellow) than incongruent conditions (sweet-yellow and sour-pink) in both tasks. Specifically, yellow font colors facilitated the categorization of sour taste words compared to pink font colors, whereas sweet taste words facilitated the discrimination of pink font colors compared to sour taste words. These results provide further evidence for the congruency effect of color-taste correspondence in facilitating the processing of taste-related words and colors. Furthermore, the congruency effect was shown to operate bidirectionally, influencing both the conceptual meaning of tastes and perceptual color perception. This study highlights the significant interference effect of color-taste correspondence on cognitive processing as assessed by the Stroop task.

KEYWORDS

color-taste correspondence, Stroop, congruency effect, sweet-pink, sour-yellow

1 Introduction

Taste plays a crucial role in our daily lives, not only for identifying different foods, but also for experiencing the pleasure of gourmet flavors. Taste perception is known to be psychologically sensitive and can be influenced by various factors, such as food color (Spence et al., 2015). Recent studies have found that color can significantly affect how people perceive the taste of foods and beverages, and that different colors can enhance or alter the perception of different taste qualities (Stillman, 1993; Fiszman and Spence, 2015). For example, red and pink colors can enhance sweetness and decrease bitterness, while green and blue colors can have the opposite effect (Maga, 1974; Hidaka and Shimoda, 2014). This process of integrating color perception into the evaluation of taste suggest that humans have developed certain associations between color and taste.

Crossmodal correspondence between color and taste refers to the phenomenon that people consistently associate the basic tastes (i.e., sweet, bitter, salty, sour, and umami) with specific colors (for a review, see: Spence and Levitan, 2021). Previous studies have found that pink and red are commonly associated with a sweet taste; yellow and green are associated with a sour taste; black, green, purple, and brown have been linked to a bitter taste; and blue, white, and gray have been associated with a salty taste (Koch and Koch, 2003; Spence et al., 2015; Woods and Spence, 2016; Saluja and Stevenson, 2018; Chen et al., 2021). These color-taste associations have been observed using both actual food samples and taste words. Saluja and Stevenson (2018) examined color-taste correspondences using actual tastes, and they found that the color-taste mappings were similar to those obtained using taste words. Moreover, color-taste correspondences are observed in different populations and are influenced by cultural and contextual factors (Wan et al., 2014).

Color-taste correspondence can be explained by the statistical learning of co-occurrences in the environment, where people associate certain colors with certain tastes based on the colors of foods and beverages they encounter (Spence et al., 2015; Higgins and Hayes, 2019; Spence, 2019). This learned color-taste mapping information allows for the efficient integration of colors and tastes (Piqueras-Fiszman and Spence, 2011; Spence, 2011; Parise et al., 2013; Chen et al., 2015). For instance, participants made fewer errors in discriminating the taste of real solutions when they were colored according to a color-taste mapping in congruent conditions than incongruent conditions (Zampini et al., 2007). Velasco et al. (2015) examined the congruency effect of color and flavor label of crisp packages using visual search and Go/No-go tasks, and they found that participants responded faster when the color of the package was congruent with the flavor label (e.g., lemon flavor in yellow and tomato flavor in red) than incongruent conditions. However, to our knowledge, no study has examined the congruency effect of color-taste correspondence using the Stroop task.

The Stroop word task is a widely used neuropsychological test that can probe attentional control and cognitive processing to assess the strength and automatic processing of implicit associations, known as the Stroop effect (Stroop, 1935; Richter and Zwaan, 2009; Scarpina and Tagini, 2017). In a classic Stroop task, participants are presented with words that name colors, but whose font color is different. Participants are instructed to either read the word or name the font color, which requires them to inhibit cognitive interference. Cognitive interference occurs when the processing of one stimulus attribute, such as the meaning of the word, influences the concurrent processing of another attribute, such as the font color (Stroop, 1935). The Stroop task relies on the strong overlearned tendency of experienced readers to automatically attend to the meaning of a word, making it challenging to ignore the meaning of the word when instructed to focus on the font color. Longer response times and lower accuracy in naming the font color may indicate difficulties in inhibiting the automatic tendency to categorize the word's meaning. This interference effect, known as the Stroop cost, is consistently observed and demonstrates the robustness of the task. Thus, the Stroop word categorization task can be a useful tool to test the automaticity and strength of color-taste correspondence, and it has been used in several previous studies investigating the implicit associations (Xiao et al., 2014; Chen et al., 2023).

The current study aimed to examine the congruency effect of color-taste correspondence using the Stroop task. In a previous study, sweet-pink (86.59%) and sour-yellow (80.49%) correspondences were the most chosen associations in a sample of Japanese participants (with a chance level of 20% in Chen et al., 2021). Thus, the sweet-pink and sour-yellow correspondences were selected to be tested using the Stroop task. By manipulating the levels of font color and taste words with congruent and incongruent color-taste pairs, this study examined whether the interference effect of color-taste correspondences would influence behavioral performance with response times and accuracy. Ten words (5 with sweet food names and 5 with sour food names) in ten font colors (5 levels of pink and yellow colors) were presented as visual stimuli. Through the two Stroop word tasks, we explored the strength of sweet-pink and sour-yellow correspondences in both directions with the effect of font color on words taste categorization and the effect of words meaning on font color discrimination. Specifically, participants were presented with color-word stimuli in a modified Stroop task paradigm, where they were required to categorize the taste of the presented words as sweet or sour while ignoring the font colors (Stroop word taste categorization task) and to identify the font color of the presented words as pink or yellow while ignoring the semantic content with taste (Stroop word color discrimination task). We hypothesized that a congruency effect of color-taste correspondence could influence both the taste word categorization and font color processing.

2 Methods

2.1 Participants

Thirty-eight Japanese undergraduate students (23 males and 15 females, *M*age = 20.4, *SD* = 1.5) took part in this study. All of the participants had normal or corrected-to-normal visual acuity and normal color vision. The sample size was set *a priori* at 33 based on a target of 0.8 power with a medium effect size (Cohen's $d = 0.5$) using power analysis (G*Power 3.0; Faul et al., 2007). We collected additional data in response to some participants' poor performance in the task (e.g., error rate > 20%). This experiment was approved by the institutional review board of Waseda University, and conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki.

2.2 Apparatus and stimuli

The experiment was programmed in E-Prime 3.0 (Psychology Software Tools[1]). The stimuli were displayed on a 24-inch LCD monitor (EIZO FG2421, EIZO Corp, Hakusan, Japan), with a resolution of 1920 × 1,080 pixels and a refresh rate of 100 Hz. Participants viewed the monitor at a distance of approximately 60 cm.

Ten Japanese words of food name were used as visual stimuli to rank sweetness and sourness. Five of the words were typically associated with two levels of sweetness [high: チョコレート (chocolate), あんこ (red bean paste); low: パフェ (parfait), ドーナツ (donut), 砂糖 (sugar)], while the remaining five are associated with two levels of sourness [high: うめぼし (dried plum), グレップフルーツ (grapefruit); low: パイナップル (pineapple), キウイ (kiwi), お酢 (vinegar)]. The food names were selected from the top twenty foods

1 http://pstnet.com/products/e-prime/

TABLE 1 Color values in the L*a*b* color space.

ID	L*	a*	b*	Color
P5	74.99	42.02	−2.64	
P4	77.44	34.74	4.60	
P3	78.59	30.61	10.35	
P2	80.10	26.34	14.84	
P1	81.27	22.33	20.71	
Y1	81.34	20.13	27.77	
Y2	82.86	15.21	35.58	
Y3	84.37	11.35	39.72	
Y4	85.21	9.79	39.88	
Y5	87.38	2.52	54.28	

based on the ranking of sweet[2] and sour foods[3] obtained from the Japanese "Everyone's Ranking" website. The word stimulus was presented in 40 Pt MS gothic font in the center of the screen. Seven of the words were two to four letters long, and three of the words were in six and eight letters. Sour words are five letters longer than sweet words in total. One Japanese letter sustained about 1.3° of visual angle.

The font colors of the word stimuli were 10 colors rendering gradually from pink (P5 in Table 1) to yellow (Y5 in Table 1). They were grouped as pink font colors (high level pink: P5 and P4; low level pink color: P3, P2, and P1 in Table 1), and yellow font colors (high level yellow: Y5 and Y4; low level yellow color: Y3, Y2, and Y1 in Table 1). The ten colors were measured by PR-655 (Photo Research, Chatsworth, CA, USA), and each color was measured 10 times and the average was calculated. The color information is shown in Table 1.

2.3 Procedure

The experiment was carried out in a laboratory with dimmed lighting condition (1 lux on the wall). All participants performed two Stroop word categorization tasks. One was the Stroop word taste categorization task, and participants were asked to categorize the taste of the word stimulus as sweet or sour. The other task was the Stroop word color categorization task, and participants were asked to discriminate the font color of the word stimulus as pink or yellow. At the beginning of each task, written instructions were provided on the computer screen. During the experiment, one of the word stimuli was presented in one of the 10 font colors. The background of the screen was always gray ($L^* = 20.24$, $a^* = 0.33$, $b^* = 0.36$). In the Stroop word taste categorization task, participants were asked to categorize the word stimulus as a "sweet" or "sour" word by pressing a labeled key (e.g., *z* or *m*) on the keyboard with their two index fingers. In the Stroop word color categorization task, participants were asked to discriminate the font color of the word stimulus as pink or yellow by pressing a labeled key (e.g., *z* or *m*). Key assignments were counterbalanced between participants and tasks. Participants were required to respond as quickly and accurately as possible. At the beginning of each trial, a fixation cross appeared for 300 ms, and then a target word stimulus appeared until a response was made. Trials were separated by an inter-stimulus interval (ISI) of 500 ms (see Figure 1). The experimental task had a 2 (taste of words: sweet vs. sour) × 5 (words for each taste group) × 2 (font color groups: pink vs. yellow) × 5 (colors for each font group) × 2 repeated measures design, resulting in a total of 200 trials. Twenty practice trials with feedback preceded the experiment. The main experiment was divided into 5 blocks of 40 trials each. At the end of each block, participants had a self-timed break. Each experiment of the Stroop word categorization task lasted for 5 ~ 10 min. Participants were randomly instructed to either take the Stroop word taste categorization task first or the Stroop word color categorization task first, with a 5-min rest period in between the two tasks. The entire experiment lasted for approximately 20 ~ 30 min.

2.4 Data analysis

Data from three participants who made more than 20% errors on either task were excluded. Thus, data from 35 participants were used for the data analysis. Response times (RTs) for correct trials and errors were recorded. Data were analyzed using the R.4.0.2 statistical software (R Core Team, 2020). Statistical analyses on response times and response accuracy were computed with generalized linear mixed-effect model (GLMM) using the function *glmer* from the lme4 package in R (Bates et al., 2015). The distributions of response times are positively skewed and not normally distributed; therefore, generalized mixed-effects models were used to obtain unbiased parameters without the need for data transformation (Lo and Andrews, 2015). We used *fitdistrplus* package to seek the best distribution fit to the present RT data (Delignette-Muller and Dutang, 2015), and selected the inverse gaussian distribution (Johnson et al., 1995). Thus, inverse gaussian distribution with identity link function was specified as the distribution of RTs in the GLMM models. For the error analysis, we used a GLMM with binomial distribution and logit link function. The main interest is the effect of congruency condition (congruent vs. incongruent) and the interactions with levels of taste and font colors on RTs and errors, thus, the fixed factors were congruency condition (congruent vs. incongruent), taste (sweet vs. sour), taste level (high vs. low), font color (pink vs. yellow), font color level (high vs. low), and the interactions. We included random intercept by participant (the random effects were not maximum, so to simplify the model we restricted to a random effect of participant on the intercept for all analysis; Barr et al., 2013; Meteyard and Davies, 2020). To test the main and interaction effects, we compared models with and without that fixed effect of interest using likelihood ratio tests. The *Anova* function (using type III Wald chi-square test) from the *car* package (Fox et al., 2013) was used to test the significance of fixed factors. Significant effects were compared using the *emmeans* package (Lenth et al., 2018). We ran a post-hoc analysis using paired sample *t*-test with Bonferroni correction to determine the simple main effects. Further, Bayes Factors (BF10) were used to determine whether there was support in favor of the alternative (H1) or null (H0) hypotheses (Morey and Rouder, 2018). A value of 1 indicates that null and

2 https://ranking.net/rankings/best-sweet-foods
3 https://ranking.net/rankings/best-sour-foods

10.3389/fpsyg.2024.1250781

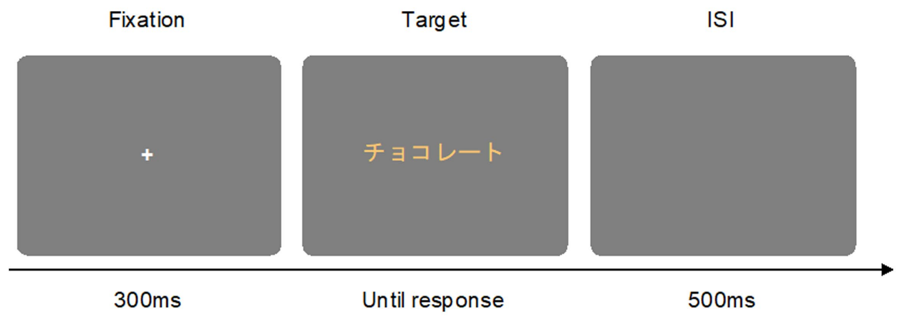

FIGURE 1
Example of the trial sequence. Participants were asked to categorize the taste of the word stimulus [e.g., "チョコレート(chocolate)"] as sweet or sour by pressing labeled keys in the taste categorization task. Participants were asked to categorize the font color of the word stimulus [e.g., "チョコレート (chocolate)" in yellow font color] as pink or yellow by pressing labeled keys in the color categorization task.

alternative hypotheses are equally likely, larger values indicate that the data support the alternative hypothesis, and smaller values indicate that the data support the null hypothesis.

3 Results

3.1 Results of the Stroop word taste categorization task

3.1.1 Response times

Error trials (5.31%) and those longer than 3,000 ms and shorter than 100 ms (0.35%) were removed. Trials in which the response times fell out of the outliner (mean RT ± 2.5 sd) for each participant in each color and taste condition were excluded from the data analysis (3.38%). Results of GLMM analysis are shown in Table 2. The GLMM model analysis revealed a main effect of congruency condition, with strong evidence for faster RT for congruent trials (mean RT = 585.16 ms, *sd* = 160.94) than for incongruent trials [mean RT = 595.09 ms, *sd* = 170.76; $x^2(1) = 61.06$, $p < 0.001$]. A significant interaction effect between congruency condition and font color was observed, $x^2(1) = 63.00$, $p < 0.001$. Multiple comparison analysis showed that when categorizing the sour food names, participants responded faster to yellow (congruent condition; 609.38 ms, *sd* = 96.48) than to pink font color [incongruent condition; 623.70 ms, *sd* = 98.90; $t(34) = 3.02$, Bonferroni corrected $p = 0.0096$, Cohen's $d = 0.51$, BF10 = 7.98]. For categorization of the sweet food names, there was no significant difference in RTs between pink (congruent condition; 563.26 ms, *sd* = 74.65) and yellow (incongruent condition; 571.13 ms, *sd* = 78.09) font colors [$t(34) = 1.91$, Bonferroni corrected $p = 0.12$, Cohen's $d = 0.32$, BF10 = 0.93; see Figure 2]. There was no significant interaction between congruency condition and levels of font color, $x^2(1) = 0.56$, $p = 0.45$, and no significant three-way interaction between congruency condition, font color, and levels of font color, $x^2(1) = 1.87$, $p = 0.17$.

3.1.2 Accuracy

Participants made 5.31% of the errors in all trials. The GLMM analysis showed no significant effect of the congruency condition, $x^2(1) = 2.30$, $p = 0.13$. Thus, there was no significant difference on the effect of font color on taste categorization by error rate.

TABLE 2 Summary of fixe effects for RTs in Task 1.

Predictor	Estimate	SE	*t* value	*p*
(Intercept)	595.23	20.60	28.89	<0.001***
Congruency	48.94	6.26	7.81	<0.001***
Font color	38.17	6.38	5.98	<0.001***
Color level	−1.55	5.55	−0.28	0.78
Congruency × Font color	−75.04	9.45	−7.94	<0.001***
Congruency × Color level	6.41	8.52	0.75	0.45
Font color × Color level	1.97	8.67	0.23	0.82
Congruency × Font color × Color level	−17.34	12.67	−1.37	0.17

'***' < 0.001; '**' < 0.01; '*' < 0.05; '+' < 0.1.

3.2 Results of the Stroop word color categorization task

3.2.1 Response times

Error trials (11.57%) and trials in which response times fell out of the outline (mean RT ± 2.5 sd) for each participant in each color and taste condition were excluded from data analysis (3.44%). There was no trial with response times slower than 3,000 ms and faster than 100 ms.

Results of GLMM analysis were shown in Table 3. The model analysis showed a significant main effect of congruency condition, that RTs in congruent conditions (mean RT = 535.93 ms, *sd* = 177.49) were faster than incongruent conditions [mean RT = 549.53 ms, *sd* = 188.47, $x^2(1) = 51.42$, $p < 0.001$]. A significant interaction effect between congruency condition and taste was observed, $x^2(1) = 64.26$, $p < 0.001$. Further analysis showed that for discriminating pink font color, participants responded faster in sweet taste words (congruent condition; 519.36 ms, $sd = 79.13$) than sour taste words [incongruent condition; 537.67 ms, $sd = 96.58$; $t(34) = 2.88$, Bonferroni corrected $p = 0.014$, Cohen's $d = 0.49$, BF10 = 5.93]. For discriminating yellow font colors, there was no significant difference on RTs between sweet (incongruent condition; 561.96 ms, $sd = 94.59$) and sour [congruent condition; 559.95 ms, $sd = 101.89$, taste words, $t(34) = 0.33$, Bonferroni corrected $p > 1$, Cohen's $d = 0.06$, BF10 = 0.19; Figure 3]. There was a marginal significant effect between congruency condition and level of taste, $x^2(1) = 3.64$, $p = 0.056$, and a significant three-way interaction between congruency condition, level of taste, and taste, $x^2(1) = 6.09$, $p = 0.013$.

3.2.2 Accuracy

Participants made 11.57% of the error trials, indicating some difficulty in discriminating the low-level pink and yellow font colors. The GLMM model analysis showed no significant effect of congruency condition on errors, $x^2(1) = 1.84$, $p = 0.18$. Thus, there was no significant difference on the effect of taste on color discriminating by error rate.

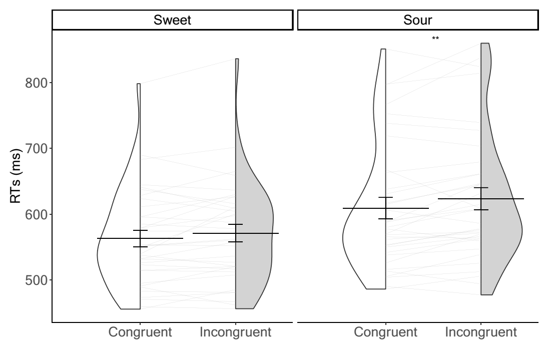

FIGURE 2
Violin plot showing the distribution of individual RTs in categorizing sweet and sour tasted words in congruent (white color) and incongruent (gray color) conditions. The horizontal line represents the mean. Error bars represent the standard errors of the mean (**$p < 0.01$, Bonferroni corrected).

3.3 Comparison between the two tasks

The GLMM model analysis showed that there was no significant interaction effect between the two tasks and the congruency conditions, on both response times, $x^2(1) = 0.13$, $p = 0.72$, and errors, $x^2(1) = 3.51$, $p = 0.06$. Thus, the color-taste congruency effect showed little difference in taste categorization and font color discrimination.

4 Discussion

In two Stroop-word categorization tasks, we manipulated the font color (pink and yellow) and the taste of food names (sweet and sour) to examine the congruency effect of associations between color and taste (pink-sweet and sour-yellow). The results showed that color-taste congruency influenced the speed of categorization of both taste-related conceptual words and perceptual font colors. Specifically, yellow font color facilitated the categorization of sour taste words compared to pink font color, while sweet taste words facilitated the discrimination of pink font color compared to sour taste. The high and low levels of taste words and font colors failed to show any interaction effect with the congruency effect. These findings provided further evidence for the congruency effect of color-taste correspondences on cognitive processing and behavioral performance (Velasco et al., 2015; Chen et al., 2021).

When participants were instructed to categorize food names as sweet or sour, the font colors were processed simultaneously, which interfered with the taste process. Our participants responded faster when the font color was congruent with the taste category (sweet-pink/sour-yellow) than when it was incongruent (sweet-yellow/sour-pink). Specifically, sour-tasting words in yellow font colors were recognized faster than those in pink font colors, indicating a congruency effect of sour-yellow correspondences on word taste categorization. Meanwhile, when participants were asked to discriminate the font color as pink or yellow, the associated taste (sweet or sour) of the words was also automatically activated, that discriminating font colors in congruent conditions (pink-sweet/yellow-sour) was faster than in incongruent conditions (pink-sour/yellow-sweet). Specifically, a sweet-related word facilitated the discrimination of pink font color compared to sour-related words, suggesting a sweet-pink correspondence. Thus, congruent color-taste correspondences facilitated perceptual color discrimination. Taken together, the results indicate the congruency

TABLE 3 Summary of fixed factors for RTs in Task 2.

Predictor	Estimate	SE	*t* value	Pr (>|z|)
(Intercept)	547.86	18.79	29.16	<0.001***
Congruency	43.23	6.03	7.71	<0.001***
Taste	44.58	6.27	7.11	<0.001***
Taste level	8.77	5.29	1.66	0.098+
Congruency × Taste	−68.96	8.60	−8.02	<0.001***
Congruency × Taste level	−14.18	7.43	−1.91	0.056+
Taste × Taste level	−22.36	7.82	−2.86	0.004**
Congruency × Taste × Taste level	25.64	10.39	2.47	0.01*

'***' < 0.001; '**' < 0.01; '*' < 0.05; '+' < 0.1.

effect of an automatically activated color-taste correspondence that modulates the processing of both conceptual word taste categorization and perceptual font color discrimination.

Indeed, it is worth noting that foods in the natural environment that represent sweet and sour tastes often have colors that differ from the commonly associated pink and yellow. For example, dried plum, a popular sour-tasting tsukemono (pickled vegetables) in Japan, is usually red in color. In contrast, chocolate, which is sweet and bitter, are predominantly brown in color. Similarly, donuts, which are known for their sweetness, are typically found in brown and yellow/orange colors. These examples illustrate how learned color-taste correspondences can override natural color associations, potentially influencing the perception and behavior.

Previous research suggests that individuals learn the relationships between sensory features/dimensions across different modalities, leading to the establishment of crossmodal correspondences (Spence, 2011). These correspondences aid in the efficient binding and processing of multisensory information, allowing the brain to efficiently integrate and combine sensory inputs to create a coherent perceptual environment (Parise et al., 2013). In particular, color-taste correspondences can be acquired through mere exposure to the co-occurrence of tastes and colors in foods and beverages (e.g., sweet-pink from the image of ripe and sweet red fruits, and sour-yellow from the yellow lemon). Importantly, studies have shown that similar color-taste correspondences exist across different cultural backgrounds, and some are influenced by cultural contexts (Wan et al., 2014). For example, Japanese individuals show high congruency for sweet-pink (86.59%) and sour-taste (80.49%) correspondences (at a chance level of 20%), which may be due to the regional food behaviors.

Further, color-taste correspondence can also be mediated by the emotional responses from those colors and tastes (e.g., hedonics; Velasco et al., 2015). For example, the color pink is often associated with happiness and romanticism, which aligns with the positive emotions evoked by the sweet taste. Similarly, the color yellow is associated with stimulation and excitement, which aligns with the sensory experience of sour tastes. These overlapping emotional experiences between colors and tastes could contribute to the observed correspondences between them. Thus, the emotional correspondence hypothesis offers a potential explanation for color-taste correspondence.

One limitation of the current study is that it focused only on the sweet-pink and sour-yellow correspondences. Future research could extend this by testing the Stroop effect of correspondences between other basic tastes and colors. Thus, the colors implicitly associated with each basic taste and their strength can be revealed using the Stroop task. Another limitation is the selection of food names representing different levels of sweet and sour tastes from online ranking websites. Future studies could adopt a more systematic approach to examine and rank the basic tastes of different foods to ensure a more comprehensive representation of taste profiles. Moreover, the levels of tastes (i.e., high and low levels of sweetness and sourness) were not well controlled. The main interest of this study is to examine the congruency effect, we did not ask participants to rank the sweetness/sourness of the taste. Future studies could further explore the strength of these associations using the Stroop effect, such that the higher level of color-taste correspondence, the stronger the congruency effect. At last, participants reported whether they have abnormal color vision subjectively, future study should use the Ishihara color vision test and a chemosensory disorder questionnaire to assess the perception of color and taste. Besides, there is no control over the luminance of the chromaticity, future studies need to use isoluminant colors. Our participants were all university students, future studies should consider the effect of age difference on the formation and congruency effect of color-taste correspondence.

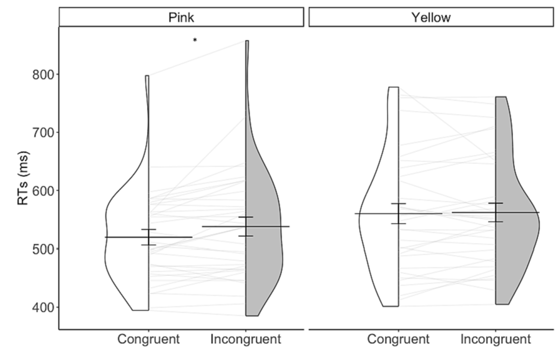

FIGURE 3
Violin plot showing the distribution of individual RTs in discriminating pink and yellow font colors in congruent (white color) and incongruent (gray color) conditions. The horizontal line represents the mean. Error bars represent the standard errors of the mean (*$p < 0.05$, Bonferroni corrected).

In conclusion, the present study demonstrated a congruency effect of color-taste correspondences (pink-sweet and yellow-sour), on both word taste categorization and font color discrimination tested by the Stroop tasks. This highlights the congruency effect of color-taste correspondences in influencing the cognitive processing of perceptual and conceptual information. Future research could investigate the statistical nature of these correspondences by examining the learning effects and training participants to establish color associations with novel tastes.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Ethics statement

The studies involving humans were approved by The institutional review board of Waseda University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NC coded the experiment. YY and MK conducted the experiments. YY and NC analyzed the data, confirmed the statistical analyses, and drafted the manuscript. KW provided the critical revisions on the introduction and discussion. All authors designed the study, read and approved the final manuscript.

Funding

This work was supported by JSPS Grants-in-Aid for Scientific Research (19K20591, 22KJ2877, 22H00090) from Japan Society for the Promotion of Science.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using {lme4}. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Chen, N., Nakamura, K., and Watanabe, K. (2023). An automatic red-female association tested by the Stroop task. *Acta Psychol.* 238:103982. doi: 10.1016/j.actpsy.2023.103982

Chen, N., Tanaka, K., and Watanabe, K. (2015). Color-shape associations revealed with implicit association tests. *PLoS One* 10:e0116954. doi: 10.1371/journal.pone.0116954

Chen, N., Watanabe, K., and Wada, M. (2021). People with high autistic traits show fewer consensual crossmodal correspondences between visual features and tastes. *Front. Psychol.* 12:714277. doi: 10.3389/fpsyg.2021.714277

Delignette-Muller, M. L., and Dutang, C. (2015). Fitdistrplus: an R package for fitting distributions. *J. Stat. Softw.* 64, 1–34. doi: 10.18637/jss.v064.i04

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*power 3: a exible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fiszman, B. P., and Spence, C. (2015). "Colour correspondences in chemosensation: the case of food and drinks" in *Nutrition and sensation*. ed. A. L. Hirsh (Boca Raton: CRC Press), 139–158.

Fox, J., Friendly, M., and Weisberg, S. (2013). Hypothesis tests for multivariate linear models using the car package. *R J.* 5, 39–52. doi: 10.32614/RJ-2013-004

Hidaka, S., and Shimoda, K. (2014). Investigation of the effects of color on judgments of sweetness using a taste adaptation method. *Multisens. Res.* 27, 189–205. doi: 10.1163/22134808-00002455

Higgins, M. J., and Hayes, J. E. (2019). Learned color taste associations in a repeated brief exposure paradigm. *Food Qual. Prefer.* 71, 354–365. doi: 10.1016/j.foodqual.2018.08.013

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions*, vol. *289*. New York: John Wiley & Sons.

Koch, C., and Koch, E. C. (2003). Preconceptions of taste based on color. *J. Psychol.* 137, 233–242. doi: 10.1080/00223980309600611

Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2018). Emmeans: estimated marginal means. AKA least-squares means, R package version 1.5.0. Available at: https://rdrr.io/cran/emmeans/.

Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front. Psychol.* 6:1171. doi: 10.3389/fpsyg.2015.01171

Maga, J. A. (1974). Influence of color on taste thresholds. *Chem. Senses* 1, 115–119. doi: 10.1093/chemse/1.1.115

Meteyard, L., and Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *J. Mem. Lang.* 112:104092. doi: 10.1016/j.jml.2020.104092

Morey, R. D., and Rouder, J. N. (2018). BayseFactor: computation of bayes factors for common designs. R package version 0.9.12–4.2. Available at: https://cran.r-project.org/web/packages/BayesFactor/index.html

Parise, C. V., Harrar, V., Ernst, M. O., and Spence, C. (2013). Cross-correlation between auditory and visual signals promotes multisensory integration. *Multisens. Res.* 26, 307–316. doi: 10.1163/22134808-00002417

Piqueras-Fiszman, B., and Spence, C. (2011). Crossmodal correspondences in product packaging. Assessing color–flavor correspondences for potato chips (crisps). *Appetite* 57, 753–757. doi: 10.1016/j.appet.2011.07.012

R Core Team. (2020). R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Richter, T., and Zwaan, R. A. (2009). Processing of color words activates color representations. *Cognition* 111, 383–389. doi: 10.1016/j.cognition.2009.02.011

Saluja, S., and Stevenson, R. J. (2018). Cross-modal associations between real tastes and colors. *Chem. Senses* 43, 475–480. doi: 10.1093/chemse/bjy033

Scarpina, F., and Tagini, S. (2017). The stroop color and word test. *Front. Psychol.* 8:557. doi: 10.3389/fpsyg.2017.00557

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C. (2019). On the relationship (s) between color and taste/flavor. *Exp. Psychol.* 66, 99–111. doi: 10.1027/1618-3169/a000439

Spence, C., and Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *i-Perception* 12:20416695211018223. doi: 10.1177/20416695211018223

Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., et al. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour* 4, 1–17. doi: 10.1186/s13411-015-0033-1

Stillman, J. A. (1993). Color influences flavor identification in fruit-flavored beverages. *J. Food Sci.* 58, 810–812. doi: 10.1111/j.1365-2621.1993.tb09364.x

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Velasco, C., Woods, A., Deroy, O., and Spence, C. (2015). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Qual. Prefer.* 41, 151–158. doi: 10.1016/j.foodqual.2014.11.010

Wan, X., Woods, A. T., van den Bosch, J. J., McKenzie, K. J., Velasco, C., and Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Front. Psychol.* 5:1365. doi: 10.3389/fpsyg.2014.01365

Woods, A. T., and Spence, C. (2016). Using single colors and color pairs to communicate basic tastes. *i-Perception* 7:2041669516658817. doi: 10.1177/2041669516658817

Xiao, X., Dupuis-Roy, N., Yang, X. L., Qiu, J. F., and Zhang, Q. L. (2014). The taste-visual cross-modal Stroop effect: an event-related brain potential study. *Neuroscience* 263, 250–256. doi: 10.1016/j.neuroscience.2014.01.002

Zampini, M., Sanabria, D., Phillips, N., and Spence, C. (2007). The multisensory perception of flavor: assessing the influence of color cues on flavor discrimination responses. *Food Qual. Prefer.* 18, 975–984. doi: 10.1016/j.foodqual.2007.04.001