Mood Effects on the Decoding of Emotional Voices

Alda Troncone[1,*], Davide Palumbo[1], and Anna Esposito[1,2]

[1] Department of Psychology, Second University of Naples, Caserta, Italy
alda.troncone@unina2.it,
da.palumbo@libero.it
[2] International Institute for Advanced Scientific Studies (IIASS), Vietri sul Mare, Italy
iiass.annaesp@tin.it

Abstract. This study examines the effect of mood induction on the decoding of emotional vocal expressions. An adequate sample of 145 students (71 females, 74 males; mean age = 23.37 ± 2.05) was recruited at the Second University of Naples (Italy). Subjects were randomly assigned to one of three (sad, fear or neutral) emotion conditions induced by viewing short movies. The results showed a significant general decrease in the decoding accuracy in the mood induction conditions when compared to the accuracy of the participants who did not received such mood induction. Post hoc analyses revealed that recognition of emotional vocal voices conveying anger was especially impaired by mood induction conditions. No findings consistent with mood congruity theory were observed. This study contributes to emotion regulation research by showing differences in emotion decoding tasks by voices due to mood induction procedures, as already observed in studies exploiting the decoding of emotional faces.

Keywords: mood induction, emotional vocal expressions, emotional voices decoding.

1 Introduction

Mood is an ever-present and influential feature of our mental lives. There is much evidence in literature supporting the notion that mood affects cognitive processes, such as memory [3], executive functions (e.g. working memory) [11] and the decoding of emotional facial emotion expressions [10].

With respect to how mood affects emotion recognition, some empirical studies explain the effect of mood on the ability to decode emotions through mood-congruity theories [2], [12-13]. The mood-congruity theories state that a person's mood exerts a congruity effect on social judgments. In other words, subjects in a specific mood would be faster and more accurate in decoding expressions of emotions congruent with their mood, and slower and less accurate in decoding incongruent emotions

* To contact the authors, please write to: Alda Troncone, Department of Psychology, Second University of Naples, Viale Ellittico 31, 81100, Caserta, Italy.

S. Bassis et al. (eds.), *Recent Advances of Neural Network Models and Applications*,
Smart Innovation, Systems and Technologies 26,
DOI: 10.1007/978-3-319-04129-2_32, © Springer International Publishing Switzerland 2014

(e.g. subjects with a sad mood are more accurate in identifying expressions of sadness than subjects with a different mood). The validity of mood congruity theories is controversial and not supported by all experimental data [5], [14].

These aforementioned studies focused on the ability of subjects to recognize certain emotions through tasks decoding facial expressions. To the date, direct effects of induced emotional states on the ability to decode emotional vocal voices have not been investigated.

The goal of this study is to contribute to emotion regulation research by investigating possible similarities and/or differences in the task of emotion decoding, due to mood induction procedures, as has already been done for the decoding of emotional faces. In particular, it aims at understanding how sadness, fear, neutral moods affect, in a healthy subject sample, the ability to decode 5 out of the 6 basic emotions according to Ekman & Friesen's definition [6]. In addition it further inspects whether the current induced mood (in our case sadness or fear) boosts or hinders the recognition of mood-congruent and mood incongruent voices compared to the neutral induced mood.

The research questions are:

1. Does mood induction affect subjects' ability to decode emotional vocal expressions?
2. Are subjects with an induced mood of sadness (or fear) more accurate in recognizing congruent emotional voices?

In order to answer the aforementioned questions the decoding accuracy obtained by subjects participating to the mood induction procedure was compared with the performance of no mood induced subjects involved in the decoding of the same emotional voices in a previous study of ours [7-8].

Along with mood, the socioeconomic status (SES) of subjects was considered. In spite of the well-known effects of SES on cognitive performance [4], [15] there is a lack of data about the influence of SES on the decoding of emotional expressions.

2 Method

2.1 Subjects

The study was carried out on an adequate sample of 154 university students recruited at the Second University of Naples (Italy), 9 of which were excluded from the study because their socio-demographical characteristics did not fit those of the majority of subjects (e. g. ages and marital status).

The final sample consisted of 145 subjects (74 males, 71 females) with mean age of 23.37 years (SD= 2.05). The subjects are mainly undergraduate (81.4%), full time students (77.1 %), from the Faculties of Psychology (66.2%), Political Science (8.3%) and others (25.5%).

The control group was made of 30 adults, students at the Faculty of Science, in Salerno, aged between 23 to 29 years.

2.2 Materials

Mood Induction Stimuli

The mood induction procedure consisted in letting the participants watch 3 short movies believed to induce a fear, sadness and a neutral mood. The movies were selected among the highest rated as being effective in inducing mood from a previously validated database of films [8].

The fear induction movie lasts 25 seconds and starts by showing frames of a security camera monitoring an entrance. Suddenly a screaming witch-like face appears right in front of the camera. The film used for inducing sadness, lasts 79 seconds, features a series of photographs of babies and young children, sometimes interacting with their mothers, overlaid by written text about the denial of life connected to abortion and has a soft, placid but slightly sad soundtrack.

The neutral induction movie mood lasts 30 seconds, and shows photographs of household furniture.

Emotional Voice Decoding Task

In order to assess their ability to decode emotional feelings, the subjects were asked to participate at an audio test where they listened to 20 vocal emotional stimuli, selected by one of the authors from a database of emotional voices already assessed and published in literature [7-8]. The "emotional voices" are Italian sentences of short duration, with a semantic content not related to the expressed emotion. They were extracted from video-clips acted by famous Italian actors/actresses expressing vocal emotional expressions of joy, sadness, fear, surprise, and anger. Generally, the stimuli used in these experiments are collected asking an actor/actress to produce a given sentence with different emotional contents. In this respect, the actor/actress is portraying the requested emotional state from scratch, in a laboratory setting, without an external and internal context of reference that is normally present when making a movie. The absence of a frame of reference may make such stimuli less ecological than those extracted from movies since the latter are more close to real social stimuli both because the audio track is affected by the environmental noise of the scene from which it is extracted and because the actor/actress was not asked to portray an emotionally coloured sentence from scratch, rather his/her coloured vocal expression is produced in the context of the movie scene. A detailed description of the stimuli exploited in this work is reported in Esposito [9]. Participants were requested to listen and label the sentences attributing to them one of the five emotional labels reported above.

The Socioeconomic Status (SES)

The assessment of the participants' socioeconomic status (SES) was made using the Barratt Simplified Measure of Social Status (BSMSS) [1]. This simplified measure derives from an update of the pioneering work of Hollingshead [1] in devising a simple measure of Social Status based on marital status, current employment status (former status for retirees), level of education and occupational prestige. This index attributes parental SES, and a combination of their educational level and work

activity, to the students. The obtained SES's values varies from 8 to 66 giving rise to three different near-equal groups of students belonging to a low (8-27 points), medium (28-47 points), high (48-66 points) SES category.

2.3 Procedure

Each participant first filled in and signed a consent form providing also her/his general and socio-economic information. Generally within a week after the consent participants made an appointment with the experimenter to undergo the mood induction and the subsequent emotional voice decoding task.

A suitable neutral setting was created in the laboratory for these two tasks (free of distractions and disturbing events), each subject, after being informed of the ongoing experiment, was randomly assigned to one of the three mood induction conditions (fear, sad and neutral) and was asked to watch and listen through headphones the corresponding movie on a 13-inch computer screen.

Immediately after the mood induction procedure, the subject underwent the emotional voice decoding task, where the 20 emotional vocal expressions were randomly presented through headphones. For each stimulus she/he had to select the emotional label she/he wanted to attribute to it by crossing the right box on an answer grid reporting the five emotional labels. The subjects were allowed to listen to the stimuli as many times as they needed before selecting their answer.

The above reported experimental set-up can rise two methodological concerns. First, no assessment of the subject's emotional intelligence was made before the mood induction procedure. This concern is circumvented by the high number of the involved subjects (145) ensuring that outliers line up at the boundaries of the answer distribution. In addition, the effectiveness of the induction procedure was not assessed in order to avoid interrupting and/or interfering with the induced emotional state, thereby accelerating its physiological dissolution.

The subjects of the control group underwent to the same decoding task of the subjects involved in the mood induction experiment except that their choice was made among 8 other than 5 emotional labels (in addition to joy, fear, anger, surprise, sadness, they can also choose among the labels irony, "other", and "no emotion"). However, the frequency of answers such as "other" and "no emotion" was very low we were able to consider their response for the comparison with the mood induced groups.

3 Results

The average SES value was 34.82 (SD=12.92), which can be categorized as a mid SES level. In order to check if it affected the emotion decoding task, a one-way ANOVA was conducted on each of the five emotional labels considered, using the number of correct answers to emotional voices as the dependent variable. No SES effects were found in the decoding of stimuli conveying joy ($F(69,140)=1.005$, p=n.s.), fear ($F(69,140)=.938$, p=n.s.), anger ($F(69,140)=.901$, p=n.s.), surprise ($F(69,140)=1.388$, p=n.s.) and sadness ($F(69,140) = .888$, p=n.s.)

To explore the effects of the mood induction procedure a chi-square test (α=.001) was conducted comparing the percentage of correct answers to the emotional voice decoding task obtained from subjects involved in the experiments without (control) and with mood induction (sad, fear, neutral) procedures. The results showed a significant decrease in the emotion decoding accuracy for the latter group. This decrease affected subjects in induced moods of sadness ($\chi 2$ =58.82, df=4, p<.001), fear ($\chi 2$ =51.87, df=4, p<.001) and in the induced neutral mood ($\chi 2$ = 66.49, df=4, p<.001) when compared with the control group's performance. The confusion matrices showing the percentages of the decoding accuracy are reported in Tables 1, 2 and 3 respectively. In these Tables, the grey and white rows report the percentages of correct answers obtained by the mood induced and control group respectively in each of the three different conditions, and for each emotional labels. Figure 1 provides an overall view on how the percentages of correct answers differ in the control and in the three different mood induction conditions. A subsequent Bonferroni post hoc testing (adjusted α = 0.00025) indicated that the performance decrease in the three mood induction conditions was principally concerned with the decoding of voices expressing anger.

Table 1. The confusion matrix reporting the percentages of accuracy (on the diagonal) in the emotional voice decoding task obtained by subjects participating (grey rows) or not (white rows) to the **sadness** mood induced condition

emotion to identify	% answers					
	Joy	fear	anger	surprise	sadness	2 Crit= 21.5
joy	**58**					0.77
	51.7					
fear		**37.2**				11.35
		64.2				
anger			**33.5**			**28.44**
			81.7			
surprise				**42**		12.01
				71.2		
sadness					**48.4**	6.25
					69.2	

Table 2. The confusion matrix reporting the percentages of accuracy (on the diagonal) in the emotional voice decoding task obtainedby subjects participating (grey rows) or not (white rows) to the **fear** mood induced condition

emotion to identify	% answers					
	Joy	fear	anger	surprise	sadness	2 Crit= 21.5
joy	**54.9**					0.2
	51.7					
fear		**41.6**				7.95
		64.2				
anger			**31.9**			**30.35**
			81.7			
surprise				**44.6**		9.94
				71.2		
sadness					**54**	3.34
					69.2	

Table 3. The confusion matrix reporting the percentages of accuracy (on the diagonal) in the emotional voice decoding task obtained by subjects participating (grey row) or not (white rows) to the **neutral** mood induced condition

emotion to identify	% answers					
	Joy	fear	anger	surprise	sadness	2 Crit= 21.52
joy	**54.8**					0.18
	51.7					
fear		**36.2**				12.21
		64.2				
anger			**28.8**			**34.25**
			81.7			
surprise				**39.4**		14.24
				71.2		
sadness					**49.5**	5.61
					69.2	

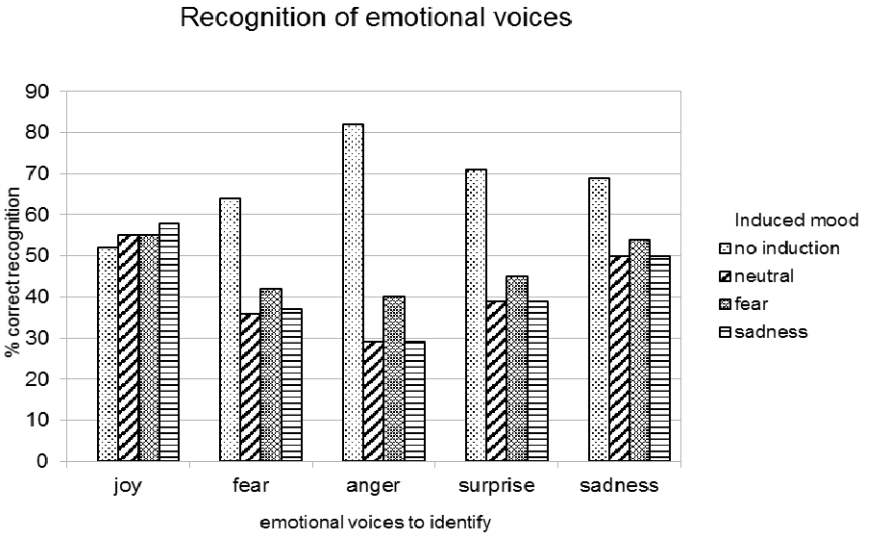

Fig. 1. The percentage of emotion's correct labeling by the control group (grey bars) and by the subjects participating to the neutral (black bars), fear (green bars) and sadness (blue bars) mood induced conditions

In addition, the obtained findings suggest that the current mood (sadness or fear) neither boosts or hinders the decoding of mood-congruent or mood incongruent voices compared to neutral mood. The fear mood induced subjects not to decode fear stimuli better than neutral mood induced ones ($\chi 2$=.805, df=1, p=n.s); the sadness mood induced subjects not to decode sad stimuli better than the neutral mood induced ($\chi 2$= 0.024, df=1, p=n.s) and the subject's ability to decode incongruent stimuli (joy) does not worsen ($\chi 2$=0.186, df=1, p=n.s).

4 Discussion

It is worth to note that there was a difference between the control and mood induced group experimental design. The subjects in the first group were forced to choose among 8 emotional labels (joy, fear, anger, surprise, sadness, irony, "other", and "no emotion") whereas those in the second one were forced to choose among only 5 labels (joy, fear, anger, surprise, sadness). This difference may have also affected the subjects' performance. Nevertheless, they can be considered as a trend in the distribution of the correct answers and interestingly suggest that any typology of mood induction procedure (even the neutral states) affects the subjects' judgment and their correct decoding performance. Speculating on this difference, the decrease in performance can be attributed to a cognitive load caused by the movie processing task regardless of the changes in the emotional feelings. We suppose that subjects in the no mood induced group are more accurate in decoding emotional voices than subjects in the mood induced groups because their cognitive processes are completely engaged by only one task.

Even if we are aware that the comparison of the present study with the available literature is limited by the specificity of the methodology applied (recognition of emotional voices vs emotional faces), our findings are in line with Chepenik et al. [5] who found, in a sample of healthy adults, a general worsening of ability in recognize emotions in subjects with induced mood of sadness.

In addition, the absence of mood congruity influences does not support mood congruity theories but may also have been affected by the experimental design giving the subjects the chance to listen to the emotional voices as many times as they needed before selecting their answer. This procedure may have favoured the dissolution of the induced mood and eliminated congruence effects.

The reported results are considered as pilot requiring support from more experimental data and the specific impairment in the recognition of voices expressing anger need to be supported by further investigations. An analysis of the errors can also help clarify the nature of this impairment.

References

1. Barratt, W.: The Barratt Simplified Measure of Social Status (BSMSS) measuring SES. Indiana State University (2006), http://wbarratt.indstate.edu/socialclass/Barratt_Simplifed_Measure_of_Social_Status.pdf (unpublished manuscript) (retrieved)
2. Bouhuys, A.L., Bloem, M.G., Groothuis, T.G.G.: Induction of Depressed and Elated Mood by Music Influences the Perception of Facial Emotional Expressions in Healthy Subjects. Journal of Affective Disorders 33, 215–226 (1995)
3. Bower, G.H., Forgas, J.: Handbook of Affect and Social Cognition. Erlbaum, Mahwah (2001)
4. Caro, D.H.: Socio-economic Status and Academic Achievement Trajectories from Childhood to Adolescence. Canadian Journal of Education 32(3), 558–590 (2009)

5. Chepenik, L.G., Cornew, L.A., Farah, M.J.: The Influence of Sad Mood on Cognition Emotion 7(4), 802–811 (2007)
6. Ekman, P., Friesen, W.V., Hager, J.C.: The Facial Action Coding System, 2nd edn. Salt Lake City, Weidenfeld & Nicolson (2002)
7. Esposito, A., Riviello, M.T., Di Maio, G.: The COST 2102 Italian Audio and Video Emotional Database. In: Apolloni, B., Bassis, S., Morabito, C.F. (eds.) Frontiers in Artificial Intelligence and Applications, vol. 204, pp. 51–61 (2009), http://www.booksonline.iospress.nl/Content/View.aspx?piid=14188
8. Esposito, A., Riviello, M.T.: The New Italian Audio and Video Emotional Database. In: Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.) Second COST 2102. LNCS, vol. 5967, pp. 406–422. Springer, Heidelberg (2010)
9. Esposito, A.: The Perceptual and Cognitive Role of Visual and Auditory Channels in Conveying Emotional Information. Cogn. Comput. J. 1(2), 268–278 (2009)
10. Langenecker, S.A., Bieliauskas, L.A., Rapport, L.J., Zubieta, J.K., Wilde, E.A., Berent, S.: Face emotion perception and executive functioning deficits in depression. Journal of Clinical and Experimental Neuropsychology 27, 320–333 (2005)
11. Marvel, C.L., Paradiso, S.: Cognitive and Neurological Impairment in Mood Disorders. Psychiatric Clinics of North America 27, 19–36 (2004)
12. Niedenthal, P.M., Halberstadt, J.B., Margolin, J., Innes-Ker, A.H.: Emotional State and The Detection of Change in Facial Expression of Emotion. European Journal of Social Psychology 30, 211–222 (2000)
13. Niedenthal, P.M., Brauer, M., Halberstadt, J.B., Innes-Ker, A.H.: When Did Her Smile Drop? Facial Mimicry and the Influences of Emotional State on the Detection of Change in Emotional Expression. Cognition and Emotion 15(6), 853–864 (2001)
14. Schmid, P.C., Schmid Mast, M.: Mood Effects on Emotion Recognition. Motive Mot. 34, 288–292 (2010)
15. Sirin, S.R.: Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. Review of Educational Research 75(3), 417–453 (2005)