**Results**

***Main Analyses with Raw Scores for All Stimuli (Happy, Sad, Angry, and Scared)***

To understand the differences in responses to each emotion and type of stimulus, across age groups, we conducted a series of mixed ANOVAs. In all analyses, age group (primary school, middle school, high school, university students) was the between-groups variable, while the emotion of the stimulus (happy, sad, angry, scared) was the within-groups variable. Separate ANOVAs were conducted for the 3 types of stimuli – vocal bursts, music with congruent lyrics and melody, and music with incongruent lyrics and melody. For cases in which the assumption of sphericity was violated, Greenhouse-Geisser corrections were used. Our dependent variable was participants’ recognition of the emotion presented. In the case of the mismatched music, because the four target emotions were conveyed in the lyrics, answering with the emotion of the lyrics was coded as “correct”, while all other answers were “incorrect”.

*Incongruent Melody and Lyrics*

As predicted, a significant main effect of age was found, *F*(3, 168) = 3.70, *p* = .013, η2*p*= .06. Least significant difference (LSD) follow up tests showed that both primary (*M* = .45) and senior students (*M* = .42) were more likely to label the music using the lyrics than university students (*M* = .14), *p*s ≤ .012, although middle school students (*M* = .34) scored similarly to all age groups, *p*s ≥ .051. Additionally, the main effect of emotion was significant, *F*(3, 504) = 8.97, *p* < .001, η2*p*= .051. Overall, participants were more inclined to choose the emotion of the lyrics for the music with sad lyrics and happy melody (*M* = .54) than the music with happy lyrics and sad melody (*M* = .36), angry lyrics and happy melody (*M* = .28) and scared lyrics and happy melody (*M* = .32), *p*s ≤ .001.

Finally, the emotion by age group interaction was significant, *F*(9, 504) = 4.71, *p* < .001, η2*p*= .08. For the happy incongruent music, primary school children (*M* = .59) were more likely to use the lyrics to determine emotion than middle school (*M* = .27), senior school (*M* = .32) and university (*M* = .28) students, *p*s ≤ .009. For the sad incongruent music, both primary (*M* = .71) and middle school (*M* = .69) children were equally likely to use the lyrics to determine emotion, *p* = .911. These younger age groups were more likely to use the lyrics than both senior school (*M* = .37) and university students (*M* = .40), *p*s ≤ .012, and there were no differences between senior school and university students, *p* = .696.

Contrary to predictions, the angry and scared incongruent music did not follow the pattern outlined in previous research (e.g. Morton & Trehub, 2007). For the angry incongruent music, primary (*M* = .26), middle (*M* = .23) and university (*M* = .21) students all scored similarly, *p*s ≥ .622. The low means for these variables suggest that these age groups did not use the lyrics to determine emotion for music with angry lyrics. Senior school students (*M* = .42) were more likely to use the lyrics to determine emotion than university students for this music, *p* = .024, although were as likely to use the lyrics as the other two age groups, *p*s ≥ .089. Finally, for the scared incongruent music, senior school students (*M* = .55) were more likely to use the lyrics to determine emotion than primary (*M* = .24), middle (*M* = .15) and university (*M* = .34) students, *p*s ≤ .017. The primary school, middle school, and university students scored similarly, *p*s ≥ .103. These results suggest that older adolescents are more likely to use lyrics to determine emotion when lyrics and melody are incongruent than other age groups, for angry and scared music, while for music with happy and sad incongruent lyrics and melody, young children use the lyrics more than adolescents do.

*Congruent Melody and Lyrics*

For music with congruent lyrics and music, a significant main effect of age was found, *F*(3, 168) = 3.45, *p* = .018, η2*p*= .058. Primary school children (*M* = .58) performed similarly to middle school children (*M* = .65), *p* = .160, and worse than senior school (*M* = .67) and university (*M* = .72) students, *p*s ≤ .030. Middle schoolers performed similarly to all other age groups, *p*s ≥ .154, and senior school and university students scored similarly, *p* = .187. A significant main effect of emotion was also found, *F*(2.64, 442.78) = 78.89, *p* < .001, η2*p*= .32. Post-hoc comparisons showed significant differences between all 4 emotions, *p*s ≤ .017, such that participants were more likely to recognise the sad music (*M* = .94; *p*s < .001) than the happy (*M* = .77), angry (*M* = .65) and scared music (*M* = .27).

For congruent music, a significant emotion by age group interaction was found, *F*(7.91, 442.78) = 2.45, *p* = .014, η2*p*= .04. There were no age effects for the happy music (*M*s range from .61-.66). For the sad music, primary school children (*M* = .82) scored significantly lower than university students (*M* = 1.0), *p* = .048, but similarly to middle (*M* = 1.0) and senior (*M* = .95) students, *p*s ≥ .087. For the angry music, primary school children (*M* = .56) performed significantly worse than middle school (*M* = .81), senior school (*M* = .89) and university (*M* = .81) students, *p*s ≤ .016. All other age groups performed similarly, *p*s ≥ .27. Finally, for the scared music, all participants scored poorly. University students (*M* = .43) performed better than both middle (*M* = .19) and senior (*M* = .19) school students, *p*s ≤ .016, but similarly to primary school students, (*M* = .29). There were no other differences between age groups.

These data indicate that while all age groups scored similarly for the happy music, young children tended to score lower than university students for the sad music and worse than all age groups for the angry music. The mean scores for the scared music were low, with participants showing low agreement for the emotion category conveyed by the music.

*Vocal Bursts*

Primary school children correctly recognised 87.9% of vocal bursts, middle school children recognised 83.7%, senior school students recognised 88.5%, and university students recognised 92.4%. Contrary to our hypothesis, no main effect of age group was found, *F*(3, 168) = .74, *p* = .531, η2*p* = .013.

A significant main effect of emotion was found for vocal bursts, *F*(2.22, 372.19) = 18.35, *p* < .001, η2*p* = .10. Participants were significantly more likely to recognise happy (*M* = .92) than angry (*M* = .89), *p* = .008, or scared (*M* = .81), *p* < .001. There were no differences between recognition of happiness and sadness (*M* = .91), *p* = .118. Similarly, participants were more likely to recognise sad and angry than scared, *p*s < .001, Overall, this suggests that fear was the most difficult emotion to recognise from vocal bursts, although all vocal bursts were recognised in more than 80% of trials. No significant interaction between age and emotion of the vocal bursts was found, *F*(6.65, 372.19) = 1.34, *p* = .235, η2*p* = .023.

*Emotion Lexicon*

Differences between age groups on the number of emotion words produced was examined in a one-way ANOVA. On average, primary school children produced 7.74 (*SD* = 2.22) words, middle school children produced 8.50 (*SD* = 4.78) words, senior school participants produced 8.63 (*SD* = 4.34) words, and university students produced 9.53 (*SD* = 3.43) words. No significant main effect for age groups was found, *F*(3, 168) = 1.48, *p* = .222. Additionally, the size of emotion lexicon was not correlated with correctly recognising congruent music, *r* = .02, *p* = .799, nor using the lyrics to determine emotion for incongruent music, *r* = -.04, *p* = .640.

***Incongruent Music Stimuli with Raw Scores for Happy, Sad, and Scared***

Incongruent music stimuli were examined again with the raw scores, instead coding the melody as “correct” rather than the lyrics. The 4 (Age Group: primary, middle, senior, university) x 3 (Emotion: happy lyrics/sad melody, sad lyrics/happy melody, scared lyrics/happy melody) mixed design ANOVA showed no main effect of emotion, *F*(2, 336) = .265, *p* =.768, η2*p*= .002, but a main effect of age, *F*(3, 168) = 5.10, *p* = .002, η2*p*= .083. Least significant difference (LSD) follow up tests showed that university students (*M* = .53) were significantly more likely to answer with the emotion of the melody than primary (*M* = .34), middle, (*M* = .32), and senior students (*M* = .36), *p*s ≤ .003. Finally, the emotion by age group interaction was significant, *F*(6, 336) = 4.47, *p* < .001, η2*p*= .074. For the happy incongruent music, university students (*M* = .61) were more likely to answer with the emotion of the melody than primary (*M* = .27), middle (*M* = .32), and senior (*M* = .41) students, *p*s ≤ .031. For the sad incongruent music, university (*M* = .52) and senior (*M* = .50) students were more likely to answer with the emotion of the melody than both primary (*M* = .25) and middle school (*M* = .20) students, *p*s ≤ .011. Finally, for the scared incongruent music, senior school students (*M* = .18) were less likely to use the melody to determine emotion than primary (*M* = .51), middle (*M* = .43) and university (*M* = .48) students, *p*s ≤ .021.

Table S1

*Validation Sample Confusion Matrix*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Song | Stimulus | Happy | Sad | Angry | Scared |
| Happy lyrics, sad melody | Text | **75.0** | 25.0 | 0.0 | 0.0 |
| Sad lyrics, happy melody | Text | 1.8 | **70.9** | 9.1 | 18.2 |
| *Angry lyrics, happy melody* | *Text* | *11.1* | *22.2* | ***24.1*** | *42.6* |
| Scared lyrics, happy melody | Text | 14.3 | 3.6 | 5.4 | **76.8** |
|  |  |  |  |  |  |
| Happy lyrics, sad melody | Melody | 26.2 | **66.7** | 2.4 | 4.8 |
| Sad lyrics, happy melody | Melody | **72.7** | 22.7 | 2.3 | 2.3 |
| *Angry lyrics, happy melody* | *Melody* | ***55.3*** | *4.3* | *31.9* | *8.5* |
| Scared lyrics, happy melody | Melody | **71.7** | 4.3 | 8.7 | 15.2 |
|  |  |  |  |  |  |
| Happy | Text | **87.5** | 7.1 | 1.8 | 3.6 |
| Sad | Text | 3.6 | **82.1** | 3.6 | 10.7 |
| Angry | Text | 3.6 | 10.7 | **21.4** | 64.3 |
| *Scared* | *Text* | *11.3* | *37.7* | *39.6* | ***11.3*** |
|  |  |  |  |  |  |
| Happy | Melody | **35.6** | 6.7 | 15.6 | 42.2 |
| Sad | Melody | 2.3 | **93.0** | 2.3 | 2.3 |
| Angry | Melody | 7.0 | 4.7 | **74.4** | 14.0 |
| *Scared* | *Melody* | *35.7* | *7.1* | *33.3* | ***23.8*** |

*Note. Bolded results indicate correct answers in each column. Italicised rows indicate that the stimulus was removed from the main analyses*