ObrelaJOURNAL Dialogues in Clinical Neuroscience & Mental Health

e-ISSN: 2585-2795 • p-ISSN: 2654-1432
DOI: 10.26386/obrela.v5i3.238
p. 113-119

Can emotional music and faces influence our worldview? A simultaneously combined bimodal paradigm

Karvounopoulos, N., Christatou, S., Karadima, V., Afentouli, P., & Louka, P.[1]

1. Mediterranean College, School of Psychology

Abstract

Individuals' perceptions of valence ambiguous stimuli can be affected by other emotional stimuli introduced through their modalities. However, there is evidence that modalities influence each other and as a result the effect of the emotional stimuli received by one modality may dominate. Affective priming and mood induction procedures are two widely used paradigms that investigate the effect of emotional stimuli on the individuals. This study combines them into a cross-modal paradigm in which emotional faces are utilised as primes and emotional music as a mood induction cue. A convenience gender balanced sample of 54 individuals (M=28.98years, SD=11.10years) evaluated a set of valence ambiguous images derived from the GAPED database, once preceded by a happy and once by a sad face, derived from the FACES database. During the experiment, participants were also exposed to either the happy song "Barber - Adagio for Strings", or to the sad one "Mozart - Flute Concerto in D Major". Indeed, a modality dominance effect was established in this study, as merely the main effect of emotional faces was significant $F(1,52) = 18.841$, $p < .001$, $\eta^2 = .17$ supported by a large effect size. As a result, it could be argued that the visual stimuli dominate the other modalities based on the emotional effect and that affective priming produces more intense affective effects than the mood induction procedure. However, the above results should be cautiously interpreted, because of the small sample size as well as the presence of outliers.

Keywords

mood induction, affective priming, ambiguous visual stimuli, modality dominance, bimodal asymmetry, time perception

Corresponding author: Nikolaos Karvounopoulos (Καρβουνόπουλος Νικόλαος N.Karvounopoulos@mc-class.gr)

Introduction

Emotions are "brief, involuntary, full-system, patterned responses to internal and external stimuli" [1, p. 6] and their efficient identification is a basic feature of everyday life [2]. Emotions are able to induce a mood which can also turn into an emotion; emotions differ from moods because the latter last longer, are less intense and thus, they cannot be easily perceived just like the factor that caused them [3]. Emotional environmental cues differentiate based on the arousal dimension; the level of physiological activation (low/high), and the valence dimension; whether an emotion is negative or positive [4], and they are introduced through individuals' senses influencing their feelings and behaviours [5]. Specifically, facial expressions are omnipresent visual stimuli conveying fundamental emotional information which can be immediately detected, influencing individuals' emotional states [6]. According to Ekman (1972) [7], the six basic facial expressions (sadness, happiness, disgust, fear, anger, and surprise) are globally associated to a specific meaning thus, they can convey desired emotional information [8].

It is well-documented that when an emotional stimulus is perceived, its valence is automatically and unintentionally activated; a well-investigated phenomenon through Affective Priming (AP) [9]. AP studies highlight the "unintentional influence of a first evaluative response (to the prime) on subsequent processing (of the target)" [10, p. 70]. Specifically, in one AP paradigm, participants evaluate neutral target stimuli preceded by briefly presented emotional stimuli (primes) and consequently, their behaviours are biased towards prime's valence [11]. For instance, emotional faces used as primes influenced the Chinese ideograms' semantics used as neutral target stimuli [12, 13]. Specifially, Winkielman & Zajonc (1997) [13] stated that participants rated the ideographs preceded by happy faces significantly higher than those preceded by angry ones. An AP effect also occurs when a response to a target is facilitated under the valence congruency of prime and target stimuli [4]. Specifically, Aguado et al. (2007) [9] utilised pictures (11.29×16.95cm) of happy faces to prime a valence positive affect and an angry or a fearful face to prime a valence negative affect, and valence negative and positive words as targets. They stated that the reaction times in the congruent condition (e.g. happy face – happy word) were significantly lower than the incongruent one, concluding that AP occurs whether the activation of prime's valence is either intentional or unintentional. However, because the images containing scenes are stated to exert a stronger emotional influence than those containing faces, theorists argue that even though faces convey emotional information, it is not mandatory to elicit particular emotions on individuals [14].

Additionally, numerous cross-modal studies have indicated that people naturally communicate and receive affective information from multiple sensory modalities, providing evidence of auditory and visual stimuli interactions [4, 15]. For instance, de Gelder and Vroomen (2000) [16] in their experiments exposed participants to all the possible combinations of emotional recordings of sentences (with either a happy or a sad tone) via headphones and emotional black-and-white faces (sad or happy) via a computer screen. The photos lasted approximately 350ms during merely the last word of the recorded sentence. Participants were instructed to attend to one modality each time and had to choose whether the emotion of the attended modality was happy or sad. They concluded that the emotional stimuli from the unattended modality biased participants' perceptions of the attended one, however, participants were less biased from the audio stimuli when they judged the faces than vice versa. Similarly, Hietanen et al. (2004) [15] replicated the cross-modal interaction having synchronized the stimuli duration and furthermore provided evidence of an additive effect when the bimodal stimuli were valence congruent and were recognized more accurately and rapidly. Likewise, more recently Ramsay (2016) [17] utilised coloured images and recorded phrases lasting 10 seconds, instructing participants to emotionally evaluate merely the stimuli of the one modality ignoring the other, and he stated that they rated the stimuli of the attended modality as expected in accordance to their valence (happy stimuli were rated higher than valence ambiguous, which were rated higher than the sad ones). He also stated that a bimodal interaction existed. Thus, participants' evaluations were not merely affected by one modality, when each one was attended separately, however, he called other researchers to utilise different auditory stimuli like music.

Theorists have provided important results that individuals can be also emotionally influenced by music, which can be described as a language to communicate emotions to the listeners [18, 19]. Indeed, it is stated that people are strongly emotionally connected with music, which can change their emotions and interfere in the processing of subsequent emotional information received from other sensory modalities [20]. For instance, Tay & Ng (2019) [21] primed short song extracts via headphones to the participants before their exposure to valence neutral black and white pictures which they had to evaluate provided a list of 28 emotional words. Indeed, significantly more negative valence words were selected after participants' exposure to minor mode extracts, supporting previous studies associating them with negative feelings opposite to those in major mode. However, music can also temporarily elicit certain mood effects thus, it is also utilised in Mood Induction Procedures (MIPs), which are divided into simple; when merely one mood induction technique is utilised, and combined; when two or more ones are utilised in one paradigm [22].

Furthermore, there is evidence of a cross-modal additive effect, where participants showed higher accuracy when the same emotional type of the bimodal stimuli were congruent, than when presented in one sensory [23, 24], highlighting women's advantage both unimodally and bimodally [25]. The enhancement of valence ratings in combined congruent conditions have also been confirmed in fMRI studies [26, 27], however, it is stated that these results cannot be generalized across emotional categories [28]. Additionally, Sanchez et al. (2014) [29] focused on the effect of mood on the attentional deployment, performed a combined MIP. They utilised both emotional images and songs, which among others, they selected "Barber - Adagio for Strings" and "Mozart - Flute Concerto in D Major" to induce a sad and a happy mood, respectively to congruently support images' mood induction. The trials started with a black screen, followed by a second one containing a white fixation cross in the middle, both lasting 500ms. Participants were exposed to two face images; one neutral and one emotional and eventually, in the happy mood condition they showed both a happy mood elevation and an attentional deployment increase. Accordingly, because multiple inductions cause an additive contribution to a mood, several combined MIPs have successfully been applied in many studies, supporting this phenomenon [22]. However, recent studies question the above results, providing evidence that MIPs' effects do not last long and besides that participants answer the specific questionnaire immediately after MIP ends, the emerged results will be contaminated due to a diminished MIP effect during the whole process of completing the questionnaire [30, 31].

An intriguing phenomenon of modality dominance in which vision can eliminate the emotional auditory effect and vice versa has been reported, thus, it is suggested that the paradigm's task may influence which modality dominates [32]. Specifically, more than five decades ago, it has been stated that the visual stimuli dominated the auditory ones when presented simultaneously in an attention measuring paradigm, even when the auditory stimuli's intensity was twice that of the visual ones [33]. In addition, there is evidence that the visual dominance is also replicated in terms of affect, in bimodal paradigms measuring emotional identification accuracy [23, 24, 34]. For instance, Collignon et al. (2008) [25] stated that participants were more prone to be influenced by their vision when the bimodal stimuli; video clips of actors portraying angry and disgust facial expressions combined with the same emotional voices, were incongruent. Additionally, they stated that participants performed better when they were focused on the visual modality than the auditory one. Furthermore, they deliberately chose dynamic visual stimuli based on previous neuroimaging studies indicating their advantage on static faces to activate the specific brain regions which are implicated in the analysis of facial expressions, such as the amygdala the insula and the posterior superior temporal sulcus. Thus, they attributed their findings to the higher intensity of the dynamic visual stimuli compared to the static ones. Additionally, in their study Takagi et al. (2015) [34], who incorporated both congruent and incongruent conditions in which participants were instructed to attend once merely to the visual stimuli (faces in video clip), once merely to the audio one (voice) and once none of them and to emotionally judge both stimuli, they also indicated a visual dominance, in which more emotional categories' perceptions were dominated by the facial expressions.

Consequently, the literature is contradictory on whether vision dominates or it just biases hearing in terms of shaping individuals' emotional responses, and this study aims to explore this phenomenon. Hence, based on Tay & Ng [21], it adopts valence neutral pictures as target stimuli, approaching the same research question, whether AP can influence individuals' evaluations on valence ambiguous visual stimuli. However, here the selected images were coloured, like Ramsay (2016) [17] did, providing evidence based on a more realistic paradigm. Additionally, this study adopts emotional faces as primes, which are simultaneously combined with emotional music utilised as a MIP, examining both their effect separately and their cross-modal interaction. Thus, the hypotheses are: there will be a main effect of music's valence on the perception of valence ambiguous visual stimuli (H_1), there will be a main effect of faces' valence on the perception of valence ambiguous visual stimuli (H_2), and there will be an interaction between the valence of music and faces on the perception of valence ambiguous visual stimuli (H_3). However, while most studies [15, 16, 17, 29] utilise two affective stimuli and participants evaluate the second, we utilise three and participants evaluate the third after their exposure to the two others.

The previous years, numerous studies have been conducted, focused on individuals' time perceptions because of their exposure to emotional stimuli of faces and music [35], suggesting that time accelerates when individuals are satisfied [36]. Emotional music is categorised as dynamic stimuli because it involves various parameters evolving through time [37]. For instance, in terms of tempo, Fang (2015) [38] suggested that slow music extends the customer's stay in a restaurant, as they perceive the time passing slowly. In terms of valence, Bisson et al. (2009) [39] showed that participants perceived Bach's happy song lasting longer compared to Barber's Adagio for Strings. In their study, Droit-Volet et al. (2013) [37] concluded that when music is fast or has positive valence leads participants to interpret it as having a shorter duration. Additionally, the literature posits that individuals perceive high-arousal negative valence faces to last longer than low-arousal valence ambiguous ones [37]. Many studies have also revealed evidence that when individuals process time in one modality, they are influenced by other modalities' signals [40]. Two fundamental paradigms explore this phenomenon; the prospective in which participants a priori know that they will be asked about their time perceptions and the retrospective in which they do not [36]. This study also explores in a prospective paradigm whether individuals' time perceptions are influenced by music's valence when controlling for the same visual task thus, it is hypothesised that the valence of music will have an effect on individuals' time estimations (H_4).

Method

Design: A mixed design was adopted in this study, with two Independent Variables (IVs), each one consisted of two conditions; valence of music (between-subject) and valence of faces (within-subject) both consisted of happy and sad conditions, and one Dependent Variable (DV); valence ambiguous images evaluation. Thus, each participant was exposed to either a happy or a sad music stimuli utilised as a MIP, and both to happy and sad faces utilised as primes. Additionally, a second DV is participants' time estimations influenced by merely one IV; The valence of music (between-subject).

Participants: A convenience sample of 27 women (M=31.37 years, SD=12.30years) and 27 men (M=26.59 years, SD=9.39 years) (N=54 people M=28.98 years, SD=11.10 years) was chosen, constructed by peers, undergraduate students and family who volunteered (see Appendix, p. 22). All participants were healthy with normal or corrected to normal vision and hearing and did not suffer from any any mental disorder.

Materials: Similarly, to Sanchez et al. (2014) [29]., the music pieces "Barber - Adagio for Strings" and "Mozart - Flute Concerto in D Major" were utilised for the sad and happy mood induction, respectively. Samuel Barber's Adagio for Strings, Op 11 has been widely used in sad MIPs, and has recently been favoured among others for its effectiveness [41]. Happy, sad and neutral faces were selected from FACES database [42], which contains images from people with a wide age range, categorised in six facial expressions (neutral, angry, happy, fearful, disgusted and sad). The valence ambiguous target stimuli were coloured images selected from Geneva Affective Picture Database [GAPED, 43]. GAPED is a widely used database [44] and consists of 730 images categorised as neutral, negative or positive depending on their emotional content. Based on Tay & Ng (2019) [21], the emotional evaluations of the neutral target stimuli were given in a 7-point Likert-type scale questionnaire graded from 1 (totally negative) to 7 (totally positive), where 4 was neutral (see Appendix, p. 57). The experiment took place in a college PC lab. The music stimuli were played via speakers and a decibel meter application was used by a smartphone to ensure that all participants were exposed to the music stimuli at approximately the same volume (70db). Power Point was utilised to create and present the visual task which according to Aguado et al. (2007) [9] contained 11.29×16.95cm target pictures centred in the middle of the screen.

Procedure: Participants evaluated the target images in 60 trials. Each trial consisted of 4 slides; a black screen presented until the "space key" was pressed, a black screen with a white cross as a fixation stimulus in its centre presented for 500ms, an image of either an emotional (happy or sad) or a valence ambiguous face presented for 250ms, and a valence ambiguous target image presented for 2 seconds. Participants were instructed to answer fast so that they could not scrutinize their emotional state. 20 valence ambiguous target stimuli were evaluated once preceded by a happy primed face and once with a sad one to control any unwanted emotional effects of the target stimuli. The sum of these evaluations was each participant's score for each condition ranging from 20 (1x20) to 140 (7x20). Because the identification of this experiment's purpose would be relatively obvious, 20 different valence neutral target stimuli preceded by valence neutral faces were added, without being included in the data analysis. Additionally, the time perception question was also a distracting cue, because participants were invited to take part in an experiment measuring their performance in a dual-task procedure (see Appendix, pp. 45-55), in which they had to evaluate the valence of the scenes of the images, but not the faces' and to estimate the time that this experiment lasted. They were also instructed to switch off their phones and to not obtain watches. Two trials containing valence neutral faces and valence neutral target images that were not included in the experiment were used as examples so that participants understand the procedure before the experiment begins. As soon as participants' exposure to the trials was completed, and they had no questions about the procedure, they were instructed that they will hear a song to remain focus on the task, eliminating the environmental noises. Thus, they were first randomly exposed to the happy or the sad song and then the procedure begun. The target image stimuli were presented in the same randomized order and the songs were played at approximately 70db. To secure that the MIP effect will last through the whole process [30, 31], the music played until the evaluation of the last image and the following time estimation answer was given. Songs' duration was satisfactory to last if the time needed for participants to be done however, in case where participants were late whenever a song ended, it was instantly automatically replayed.

Ethical Consideration: In this study, three forms were also utilised; a participant information sheet introducing participants to this study (see Appendix, pp. 54-55), a consent form ensuring anonymity (see Appendix, p. 56) and a debrief sheet, which revealed the real purpose of this experiment and informed participants about their right to withdraw (see Appendix, p. 58). This study was accepted from the University of Derby.

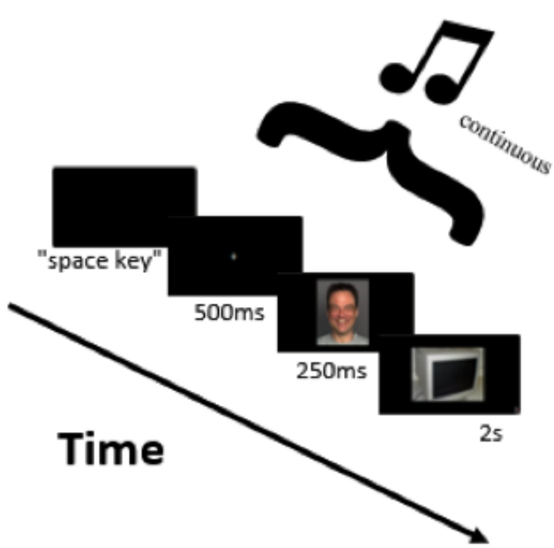

Fig. 1 A simultaneous cross-modal trial combining the priming of an emotional face and a music MIP.

Results

In the emotional evaluation data, four participants were displayed as outliers in the boxplots (see Appendix, pp. 32-36) thus, because neither did the transformations proposed by Field (2018) [45] effectively absorb them, nor equivalent non parametric test exists to explore the interaction between the two IVs on the DV, the level of significance became more strict (a=0.01) and data analysis continued with the original scores and statistical test.

Table 1. The means of the valence ambiguous image evaluation scores (with standard deviations)

	Happy Face	Sad Face	Total
Happy Music	77.78 (16.04)	69.22 (17.61)	73.50 (17.23)
Sad Music	82.96 (15.92)	61.89 (16.38)	72.43 (19.21)
Total	80.37 (16.04)	65.56 (17.25)	

Utilising the 2(Face Valence) x 2(Music Valence) Factorial Independent Measures ANOVA test, a significant main effect of the face valence on the Valence Ambiguous Image Evaluation (VAIE) was found $F(1,52) = 18.841$, $p<.001$, $\eta^2=.17$ with the mean evaluation of the target valence ambiguous images were significantly higher when preceded by happy faces than by sad faces (80.37>65.56) (see Appendix, p. 29). However, there was neither a significant main effect of music valence on VAIE $F(1,52)=.138$, $p=.714$, nor a significant interaction between music valence and face valence on VAIE $F(1,52)=3.363$, $p=.072$ (see Appendix, pp. 40-41). In other words, participants evaluated the valence ambiguous images biased by the emotions of the primed faces preceded, but not from music or its interaction with the faces. In addition, the effect size of face valence on VAIE is Large [46] (see Appendix, p. 47).

Due to the existence of one outlier in the data of participants' time perceptions, the non-parametrical Mann-Whitney U Test was utilised and showed no statistical significant results Mann-Whitney U (n1=27; n2=27)=334.5, z=-0.525, p=.599, two tailed (see Appendix, pp. 51-52). Thus, H_4 is rejected.

Discussion

Individuals daily receive emotional information from their senses, such as faces or songs which according to the literature they can bias their evaluations on separate valence ambiguous ones [12, 13, 21]. When individuals are simultaneously exposed to both emotional visual and audio stimuli these stimuli can interact and bias their evaluations on these stimuli [15, 16, 17]. It is even stated that the emotional visual stimuli can dominate the visual ones [34]. however, contradictions exist on whether vision dominates or just biases hearing in terms of shaping individuals' emotional responses. Hence, this study explored the main effect of emotional faces and music, as well as their interaction on valence ambiguous images. The results showed that merely the main effect of the emotional faces was significant where specifically, the mean of the scores of the VAIE was higher when preceded by a happy face than by a sad face, supporting the previous literature for the contradictory effect between valence positive and negative faces [13]. As a result, the emotional faces efficiently influenced participants' evaluations on the valence ambiguous target images containing scenes, providing salient evidence of their ability to emotionally influence the individuals contrary to previous literature [14]. Additionally, in this study participants' evaluations were neither influenced by emotional music nor by its interaction with the emotional faces, contrary to previous studies, which demonstrated an interaction between the emotional recorded phrases and emotional faces [15, 16, 17]. Thus, this study could support the cross-modal emotional dominance effect in which due to its visual task, vision dominated the auditory perception and emotional faces outcompeted emotional music's effect [34]. Specifically, we support the results of Collignon et al. (2008) [25] and Takagi et al. (2015) [34] who found that participants were more biased from the dynamic emotional visual stimuli than the recorded phrases, extending the literature on static stimuli, too, despite their disadvantage advocated by previous neuroimaging studies as stated by Collignon et al. (2008) [25].

Analytically, this study utilised emotional music as a MIP and emotional faces as primes. Given the fact that merely the AP was effective, this study could sceptically support that AP is more effective to bias individuals' evaluations on valence ambiguous target stimuli than MIPs when they are simultaneously combined. Furthermore, it could be also implied that the emotions elicited by the AP were more crucial to bias individuals evaluations rather than their moods, supporting the literature that emotions are more intense than moods [3]. As a result, this argument could justify why AP was more efficient than MIP to bias individual's affect. However, these connotations should be critically argued due to this paradigm's design, which utilised the visual stimuli both as primes and as target stimuli. Thus, merely one modality (vision) was intentionally in depth attended and this could obtain an advantage to visual stimuli's emotional effect independently of whether an AP or a MIP was served. Specifically, Collignon et al. (2008) [25] in their incongruent bimodal paradigm stated that participants had a more accurate performance when they were focused on the visual stimuli than the auditory ones. Thus, to replicate the emotional modality dominance based on this combined cross-modal paradigm, future studies could control the target stimuli's modality, utilising both auditory and visual target stimuli, where participants will be instructed to once evaluate the audio target stimuli and once the visual one, as shown in Ramsay's (2016) [17] study, or they could also control the modalities that the AP/MIP will serve.

This study included a prospective time estimation paradigm and data analysis showed no statistically significant differences on participants' time evaluations depending on the music stimulus they were exposed to. These results contrast with previous studies suggesting that tempo [37, 38] and valence [37, 39] can affect individuals' time perceptions. Specifically, even though we utilised the same sad song to

Bisson et al. (2009) [39], it could not exert statistically important effects on the participants' estimations compared to the happy song. However, this prospective paradigm also included a visual task and as a result, participants also received visual information that had to evaluate. Thus, based on the literature suggesting that other modalities may influence time perception estimations [40], we could interpret these results as being contaminated by the stronger effect of the emotions elicited by the faces [3], neutralising music's emotional effect. Because participants were exposed both to happy and sad faces, but merely to a sad or a happy song, the valence of the song could not bias participants' emotional state. Therefore, these results could support that the emotional faces were the only stimuli that influenced participants emotional states.

However, some limitations exist. The results of this study should be judged cautiously because of the outliers, which scores were not adjusted. Additionally, the participants listened to the music via speakers and not via headphones like previous studies [16, 21], potentially influencing music's effect. Headphones may enhance the attention to music and decrease the outside noises functioning as a confounding variable. Actually, the interaction between music and faces was not far from reaching significance ($p=.07$). Additionally, because its observed power was .44 (see Appendix, p. 40) failing to reach the recommended .8 [46], a bigger sample could have obtained significance. However, a previous study supports vision dominance in terms of attention deployment, despite the fact that the intensity of the auditory stimuli was twice that of the visual ones [33]. Furthermore, the analysed target stimuli were displayed two times, once preceded by a happy and once by a sad face, securing that the evaluations were biased due to the IVs and not the target image itself. Thus, participants could remember the first score and replicate it, altering the mean differences. A procedure with more trials could minimize this phenomenon. Thus, future studies replicating this one should consider the above cues carefully. Additionally, the previous studies which demonstrated an interaction between the emotional audio stimuli and emotional faces utilised emotional recorded phrases rather than emotional songs [15, 16, 17]. An implication in this circumstance may be the fact that the emotional instrumental music may not have such a strong emotional impact compared to the recorded phrases. Thus, taken together the dominance of vision may be modulated by the utilised audio sample, something that future studies could also investigate.

Conclusion

This study states that when individuals are simultaneously exposed to both emotional visual and audio stimuli evaluate the valence ambiguous visual stimuli (photos of real-life scenes) merely biased from the preceded emotional visual stimuli (happy and sad photos of faces) rather than the continuous presence of the emotional audio stimuli (happy or sad songs). As a result, when individuals evaluate stimuli via vision are biased from the emotional stimuli preceded by the same modality independently of whether they are exposed to a happy or a sad song. Additionally, this evidence is supported from the time perception data that also exhibit no statistically significant effect of the songs on individuals time perceptions.

References

1. Linehan, M. M. (2014). *DBT skills training manual*. (2nd ed.). New York: Guilford Press.
2. Marin, M. M., Gingras, B., & Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion, 12*, 618-631.
3. Eysenck, M. W., & Keane, M. T. (2015). *Cognitive psychology: A student's handbook*. Hillsdale: Lawrence Erlbaum Associates.
4. Kang, C., Ye, N., Zhang, F., Wu, Y., Jin, G., Xie, J., & Du, L. (2021). The cross-modal affective priming effect: Effects of the valence and arousal of primes. *Social Behavior and Personality: An international journal, 49*(12), e10202. https://doi.org/10.2224/sbp.10202.
5. Schreuder, E., Erp, J. V., Toet, A., & Kallen, V. L. (2016). Emotional responses to multisensory environmental stimuli: A conceptual framework and literature review, *SAGE Open, 6*(1), 1-19. https://doi.org/10.1177/2158244016630591.
6. Andrews, V., Lipp, O. V., Mallan, K. M., & König, S. (2011). No evidence for subliminal affective priming with emotional facial expression primes. *Motivation and Emotion, 35*, 33-43.
7. Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (Vol. 19). Lincoln: University of Nebraska Press.
8. Carminati, M. N., & Knoeferle, P. (2013). Emotional priming of sentence comprehension: effects of a speaker's static emotional expression and listener age. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1976-1981). Austin, TX: Cognitive Science Society.
9. Aguado, L., García-Gutiérrez, A., Castañeda, E., & Saugar, C. (2007). Effects of prime task on affective priming by facial expressions of emotion. *The Spanish Journal of Psychology, 10*, 209–217. doi:10.1017/ S1138741600006478.
10. Klauer, K. C. (1998). Affecting priming. *European Review of Social Psychology, 8*, 63–107.
11. Dannlowski, U., Ohrmann, P., Bauer, J., Kugel, H., Arolt, V., Heindel, W., & Suslow, T. (2007). Amygdala reactivity predicts automatic negative evaluations for facial emotions. *Psychiatry Research: Neuroimaging, 154*(1), 13e20. https://doi.org/10.1016/j.pscychresns.2006.05.005.
12. Heerdink, M. W., Van Kleef, G. A., Homan, A. C., & Fischer, A. H. (2015). Emotional expressions as cues of rejection and acceptance: Evidence from the affect misattribution paradigm. *Journal of Experimental Social Psychology, 56*, 60–68. https://doi.org/10.1016/j.jesp.2014.09.004.
13. Winkielman, P., Zajonc, R. B., Schwarz, N. (1997). Subliminal affective priming resists attributional interventions. *Cognition and Emotion, 11*(4), 433– 65. https://doi.org/10.1080/026999397379872.
14. Gerdes, A. B. M., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion

cues in multiple domains. *Frontiers in Psychology, 5*, 1-12. https://doi.org/10.3389/fpsyg.2014.01351.
15. Hietanen, J. K., Leppänen,, J. M., Illi, M., & Surakka, V. (2004). Evidence for the integration of audio-visual emotional information at the perceptual level of processing. *European Journal of Cognitive Psychology, 16*, 769-790.
16. de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and eye. *Cognition and Emotion, 14*, 289–311. https://doi.org/10.1080/026999300378824.
17. Ramsay, H. J. (2016). Bidirectional cross-modal influence on emotion ratings of auditory and visual stimuli. *Berkeley Scientific Journal, 21*(1), 70-75.
18. Logeswaran, N., & Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neuroscience Letters, 455*, 129–133. https://doi.org/10.1016/j.neulet.2009.03.044.
19. Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68*, 111–141. https://doi.org/10.1016/S0010-0277(98)00043-2.
20. Lense, M. D., Gordon, R. L., Key, A. P. F., & Dykens, E. M. (2014). Neural correlates of cross-modal affective priming by music in Williams syndrome. *Social Cognitive and Affective Neuroscience, 9*(4), 529–537. https://doi.org/10.1093/scan/nst017.
21. Tay, R. Y. L., & Ng, B. C. (2019). Effects of affective priming through music on the use of emotion words. *PloSone, 14*(4), 1–26. https://doi.org/10.1371/journal.pone.0214482.
22. Jallais, C., & Gilet, A. L. (2010). Inducing changes in arousal and valence: Comparison of two mood induction procedures. *Behavior Research Methods, 42*, 318–325. http://dx.doi.org/10.3758/BRM.42.1.318.
23. Föcker, J., Gondan, M., & Röder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychologica, 137*, 36–47. doi:10.1016/j.actpsy.201 1.02.004.
24. Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing for recognizing multi-modal emotional stimuli? *Motivation and Emotion, 35*, 192–201. doi: 10.1007/s11031-011-9206-0.
25. Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2009). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia, 48*(1), 220–225.
26. Baumgartner, T., Lutz, K., Schmidt, K. F., & Jäncke, L. (2006). The emotional power of music: How music enhances the feeling of affective pictures. *Brain Research, 1075*, 151–164. doi:10.1016/j.brainres.2005.12.065.
27. Park, J.-Y., Gu, B.-M., Kang, D.-H., Shin, Y.-W., Choi, C.-H., Lee, J.-M., & Kwon, J. S. (2010). Integration of cross-modal emotional information in the human brain: An fMRI study. *Cortex, 46*(2), 161-169.
28. Marin, M. M., Gingras, B., & Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion, 12*, 618-631.
29. Sanchez, A., Vázquez, C., Gomez, D., & Joormann, J. (2014). Gaze-fixation to happy faces predicts mood repair after a negative mood induction. *Emotion, 14*, 85-94. https://doi.org/10.1037/a0034500.
30. Gillies, J. C. P., & Dozois, D. J. A. (2021). How long do mood induction procedure (MIP) primes really last? Implications for cognitive vulnerability research. *Journal of Affective Disorders, 292*, 328-336. https://doi.org/10.1016/j.jad.2021.05.047.
31. Kuijsters, A., Redi, J., de Ruyter, B., & Heynderickx, I. (2016). Inducing sadness and anxiousness through visual media: Measurement techniques and persistence. *Frontiers in Psychology, 7*, 1141. https://doi.org/10.3389/fpsyg.2016.01141.
32. Sandhu, R., & Dyson, J. B. (2012). Re-evaluating visual and auditory dominance through modality switching costs and congruency analyses. *Acta Psychologica, 140*, 111–118.
33. Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 16*(2), 409–412.
34. Takagi, S., Hiramatsu, S., Tabei, K.-I., & Tanaka, A. (2015). Multisensory perception of the six basic emotions is modulated by attentional instruction and unattended modality. *Frontiers in Integrative Neuroscience, 9*, 1-10. https://doi.org/10.3389/fnint.2015.00001.
35. Droit-Volet, S. (2013). Time perception, emotions and mood disorders. *Journal of Physiology – Paris, 107*(4), 255–264. http://dx.doi.org/10.1016/j.jphysparis.2013.03.005.
36. Ansani, A., Marini, M., Mallia, L., & Poggi, I. (2021). Music and time perception in audiovisuals: Arousing soundtracks lead to time overestimation no matter their emotional valence. *Multimodal Technologies and Interaction*, *5*(11), 68. https://doi.org/10.3390/mti5110068.
37. Droit-Volet, S., Ramos, D., Bueno, J. L., & Bigand, E. (2013). Music, emotion, and time perception: The influence of subjective emotional valence and arousal? *Frontiers in Psychology, 4*, 417. https://doi.org/10.3389/fpsyg.2013.00417.
38. Fang, Z. (2015). The Study on the effect of background music on customer waiting time in restaurant. *The Open Cybernetics & Systemics Journal, 9*, 2163–2167. doi.10.2174/1874110X01509012163.
39. Bisson, N., Tobin, S., & Grondin, S. (2009). Remembering the duration of joyful and sad musical excerpts. *Neuroquantology, 7*(1), 46–57.
40. Chen, K.-M., & Yeh, S.-L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica, 130*, 225-234. https://doi.org/10.1016/j.actpsy.2008.12.008.
41. Lawrie, L., Jackson, M. C., & Phillips, L. H. (2019). Effects of induced sad mood on facial emotion perception in young and older adults. *Aging Neuropsychology and Cognition, 26*(3), 319–335. https://doi.org/10.1080/13825585.2018.1438584.
42. Ebner, N., Riediger, M., & Lindenberger, U. (2010). FACES - A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior research Methods, 42*, 351-362. doi:10.3758/BRM.42.1.351.
43. Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behavior Research Methods, 43*, 468-477. doi:10.3758/s13428-011-0064-1.
44. Lakens, D., Fockenberg, D. A., Lemmens, K. P. H., Ham, J., & Midden, C. J. H. (2013). Brightness differences influence the evaluation of affective pictures. *Cognition and Emotion, 27*(7), 1225–1246. https://doi.org/10.1080/02699931.2013.781501.
45. Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). London: Sage.
46. Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.