RADBOUD UNIVERSITY NIJMEGEN

BACHELOR DEGREE THESIS IN ARTIFICIAL INTELLIGENCE

Emotion Stroop task: Recognizing emotions from face and voice

---

Author:
Janne Weijkamp
Janneweijkamp@student.ru.nl
studentnumber: 0733318

Supervisors:
Dr. M. Sadakata
Dr. I.G. Sprinkhuizen-Kuyper

13-08-2013

Emotion Stroop Task
Recognizing Emotions from Face and Voice

Janne Weijkamp

Abstract

To get insights into recognition of emotions from face and voice, we developed a special form of the Stroop task (the Emotion Stroop task). Instead of reading words and naming colors, the tasks are to recognize emotions either from faces (visual) or voices (auditory). Because musical expertise has been shown to positively correlate with recognition of emotion in speech prosody, we examined the difference in performance on the Emotion Stroop task between musicians and non-musicians. We found the Stroop effect on the Emotion Stroop task, as well as the Interference effect. Furthermore, people were faster and more correct on recognizing emotions from face than from voice. Results also showed that when participants had to ignore the face and judge emotion in the voice, musicians were more correct than non-musicians.

1 Introduction

In human-robot interaction there is an increased interest in building robots that can interact with people in a social, life-like manner [1, 2]. Some examples of robotic applications where social interaction with the human is important are: robotic nursemaids for elderly people, robot pet for children or a therapy robot for autistic children [2].

Van Breemen et al. [1] developed the “iCat”: a robotic research platform for studying social human-robot interactions. The platform is a desktop user-interface robot than can express emotions by facial expression. The iCat can generate different facial expressions (happy, surprise, angry, sad). Philips Research (Eindhoven, the Netherlands) made the robot available to stimulate research in the field of social human-robot interaction.

Dautenhahn et al. [2] designed a minimally expressive robot, called “KASPAR”, which can be used for human-robot interaction research. They show how the robot can be used in robot-assisted play for children with autism. The robot has minimal expressive facial expression to not overwhelm the children with social cues, but allow them to learn how to interpret a few basic emotions.

To improve social interaction between humans and robots it is important for the robot to understand the emotion of the human, as well as for the human to understand the emotion expressed by the robot. A lot of research has been done on building systems to recognize human emotions from face and/or voice [3, 4, 5, 6, 7, 8, 9]. This bachelor project aimed to find out more about how humans deal with recognition of emotions from face and voice. More specifically, the goal was to find out if people are better in recognizing emotions from either face or voice. This could give us a notion of which modality a robot should use to express emotions to a human (face or voice). Furthermore, this project aimed to find out if there exist differences between groups of people in which modality they find easier (face or voice). If such a difference exist, it could be possible to adjust robots to specific user groups.

De Gelder & Vroomen [10] used a bimodal perception situation, in which varying degrees of discordance can be created between the affects expressed in a face or voice. To get insight into how we integrate emotional information from face and voice they did three experiments. In their experiments they used sad/happy voices and faces on a continuum between sad/happy. In the first experiment participants were presented with a face and a voice at the same time randomly combined, and asked to indicate whether the person was happy or sad. Results showed that identification of the emotion in the face is biased in the direction of the simultaneously presented voice. In the second experiment participants were presented with the same stimuli but now instructed to judge the face and ignore the voice. Results showed again the identification of the emotion in the face is biased in the direction of the simultaneously presented voice. In the last experiment the participants had to judge the voice and ignore the face. There the results showed the reverse effect, identification of the emotion in the voice is biased in the direction of the simultaneously presented face.

Sadakata et al. [11] showed that musicians have higher sensitivity when comparing small differences in linguistic timing information and spectral information. Musicians also had an increased ability in learning and identifying linguistic timing information.

Musical expertise has also been shown to positively correlate with recognition of emotion in speech prosody [12]. Therefore we will examine the difference in performance on the Emotion Stroop task between musicians and non-musicians.

Our predictions following these findings are that musicians might experience a stronger interference effect of voice information when presented with incongruent materials because they catch the detailed information of voice more than non-musicians. If we find a difference between musicians and non-musicians in how they deal with recognition of emotion, it indicates that we could adjust robots to certain user-groups. If, for example, a user-group would be better in perceiving emotions from voice than from face, we should adjust the robot to express emotions by voice.

1.1 Original Stroop task

The original Stroop task, named after John Ridley Stroop [13], looks into the interference of two cognitive processes: naming colors and reading words. In the test participants are presented with names of colors (e.g. “blue”, “green”, or “red”), printed in the ink of either the same or a different color.

In one experiment, participants were presented with a set of words where the names of colors and the ink they are printed in are congruent (e.g. “red” printed in red ink), and a set of words were the names of the colors and the ink they are printed in are incongruent (e.g. “red” printed in blue ink). The task for the participants was to name the color of the words. Results showed that people were significantly faster on congruent trials than on incongruent trials. This means that people were distracted by (and cannot ignore) incongruent word meaning. This is called the Stroop effect (see Figure 1.1).

In another experiment, participants had to do two tasks. In one task they were presented with a set of words where the names of colors are printed in black, and a set of words were the names of the colors and the ink they are printed in are incongruent. The task for the participants was to read the words. Results showed people were hardly distracted by the incongruent ink of the words. This means that there is no interference of color of the word on reading words. In the other task, participants were presented with a set of squares printed in different colors, and a set of words were the names of the colors and the ink they are printed in are incongruent. The task for the participants was to name the colors. Results showed that people were strongly distracted by the name of the color (word meaning). This means that there is an interference of word reading on naming colors. The two tasks together show that people were significantly more distracted by word meaning when naming colors than by color of the words when reading words. This is called the interference effect (see Figure 1.1).

There are several explanations for the Stroop effect and the interference effect. The discussion about why the interference effect is occurring is still continuing. The most generally accepted and often used in books as the only explanation involves automaticity. Automatic processes are faster and require less attention than controlled processes and in the same time are (almost) impossible to suppress. This paper is not trying to find out the explanation for the Stroop effect, but focusses on finding out if a similar process occurs with different stimuli. To still be able to explain the results in understandable language the automaticity explanation will be used, but be aware that this is not the only possible explanation for the effect.

1.2 Emotion Stroop task

In this bachelor thesis a special form of the Stroop task (the Emotion Stroop task) was developed. Instead of reading words and naming colors, on the Emotion Stroop task the tasks are recognizing emotions from faces (visual) and voices (auditory). Thus, on the Emotion Stroop task there are two modalities (sound/vision), while in the original Stroop task there is only one (vision).

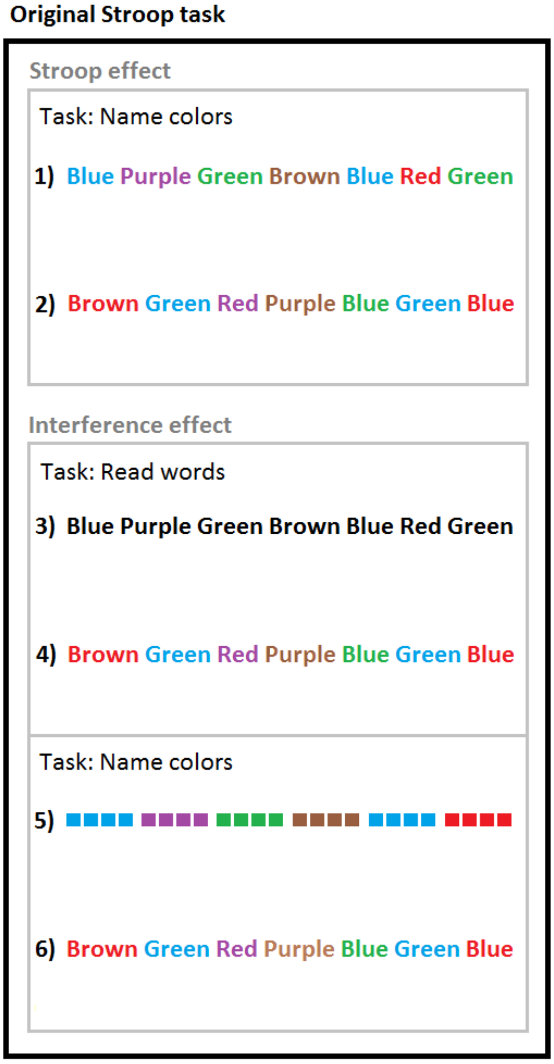
In the Emotion Stroop task, participants are presented with faces and voices with different emotions. In Figure 1.2 is displayed how the Emotional Stroop task relates to the original Stroop task. Before carrying out the experiment it was not yet clear which process was more automated (or faster); emotion recognition from face or emotion recognition from voice. In that way, reading words on the original Stroop task could have been related to recognizing emotions from voice on the Emotion Stroop task and naming colors on the original Stroop task to recognizing emotions from face on the Emotion Stroop task.

The next chapter, Method, design and procedure, describes the details about the tasks in the Emotion Stroop task.

1.3 Research questions

The Emotional Stroop task aimed to find out more about the cognitive processes of emotion recognition from face and voice. Furthermore, the goal was to find out if there are differences between groups (musicians and non-musicians) in emotion recognition from face and voice. Hence, this project aimed to answer the following questions:

1. Can we find the Stroop effect on the Emotion Stroop task?
2. Can we find the Interference effect on the Emotion Stroop task?
3. Are emotions easier to recognize from a face or a voice?
4. Are these effects the same for musicians?



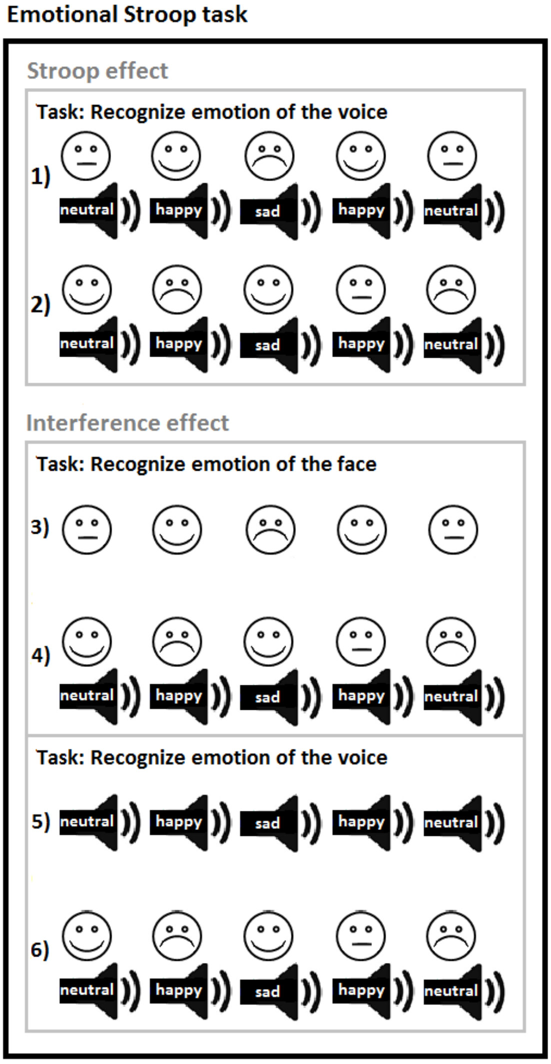


Figure 1.1: The Stroop effect and interference effect in the original Stroop task. The Stroop effect means that we are faster on naming the colors of line 1 than 2, because we are distracted by the incongruent word meaning. The interference effect means that we are more distracted by word meaning when naming colors, than by colors when reading words. There is a small difference in reaction time between reading words of line 3 or 4, while there is a significant bigger difference between naming colors of line 5 and 6.

Figure 1.2: The Stroop effect and interference effect in the Emotions Stroop task. The Stroop effect means that we are faster on recognizing the emotions in the voices of line 1 than 2, because we are distracted by the incongruent emotion on the face. The interference effect means that we are more distracted by emotion on the face when recognizing emotion in the voice, than by emotion in the voice when recognizing emotion on the face. There is a small difference between recognizing emotion of the faces of line 3 or 4, while there is a significant bigger difference between recognizing emotions of the voices of line 5 and 6.

2 Method, design and procedure

2.1 Method

Participants. Sixteen musicians (mean age: 29.25) and sixteen non-musicians (mean age: 21.81) were asked to volunteer in this experiment. The criteria for being a musician were: more than 5 years of formal musical lessons (with a teacher), all practicing instruments, actively practicing the instrument(s) more than 2.5 hours per week. The criteria for being a non-musician were: less than 2 years of formal musical lessons (with a teacher), not practicing any instrument(s) for the last 2 years.

Visual materials. Twelve black-and-white photographs of faces with sad, happy and neutral emotions were used. The photographs are from four different people (two women and two men). See appendix A for more information about the visual material.

Auditory materials. Twelve humming sounds of voices (with an average duration of ± 600 ms) with a sad, happy and neutral emotion were used. The voices are from four different people (two women and two men) with for every person three different emotions. See appendix A for more information about the auditory material.

2.2 Design and procedure

An experiment was constructed using an open source application called Psychopy [14].

The experiment consisted of four tasks: a face task, a voice task, a focus-on-face task and a focus-on-voice task. These tasks consisted of three different kinds of trials. On a *visual trial,* a fixation cross was presented for 1500 ms, and one of the twelve faces was presented at the same location for 600 ms directly after. On an *auditory trial* one of the twelve voices was presented. On a *bimodal trial* a fixation cross was presented for 1500 ms. Directly after, one of the twelve faces was presented for 600 ms together with one of the twelve voices. The face and voice where chosen randomly by the program, therefore, bimodal trials were either congruent (e.g. happy face + happy voice) or incongruent (e.g. happy face + sad voice).

In the *face task,* two times 12 visual trials (12 faces) were presented in a random order. Participants had to make a forced choice on whether the emotion expressed on the face was happy, neutral or sad. Reaction time was measured from the onset of the picture.

In the *voice task,* two times 12 auditory trials (12 voices) were presented in a random order. Participants had to make a forced choice on whether the emotion expressed in the voice was happy, neutral or sad. Reaction time was measured from the onset of the sound.

In the *focus-on-face task,* two times 72 bimodal trials (2 genders * 6 faces * 6 voices) were presented in a random order. The six faces of the men were combined with the six voices of the men and the six faces of the women were combined with the six voices of the women. Participants had to make a forced choice on whether the emotion expressed on the face was happy, neutral or sad. Reaction time was measured from the onset of stimuli. At half of the task (after 72 trials) there was a break.

In the *focus-on-voice task,* two times 72 bimodal trials (2 genders * 6 faces * 6 voices) were presented in a random order. The six faces of the men were combined with the six voices of the men and the six faces of the women were combined with the six voices of the women. Participants had to make a forced choice on whether the emotion expressed in the voice was happy, neutral or sad. In this task participants were instructed to keep looking at the faces. Reaction time was measured from the onset of the stimuli. At half of the task (after 72 trials) there was a break.

To avoid learning effects, the first two tasks were f*ocus-on-face task* and f*ocus-on-voice* task (counterbalanced) and the last two tasks were the f*ace task* and the *voice task* (counterbalanced).

Participants had to respond by pressing one of the three buttons: happy, sad or neutral (See figure 2.1). Furthermore, voices were presented to the participants through headphones.

Before every task, instructions about the task were presented on the screen and participants were encouraged to ask questions if something was unclear. Participants were instructed to respond as accurate and as fast as they could. Before the first task, participants were presented with twelve practice trials. To avoid learning effects, the faces and voices used in the practice trials were different from those used in the experimental trials.

Since it is possible to ignore the face, participants were instructed to not close their eyes and to keep looking at the faces in the *focus-on-voice task*. As an extra control to make sure people would not ignore the faces, the position of the fixation cross (and following presented face) changed between trials. The position was for every trial randomly selected out of the two possible positions. See figure 2.2 for these two positions. To not influence participants in their responses by pressing buttons (which were aligned horizontally) the positions are only changing vertically and not horizontally.

After the experiment, participants were asked to fill in a questionnaire about their musical background. (See Appendix B for the questionnaire)

In total, the experiment took between 20 and 30 minutes, depending on how fast people were and how long the instructions took. Participants were tested in a soundproof room.

Figure 2.1: Keyboard with special buttons for happy, neutral and sad.

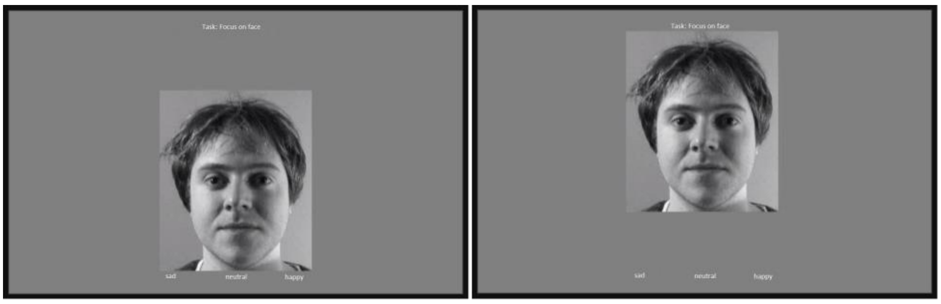


Figure 2.2: Two possible positions of presenting faces, as an extra control to make sure people do not ignore the face while responding to voice.

Results

Responses with a reaction time longer than 3 standard deviations from the mean were identified as outliers (musicians: 2581ms, non-musicians: 2414ms). Identified outliers (less than 1.6% of the data) were discarded from the analyses. Overviews of the data are plotted in Figure 3.1 for Reaction times and Figure 3.2 for Correct response rates.

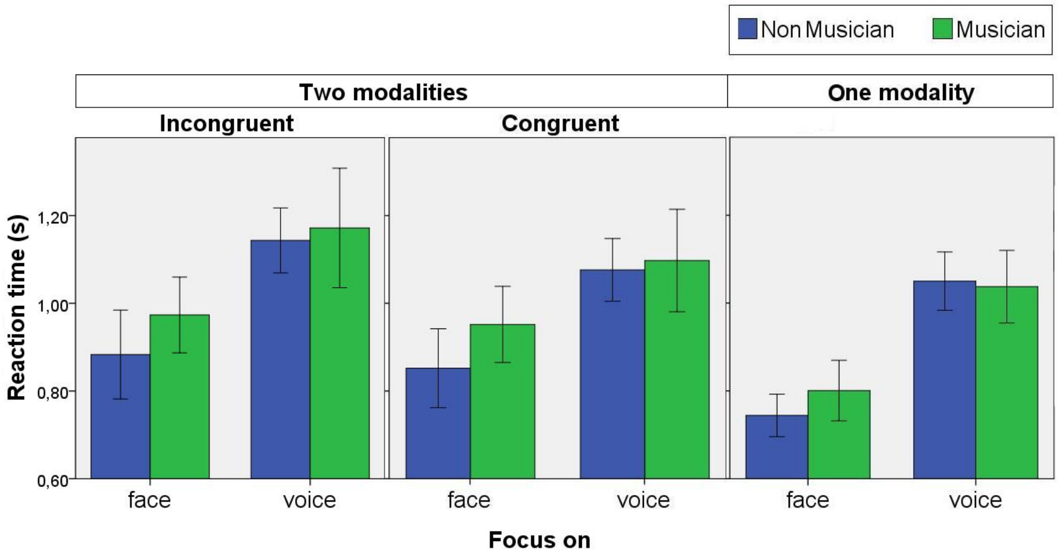


Figure 3.1: Estimated marginal means of Reaction time on Emotion Stroop task for Incongruent, Congruent and control condition (One Modality) separately plotted for musicians and non-musicians.

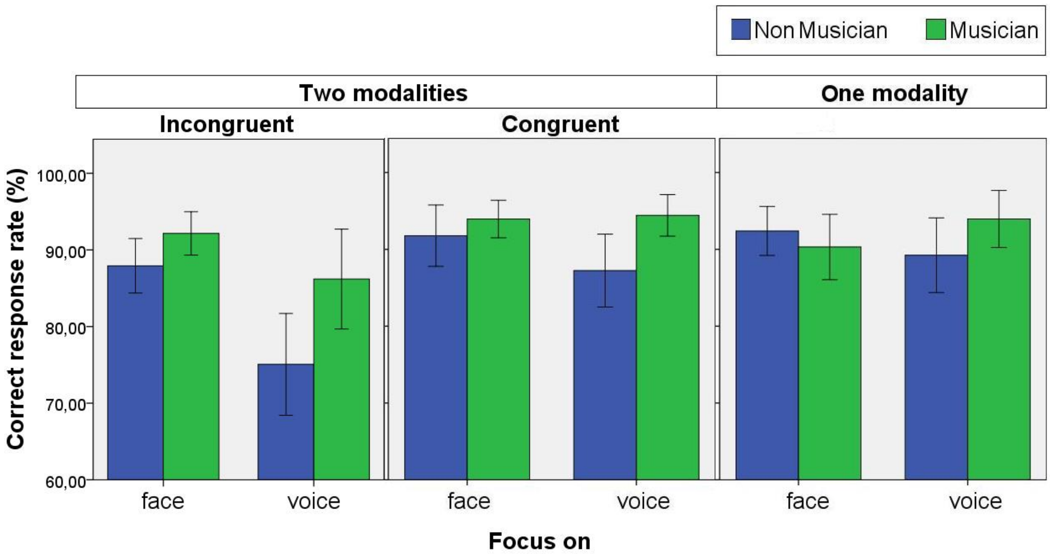


Figure 3.2: Estimated marginal means of Correct response rate on Emotion Stroop task for Incongruent, Congruent and control condition (One Modality) separately plotted for musicians and non-musicians.

2.3 Cost-analysis: Interference effect

The first analysis looks into the interference effect on the Emotion Stroop task. For this, a cost-analysis was done, for which four difference variables were calculated:

	Reaction time		Correct response rate
1	Reaction time difference between incongruent trials from *Focus-on-face* and *Face task* conditions	3	Correct response rate difference between incongruent trials from *Focus-on-face* and *Face task* conditions
2	Reaction time difference between incongruent trials from *Focus-on-voice* and *Voice task* conditions	4	Correct response rate difference between incongruent trials from *Focus-on-voice* and *Voice task* conditions

Table 3.1: Calculated difference variables for cost-analysis. The Interference effect relates to comparing 1 with 2, and 3 with 4 in the table.

A MANOVA-Repeated measures analysis was performed with Reaction time and Correct response rate as dependent variables, Modality (face/voice) as within-subjects independent variable, and Musical expertise (musician/non-musician) as between-subjects independent variable. For SPSS output of this analysis, see Appendix C.

Multivariate tests showed a main effect of Modality ($F(2,29) = 13.897$, $p<0.0005$; Wilk's $\Lambda = .511$, partial $\eta^2 =.489$), and a main effect of Musical expertise ($F(2,29) = 4.806$, $p=.016$; Wilk's $\Lambda = .751$, partial $\eta^2 = .249$).

The test of between-subject effects showed that musicians tend to give more correct responses than non-musicians ($F(1,30) = 8.595$, $p=.006$; partial $\eta^2 = .223$), while there was no significant difference in Reaction time. This means that musicians were better in focusing on emotion of instructed modality than non-musicians when an extra (incongruent) modality is introduced.

The univariate tests showed a significant main effect of Modality on Correct response rate ($F(1,30) = 21.017$, $p<.0005$; partial $\eta^2 = .412$), while no significant difference in Reaction time. This means, adding a face with an incongruent emotion when responding to an emotion of a voice is more distracting, than adding a voice with an incongruent emotion when responding to an emotion of a face. This suggests that recognizing emotions from face is a more automated or faster process than recognizing an emotion from a voice. Figure 3.3 displays this (Stroop) interference effect in a more visible way.

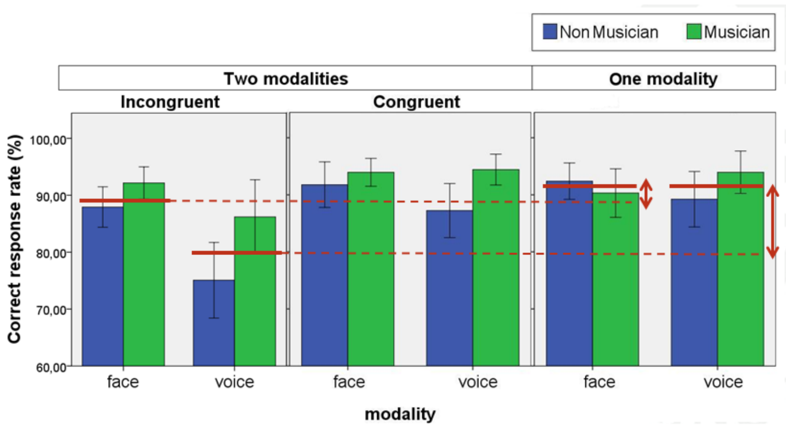


Figure 3.3: Estimated marginal means of Correct response rate on Emotion Stroop task for Incongruent, Congruent and control condition (One Modality) separately plotted for musicians and non-musicians. The difference between the red arrows displays the Interference effect. The error bars show standard errors.

2.4 2x2x2 Mixed MANOVA

The second analysis looks into the Stroop effect, the effect of Modality and the effect of Musical expertise. For this, a MANOVA-Repeated measures analysis was performed with Reaction time and Correct response rate as dependent variables, Modality (face/voice) and Congruency (incongruent/congruent) as within-subjects independent variables, and Musical expertise (musician/non-musician) as between-subjects independent variable. In this analysis only the data from the *focus-on-face* task and *focus-on-voice* task were used. For the reason that in the task with only one modality (face task or voice task) the variable Congruency is not applicable. For the SPSS output of this analysis, see Appendix D.

The multivariate tests showed a significant interaction effect between Modality and Congruency ($F(2,29) = 8.400$, $p=.001$; Wilk's $\Lambda = .633$, partial $\eta^2= .367$), and a main effect of Musical expertise ($F(2,29) = 3.542$, $p= .042$; Wilk's $\Lambda = .804$, partial $\eta^2= .196$).

The univariate tests showed a significant interaction effect between Modality and Congruency on Reaction time ($F(1,30) = 12.314$, $p=.001$; partial $\eta^2 = .291$), as well as on Correct response rate ($F(1,30) = 13.482$, $p=.001$; partial $\eta^2 = .310$). See Figure 3.4. This means that the effect of Modality(face faster and more correct than voice) is different for Congruent than for Incongruent trials. In Figure 3.4 you can see that the effect of Modality is bigger for incongruent trials.

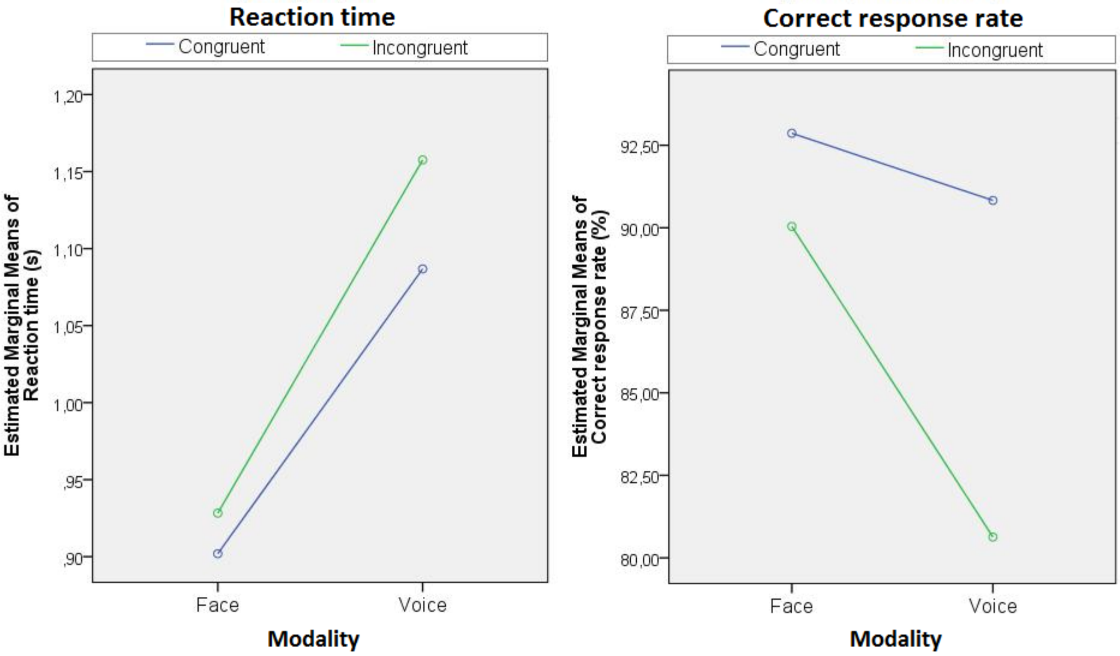


Figure 3.4: Plots displaying the interaction effect between Modality and Congruency. On the left the Estimated marginal means of Reaction time on Emotion Stroop task for recognizing emotion of face and voice, separately plotted for Incongruent and Congruent conditions. On the right the same kind of plot, with the Estimated marginal means of Correct response rate. They show that the effect of Modality is different for Congruent than for Incongruent trials.

2.4.1 Musical expertise

The test of between-subjects effects showed a significant main effect of Musical expertise on Correct response rate ($F(1,30)= 7.108$, $p=.012$; partial $\eta^2= .192$), while there was no significant difference in Reaction time. So, musicians are significantly more correct than non-musicians. The Pairwise Comparisons (see Appendix D, Table 11), Univariate Tests (See Appendix D, table 12), and Estimates (see Appendix D, Table 10) were used to find out exactly were musicians and non-musicians are different.

These showed that musicians were more correct on congruent trials from *Focus-on-voice* task ($F(1,30)= 6.921$, $p=.013$; partial $\eta^2= .187$), as well as on incongruent trials from *Focus-on-voice* task Congruency and Stroop effect ($F(1,30)= 5.734$, $p=.023$; partial $\eta^2= .160$)., while no significant difference between musicians and non-musicians was found on the *Focus-on-face* task. This effect is displayed in Figure 3.6. This indicates that musicians are better in focusing on the emotion from a voice when they are presented with a face and a voice.

2.4.2 Congruency and Stroop effect

For the Stroop effect we looked into the difference on *Focus-on-voice* task between congruent and incongruent trials. Since a significant interaction effect between Modality and Congruency is found, the Pairwise Comparisons (see Appendix D, Table 5) and Estimates (see Appendix D, Table 4) were used.

This showed the Stroop effect: people were significantly faster on congruent trials than on incongruent trials when recognizing emotions from a voice (1.087s compared to 1.158s; $p<.0005$), and also significantly more correct on congruent trials than on incongruent trials when recognizing emotions from a voice (90.83% compared to 80.631; $p<.0005$). It also showed the effect on the *Focus-on-face* task: people were significantly faster on congruent trials than on incongruent trials when recognizing emotions from a face (.902s compared to .928s; $p<.0005$), and also significantly more correct on congruent trials than on incongruent trials when recognizing emotions from a voice (90.83% compared to 80.631; $p<.0005$). These effects are displayed in Figure 3.5 for Reaction time and Figure 3.6 for Correct response rate.

In summary, the Stroop effect is found on Reaction time as well as on Correct response rate. Furthermore, people are in general significantly faster and more correct on congruent trials (when the emotion of the face and voice are matching), than on incongruent trials (when the emotion of the face and voice are not matching).

2.4.3 Modality: Are emotions easier to recognize from a face or a voice?

Since a significant interaction effect between Modality and Congruency is found, the Pairwise Comparisons (see Appendix D, Table 8) and Estimates (see Appendix D, Table 7) were used. People were significantly faster when recognizing emotions from a face than recognizing an emotion from a voice on congruent trials (.902s compared to 1.087s; $p<.0005$), while there was no significant difference in correct response rate. Furthermore, people were also significantly faster when recognizing emotions from a face than recognizing an emotion from a voice on incongruent trials (.928s compared to 1.158s; $p<.0005$), as well as more correct (90.04% compared to 80.63%; $p<.0005$). These effects are displayed in Figure 3.5 for Reaction time and Figure 3.6 for Correct response rate. In summary, people are significantly faster and more correct in recognizing emotions from a face than from a voice.

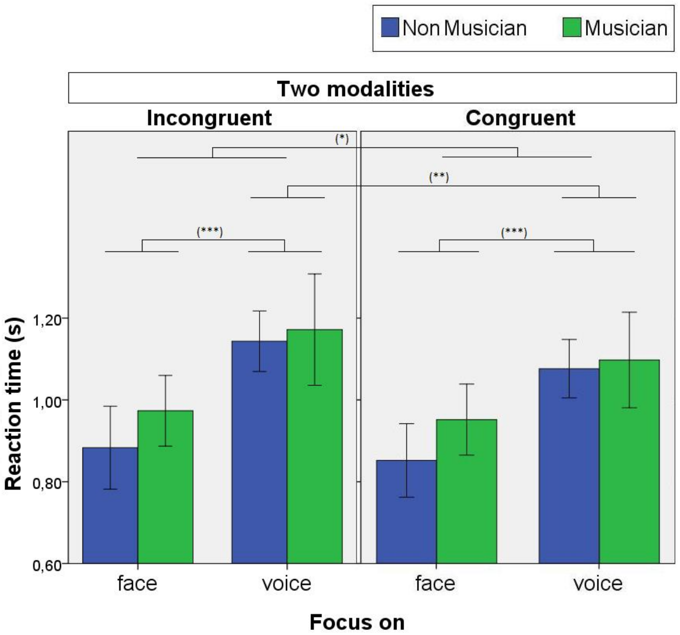


Figure 3.5: Estimated marginal means of Reaction time on Emotion Stroop task for Incongruent, Congruent conditions, separately plotted for musicians and non-musicians. (*) Displays that people are significantly faster on congruent trials than on incongruent trials. (**) Displays the Stroop effect on Reaction time. (***) Displays the effect of Modality. People are significantly faster in recognizing emotions from a face than a voice. The error bars show standard errors.

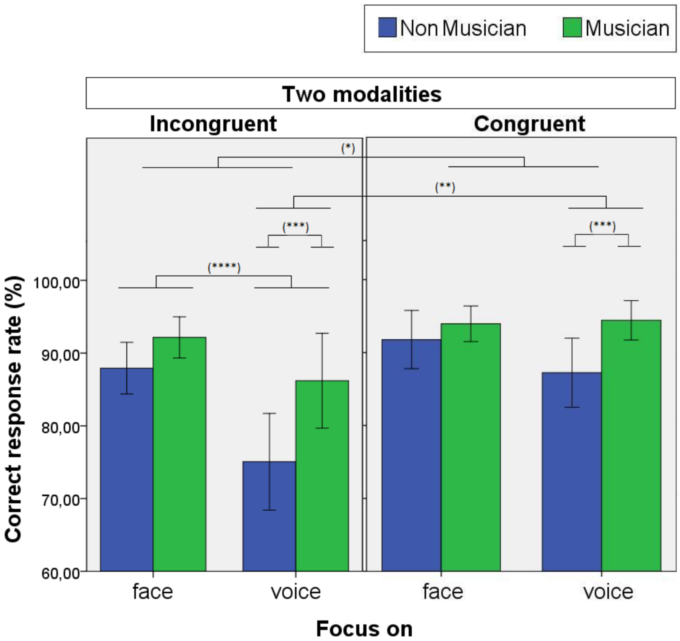


Figure 3.6: Estimated marginal means of Correct response rate on Emotion Stroop task for Incongruent, Congruent conditions, separately plotted for musicians and non-musicians. (*) Displays that people are significantly more correct on congruent trials than on incongruent trials. (**) Displays the Stroop effect on Correct response rate. (***) Displays the effect of musical expertise. Musicians are significantly more correct of the Focus-on-voice trials. (****) Displays the effect of Modality. People are more correct in recognizing emotions from a face than a voice on incongruent trials. The error bars show standard errors.

2.5 Extra results

2.5.1 Emotions

To look into the effect of emotions, a MANOVA-Repeated measures analysis was performed with Reaction time and Correct response rate as dependent variables, Modality (face/voice) and Emotion (happy/neutral/sad) as within-subjects independent variables, and Musical expertise (musician/non-musician) as between-subjects independent variable. In this analysis only the data for one modality (*Face task* and *Voice task*) were used. For SPSS output of this analysis, see Appendix E.

Multivariate tests showed a main effect of Modality ($F(2,29) = 75.708$, $p<0.0005$; Wilk's $\Lambda = .161$, partial $\eta^2 =.839$), and a main effect of Emotion ($F(2,29) = 3.902$, $p=.013$; Wilk's $\Lambda = .634$, partial $\eta^2 = .366$).

The univariate tests showed a significant main effect of Modality on Reaction time ($F(1,30) = 144.230$, $p<.0005$; partial $\eta^2 = .828$), while no significant difference in Correct response rate. This is the effect that people were faster in recognizing emotion from face than from voice. Furthermore, it showed a significant main effect of Emotion on Reaction time ($F(1,30) = 8.107$, $p=.001$; partial $\eta^2 = .213$), as well as on Correct response rate ($F(1,30) = 3.299$, $p=.044$; partial $\eta^2 = .099$). Pairwise comparisons and Estimates (see Appendix E, table 4 & table 5) showed that reaction time was different for emotions (happy: .889s, neutral: .894s, sad: .945s). People were significantly slower in recognizing sad emotions compared to happy and neutral emotions. Figure 3.7 shows these effects.

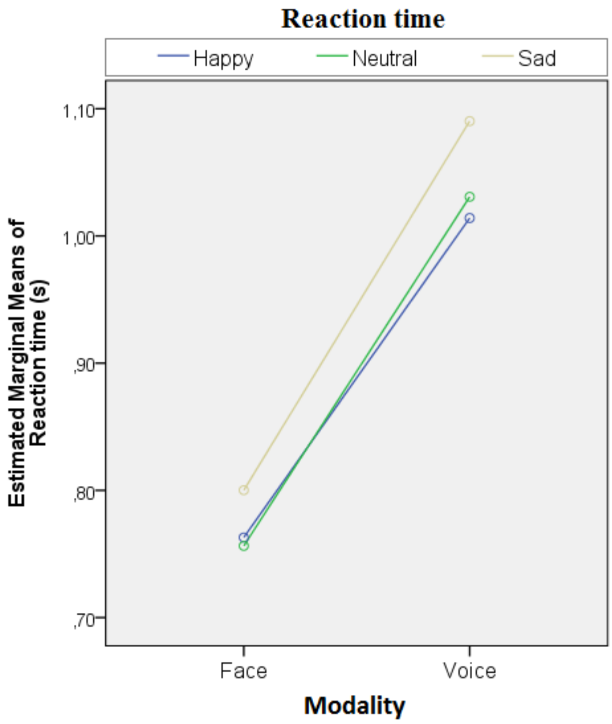


Figure 3.7: Estimated marginal means of Reaction time (s) on Face and Voice task, separately plotted for the three emotions categories: happy, neutral and sad. It shows again that people are faster in recognizing emotions from a face than from a voice. Is also shows that people are slower on recognizing sad emotions than happy or neutral emotions.

2.5.2 Congruent and incongruent emotions

Intuitively, a sad face with a happy voice is more incongruent than a sad face with a neutral voice (here we will refer to this as half(in)congruent). To find out if we can find this back in the data, a MANOVA-Repeated measures analysis was performed with Reaction time and Correct response rate as dependent variables, Congruency (congruent /incongruent /half(in)congruent) as within-subjects independent variable (See table 3.2). Furthermore, and Musical expertise (musician/non-musician) was added as between-subjects independent variable.

Congruent	Incongruent	Half(in)congruent
happy face + happy voice neutral face + neutral voice sad face + sad voice	happy face + sad voice sad face + happy voice	neutral face + happy voice neutral face + sad voice neutral voice + happy face neutral voice + sad face

Table 3.2: Showing congruent, incongruent and half(in)congruent combinations.

Multivariate tests showed a main effect of Congruency ($F(4,27) = 18.124$, $p<0.0005$; Wilk's $\Lambda = .271$, partial $\eta^2 = .729$), and a main effect of Musical expertise ($F(2,29) = 3.646$, $p=.039$; Wilk's $\Lambda = .799$, partial $\eta^2 = .201$).

The univariate tests showed a significant main effect of Congruency on Reaction time ($F(2,60) = 33.968$, $p<.0005$; partial $\eta^2 = .531$), as well as on Correct response rate ($F(2,60) = 34.770$, $p<.0005$; partial $\eta^2 = .537$). Pairwise Comparisons and Estimates (see Appendix F, table 4 & 5) showed that Reaction time and Correct response rate was significantly different on between all three levels of Congruency. People were fastest and most correct on congruent trials, and slowest on incongruent trials. People were significantly faster and more on half(in)congruent trials than on incongruent trials, and significantly slower and less correct on half(in)congruent trials than on congruent trials. It shows that the intuitive thought that, for example, a sad face with a happy voice is more incongruent than a sad face with a neutral voice is also found in the data. Figure 3.8 displays this effect.

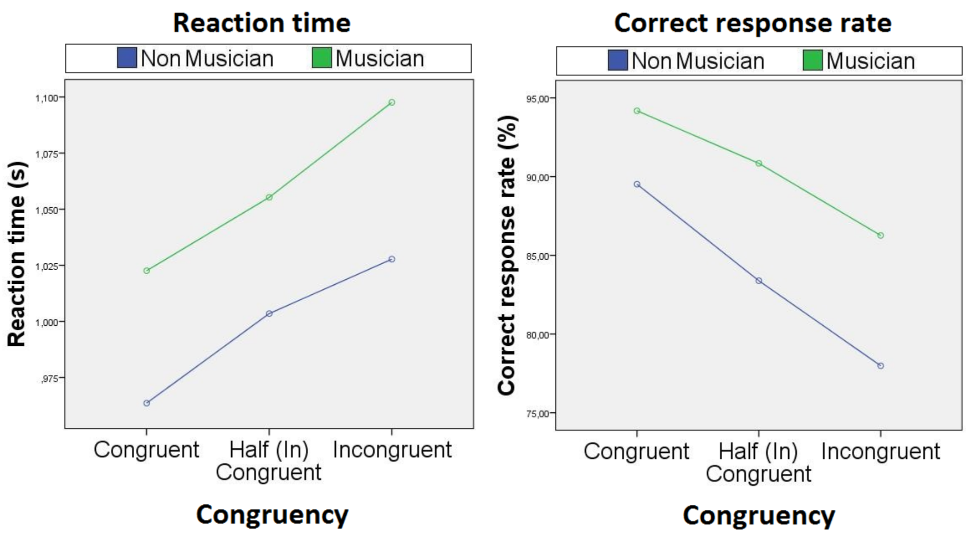


Figure 3.8: Plots displaying the Congruency effect for three different levels of Congruency (see Table 3.2). On the left the Estimated marginal means for Reaction time and on the right for Correct response rate. They show that the intuitive thought that a sad face with a happy voice is more incongruent than a sad face with a neutral voice is also found in the data.

3 Conclusion

The first research question in this project was if the Stroop effect can be found on the Emotional Stroop task. Results showed that the Stroop effect is found on Reaction time and on Correct response rate. When recognizing the emotion in the voice people were more correct when the simultaneously presented face had the same emotion as the voice. In general, people are faster and more correct on congruent trials (when the emotion of the face and voice are matching), than on incongruent trials (when the emotion of the face and voice are not matching).

The second research question in this project was if the Interference effect can be found on the Emotional Stroop task. Results showed that the Interference effect was found on Correct response rate, while there was no significant difference in Reaction time. That means adding a face with an incongruent emotion when responding to an emotion of a voice is more distracting than adding a voice with an incongruent emotion when responding to an emotion of a face. This suggests that recognizing emotions from face is a more automated or faster process than recognizing an emotion from a voice.

The third research question in this project was if emotions are easier to recognize from a face or a voice. Results showed that people were significantly faster and more correct in recognizing emotions from a face than from a voice.

Finally, this project aimed to find out if there are differences between musicians and non-musicians on performance on the Emotion Stroop task. Results showed that when participants had to ignore the face and judge emotion in the voice, musicians were more correct than non-musicians, while there was no significant difference in Reaction time.

These findings suggest that musicians are better in focusing on the emotion of the voice when presented with a face and a voice. The same effect was not found for face, which indicates that musicians are not just better in focusing on one modality.

4 Discussion

This bachelor project aimed to find out more about how humans deal with recognizing emotions from face and voice, and if there are differences between musicians and non-musicians in emotion recognition. For this, we developed the Emotion Stroop task.

Results indicate that recognition of emotion from face is a more automated process than recognition of emotion from voice. For human-robot interaction, when making socially interactive robots, these results indicate that it might be better to invest in robots that express emotions by face instead of by voice. Results also indicate that musicians are better in focusing on the emotion of the voice when presented with a face and a voice. Further research is necessary to find out if robots should really be adjusted to the user group in terms of modality in which they express emotion.

4.1 Implications

There are a several other explanations for the effects that were found. They will be discussed in the next paragraphs.

4.1.1 Database bias

We choose to make the databases ourselves, because we could not find a professional database that was fitting exactly with what we needed [15]. On the Emotion Stroop task we wanted to flash a face and together with that present a short sound fragment of a voice. Therefore, we could not use a database with emotional speech, like the German emotional speech database [16], because the sound fragments in those databases are too long. Moreover, it had our preference to have humming sounds instead of spoken words, to make the combination of a non-moving face together with the voice as realistic as possible. Considering that humming sounds are sounds you can make with a non-moving, closed mouth. For this reason we also choose to let the actors express emotion on the pictures with their mouth closed. The inconvenience of our database is that it is not extensively tested and/or validated. The results of testing the database showed that emotions in faces and voices were not always perfectly clear, which could induce a bias on the Emotion Stroop task. The voice database could be less expressive than the face database, resulting in participants being better in recognizing emotions from faces. In the same time this is a really difficult problem. Considering that if you would make a database in which the emotions of the faces and voices are recognizable with the same ease, you would not find any difference between face and voice. There are much more aspects in the database that could influence the resulting findings (e.g. duration of stimuli or emotions that are used), but the remaining question will always be: Are people better in recognizing emotions from face or are we better in expressing emotions by face?

4.1.2 Group bias

The conditions of the experiment were the same for musicians and non-musicians. Therefore, differences could only be explained from differences between the two groups.

Our assumption is that the difference that is found between musicians and non-musicians is related to their difference in musical expertise. Enhancement of emotion recognition from voice might be a consequence of musical training, but groups were not randomized so a causal relationship cannot be made. It might be the case that people who have more sensitive hearing become musicians, and their sensitive hearing is the cause of being better in emotion recognition from voice.

Another explanation for the found effect could be related to the age difference between musicians and non-musicians that were tested in the experiment (musicians: 29.25 compared to non-musicians: 21.81).

4.1.3 Modality bias

Another explanation could be in the difference between the two modalities. Even though we tried to control for the possibility that in the bimodal trials participants would ignore one of the two modalities, this difference could still induce a bias.

4.2 Future research

This project was a first step in finding out more about emotion recognition. Besides giving some promising insights it also raised a lot of new questions.

- Will the effects, found on the Emotion Stroop task, be the same if we use a database with emotional robot faces and robot voices?
- Are the effects consistent when including more emotions (e.g. anger, fear, disgust, boredom)?
- Are the effects the same if we use a professional database?
- Is musical expertise enhancing recognition of emotions from voice, or is there another variable Z that is influencing recognition of emotion from voice and becoming a musician?

Acknowledgements

I would like to thank Makiko Sadakata for giving me the opportunity to do this project and for all her supervision. Furthermore I would like to thank Ida Sprinkhuizen-Kuyper for supervision and help to make the connection to Robotics. Also, I would like to thank all the participants that acted for the database, tested the database and did the Emotion Stroop task.

References

[1] van Breemen, A., Yan, X., & Meerbeek, B. (2005). iCat: an animated user-interface robot with personality. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems* (pp. 143-144). ACM.

[2] Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., & Blow, M. (2009). KASPAR–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, *6*(3-4), 369-397.

[3] Ekman, P., & Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial movement. *CA: Consulting Psychologists Press Palo Alto, 12*, 271-302.

[4] Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., & Narayanan, S. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 205-211).

[5] Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using hidden Markov models. *Speech communication*, *41*(4), 603-623.

[6] ten Bosch, L. (2003). Emotions, speech and the ASR framework. *Speech Communication*, *40*(1), 213-225.

[7] Pantic, M., & Rothkrantz, L. J. (2003). Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, *91*(9), 1370-1390.

[9] Chen, L. S., & Huang, T. S. (2000). Emotional expressions in audiovisual human computer interaction. In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on* (Vol. 1, pp. 423-426). IEEE.

[10] De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289-311.

[11] Sadakata, M., & Sekiyama, K. (2011). Enhanced perception of various linguistic features by musicians: a cross-linguistic study. *Acta psychologica*, *138*(1), 1-10.

[12] Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021.

[13] Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of experimental psychology, 18(6), 643.

[14] Peirce, J.W. (2007). PsychoPy - psychophysics software in Python. Journal of Neuroscience Methods 162:8-1

[15] Ververidis, D., & Kotropoulos, C. (2003). A review of emotional speech databases. In *Proc. Panhellenic Conference on Informatics (PCI)* (pp. 560-574).

[16] Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A database of German emotional speech. In *Interspeech* (pp. 1517-1520).

[17] Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, *93*, 1097.

[18] Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, *70*(3), 614.

[19] Kimball, S., Mattis, P. (2009).Gimp: GNU Image Manipulation Program (Version 2.6). Software available at *www.gimp.org*

[20] Boersma, P., & Weeninkpraat, D. (2013). Praat: Doing phonetics by computer (Version 5.3.40). Software available at *http://praat.org/*

Appendix A: Making the database

A.1 Introduction

This appendix describes the making of the database used in the Emotion Stroop-task that was designed for this thesis.

Six participants (3 male and 3 female) between 20 and 28 years old, were asked to model for the database. Pictures were taken of their face with a sad, neutral, and happy expression. Sound recordings were made of their voice with a sad, neutral, and happy expression.

The pictures and sounds were edited. After which, the best pictures and sounds were selected, and used in an experiment. The experiment was used to test how sad, neutral or happy people rated the emotions. The tests were run on six participants.

Out of the six actors, four (two male and two female) were selected, regarding to their results, for the faces database. And four (two male and two female) were selected, regarding to their results, for the voices database. The selection of faces and voices was done independently, so the voice of one actor could be recombined with the face of another actor. Although, a match in the gender of the voice and face was respected.

A.2 Choice of emotions

The emotions happy and sad were chosen, because in face and in voice they are easy to discriminate. The features in a face (e.g. shape of the mouth, size of the eyes) are very different for a happy face than for a sad face. The features in a voice (e.g. pitch, energy, formant) [17, 18] are also very different for a happy voice or a sad voice.

The neutral emotion was added for three reasons. First of all, the neutral emotion could be used as a measure for neutrality of a face or a voice. Some people might have a naturally more happy or sad looking/sounding face or voice. In the testing of the database this was already tested for. The second reason was a more intuitive one. In the Emotion Stroop task people are presented with a non-moving face together with a humming voice. This is not what we see in real-life when interacting with people, because then a face is moving when we are talking. This raised the question if on the Emotion Stroop task the combination of a static happy face together with a sad voice would be too peculiar for people. In the way that when presenting people with a static happy face combined with a humming sad voice, might be too far away from s real life situation. This could lead them to not be distracted of confused by the other modality, because what they perceive is just absurd. While, when people are presented with, for example, a neutral face combined with a sad voice, it could be more realistic. This could lead the participants to be confused by the combination of face and voice. The last point is about human-robot interaction. Often, when making socially interactive robots, robots are expressing emotion by only one modality (face, voice or gestures). When a robot is made to, for example, only express emotion by face, often the voice is kept neutral. Therefore, it would be interesting to look at how people react to such stimuli.

A.3 Recording the data

For making the database, six people were asked to act (later defined as actors). For every actor several pictures were taken of their face and several recordings of their voice was made.

Pictures were taken with a Canon HF10 camera. Light conditions were kept the same for all the photograph sessions of the actors. For every actor at least three pictures per emotion were taken.

To achieve high quality sounds, the recordings were done in a recording studio of the Radboud University Nijmegen. For recording the participants were sitting in front of a table with a microphone on it. For every participant at least six sounds per emotions were recorded, three long ones (approximately 1.4 sec) and three short ones (approximately 0.5 seconds).

A.4 Editing the data

Pictures were edited with Gimp 2.6 [19]. The pictures were cropped, to ensure that all the faces had approximately the same size. Furthermore, colors were switched to black and white and if necessary the pictures were rotated so the faces would be vertically aligned.

Sounds were edited with Praat 5.3.40 [20]. The whole recording session of one person was taped in one recording.

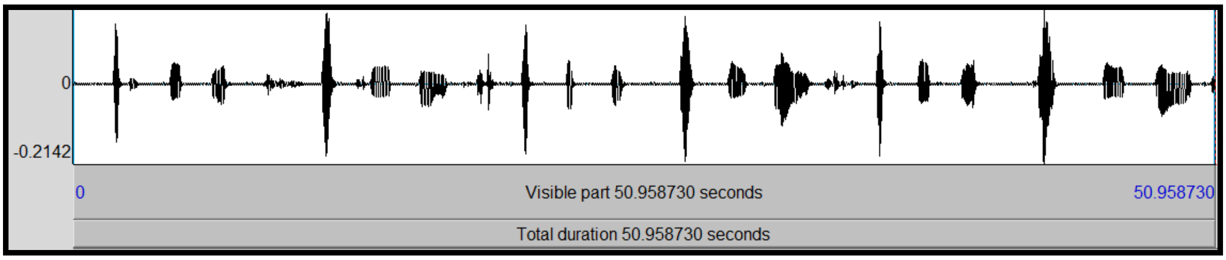


Figure A .1: Example of a recording session

First all the individual sound fragments that were good were cut out.

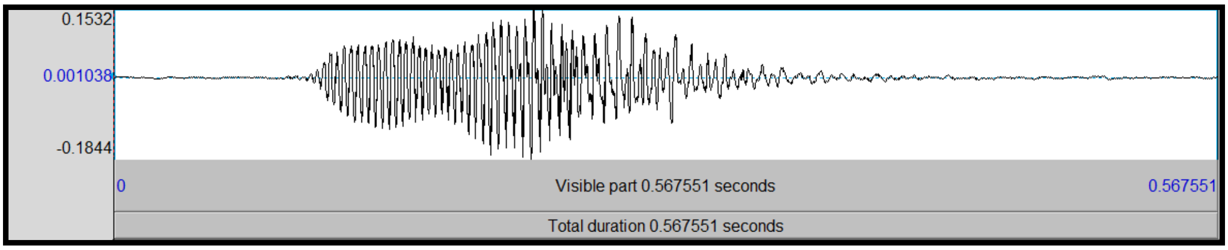


Figure A.2: One individual sound fragment cut out of the whole recording

Then, the functions 'Move start of selection to nearest zero cross section' and 'Move end of selection to nearest zero cross section' from the Praat program were used to avoid click sounds in the beginning and ending of the sound fragment.

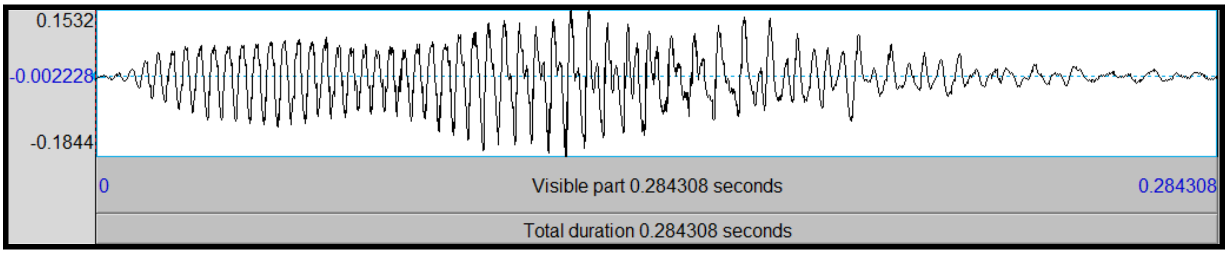


Figure A.3: Sound fragment after using the functions 'Move start of selection to nearest zero cross section' and 'Move end of selection to nearest zero cross section'. The beginning en ending of the fragment are now at the zero crossing.

After a script was ran to generate a fade-in and fade-out to the sound fragment.

```
#This script generates a fade-in and fade-out to the selected Sound
# Using the cosine-function squared in [0, 0.5*pi]
# The variable 't' (next line) determines the window of the fading
t = 0.005
ft = Get finishing time
Formula… if (x > (‘ft’ – ‘t’)) then self * (1-(cos((0.5*pi * ((‘ft’ – x)/’t’)))^2)) else self fi
```

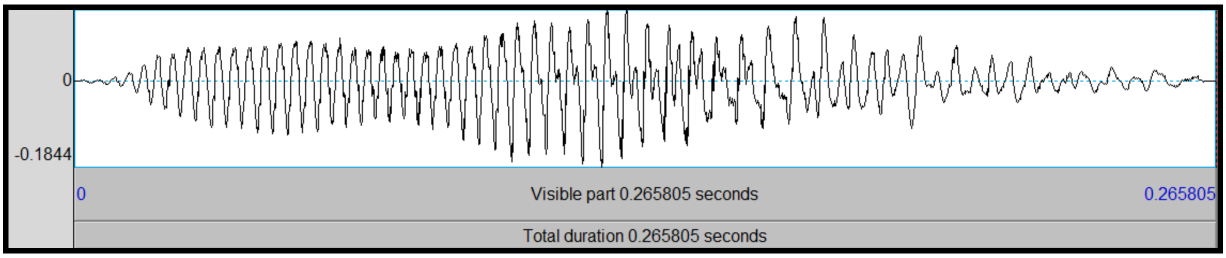


Figure A.4: Sound fragment after running the script for fade in and out.

In the end, the "To Intensity…"-function from Praat was used to convert all the sound fragments to an intensity around 70 dB.

A.5 Testing

The most expressive and qualitatively best pictures and sounds were selected. These sounds and pictures were used for testing. A computer experiment was programmed using PsychoPy. Before testing, a pilot of the experiment was done with three people. Comments on the instructions and tasks were used to improve the experiment.

Over participants the order of task was counterbalanced. The experiment consisted of two tasks; a face task and a voice task.

In the *face task,* two times 18 visual trials (18 faces) were presented in a random order. Participants had to rate the emotion expressed on a nine point scale, where 1 was sad 5 was neutral and 9 was happy (see Figure 6.5). There was no time limit for answering.

In the *voice task,* two times 27 auditory trials (27 voices) were presented in a random order. On every trial, a voice was played twice. After that, participants had to rate the emotion expressed on a nine point scale, where 1 was sad 5 was neutral and 9 was happy. There was no time limit for answering.

In total, six people participants tested the database.

sad				neutral				happy
1	2	3	4	5	6	7	8	9

Figure A.5: The nine point scale that was used in the experiment for testing the database.

A.6 Results

Results are shown in Table 6.1 and 6.3. Summaries of the results are given in Table 6.2 and 6.4.

Face	Pp 1	Pp2	Pp3	Pp4	Pp5	Pp6	Modus	Average
MaleActor1-happy1.jpg	7	9	6	7	8	7	7	7.416667
	8	9	7	7	7	7		
MaleActor2-happy1.jpg	8	9	6	9	8	8	8	7.916667
	8	9	6	8	7	9		
MaleActor3-happy1.jpg	8	9	7	8	7	8	8	7.75
	9	9	6	8	7	7		
MaleActor1-neutral1.jpg	5	5	5	5	5	5	5	5
	5	5	5	5	5	5		
MaleActor2-neutral1.jpg	5	4	3	5	5	4	5	4.5
	4	5	4	5	5	5		
MaleActor3-neutral1.jpg	5	6	4	6	5	5	5.5	5.25
	5	5	4	6	6	6		
MaleActor1-sad1.jpg	2	1	1	1	3	3	1	1.666667
	1	1	3	1	2	1		
MaleActor2-sad1.jpg	2	1	3	2	3	2	3	2.25
	1	2	3	3	3	2		
MaleActor3-sad1.jpg	3	3	4	3	3	3	3	3.333333
	3	5	4	3	4	2		
FemaleActor1-happy1.jpg	7	9	6	7	7	9	7	7.5
	7	9	6	8	7	8		
FemaleActor2-happy1.jpg	7	9	7	7	7	6	7	6.916667
	6	8	7	6	6	7		
FemaleActor3-happy1.jpg	7	9	8	8	7	7	7	7.666667
	8	9	7	7	7	8		
FemaleActor1-neutral1.jpg	3	1	3	5	5	5	5	3.833333
	2	4	4	5	4	5		
FemaleActor2-neutral1.jpg	5	5	6	6	5	5	5	5.083333
	5	6	4	5	4	5		
FemaleActor3-neutral1.jpg	3	5	2	4	4	4	4	3.833333
	3	4	4	4	5	4		
FemaleActor1-sad1.jpg	1	1	1	2	2	1	1	1.333333
	1	1	1	1	3	1		
FemaleActor2-sad1.jpg	5	4	2	2	5	2	2	3.416667
	3	4	4	2	5	3		
FemaleActor3-sad1.jpg	3	2	3	3	4	2	3	3.166667
	3	4	3	4	3	4		

Table A.1: Results of database testing. Six participants (pp1 t/m pp2) tested the database. They rated all the faces twice. The scale was a nine point scale, where 1 was sad 5 was neutral and 9 was happy.

Face	Error modus	Error average
MaleActor1	2	1.75
MaleActor2	3	2.83
MaleActor3	3.5	3.83
FemaleActor1	2	3
FemaleActor2	3	4.58
FemaleActor3	4	4.67

Table A.2: Summary of the results of testing the database. The "Error modus" is calculated by the distance between the perfect scores (9 for happy, 5 for neutral and 1 for sad) and the real modus. The "Error average" is calculated by the distance between the perfect scores and the real average.

Voice	Pp1	Pp2	Pp3	Pp4	Pp5	Pp6	Modus	Average
MaleActor1-happy1.wav	6	9	7	8	6	7	7	7.1666667
	8	8	6	7	6	8		
MaleActor1-happy2.wav	7	7	6	6	7	7	7	6.75
	8	7	6	7	5	8		
MaleActor2-happy1.wav	8	8	7	8	7	8	8	7.75
	8	8	7	8	7	9		
MaleActor2-happy2.wav	7	7	4	6	6	7	7	6.5
	5	7	6	7	8	8		
MaleActor3-happy1.wav	8	8	5	7	6	5	7	6.8333333
	9	7	6	7	7	7		
MaleActor1-neutral1.wav	5	5	3	4	5	5	5	4.3333333
	5	4	5	3	5	3		
MaleActor1-neutral2.wav	5	4	4	5	4	4	4	4.3333333
	5	4	4	4	5	4		
MaleActor2-neutral1.wav	5	5	5	4	5	5	5	4.9166667
	5	5	5	5	5	5		
MaleActor2-neutral2.wav	5	5	4	6	6	4	5	5
	5	5	5	5	5	5		
MaleActor3-neutral1.wav	5	4	4	6	5	3	5	4.5833333
	5	5	4	4	5	5		
MaleActor1-sad1.wav	1	1	5	3	6	2	3	3
	2	3	4	3	4	2		
MaleActor2-sad1.wav	1	3	4	3	3	2	3	2.25
	2	1	2	3	2	1		
MaleActor2-sad2.wav	1	1	2	3	4	1	1	2.25
	3	1	2	3	5	1		
MaleActor3-sad1.wav	5	3	4	6	4	2	4	4.3333333
	2	5	4	7	4	6		
FemaleActor1-happy1.wav	9	8	6	9	7	8	8	7.5
	8	8	6	7	6	8		
FemaleActor2-happy1.wav	6	9	7	7	8	9	7	7.4166667
	7	8	7	7	6	8		
FemaleActor3-happy1.wav	8	9	9	8	8	7	9	8.0833333
	9	9	6	7	8	9		
FemaleActor1-neutral1.wav	5	5	4	5	7	5	5	4.8333333
	5	6	5	5	5	1		
FemaleActor2-neutral1.wav	5	5	4	7	8	5	5	5.4166667
	5	5	5	6	6	4		
FemaleActor2-neutral2.wav	5	5	5	6	6	6	5	5.25
	5	6	4	5	5	5		
FemaleActor3-neutral1.wav	5	8	4	5	4	5	5	5

	5	6	4	6	3	5		
FemaleActor1-sad1.wav	2	1	3	2	7	1	1	2.5
	3	1	2	3	4	1		
FemaleActor1-sad2.wav	1	1	3	2	5	1	1	2.1666667
	3	1	3	2	3	1		
FemaleActor2-sad1.wav	2	3	3	4	4	1	2	2.5833333
	2	2	4	2	3	1		
FemaleActor2-sad2.wav	2	1	3	2	2	1	1	1.8333333
	1	1	3	2	3	1		
FemaleActor3-sad1.wav	3	4	3	4	4	6	4	4
	5	5	4	3	4	3		

Table A.3: Results of database testing. Six participants (pp1 t/m pp2) tested the database. They rated all the voices twice. For sound fragments with the same emotion and from the same actor we selected the best one for comparison with the other actors (only black ones were compared in the end).

Voice	Error modus	Error average
MaleActor1	3.5	4.50
MaleActor2	1	2.50
MaleActor3	5	5.92
FemaleActor1	1	2.83
FemaleActor2	2	2.67
FemaleActor3	3	3.92

Table A.4: Summary of the results of testing the database. The "Error modus" is calculated by the distance between the perfect scores (9 for happy, 5 for neutral and 1 for sad) and the real modus. The "Error average" is calculated by the distance between the perfect scores and the real average.

A.7 Conclusion

Selection of best faces and voices was done using the summary tables (Table 6.2 and 6.4). An average was not always giving the best reflection of the data. That is why the modus is also calculated.

For example, in Table 6.1 for the scores given to FemaleActor1-neutral1.jpg, the average and modus are very different. The picture was mostly rated with a 5, which is the perfect score. On the other hand the average is much further from the perfect score. In this case this is maybe related to a participant that on accident clicked the wrong answer (pp2 rated one time with 1 and the other time with 4), but it can also be related to which features participants are mostly looking at. When looking mostly at the eyes, this could give a different image of the expressed emotion than when looking at the mouth. Besides average and modus also the raw data was taken into account.

Finally, four faces were chosen (two male and two female): the faces of FemaleActor1, FemaleActor2, MaleActor1 and MaleActor2 (see Figure 6.6 and 6.7). And four voices were chosen (two male and two female): the best voices (black instead of grey in Table 6.3) of FemaleActor1, FemaleActor2, MaleActor1 and MaleActor2.

Figure A.6: Faces of the men that were selected for testing. Upper two rows were selected for the database

Figure A.7: Faces of the women that were selected for testing. Upper two rows were selected for the database

Appendix B: Questionnaire

Radboud University Nijmegen

Questionnaire

Name: ….………………………………………………………………………………………………….

Date of birth: ………………………………………………………………………………………..

Gender: ……………………………………………………………………………………………...

Nationality: ………………………………………………………………………………………….

Study: ………………………………………………………………………………………………….

Date: ………………………………………………………………………………………………...

1) *How many years of musical training (with a teacher) do you have?*

   ………………………………………………………………………………………………………

2) *Which instrument do you play? Which style of music?*

   ………………………………………………………………………………………………………

3) *How old were you when you had musical training?*

   ………………………………………………………………………………………………………

4) *How many years did you practice music? How many hours per week on average?*

   ……………………………………………………………………………………………………

5) *How many hours per week did you practice in the last two years on average?*

   ………………………………………………………………………………………………………

6) *How many hours per week do you listen to music on average? Which style of music?*

   ………………………………………………………………………………………………………

7) *Question or remarks?*

   ………………………………………………………………………………………………………

Appendix C: SPSS output cost-analysis

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,686	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Wilks' Lambda	,314	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Hotelling's Trace	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Roy's Largest Root	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
	Musical expertise	Pillai's Trace	,249	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Wilks' Lambda	,751	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Hotelling's Trace	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Roy's Largest Root	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
Within Subjects	Modality	Pillai's Trace	,489	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Wilks' Lambda	,511	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Hotelling's Trace	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Roy's Largest Root	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
	Modality * Musical expertise	Pillai's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Wilks' Lambda	1,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Hotelling's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Roy's Largest Root	,000	,007[a]	2,000	29,000	,993	,000	,014	,051

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,686	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Wilks' Lambda	,314	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Hotelling's Trace	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Roy's Largest Root	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
	Musical expertise	Pillai's Trace	,249	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Wilks' Lambda	,751	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Hotelling's Trace	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Roy's Largest Root	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
Within Subjects	Modality	Pillai's Trace	,489	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Wilks' Lambda	,511	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Hotelling's Trace	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Roy's Largest Root	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
	Modality * Musical expertise	Pillai's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Wilks' Lambda	1,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Hotelling's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Roy's Largest Root	,000	,007[a]	2,000	29,000	,993	,000	,014	,051

a. Exact statistic
b. Computed using alpha = ,05
c. Design: Intercept + Musical expertise
Within Subjects Design: Modality

Table 2: Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Modality	Reaction time	Sphericity Assumed	,028	1	,028	1,428	,241	,045	1,428	,212
		Greenhouse-Geisser	,028	1,000	,028	1,428	,241	,045	1,428	,212
		Huynh-Feldt	,028	1,000	,028	1,428	,241	,045	1,428	,212
		Lower-bound	,028	1,000	,028	1,428	,241	,045	1,428	,212
	Correct response rate	Sphericity Assumed	1491,642	1	1491,642	21,017	,000	,412	21,017	,993
		Greenhouse-Geisser	1491,642	1,000	1491,642	21,017	,000	,412	21,017	,993
		Huynh-Feldt	1491,642	1,000	1491,642	21,017	,000	,412	21,017	,993
		Lower-bound	1491,642	1,000	1491,642	21,017	,000	,412	21,017	,993
Modality * Musical expertise	Reaction time	Sphericity Assumed	,000	1	,000	,011	,918	,000	,011	,051
		Greenhouse-Geisser	,000	1,000	,000	,011	,918	,000	,011	,051
		Huynh-Feldt	,000	1,000	,000	,011	,918	,000	,011	,051
		Lower-bound	,000	1,000	,000	,011	,918	,000	,011	,051
	Correct response rate	Sphericity Assumed	,037	1	,037	,001	,982	,000	,001	,050
		Greenhouse-Geisser	,037	1,000	,037	,001	,982	,000	,001	,050
		Huynh-Feldt	,037	1,000	,037	,001	,982	,000	,001	,050
		Lower-bound	,037	1,000	,037	,001	,982	,000	,001	,050
Error(Modality)	Reaction time	Sphericity Assumed	,593	30	,020					
		Greenhouse-Geisser	,593	30,000	,020					
		Huynh-Feldt	,593	30,000	,020					
		Lower-bound	,593	30,000	,020					
	Correct response rate	Sphericity Assumed	2129,202	30	70,973					
		Greenhouse-Geisser	2129,202	30,000	70,973					
		Huynh-Feldt	2129,202	30,000	70,973					
		Lower-bound	2129,202	30,000	70,973					

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,686	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Wilks' Lambda	,314	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Hotelling's Trace	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
		Roy's Largest Root	2,188	31,724[a]	2,000	29,000	,000	,686	63,448	1,000
	Musical expertise	Pillai's Trace	,249	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Wilks' Lambda	,751	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Hotelling's Trace	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
		Roy's Largest Root	,331	4,806[a]	2,000	29,000	,016	,249	9,612	,753
Within Subjects	Modality	Pillai's Trace	,489	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Wilks' Lambda	,511	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Hotelling's Trace	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
		Roy's Largest Root	,958	13,897[a]	2,000	29,000	,000	,489	27,794	,996
	Modality * Musical expertise	Pillai's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Wilks' Lambda	1,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Hotelling's Trace	,000	,007[a]	2,000	29,000	,993	,000	,014	,051
		Roy's Largest Root	,000	,007[a]	2,000	29,000	,993	,000	,014	,051

a. Computed using alpha = ,05

Table 3: Tests of Between-Subjects Effects
Transformed Variable:Average

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	Reaction time	1,164	1	1,164	39,859	,000	,571	39,859	1,000
	Correct response rate	2409,885	1	2409,885	32,017	,000	,516	32,017	1,000
Musical expertise	Reaction time	,022	1	,022	,766	,388	,025	,766	,135
	Correct response rate	646,956	1	646,956	8,595	,006	,223	8,595	,810
Error	Reaction time	,876	30	,029					
	Correct response rate	2258,049	30	75,268					

a. Computed using alpha = ,05

Appendix D: SPSS output 2x2x2 mixed MANOVA

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,995	3000,012[a]	2,000	29,000	,000	,995	6000,023	1,000
		Wilks' Lambda	,005	3000,012[a]	2,000	29,000	,000	,995	6000,023	1,000
		Hotelling's Trace	206,897	3000,012[a]	2,000	29,000	,000	,995	6000,023	1,000
		Roy's Largest Root	206,897	3000,012[a]	2,000	29,000	,000	,995	6000,023	1,000
	Musical expertise	Pillai's Trace	,196	3,542[a]	2,000	29,000	,042	,196	7,083	,612
		Wilks' Lambda	,804	3,542[a]	2,000	29,000	,042	,196	7,083	,612
		Hotelling's Trace	,244	3,542[a]	2,000	29,000	,042	,196	7,083	,612
		Roy's Largest Root	,244	3,542[a]	2,000	29,000	,042	,196	7,083	,612
Within Subjects	Modality	Pillai's Trace	,553	17,971[a]	2,000	29,000	,000	,553	35,942	1,000
		Wilks' Lambda	,447	17,971[a]	2,000	29,000	,000	,553	35,942	1,000
		Hotelling's Trace	1,239	17,971[a]	2,000	29,000	,000	,553	35,942	1,000
		Roy's Largest Root	1,239	17,971[a]	2,000	29,000	,000	,553	35,942	1,000
	Modality * Musical expertise	Pillai's Trace	,114	1,857[a]	2,000	29,000	,174	,114	3,715	,355
		Wilks' Lambda	,886	1,857[a]	2,000	29,000	,174	,114	3,715	,355
		Hotelling's Trace	,128	1,857[a]	2,000	29,000	,174	,114	3,715	,355
		Roy's Largest Root	,128	1,857[a]	2,000	29,000	,174	,114	3,715	,355
	Congruency	Pillai's Trace	,649	26,803[a]	2,000	29,000	,000	,649	53,606	1,000
		Wilks' Lambda	,351	26,803[a]	2,000	29,000	,000	,649	53,606	1,000
		Hotelling's Trace	1,848	26,803[a]	2,000	29,000	,000	,649	53,606	1,000

	Roy's Largest Root	1,848	26,803[a]	2,000	29,000	,000	,649	53,606	1,000
Congruency * Musical expertise	Pillai's Trace	,059	,914[a]	2,000	29,000	,412	,059	1,827	,192
	Wilks' Lambda	,941	,914[a]	2,000	29,000	,412	,059	1,827	,192
	Hotelling's Trace	,063	,914[a]	2,000	29,000	,412	,059	1,827	,192
	Roy's Largest Root	,063	,914[a]	2,000	29,000	,412	,059	1,827	,192
Modality * Congruency	Pillai's Trace	,367	8,400[a]	2,000	29,000	,001	,367	16,800	,945
	Wilks' Lambda	,633	8,400[a]	2,000	29,000	,001	,367	16,800	,945
	Hotelling's Trace	,579	8,400[a]	2,000	29,000	,001	,367	16,800	,945
	Roy's Largest Root	,579	8,400[a]	2,000	29,000	,001	,367	16,800	,945
Modality * Congruency * Musical expertise	Pillai's Trace	,040	,607[a]	2,000	29,000	,552	,040	1,215	,141
	Wilks' Lambda	,960	,607[a]	2,000	29,000	,552	,040	1,215	,141
	Hotelling's Trace	,042	,607[a]	2,000	29,000	,552	,040	1,215	,141
	Roy's Largest Root	,042	,607[a]	2,000	29,000	,552	,040	1,215	,141

a. Exact statistic
b. Computed using alpha = ,05
c. Design: Intercept + Musical expertise
Within Subjects Design: Modality + Congruency + Modality * Congruency

Table 2: Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Modality	Reaction time	Sphericity Assumed	1,373	1	1,373	35,781	,000	,544	35,781	1,000
		Greenhouse-Geisser	1,373	1,000	1,373	35,781	,000	,544	35,781	1,000
		Huynh-Feldt	1,373	1,000	1,373	35,781	,000	,544	35,781	1,000
		Lower-bound	1,373	1,000	1,373	35,781	,000	,544	35,781	1,000
	Correct response	Sphericity Assumed	1047,948	1	1047,948	14,090	,001	,320	14,090	,953

	rate	Greenhouse-Geisser	1047,948	1,000	1047,948	14,090	,001	,320	14,090	,953
		Huynh-Feldt	1047,948	1,000	1047,948	14,090	,001	,320	14,090	,953
		Lower-bound	1047,948	1,000	1047,948	14,090	,001	,320	14,090	,953
Modality * Musical expertise	Reaction time	Sphericity Assumed	,039	1	,039	1,028	,319	,033	1,028	,166
		Greenhouse-Geisser	,039	1,000	,039	1,028	,319	,033	1,028	,166
		Huynh-Feldt	,039	1,000	,039	1,028	,319	,033	1,028	,166
		Lower-bound	,039	1,000	,039	1,028	,319	,033	1,028	,166
	Correct response rate	Sphericity Assumed	284,166	1	284,166	3,821	,060	,113	3,821	,473
		Greenhouse-Geisser	284,166	1,000	284,166	3,821	,060	,113	3,821	,473
		Huynh-Feldt	284,166	1,000	284,166	3,821	,060	,113	3,821	,473
		Lower-bound	284,166	1,000	284,166	3,821	,060	,113	3,821	,473
Error (Modality)	Reaction time	Sphericity Assumed	1,151	30	,038					
		Greenhouse-Geisser	1,151	30,000	,038					
		Huynh-Feldt	1,151	30,000	,038					
		Lower-bound	1,151	30,000	,038					
	Correct response rate	Sphericity Assumed	2231,204	30	74,373					
		Greenhouse-Geisser	2231,204	30,000	74,373					
		Huynh-Feldt	2231,204	30,000	74,373					
		Lower-bound	2231,204	30,000	74,373					
Congruency	Reaction time	Sphericity Assumed	,075	1	,075	32,186	,000	,518	32,186	1,000
		Greenhouse-Geisser	,075	1,000	,075	32,186	,000	,518	32,186	1,000
		Huynh-Feldt	,075	1,000	,075	32,186	,000	,518	32,186	1,000
		Lower-bound	,075	1,000	,075	32,186	,000	,518	32,186	1,000
	Correct response rate	Sphericity Assumed	1356,050	1	1356,050	34,588	,000	,536	34,588	1,000
		Greenhouse-Geisser	1356,050	1,000	1356,050	34,588	,000	,536	34,588	1,000
		Huynh-Feldt	1356,050	1,000	1356,050	34,588	,000	,536	34,588	1,000

		Lower-bound	1356,050	1,000	1356,050	34,588	,000	,536	34,588	1,000
Congruency * Musical expertise	Reaction time	Sphericity Assumed	9,314E-6	1	9,314E-6	,004	,950	,000	,004	,050
		Greenhouse-Geisser	9,314E-6	1,000	9,314E-6	,004	,950	,000	,004	,050
		Huynh-Feldt	9,314E-6	1,000	9,314E-6	,004	,950	,000	,004	,050
		Lower-bound	9,314E-6	1,000	9,314E-6	,004	,950	,000	,004	,050
	Correct response rate	Sphericity Assumed	72,226	1	72,226	1,842	,185	,058	1,842	,260
		Greenhouse-Geisser	72,226	1,000	72,226	1,842	,185	,058	1,842	,260
		Huynh-Feldt	72,226	1,000	72,226	1,842	,185	,058	1,842	,260
		Lower-bound	72,226	1,000	72,226	1,842	,185	,058	1,842	,260
Error (Congruency)	Reaction time	Sphericity Assumed	,070	30	,002					
		Greenhouse-Geisser	,070	30,000	,002					
		Huynh-Feldt	,070	30,000	,002					
		Lower-bound	,070	30,000	,002					
	Correct response rate	Sphericity Assumed	1176,176	30	39,206					
		Greenhouse-Geisser	1176,176	30,000	39,206					
		Huynh-Feldt	1176,176	30,000	39,206					
		Lower-bound	1176,176	30,000	39,206					
Modality * Congruency	Reaction time	Sphericity Assumed	,016	1	,016	12,314	,001	,291	12,314	,924
		Greenhouse-Geisser	,016	1,000	,016	12,314	,001	,291	12,314	,924
		Huynh-Feldt	,016	1,000	,016	12,314	,001	,291	12,314	,924
		Lower-bound	,016	1,000	,016	12,314	,001	,291	12,314	,924
	Correct response rate	Sphericity Assumed	435,038	1	435,038	13,482	,001	,310	13,482	,944
		Greenhouse-Geisser	435,038	1,000	435,038	13,482	,001	,310	13,482	,944
		Huynh-Feldt	435,038	1,000	435,038	13,482	,001	,310	13,482	,944
		Lower-bound	435,038	1,000	435,038	13,482	,001	,310	13,482	,944
Modality * Congruency	Reaction time	Sphericity Assumed	,001	1	,001	,436	,514	,014	,436	,098

* Musical expertise		Greenhouse-Geisser	,001	1,000	,001	,436	,514	,014	,436	,098
		Huynh-Feldt	,001	1,000	,001	,436	,514	,014	,436	,098
		Lower-bound	,001	1,000	,001	,436	,514	,014	,436	,098
	Correct response rate	Sphericity Assumed	7,152	1	7,152	,222	,641	,007	,222	,074
		Greenhouse-Geisser	7,152	1,000	7,152	,222	,641	,007	,222	,074
		Huynh-Feldt	7,152	1,000	7,152	,222	,641	,007	,222	,074
		Lower-bound	7,152	1,000	7,152	,222	,641	,007	,222	,074
Error (Modality * Congruency)	Reaction time	Sphericity Assumed	,038	30	,001					
		Greenhouse-Geisser	,038	30,000	,001					
		Huynh-Feldt	,038	30,000	,001					
		Lower-bound	,038	30,000	,001					
	Correct response rate	Sphericity Assumed	968,040	30	32,268					
		Greenhouse-Geisser	968,040	30,000	32,268					
		Huynh-Feldt	968,040	30,000	32,268					
		Lower-bound	968,040	30,000	32,268					

a. Computed using alpha = ,05

Table 3: Tests of Between-Subjects Effects

Transformed Variable:Average

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	Reaction time	132,816	1	132,816	1203,310	,000	,976	1203,310	1,000
	Correct response rate	1004591,377	1	1004591,377	5847,077	,000	,995	5847,077	1,000
Musical expertise	Reaction time	,115	1	,115	1,042	,316	,034	1,042	,167
	Correct response rate	1221,291	1	1221,291	7,108	,012	,192	7,108	,732
Error	Reaction time	3,311	30	,110					
	Correct response rate	5154,326	30	171,811					

a. Computed using alpha = ,05

Table 4: Modality * Congruency | Estimates |

Measure	Modality	Congruency	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound

Reaction time	Face	Congruent	,902	,031	,838	,966
		Incongruent	,928	,033	,860	,996
	Voice	Congruent	1,087	,034	1,017	1,157
		Incongruent	1,158	,039	1,078	1,237
Correct response rate	Face	Congruent	92,864	1,171	90,472	95,255
		Incongruent	90,041	1,135	87,723	92,359
	Voice	Congruent	90,828	1,365	88,040	93,616
		Incongruent	80,631	2,325	75,884	85,379

Table 5: Modality * Congruency | Pairwise Comparisons |

Measure	Modality	(I) Congruency	(J) Congruency	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Reaction time	Face	Congruent	Incongruent	-,026*	,006	,000	-,039	-,014
		Incongruent	Congruent	,026*	,006	,000	,014	,039
	Voice	Congruent	Incongruent	-,071*	,014	,000	-,099	-,043
		Incongruent	Congruent	,071*	,014	,000	,043	,099
Correct response rate	Face	Congruent	Incongruent	2,823*	,985	,008	,812	4,833
		Incongruent	Congruent	-2,823*	,985	,008	-4,833	-,812
	Voice	Congruent	Incongruent	10,197*	1,870	,000	6,377	14,016
		Incongruent	Congruent	-10,197*	1,870	,000	-14,016	-6,377

Based on estimated marginal means
*. The mean difference is significant at the ,05 level.
a. Adjustment for multiple comparisons: Bonferroni.

Table 6: Modality * Congruency | Multivariate Tests |

Modality		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Face	Pillai's trace	,443	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Wilks' lambda	,557	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Hotelling's trace	,794	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Roy's largest root	,794	11,515[a]	2,000	29,000	,000	,443	23,029	,988
Voice	Pillai's trace	,580	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Wilks' lambda	,420	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Hotelling's trace	1,382	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Roy's largest root	1,382	20,035[a]	2,000	29,000	,000	,580	40,070	1,000

Each F tests the multivariate simple effects of Congruency within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

Table 6: Modality * Congruency | Multivariate Tests |

Modality		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Face	Pillai's trace	,443	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Wilks' lambda	,557	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Hotelling's trace	,794	11,515[a]	2,000	29,000	,000	,443	23,029	,988
	Roy's largest root	,794	11,515[a]	2,000	29,000	,000	,443	23,029	,988
Voice	Pillai's trace	,580	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Wilks' lambda	,420	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Hotelling's trace	1,382	20,035[a]	2,000	29,000	,000	,580	40,070	1,000
	Roy's largest root	1,382	20,035[a]	2,000	29,000	,000	,580	40,070	1,000

Each F tests the multivariate simple effects of Congruency within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.
a. Exact statistic
b. Computed using alpha = ,05

Table 7: Congruency * Modality | Estimates |

Measure	Congruency	Modality	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Reaction time	Congruent	Face	,902	,031	,838	,966
		Voice	1,087	,034	1,017	1,157
	Incongruent	Face	,928	,033	,860	,996
		Voice	1,158	,039	1,078	1,237
Correct response rate	Congruent	Face	92,864	1,171	90,472	95,255
		Voice	90,828	1,365	88,040	93,616
	Incongruent	Face	90,041	1,135	87,723	92,359
		Voice	80,631	2,325	75,884	85,379

Table 8: Congruency * Modality | Pairwise Comparisons |

Measure	Congruency	(I) Modality	(J) Modality	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
Reaction time	Congruent	Face	Voice	-,185*	,033	,000	-,252	-,117
		Voice	Face	,185*	,033	,000	,117	,252
	Incongruent	Face	Voice	-,229*	,037	,000	-,305	-,153
		Voice	Face	,229*	,037	,000	,153	,305
Correct response rate	Congruent	Face	Voice	2,035	1,292	,126	-,603	4,674
		Voice	Face	-2,035	1,292	,126	-4,674	,603
	Incongruent	Face	Voice	9,410*	2,235	,000	4,845	13,975
		Voice	Face	-9,410*	2,235	,000	-13,975	-4,845

Based on estimated marginal means

Table 7: Congruency * Modality | Estimates |

Measure	Congruency	Modality	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Reaction time	Congruent	Face	,902	,031	,838	,966
		Voice	1,087	,034	1,017	1,157
	Incongruent	Face	,928	,033	,860	,996
		Voice	1,158	,039	1,078	1,237
Correct response rate	Congruent	Face	92,864	1,171	90,472	95,255
		Voice	90,828	1,365	88,040	93,616
	Incongruent	Face	90,041	1,135	87,723	92,359

*. The mean difference is significant at the ,05 level.
a. Adjustment for multiple comparisons: Bonferroni.

Table 9: Congruency * Modality | Multivariate Tests |

Congruency		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Congruent	Pillai's trace	,511	15,175[a]	2,000	29,000	,000	,511	30,349	,998
	Wilks' lambda	,489	15,175[a]	2,000	29,000	,000	,511	30,349	,998
	Hotelling's trace	1,047	15,175[a]	2,000	29,000	,000	,511	30,349	,998
	Roy's largest root	1,047	15,175[a]	2,000	29,000	,000	,511	30,349	,998
Incongruent	Pillai's trace	,566	18,919[a]	2,000	29,000	,000	,566	37,839	1,000
	Wilks' lambda	,434	18,919[a]	2,000	29,000	,000	,566	37,839	1,000
	Hotelling's trace	1,305	18,919[a]	2,000	29,000	,000	,566	37,839	1,000
	Roy's largest root	1,305	18,919[a]	2,000	29,000	,000	,566	37,839	1,000

Each F tests the multivariate simple effects of Modality within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.
a. Exact statistic
b. Computed using alpha = ,05

Table 10: Modality * Congruency * Musical expertise | Estimates |

Measure	Modality	Congruency	MusicalExpertise	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Reaction time	Face	Congruent	Non-Musician	,852	,044	,762	,942
			Musician	,952	,044	,862	1,042
		Incongruent	Non-Musician	,883	,047	,787	,979
			Musician	,973	,047	,877	1,070
	Voice	Congruent	Non-Musician	1,076	,048	,977	1,175
			Musician	1,097	,048	,999	1,196
		Incongruent	Non-Musician	1,143	,055	1,031	1,255
			Musician	1,172	,055	1,060	1,284
Correct response rate	Face	Congruent	Non-Musician	91,780	1,656	88,398	95,161
			Musician	93,948	1,656	90,566	97,330
		Incongruent	Non-Musician	87,927	1,605	84,649	91,206
			Musician	92,155	1,605	88,876	95,433
	Voice	Congruent	Non-Musician	87,237	1,931	83,294	91,179
			Musician	94,419	1,931	90,477	98,362
		Incongruent	Non-Musician	75,065	3,287	68,351	81,779
			Musician	86,198	3,287	79,484	92,912

Table 11: Modality * Congruency * Musical expertise | Pairwise comparisons |

Measure	Modality	Congruency	(I) MusicalExpertise	(J) MusicalExpertise	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
Reaction time	Face	Congruent	Non-Musician	Musician	-,100	,062	,121	-,227	,028
			Musician	Non-Musician	,100	,062	,121	-,028	,227
		Incongruent	Non-Musician	Musician	-,090	,067	,185	-,226	,046
			Musician	Non-Musician	,090	,067	,185	-,046	,226
	Voice	Congruent	Non-Musician	Musician	-,021	,068	,759	-,161	,119
			Musician	Non-Musician	,021	,068	,759	-,119	,161
		Incongruent	Non-Musician	Musician	-,028	,078	,716	-,187	,130
			Musician	Non-Musician	,028	,078	,716	-,130	,187
Correct response rate	Face	Congruent	Non-Musician	Musician	-2,168	2,342	,362	-6,951	2,614
			Musician	Non-Musician	2,168	2,342	,362	-2,614	6,951
		Incongruent	Non-Musician	Musician	-4,227	2,270	,072	-8,864	,409
			Musician	Non-Musician	4,227	2,270	,072	-,409	8,864
	Voice	Congruent	Non-Musician	Musician	-7,183*	2,730	,013	-12,758	-1,607
			Musician	Non-Musician	7,183*	2,730	,013	1,607	12,758
		Incongruent	Non-Musician	Musician	-11,133*	4,649	,023	-20,628	-1,638
			Musician	Non-Musician	11,133*	4,649	,023	1,638	20,628

Based on estimated marginal means
a. Adjustment for multiple comparisons: Bonferroni.
*. The mean difference is significant at the ,05 level.

Table 12: Modality * Congruency * Musical expertise | Univariate Tests |

Measure	Modality	Congruency		Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Reaction time	Face	Congruent	Contrast	,080	1	,080	2,550	,121	,078	2,550	,340
			Error	,937	30	,031					
		Incongruent	Contrast	,065	1	,065	1,841	,185	,058	1,841	,260
			Error	1,064	30	,035					
	Voice	Congruent	Contrast	,004	1	,004	,096	,759	,003	,096	,060
			Error	1,126	30	,038					
		Incongruent	Contrast	,006	1	,006	,135	,716	,004	,135	,065
			Error	1,444	30	,048					
Correct response rate	Face	Congruent	Contrast	37,610	1	37,610	,857	,362	,028	,857	,146
			Error	1316,208	30	43,874					
		Incongruent	Contrast	142,971	1	142,971	3,467	,072	,104	3,467	,437
			Error	1237,081	30	41,236					
	Voice	Congruent	Contrast	412,726	1	412,726	6,921	,013	,187	6,921	,721
			Error	1788,907	30	59,630					
		Incongruent	Contrast	991,528	1	991,528	5,734	,023	,160	5,734	,640
			Error	5187,551	30	172,918					

Each F tests the simple effects of Musical expertise within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.
a. Computed using alpha = ,05

Appendix E: SPSS output emotions analysis

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,995	3086,901[a]	2,000	29,000	,000	,995	6173,802	1,000
		Wilks' Lambda	,005	3086,901[a]	2,000	29,000	,000	,995	6173,802	1,000
		Hotelling's Trace	212,890	3086,901[a]	2,000	29,000	,000	,995	6173,802	1,000
		Roy's Largest Root	212,890	3086,901[a]	2,000	29,000	,000	,995	6173,802	1,000
	Musical expertise	Pillai's Trace	,016	,241[a]	2,000	29,000	,788	,016	,482	,084
		Wilks' Lambda	,984	,241[a]	2,000	29,000	,788	,016	,482	,084
		Hotelling's Trace	,017	,241[a]	2,000	29,000	,788	,016	,482	,084
		Roy's Largest Root	,017	,241[a]	2,000	29,000	,788	,016	,482	,084
Within Subjects	Modality	Pillai's Trace	,839	75,708[a]	2,000	29,000	,000	,839	151,417	1,000
		Wilks' Lambda	,161	75,708[a]	2,000	29,000	,000	,839	151,417	1,000
		Hotelling's Trace	5,221	75,708[a]	2,000	29,000	,000	,839	151,417	1,000
		Roy's Largest Root	5,221	75,708[a]	2,000	29,000	,000	,839	151,417	1,000
	Modality * Musical expertise	Pillai's Trace	,173	3,044[a]	2,000	29,000	,063	,173	6,087	,544
		Wilks' Lambda	,827	3,044[a]	2,000	29,000	,063	,173	6,087	,544
		Hotelling's Trace	,210	3,044[a]	2,000	29,000	,063	,173	6,087	,544
		Roy's Largest Root	,210	3,044[a]	2,000	29,000	,063	,173	6,087	,544
	Emotion	Pillai's Trace	,366	3,902[a]	4,000	27,000	,013	,366	15,607	,840
		Wilks' Lambda	,634	3,902[a]	4,000	27,000	,013	,366	15,607	,840
		Hotelling's Trace	,578	3,902[a]	4,000	27,000	,013	,366	15,607	,840

		Roy's Largest Root	,578	3,902[a]	4,000	27,000	,013	,366	15,607	,840
	Emotion * Musical expertise	Pillai's Trace	,211	1,801[a]	4,000	27,000	,158	,211	7,202	,475
		Wilks' Lambda	,789	1,801[a]	4,000	27,000	,158	,211	7,202	,475
		Hotelling's Trace	,267	1,801[a]	4,000	27,000	,158	,211	7,202	,475
		Roy's Largest Root	,267	1,801[a]	4,000	27,000	,158	,211	7,202	,475
	Modality * Emotion	Pillai's Trace	,142	1,114[a]	4,000	27,000	,370	,142	4,455	,302
		Wilks' Lambda	,858	1,114[a]	4,000	27,000	,370	,142	4,455	,302
		Hotelling's Trace	,165	1,114[a]	4,000	27,000	,370	,142	4,455	,302
		Roy's Largest Root	,165	1,114[a]	4,000	27,000	,370	,142	4,455	,302
	Modality * Emotion * Musical expertise	Pillai's Trace	,044	,312[a]	4,000	27,000	,867	,044	1,248	,109
		Wilks' Lambda	,956	,312[a]	4,000	27,000	,867	,044	1,248	,109
		Hotelling's Trace	,046	,312[a]	4,000	27,000	,867	,044	1,248	,109
		Roy's Largest Root	,046	,312[a]	4,000	27,000	,867	,044	1,248	,109

a. Exact statistic
b. Computed using alpha = ,05
c. Design: Intercept + Musical expertise
Within Subjects Design: Modality + Emotion + Modality * Emotion

Table 2: Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Modality	Reaction time	Sphericity Assumed	3,548	1	3,548	144,230	,000	,828	144,230	1,000
		Greenhouse-Geisser	3,548	1,000	3,548	144,230	,000	,828	144,230	1,000
		Huynh-Feldt	3,548	1,000	3,548	144,230	,000	,828	144,230	1,000
		Lower-bound	3,548	1,000	3,548	144,230	,000	,828	144,230	1,000
	Correct response	Sphericity Assumed	1,624	1	1,624	,016	,901	,001	,016	,052

	rate	Greenhouse-Geisser	1,624	1,000	1,624	,016	,901	,001	,016	,052
		Huynh-Feldt	1,624	1,000	1,624	,016	,901	,001	,016	,052
		Lower-bound	1,624	1,000	1,624	,016	,901	,001	,016	,052
Modality * Musical expertise	Reaction time	Sphericity Assumed	,056	1	,056	2,292	,141	,071	2,292	,311
		Greenhouse-Geisser	,056	1,000	,056	2,292	,141	,071	2,292	,311
		Huynh-Feldt	,056	1,000	,056	2,292	,141	,071	2,292	,311
		Lower-bound	,056	1,000	,056	2,292	,141	,071	2,292	,311
	Correct response rate	Sphericity Assumed	558,897	1	558,897	5,464	,026	,154	5,464	,619
		Greenhouse-Geisser	558,897	1,000	558,897	5,464	,026	,154	5,464	,619
		Huynh-Feldt	558,897	1,000	558,897	5,464	,026	,154	5,464	,619
		Lower-bound	558,897	1,000	558,897	5,464	,026	,154	5,464	,619
Error(Modality)	Reaction time	Sphericity Assumed	,738	30	,025					
		Greenhouse-Geisser	,738	30,000	,025					
		Huynh-Feldt	,738	30,000	,025					
		Lower-bound	,738	30,000	,025					
	Correct response rate	Sphericity Assumed	3068,863	30	102,295					
		Greenhouse-Geisser	3068,863	30,000	102,295					
		Huynh-Feldt	3068,863	30,000	102,295					
		Lower-bound	3068,863	30,000	102,295					
Emotion	Reaction time	Sphericity Assumed	,126	2	,063	8,107	,001	,213	16,214	,950
		Greenhouse-Geisser	,126	1,941	,065	8,107	,001	,213	15,732	,945
		Huynh-Feldt	,126	2,000	,063	8,107	,001	,213	16,214	,950
		Lower-bound	,126	1,000	,126	8,107	,008	,213	8,107	,787
	Correct response rate	Sphericity Assumed	837,241	2	418,620	3,299	,044	,099	6,598	,605
		Greenhouse-Geisser	837,241	1,991	420,413	3,299	,044	,099	6,570	,603
		Huynh-Feldt	837,241	2,000	418,620	3,299	,044	,099	6,598	,605

		Lower-bound	837,241	1,000	837,241	3,299	,079	,099	3,299	,420
Emotion * Musical expertise	Reaction time	Sphericity Assumed	,044	2	,022	2,801	,069	,085	5,601	,531
		Greenhouse-Geisser	,044	1,941	,022	2,801	,071	,085	5,435	,522
		Huynh-Feldt	,044	2,000	,022	2,801	,069	,085	5,601	,531
		Lower-bound	,044	1,000	,044	2,801	,105	,085	2,801	,367
	Correct response rate	Sphericity Assumed	190,389	2	95,195	,750	,477	,024	1,500	,172
		Greenhouse-Geisser	190,389	1,991	95,602	,750	,476	,024	1,494	,171
		Huynh-Feldt	190,389	2,000	95,195	,750	,477	,024	1,500	,172
		Lower-bound	190,389	1,000	190,389	,750	,393	,024	,750	,134
Error(Emotion)	Reaction time	Sphericity Assumed	,466	60	,008					
		Greenhouse-Geisser	,466	58,216	,008					
		Huynh-Feldt	,466	60,000	,008					
		Lower-bound	,466	30,000	,016					
	Correct response rate	Sphericity Assumed	7613,413	60	126,890					
		Greenhouse-Geisser	7613,413	59,744	127,434					
		Huynh-Feldt	7613,413	60,000	126,890					
		Lower-bound	7613,413	30,000	253,780					
Modality * Emotion	Reaction time	Sphericity Assumed	,012	2	,006	,709	,496	,023	1,418	,164
		Greenhouse-Geisser	,012	1,965	,006	,709	,494	,023	1,394	,163
		Huynh-Feldt	,012	2,000	,006	,709	,496	,023	1,418	,164
		Lower-bound	,012	1,000	,012	,709	,406	,023	,709	,129
	Correct response rate	Sphericity Assumed	363,831	2	181,915	1,442	,244	,046	2,885	,297
		Greenhouse-Geisser	363,831	1,826	199,279	1,442	,245	,046	2,633	,283
		Huynh-Feldt	363,831	2,000	181,915	1,442	,244	,046	2,885	,297
		Lower-bound	363,831	1,000	363,831	1,442	,239	,046	1,442	,214
Modality * Emotion * Musical expertise	Reaction time	Sphericity Assumed	,001	2	,000	,044	,957	,001	,089	,056

		Greenhouse-Geisser	,001	1,965	,000	,044	,955	,001	,087	,056
		Huynh-Feldt	,001	2,000	,000	,044	,957	,001	,089	,056
		Lower-bound	,001	1,000	,001	,044	,834	,001	,044	,055
	Correct response rate	Sphericity Assumed	136,220	2	68,110	,540	,586	,018	1,080	,135
		Greenhouse-Geisser	136,220	1,826	74,611	,540	,570	,018	,986	,131
		Huynh-Feldt	136,220	2,000	68,110	,540	,586	,018	1,080	,135
		Lower-bound	136,220	1,000	136,220	,540	,468	,018	,540	,110
Error(Modality*Emotion)	Reaction time	Sphericity Assumed	,521	60	,009					
		Greenhouse-Geisser	,521	58,954	,009					
		Huynh-Feldt	,521	60,000	,009					
		Lower-bound	,521	30,000	,017					
	Correct response rate	Sphericity Assumed	7567,250	60	126,121					
		Greenhouse-Geisser	7567,250	54,772	138,159					
		Huynh-Feldt	7567,250	60,000	126,121					
		Lower-bound	7567,250	30,000	252,242					

a. Computed using alpha = ,05

Table 3: Tests of Between-Subjects Effects

Transformed Variable:Average

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	Reaction time	158,685	1	158,685	1801,577	,000	,984	1801,577	1,000
	Correct response rate	1605304,327	1	1605304,327	5354,009	,000	,994	5354,009	1,000
Musical expertise	Reaction time	,024	1	,024	,268	,608	,009	,268	,079
	Correct response rate	90,674	1	90,674	,302	,586	,010	,302	,083
Error	Reaction time	2,642	30	,088					
	Correct response rate	8994,966	30	299,832					

a. Computed using alpha = ,05

Table 4: Estimates

Measure	Emotion	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Reaction time	Happy	,889	,024	,840	,937
	Neutral	,894	,018	,857	,931
	Sad	,945	,027	,890	1,001
Correct response rate	Happy	90,365	1,554	87,190	93,539
	Neutral	94,358	1,577	91,137	97,578
	Sad	89,593	1,936	85,639	93,546

Table 5: Pairwise Comparisons

Measure	(I) Emotion	(J) Emotion	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
Reaction time	Happy	Neutral	-,005	,014	1,000	-,041	,031
		Sad	-,057*	,017	,005	-,099	-,015
	Neutral	Happy	,005	,014	1,000	-,031	,041
		Sad	-,052*	,016	,009	-,092	-,011
	Sad	Happy	,057*	,017	,005	,015	,099
		Neutral	,052*	,016	,009	,011	,092
Correct response rate	Happy	Neutral	-3,993	1,944	,146	-8,923	,937
		Sad	,772	2,053	1,000	-4,434	5,978
	Neutral	Happy	3,993	1,944	,146	-,937	8,923
		Sad	4,765	1,975	,067	-,244	9,774
	Sad	Happy	-,772	2,053	1,000	-5,978	4,434
		Neutral	-4,765	1,975	,067	-9,774	,244

Based on estimated marginal means
a. Adjustment for multiple comparisons: Bonferroni.
*. The mean difference is significant at the ,05 level.

Appendix F: SPSS output congruency and emotions

Table 1: Multivariate Tests[c]

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Between Subjects	Intercept	Pillai's Trace	,994	2527,172[a]	2,000	29,000	,000	,994	5054,343	1,000
		Wilks' Lambda	,006	2527,172[a]	2,000	29,000	,000	,994	5054,343	1,000
		Hotelling's Trace	174,288	2527,172[a]	2,000	29,000	,000	,994	5054,343	1,000
		Roy's Largest Root	174,288	2527,172[a]	2,000	29,000	,000	,994	5054,343	1,000
	Musical expertise	Pillai's Trace	,201	3,646[a]	2,000	29,000	,039	,201	7,292	,626
		Wilks' Lambda	,799	3,646[a]	2,000	29,000	,039	,201	7,292	,626
		Hotelling's Trace	,251	3,646[a]	2,000	29,000	,039	,201	7,292	,626
		Roy's Largest Root	,251	3,646[a]	2,000	29,000	,039	,201	7,292	,626
Within Subjects	Congruency	Pillai's Trace	,729	18,124[a]	4,000	27,000	,000	,729	72,496	1,000
		Wilks' Lambda	,271	18,124[a]	4,000	27,000	,000	,729	72,496	1,000
		Hotelling's Trace	2,685	18,124[a]	4,000	27,000	,000	,729	72,496	1,000
		Roy's Largest Root	2,685	18,124[a]	4,000	27,000	,000	,729	72,496	1,000
	Congruency * Musical expertise	Pillai's Trace	,129	1,001[a]	4,000	27,000	,424	,129	4,002	,273
		Wilks' Lambda	,871	1,001[a]	4,000	27,000	,424	,129	4,002	,273
		Hotelling's Trace	,148	1,001[a]	4,000	27,000	,424	,129	4,002	,273
		Roy's Largest Root	,148	1,001[a]	4,000	27,000	,424	,129	4,002	,273

a. Exact statistic
b. Computed using alpha = ,05
c. Design: Intercept + Musical expertise
Within Subjects Design: Congruency

Table 2: Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Congruency	Reaction time	Sphericity Assumed	,078	2	,039	33,968	,000	,531	67,935	1,000
		Greenhouse-Geisser	,078	1,665	,047	33,968	,000	,531	56,556	1,000
		Huynh-Feldt	,078	1,810	,043	33,968	,000	,531	61,474	1,000
		Lower-bound	,078	1,000	,078	33,968	,000	,531	33,968	1,000
	Correct response rate	Sphericity Assumed	1514,108	2	757,054	34,770	,000	,537	69,541	1,000
		Greenhouse-Geisser	1514,108	1,835	824,920	34,770	,000	,537	63,820	1,000
		Huynh-Feldt	1514,108	2,000	757,054	34,770	,000	,537	69,541	1,000
		Lower-bound	1514,108	1,000	1514,108	34,770	,000	,537	34,770	1,000
Congruency * Musical expertise	Reaction time	Sphericity Assumed	,001	2	,001	,581	,563	,019	1,162	,142
		Greenhouse-Geisser	,001	1,665	,001	,581	,533	,019	,967	,133
		Huynh-Feldt	,001	1,810	,001	,581	,546	,019	1,051	,137
		Lower-bound	,001	1,000	,001	,581	,452	,019	,581	,114
	Correct response rate	Sphericity Assumed	57,446	2	28,723	1,319	,275	,042	2,638	,274
		Greenhouse-Geisser	57,446	1,835	31,298	1,319	,274	,042	2,421	,263
		Huynh-Feldt	57,446	2,000	28,723	1,319	,275	,042	2,638	,274
		Lower-bound	57,446	1,000	57,446	1,319	,260	,042	1,319	,199
Error(Congruency)	Reaction time	Sphericity Assumed	,068	60	,001					
		Greenhouse-Geisser	,068	49,950	,001					
		Huynh-Feldt	,068	54,293	,001					
		Lower-bound	,068	30,000	,002					
	Correct response rate	Sphericity Assumed	1306,379	60	21,773					
		Greenhouse-Geisser	1306,379	55,064	23,725					

	Huynh-Feldt	1306,379	60,000	21,773					
	Lower-bound	1306,379	30,000	43,546					

a. Computed using alpha = ,05

Table 3: Tests of Between-Subjects Effects
Transformed Variable:Average

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	Reaction time	101,523	1	101,523	1230,446	,000	,976	1230,446	1,000
	Correct response rate	727079,443	1	727079,443	4812,084	,000	,994	4812,084	1,000
Musical expertise	Reaction time	,087	1	,087	1,055	,313	,034	1,055	,169
	Correct response rate	1109,738	1	1109,738	7,345	,011	,197	7,345	,746
Error	Reaction time	2,475	30	,083					
	Correct response rate	4532,835	30	151,095					

a. Computed using alpha = ,05

Table 4: Estimates

Measure	Congruency	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Reaction time	Congruent	,993	,028	,936	1,050
	Half(in)Congruent	1,029	,031	,967	1,092
	Incongruent	1,063	,031	1,000	1,125
Correct response rate	Congruent	91,848	1,094	89,615	94,082
	Half(in)Congruent	87,112	1,425	84,201	90,023
	Incongruent	82,122	1,690	78,671	85,572

Table 5: Pairwise Comparisons

Measure	(I) Congruency	(J) Congruency	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Reaction time	Congruent	Half(in)Congruent	-,036*	,008	,000	-,056	-,016
		Incongruent	-,070*	,010	,000	-,095	-,044
	Half(in)Congruent	Congruent	,036*	,008	,000	,016	,056
		Incongruent	-,033*	,007	,000	-,051	-,016
	Incongruent	Congruent	,070*	,010	,000	,044	,095
		Half(in)Congruent	,033*	,007	,000	,016	,051
Correct response rate	Congruent	Half(in)Congruent	4,737*	1,119	,001	1,898	7,575
		Incongruent	9,727*	1,325	,000	6,368	13,086
	Half(in)Congruent	Congruent	-4,737*	1,119	,001	-7,575	-1,898
		Incongruent	4,990*	1,036	,000	2,362	7,618
	Incongruent	Congruent	-9,727*	1,325	,000	-13,086	-6,368

	Half(in)Congruent	-4,990*	1,036	,000	-7,618	-2,362

Based on estimated marginal means
*. The mean difference is significant at the ,05 level.
a. Adjustment for multiple comparisons: Bonferroni.