

Development of an Appropriate Procedure for Inducing Emotions in Koreans and a Multimodal Emotional Response DB

Ji-Eun Park¹, Hyeong seock Oh¹, Mi-Sook Park¹ and Jin-Hun Sohn^{1*}

¹*Department of Psychology/Brain Research Institute, Chungnam National University, Daejeon, Korea*
¹*jieun711@cnu.ac.kr, ^{1*}jhsohn@cnu.ac.kr*

Abstract

This study suggests an effective and valid method of inducing emotions in Korean college students and teenagers. While 605 subjects were experiencing different types of emotions, multimodal emotional responses (physiological signals and facial, vocal, thermal and EEG signals) were acquired. Emotional response data were categorized by self-reports and partially by raters. This is the first Korean multimodal emotional database incomparable to databases in terms of its scale. For this reason, the developed database can be used for emotion research in the future.

Keywords: *Multimodal DB, Induced Emotional response*

1. Introduction

Human emotions have become a topic of great interest in many fields. How to read and imitate other emotions are important issues in emotional design, emotional marketing, and emotional robot engineering.

As such, DBs of rich emotional responses can be utilized in various emotional studies. Emotional DBs of the past have mostly involved posed expressions or imitations of expressions in photos featuring combinations of specific muscles based on the Facial Action Coding System [3]. Since posed or acted data are imitations without experience of actual emotions, they tend to be more exaggerated than actual emotional responses [7] and have different qualitative properties [1].

This study collected responses based on actual emotions. While the most natural data will be those obtained from real-life events, such moments are not only difficult to capture, but also raise ethical and technical issues. As an alternative, responses were collected after inducing emotions through various scenarios.

Unlike many previous studies that presented results in the form of still photos, this study collected video recordings, which will be more useful in future work [6] as they reveal greater details on the emotions experienced by subjects.

The second feature of this study is that collected various multimodal response data. The various possible physical reactions that can occur when you experience the emotions were collected. A facial expression alone may be adequate, but multimodal data will allow sounds and other signals to compensate for insufficient information. For the same emotional stimuli, this study simultaneously measured various modalities such as facial expression, voice, facial temperature change, brain wave, and autonomous nervous system (ANS) response including breathing and heart rate.

The third feature is that multimodal responses were synchronized to the same stimuli, so the responses facilitate to compare individual responses at specific time periods. Since emotions are assessed subjectively after they are induced, the DB contains information on the type of induced emotion and dimensional values [5] such as arousal, valence, novelty,

*Corresponding Author

approach, and potency. It helps to find a certain properties of the stimuli leading to certain physiological reactions.

Finally, this is the first database, a collection of natural emotional reaction of Korean teenagers and college students. Most prior studies of emotion were based on Western DBs due to the lack of data on Koreans' emotional responses. For emotional studies of Koreans, Korean-specific data is essential as emotions are dependent on culture and facial expressions are known to differ by race [2].

Other information covered by the database includes individual personality and the extent of depression/anxiety.

This study developed a database that improves upon existing emotional response DB's. The results are expected to serve as a valuable reference for future studies on the emotional response of Koreans.

2. Procedure

2.1 Participants

Data were collected through four rounds from different groups of participants. For the first round, 128 middle and high school students plus 172 college students were investigated. And the participants of the second, third, and fourth round were 60, 185, 57 college students respectively.

All the participants were recruited by an advertisement and they got promising 30,000 Korean won for compensation.

2.2. Procedure

The stimuli in this experiment were selected through a pilot study. Subjects were middle and high schoolers in Daejeon, and Chungnam National University students. They were asked of which drama, movies, or TV shows they usually watch, and to describe their emotional episodes of each type of emotion.

The emotion-inducing experiment was conducted four times with varying modalities and target emotions. Different stimuli were used accordingly.

The first round of experiment induced eight emotions such as anger, boredom, fear, interest, surprise, neutral, joy, and sadness. ANS response, thermal image and facial expression were recorded. In the second round, vocal expression was added to modalities, and the stimuli were changed to college students only. Anger-inducing stimuli were presented to evoke feelings of empathy among college students, who were designated as target subjects in this experiment. For the subjects to develop empathy with the characters shown in video clips, different stimuli videos were presented by gender. To attain more diversified emotion responses, disgust and cognitive surprise -inducing tasks were added. In the third round, scenes of assault were excluded for anger-inducing stimuli to prevent emotions other than anger from being mixed in the results. To allow better vocal expression of sadness, the subjects were asked to send a video message to their parents after watching the video clip. The number of active tasks were increased in the fourth round to generate more lively vocal expression engaged in natural situations. The passive procedure collects emotional responses while subjects watch video clips in front of a monitor. Under the active procedure, subjects are required to actively engage in certain actions like simple game (Figure 1). The experiment was carried out in a controlled environment of external sound and light. To prevent the order effect of emotions, the ten emotion-inducing tasks were counter balanced. After each task, the subjects self-reported on the type of emotion experienced, intensity, and emotional dimension. Each evaluation lasted for three minutes, and an additional two-minute break was given before proceeding to the next task.



Figure 1. Emotion Inducing Procedure: Passive (Left), Active (Right)

Upon arrival at the experimental site, the subjects were informed of the procedures and compensation for participation. They were asked to provide consent and agree to the use of data for research. The experiment took place in the shielding room, and the subjects were debriefed at the end.

2.2.1. Passive Procedure

When subjects enter the shielding room, their expressions tend to be stiff and awkward since they are not used to the cameras and experimental equipment. Rapport is first established, and subjects are eased into the experimental environment through a short role-playing session with the experimenter. Once the subjects appear to be comfortable with their surroundings, they are asked to watch the video clips and freely express their emotions while empathizing with the characters (Koreans may not express how they feel if uninstructed). Each video clip is two to four minutes long, and a 30-second fixation is presented at the start. Details on the video clips and their association with emotions are given in Table 1.

Table 1. Emotion Inducing Stimuli (Passive)

Emotion	Duration	Stimuli	Appropriateness (%)	Effectiveness (1~7)	Other emotions
Anger	2'31"	Bus driver is assaulted repeatedly	85.5	6.02	contempt, surprise, disgust
Boredom	3'30"	Listening repeated synthesized voice	93	5.38	anger, anxiety, disgust
Fear	3'	Horrible scenes from a movie "A Tale of Two Sisters"	93	5.41	surprise, disgust
Interest	2'23"	Optical illusion scenes	96	5.03	boredom, anxiety
Surprise	2'	Sudden a thundering noise	98	5.38	disgust, fear
Sadness	3'01"	Scene of a man crying for his father	95	6.01	pitiful
Neutral	1'14"	Reading sentences	99		

2.2.2. Active Procedure

The active procedure consists of tasks aimed at inducing more natural and effective emotions by having subjects participate in given situations involving touch and control (Table2). In the joy-inducing task, subjects were given the opportunity to win prize

money from a game of spin-the-wheel. The spinning wheel stops to point at different amounts of prize money when clicked with a mouse. Most subjects were found to experience great joy as they realized that the wheel was pointing to the highest amount of prize money.

In the fear-inducing task, subjects place their hands in a black box containing objects that cannot be seen from the outside. The box contains cold and sticky jelly, a warm and squashy wad of noodles, and a furry animal doll. Most subjects experience fear before placing their hands in the box, and then feel surprise and disgust instead of fear when groping inside. Others also expressed interest and curiosity. Another fear-inducing task was to pull on a rubber band tied around the subjects' wrists. The subjects were very afraid of the rubber band snapping and of the experimenter letting go of the rubber band.

The disgust-inducing task featured a rubbery alien doll. The doll's eyes pop out when squeezed, and the subjects would discover worms in the bloodshot eyeballs. Most subjects reported emotions of disgust, while others experienced fear.

Table 2. Emotion Inducing Stimuli (Active)

Emotion	Duration	Stimuli	Appropriateness (%)	Effectiveness (1~7)	Other emotions
Joy	2'30"	Spin the wheel	94	5.76	interest, fun
Disgust		Grasp Aliens doll	77	5.07	fear
Fear1		Black box navigation	54	5.54	surprise, disgust, interest, curiosity
Fear2		Rubber band snapping	73	5.10	surprise, interest

2.3 Data Acquisition

Emotional response data was collected over four rounds of experiment. The first round obtained data on facial expression, thermal imaging, and ANS response from middle and high school students and college students. A webcam (Logitech Quick Cam Pro9000) was set up above the monitor, and video data was recorded at 15 frames per second with 640x480 resolution. At the same time, a thermal imaging camera (NEC-Sanei H2640) was installed behind the monitor to obtain thermal imaging data at 5 frames per second. For ANS response, measurements were taken for SKT, ECG, PPG, EDA, and Respiration from sensors attached to the index finger, little finger, and ankle.

In the second round, vocal responses were collected in addition to facial expression and thermal imaging. Sound was recorded using a pin microphone and the Octopus Board developed by HCI-LAB as PCM at 16 kHz and 16bit sampling.

The third round collected EEG signals in addition to existing modalities. A Laxtha PolyG-I (8 channels) was used, and electrodes were attached to five locations (F3, F4, P3, P4, Cz). Facial expression was recorded with a Logitech HD Pro C920 at 30 frames per second.

In the fourth round, the number of active tasks were increased. Facial, vocal, EEG, and thermal data were collected from passive tasks, while only facial and vocal data were retrieved from active tasks. The method of data collection was the same as the third round.

The amount of data obtained is presented in Table 3.

2.5. Self-Report

At the end of each emotion-inducing tasks, subjects filled out a self-report questionnaire on the experienced emotions during the task. They specified the type of emotion, the scene at which the emotion was experienced, and the intensity of Active task Emotion on a scale of 1 to 7. The questionnaire contained an open-ended question on the type of emotions experienced. The same questions are repeated once more in the form of a close-ended question, requiring subjects to choose the most intense emotion from the list. Some assistance was provided for subjects facing difficulties in describing their emotions in the first open-ended question, so as to facilitate standardization of answers.

Table 3. The Amount of Collected Data

	Anger	Boredom	Disgust	Interest	Surprise	Fear	Joy*	Sadness	Fear*
face	540	539	251	538	542	544	534	542	98
voice	241	241	245	240	243	243	240	247	108
thermal	462	458	213	461	514	514	462	512	0
EEG	133	132	188	134	190	190	133	189	0
ANS	300	300	56	300	353	355	300	355	0

The subjects also reported on the dimensional value of emotions. On a scale of 1 to 7, they rated dimensions such as arousal (ranging from sleep to active), valence (ranging from negative to positive), and power (sense of control).

3. Emotion Labeling

In order to increase the objectivity of this material, the collected data were labeled based on objective evaluation. First, three graduate students found the portions of expressed emotion while watching the recorded video section. A group of 5 to 10 evaluators (undergraduate and graduate students) rated each emotional portions by observing facial expressions or listening to vocal data. The identified emotions were categorized if there was agreement between 2/3 or more of the evaluators. It was found that some emotions perceived by the evaluators were different from those experienced by the subjects.

The labeling process enabled extraction of emotions expressed at greater intensities among the naturally induced responses. As a result, there were not many scenes that evoked strong emotions. This signifies that acted data of high intensities do not occur frequently in natural situations, and that an acted response DB alone may not be sufficient in processing human emotions for HCI research. Table 4 gives the data for each labeled emotion.

4. Availability

Table 4. The Amount of Labeled Data

	Anger	Fear	Surprise	Sadness	Disgust	Joy
Face	191	60	503	107	665	548
Voice	209	71	146	94	217	137

This CNU Database can be listed the data by various combinations of retrieval cues such as gender, age, response modality, labeled emotion, induced stimuli. It will be available through a web-accessible interface after completing English version. Those are required to sign a License Agreement.

5. Discussion

This study developed an emotion-inducing method for Korean teenagers and college students, and collected natural emotional responses by implementing the new method. It is difficult to apply a standardized emotion-inducing method since emotional stimuli are heavily dependent on the periodical and cultural backgrounds of subjects. This study identified a common emotional chord from inducing emotions in college students over four years, and the accumulated know-how has helped to consistently improve emotion-inducing stimuli.

Video clips to be used in passive tasks should contain appropriate scenes to induce target emotions, and should build up to the climax. Optimal times should be set so that subjects will not become bored from lengthy procedures. This study identified appropriate session times through trial-and-error.

There are also cases when a single stimuli leads to mixed emotions. For instance, most pilot study participants answered the most anger to be felt while watching a bus driver being unjustly assaulted. However, the experiment revealed that scenes of violence and assault induced not only anger, but also other emotions such as pity towards the victim and shock. This issue was resolved by excluding assault scenes from a new scenario, in which the victim is unfairly forced to take the blame for organizational corruption.

Another difficulty of stimuli development can be traced to the need for appropriate stimuli for each modality. Even if target emotions are adequately induced, this does not necessarily lead to the desired modality data. The sadness-inducing stimuli is a video clip of a man crying for his father's death. While most subjects answered that they experienced sadness, there was no voice data indicating any such emotion. This is because sadness is associated with being quiet. To collect voice data for sadness, the subjects were asked to send a video message to their parents after watching the clip. In the video message, many subjects were close to tears and almost sobbing.

When a single stimuli is used to collect multimodal data, it may be difficult for certain pairs of modalities to co-exist. Since movement has to be controlled to obtain ANS or EEG data, collecting voice data at the same time was impossible. In addition, attaching electrodes places many constraints on producing natural facial expressions. Simultaneous measurements were attempted by maximizing the number of measurable combination pairs.

One limitation of this study is that labeling of all recorded data is incomplete. The task of tracing existing data shall be addressed in future studies. This database will be continuously improved, so as to account for more emotion types and a broader age range. Even so, it is evident that this study has contributed to the study of Koreans' emotions both qualitatively and quantitatively.

To obtain natural emotional response data from Koreans, this study developed a massive multimodal emotional response database by measuring validity and reliability, and collected various data such as facial expression, voice, ANS response, facial surface temperature, and brain wave. This study will prove valuable to psychological studies of Koreans' emotions. Further, it is expected to enhance the understanding of human emotions and have practical applications in HCI and other related fields.

Acknowledgment

This research was supported by the Converging Research Center Program through the Ministry of Science, ICT and Future Planning, Korea (2013K000332).

References

- [1] B. K. Fischer, R. Huber, J. Spilker and E. Noth, "How to Find Trouble in Communication", *Speech Comm.*, vol. 40, (2003), pp. 117-143.
- [2] H. A. Elfenbein and N. Ambady, "Universals and cultural differences in recognizing emotions", *Current Directions in Psychological Science*, vol. 12, (2003), pp. 159-164.
- [3] I. Sneddon, M. McRorie, G. McKeown and J. Hanratty, "The Belfast Induced Natural Emotion Database", *IEEE Transactions on Affective Computing*, vol. 3, no. 1, (2012), pp. 32-41.
- [4] J. Carroll and J. Russell, "Facial Expressions in Hollywood's Portrayal of Emotion," *J. Personality and Social Psychology*, vol. 72, (1997), pp. 164-176.
- [5] J. H. Sohn, "Psychological model of emotion: Categorized model, appraisal model, dimensional model", *Journal of the Ergonomics Society of Korea*, vol. 30, no. 1, (2011), pp.179-186.
- [6] M. Kamachi, V. Bruce, S. Mukaida, J. Gyoba, S. Yoshikawa and S. Akamatsu, "Dynamic Properties Influence the Perception of Facial Expressions," *Perception*, vol. 30, (2001), pp. 875-887.
- [7] T. Vogt and E. Andre, "Comparing Feature Sets for Acted and Spontaneous Speech in View of Automatic Emotion Recognition", In *Proceedings of IEEE International Conference on Multimedia & Expo, Amsterdam, The Netherlands, (2005)*.

