Audiovisual Emotional Speech of Game Playing Children: Effects of Age and Culture

Suleman Shahid[1], Emiel Krahmer[2] & Marc Swerts[2]

[1] User System Interaction, Eindhoven University of Technology, The Netherlands
[2] Communication and Cognition, Tilburg University, The Netherlands

[1]S.Shahid@tm.tue.nl, [2]{E.J.Krahmer/M.G.J.Swerts}@uvt.nl

Abstract

In this paper we study how children of different age groups (8 and 12 years old) and with different cultural backgrounds (Dutch and Pakistani) signal positive and negative emotions in audiovisual speech. Data was collected in an ethical way using a simple but surprisingly effective game in which pairs of participants have to guess whether an upcoming card will contain a higher or lower number than a reference card. The data thus collected was used in a series of cross-cultural perception studies, in which Dutch and Pakistani observers classified emotional expressions of Dutch and Pakistani speakers. Results show that classification accuracy is uniformly high for Pakistani children, but drops for older and for winning Dutch children[1].

Index Terms: Audiovisual speech, Positive and Negative emotions, Children speech

1. Introduction

It is generally assumed that children show their emotions in a more open and intense way than adults (e.g., Fabes & Martin, 1991). For young children expressing emotions is particularly important because of their low verbalization abilities. Still, we know relatively little about how children of different age groups express their emotions via audiovisual speech, and how sensitive observers are to these emotional expressions (e.g., Scherer 2002). Besides age, cultural background is another factor which is assumed to influence the expression of emotions. Recent work (e.g., Elfenbein and Ambady 2002, Elfenbein et al. 2002, Schimmack 1996) suggests that there are both important similarities and differences between cultures in the production and perception of emotions, although relatively few studies have addressed cross-cultural differences in audiovisual expression of emotion by children.

In this paper, we address the way age and culture influence the production of positive and negative emotions, zooming in on 8- and 12- year olds, from the Netherlands (a western culture) and Pakistan (a south-asian culture). To elicit natural emotions in an ethical way, we developed a very simple game in which participants have to guess whether the number of an upcoming card will be higher or lower than the previous card under discussion. Winning this game is assumed to trigger positive emotions in children, while losing it is assumed to trigger negative emotions. The use of games is a well-known and proven method to elicit emotions (see

e.g., Kaiser and Wehrle 1996), but the game used here is particularly simple and effective for both Dutch and Pakistani children.

The data collected with this game paradigm was used in a series of cross-cultural perception studies, in which Dutch and Pakistani adults look at both the Dutch and Pakistani child speakers, and have to determine on the basis of the emotional expression whether the children had just won or lost their game.

This paper is structured as follows. In section 2, we describe the game that was designed to elicit positive and negative emotions, and we report on the application of the game with both Dutch and Pakistani children. It will be seen that the method was very effective, and that inspection of the recordings suggest a number of similarities and differences between age groups and cultures. To quantify these differences, we then report on a series of cross-cultural perception studies in section 3, where adult viewers from both cultures are asked to classify winning and losing games recorded with children from both cultures. Finally, in section 4 we position our findings more broadly, and sketch various lines for future research.

2. Experiment I: Data Collection

2.1. Setup (Game design)

Figure 1: *Start of a game (L) and end of the game (R)*

To elicit positive and negative emotions in children, we developed a simple card game in which children have to guess whether an upcoming card will contain a higher or lower number than a reference number. At the start of a game (see Figure 1), players see a row of 6 cards on a computer screen five of which are turned upside down, and only the number on the first card can be seen (in our example this is a '1'). The numbers on the cards are all between 1 and 10 and players have to guess whether the next card in the row will contain a higher or a lower number (a rational choice in our example would be to guess "higher" for the second card). After guessing whether the next number will be higher or lower, the relevant card is turned around. Players receive feedback on the correctness or incorrectness of their choice via a characteristic non-speech audio sound. As soon as participants make a wrong prediction the game is over. Participants win a

Figure 2: *Representative stills for (L-R): 8 years old winning, 8 year old loosing, 12 years old winning and 12 years old loosing, for Pakistani players (Top)* and *Dutch players (Bottom)*

game if they predict all cards correctly. In that case they receive a coin, and at the end of the experimental session (consisting of multiple games) they can trade in the collected coins for a prize (and participants are led to believe that more coins give access to a bigger prize).

The games were developed using Microsoft® PowerPoint® and displayed in full screen mode. Appropriate background images were chosen for the game to give it a colorful look and to make it more attractive for children, and the animation option was used to turn cards around. During the experiment, children played six games, and could in theory win 6 coins. However, unknown to the children, each game was completely deterministic, and two different game variants were employed, one where a rational decision procedure would result in winning the game, and one in which being rational implied losing the game. Figure 1 is an example of the latter: the most probable outcome for the final card would be that its number is higher than 3, but guessing "higher" would make this a losing game. The two other losing games were: 8-3-9-7-2-1 and 7-2-8-3-9-10. Winning scenarios were 9-2-8-1-10-7, and 3-8-2-10-9-2 and 1-3-9-2-7-6. Winning and losing games were mixed, starting and ending with a variant in which children were likely to win.

2.2. Participants

In total, 88 children played the game in pairs, 48 Dutch and 40 Pakistani ones. Half of the Dutch children were around 8 years old (group 4 in the Dutch elementary school system) and the other half were around 12 years old (group 8 in the Dutch school system). In a similar fashion, half of the Pakistani kids were around 8 years old and half 12 years old. Parents gave prior written consent for their child to participate, and signed a form stating that the recordings could be used for research purposes.

2.3. Procedure

The experiment was conducted in an elementary school in Tilburg with Dutch participants and in Humaira Kamal School, Lahore with Pakistani participants. A separate room was chosen in both schools where children in self-selected pairs of the same age group were invited and asked to sit on the chairs placed in front of a desk on which a laptop computer was placed. Children participated in pairs to allow for a more natural interaction (e.g., Guering 1986, Wagner and Lee 1999). Behind the laptop, a video camera was

positioned in such a way that it could record the children's faces and the upper part of their body. The camera was adjusted to children's height before the start of each experiment. The laptop was connected with another computer, which the experimenter used to operate the game. Both in The Netherlands and in Pakistan, the experimental set-up was such that the children could not directly see the experimenter, to minimize possible interfering effects of the experimenter on the child participants.

Once the children were in the room and had chosen the appropriate seats for sitting, the experimenter welcomed them and started a small talk discussion by raising a few questions to take away the pressure ("How old are you? Do you like to play games?" Etc.). After this preliminary phase, the experimenter gave spoken instructions, telling the children about the game and the coins they could win. All the game rules outlined in the previous section were explained to the children, and when they seemed to understand the rules, the experimenter started a practice game ("So you only have to say whether the next card is higher or lower. This is just an exercise and it doesn't really count"). After this exercise, the experimenter asked the children whether they had any questions, and if not the experimenter left the children's field of vision and started the first experimental game.

During the experiment, the experimenter tried to say as little as possible, and spoke only if the children did not say anything (in which case the experimenter encouraged them to talk), or if they did not agree or got in an argument following a losing decision ("I told you we should have said 'higher'!). The experiment did not have a fixed duration because it involved a lot of discussion between the children, which varied considerably from pair to pair. On average each session lasted for approximately 10 to 12 minutes. At the end of the game session, the experimenter congratulated the children ("Very well done, you have won enough coins for the price.") and they could trade in their coins for an individual gift (a whistle or small plastic ball in the case of the Dutch participants and king size chocolates and key rings for the Pakistani participants). Finally, the experimenter thanked them and asked a new couple to come in the room.

2.4. Results and discussion

Figures 2 shows representative stills of winning and losing couples from both age categories and both cultures. In general, the game worked surprisingly well, in that all pairs of

participants indeed made the rational choices that were anticipated in most of the cases, so that each pair of children lost at least two games and won at least two games. None of the children noticed that the game was in fact a deterministic simulation. The data that was gathered offers a rich collection of audiovisual emotional child speech, consisting of both decision making dialogues, and audiovisual responses to winning or losing a game. Informal observations reveal clear differences between age groups and cultures, which we attempt to quantify in a series of cross-cultural perception experiments described below.

3. Experiment II: Perception

3.1. Stimuli

From all of the child-pairs that participated in Experiment I (with the exception of two Dutch pairs whose recordings could not be used), we selected the response to the first two of their winning games (in which they made a correct prediction for the last card) and the first two of their losing games (in which the final guess turned out to be incorrect). The stimuli were cut from the moment the final card was turned until their primary response was finished. This resulted in 88 Dutch stimuli [11 (8 year old pairs) x 11 (12 year old pairs) x 2 wins x 2 lost] and 80 Pakistani stimuli [10 (8 year old pairs) x 10 (12 year old pairs) x 2 wins x 2 lost]. Stimuli were presented in a random order, in a vision-only format to avoid participants from relying on auditory cues ("Jahoe!").

3.2. Participants

131 adults participated in one of the cross-cultural perception experiments, 71 Dutch and 60 Pakistani, with a roughly equal number of men and women.

3.3. Procedure

Four group experiments were conducted: Dutch viewers judging Dutch children [31 participants], Dutch viewers judging Pakistani children [40 participants], Pakistani viewers judging Dutch children [30 participants] and Pakistani viewers judging Pakistani children [30 participants] with essentially the same procedure for all four experiments.

For every experiment, groups of participants were invited into a quite class room where the computer screen was projected on the classroom wall using a beamer. Participants were told that they would see 88 stimuli in the case of Dutch children or 80 stimuli in the case of Pakistani children. In addition, they were instructed that the children were showing their emotions after winning or losing a game and that their task was to determine by forced whether the children had won or lost. Each stimulus was preceded by a number displayed on the screen indicating the upcoming stimulus, and followed by a six second pause during which participants could fill in their score on the answer form. The actual experiment was preceded by a short training session in which 3 clips were shown (different from the ones shown in the actual experiment) to make participants familiar with the stimuli and the experimental task. If everything was clear, the actual experiment started which lasted for approximately 17 minutes in the case of Dutch stimuli and 15 minutes in the case of Pakistani stimuli. During the experiment there was no interaction between participants and experimenter.

3.4. Statistical analysis

All tests for significance were performed using the logistic regression method with as factors Age group (levels: 8 years old, 12 years old), Game status (levels: Win, Lose) and Experiment (levels: Dutch viewers judging Dutch children, Dutch viewers judging Pakistani children, Pakistani viewers judging Dutch children and Pakistani viewers judging Pakistani kids) and with Classification (Correct, Incorrect) as the dependent variable.

3.5. Results

Table 1 summarizes the results. Overall, most classifications are correct, which indicates that participants were able to correctly determine the status of the game based on the emotional response of the children. In most cases there are no big differences between Dutch and Pakistani observers, although, interestingly, Pakistani viewers classify the Dutch children a little better than the Dutch viewers and the Dutch viewers systematically have a slightly higher percentage of correct classification for the Pakistani children. Quite substantial differences are found between Dutch and Pakistani children, however; the percentage of correct classifications for the Pakistani children are uniformly high (always more than 80% correct), whereas for the Dutch children it can be seen that the percentage of correct classification drops for the 12 years olds. Interestingly, when 12 year old Dutch children win the game, this is rather difficult to see; most viewers, especially the Dutch ones, perceive this as losing.

The logistic regression revealed a significant main effect of Age group (χ^2 (1) = 135.394, p < .001), and inspection of Table 1 indicates that overall the number of correct classifications is higher for the 8 year olds than for the 12 year olds. Also a significant main effect was found for Game status (χ^2 (1) = 181.232, p < .001), indicating that overall the percentage of correct classification is higher for losing games. And finally a significant main effect was found of Experiment (χ^2 (3) = 737.574, p < .001). When looking at Table 1 it can be seen that overall Dutch observing Dutch leads to the lowest percentage correct classification and the Dutch observing Pakistani's leads to the highest percentage of correct classification.

Given the differences in how Dutch and Pakistani children respond to winning and losing, the interaction effects are especially interesting. To begin, we found a significant interaction between Age group and Game Status (χ^2 (1) = 43.605, p < .001). This interaction can be explained by the fact that overall for 8 years old, both winning and losing are mostly recognized correctly, while for 12 years old, losing is more often classified correctly than winning. We also found a significant interactions between Age group and Experiment (χ^2 (3) = 119.516, p < .001) and between Game status and Experiment (χ^2 (3) = 180.323, p < .001), which can be explained by observing that the percentage of correct classifications drops for Dutch 12 year olds and for Dutch children that lose their game. Finally, a significant three way interaction was found between Age group, Game status and Experiment (χ^2 (15) = 1420.180, p < .001). The data show that for Dutch 8 year olds winning games are more difficult to classify correctly for both Pakistani and Dutch observers, while for Pakistani 8 year olds winning is somewhat easier to detect.

4. Discussion and Conclusion

In this paper we investigated how children of different age groups and with different cultural backgrounds signal positive and negative emotions. To collect emotional data in an ethical way, we set up a simple but very effective game in which pairs of participants have to guess whether a card will contain

Table 1: *Percentage correct classification results for Dutch and Pakistani observers judging stimuli from Dutch and Pakistani kids in 2 age groups in winning or losing game situation*

Game Players	Age Group	Game Status	Observers	
			Dutch	Pakistani
Dutch	8	Lost	86.7	83.0
		Won	70.2	74.5
	12	Lost	75.6	76.1
		Won	36.8	45.9
Pakistani	8	Lost	86.1	82.7
		Won	91.9	89.0
	12	Lost	93.0	86.0
		Won	91.5	87.3

a higher or lower number than a reference card. During the experiment, children played six games which were set up in such a way that by making "rational" choices they would win 3 games and lose 3 games. This strategy worked as intended in the vast majority of the games that were played. The same series of games was played by 48 Dutch children (in Tilburg) and 40 Pakistani children (in Lahore). In both countries, half of the children had an average age of 8 years (group 4 in the Dutch system) and half had an average age of 12 years (group 8 in the Dutch system).

The data collected in this way was used in four perception experiments, in which Dutch and Pakistani viewers (again in Tilburg and in Lahore) saw fragments (without sound) of pairs of Dutch or Pakistani children immediately after winning or losing their game. We found that overall the number of correct classifications was rather high, indicating that people could usually decide on the visual responses of the children whether they had just lost or won a game. Interestingly, for Dutch children it appears that losing is easier to detect than winning; this is true for both 8 and 12 year olds, but the differences are especially clear for the older children, where both Dutch and Pakistani observers misclassify the majority of the games that were won as games that were lost. This is the only instance where most of the classifications are incorrect, which indicates that 12 year old Dutch children barely respond visually to winning a game, although they respond visually to losing a game. This is completely different for the Pakistani children, where in all variants the classifications were correct in most of the cases, and where winning the game is somewhat easier to detect than losing, both for 8 and for 12 year old children. Interestingly, this pattern (Dutch signal negative emotions more clearly, Pakistani signal positive emotions more clearly) was also found in an analysis of how adult speakers signal emotions in a completely different setting (Shahid et al. 2007). The findings for the Dutch children are in accordance with the internalization of emotion theory, which says that younger children express their emotions more clearly than older ones (e.g., Thompson 1994), but this pattern was not found for the Pakistani children, which is in accordance with observations of Elfenbein et al. 2002 who compare the emotional expressivity of American and Indian adults. Overall, the scores of the Dutch and Pakistani observers are very similar, but the Dutch viewers score a little better than the Pakistani viewers for the Pakistani children, and the Pakistani viewers score a little better for the Dutch children. This is not what one would expect on the basis of the alleged in-group advantage (e.g., Elfenbein and Ambady 2002) which suggests that people are better at classifying the emotions of speakers belonging to their (ethnical) group. In future work we would like to further investigate this.

Various other lines of future work readily suggest themselves. To begin, participants in the perception experiments were only presented with how children responded after finding out that they won or lost their game. Inspection of the recordings revealed that there were also differences in how Dutch and Pakistani children behaved while deciding whether the upcoming card would contain a higher or lower number. One striking difference was that Pakistani children often made praying gestures (forming a cup with the hands to pray, and finishing the prayer by wiping the hands across the face) while Dutch children never prayed. In fact, this might partially explain why losing does not trigger a strong response in Pakistani children (a Pakistani proverb says that you don't complain to God if you do not get what you want). Another interesting difference is that the interpersonal distance between Pakistani children appeared to be smaller than that between Dutch children; there was more touching, holding hands, putting head on shoulder with the Pakistani than with the Dutch children. This was true both during the decision making phase, and after the game result was known. There were also striking similarities, especially for the 8 years old children. For instance, both Dutch and Pakistani 8 year old children frequently used a winning gesture, where the hands form fists and are moved downward quickly. In future work we intend to quantify these differences and similarities in non-verbal behavior between Dutch and Pakistani children, both during decision making and during the response to the outcome of the game.

5. References

Elfenbein, H. and Ambady, N. (2002). On the universality and cultural specificity of emotion recognition. *Psychological Bulletin 128*, 203-235.

Elfenbein, H., Ambady, N., Mandal, M., Harizuka, S. (2002), Cross-cultural Patterns in Emotion Research, *Emotion, 2*, 75-84.

Fabes, R. and Martin, C. (1991). Gender and age stereotypes of emotionality, *Personality and Social Psychology Bulletin*, *17*, 532-540.

Guerin, B. (1986). Mere presence effects in humans: A review. *Journal of Experimental Social Psychology, 22*, 38-77.

Kaiser, S. & Wehrle, T. (1996). Situated emotional problem solving in interactive computer games. In N. Frijda (Ed.), *Proceedings of the International Society for Research on Emotions* (pp. 276-280), Toronto.

Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms, *Speech Communication, 40*, 227-256.

Shahid, S., Krahmer, E., Swerts, M. (2007), Real vs. acted emotional speech: Comparing Caucasian and South-Asian speakers and observers, *submitted*.

Schimmack, U. (1996). Cultural influences on the recognition of emotion by facial expressions, *Journal of Cross-Cultural Psychology, 27*, 37–50.

Thompson (1994), Emotion Regulation: A Theme in Search of Definition, in: *The Development of Emotion Regulation*, pp. 25-52,

Wagner, H. & Lee, V. (1999). Facial behavior alone and in the presence of others. in P. Philippott et al. (Eds.), *The social context of nonverbal behavior* (pp. 262-286). New York: Cambridge University Press.