

Cross-Cultural Evaluations of Avatar Facial Expressions Designed by Western Designers

Tomoko Koda¹, Matthias Rehm², and Elisabeth André²

¹ Faculty of Information Science and Technology, Osaka Institute of Technology
1-79-1 Kitayama, Hirakata city, Osaka, 573-0196, Japan

² Multimedia Concepts and Applications, Faculty of Applied Computer Science, University
of Augsburg, Eichleitnerstr. 30, 86159 Augsburg, Germany
koda@is.oit.ac.jp, rehm@informatik.uni-augsburg.de,
Elisabeth.Andre@informatik.uni-augsburg.de

Abstract. The goal of the study is to investigate cultural differences in avatar expression evaluation and apply findings from psychological study in human facial expression recognition. Our previous study using Japanese designed avatars showed there are cultural differences in interpreting avatar facial expressions, and the psychological theory that suggests physical proximity affects facial expression recognition accuracy is also applicable to avatar facial expressions. This paper summarizes the early results of the successive experiment that uses western designed avatars. We observed tendencies of cultural differences in avatar facial expression interpretation in western designed avatars.

1 Introduction

Since instant messenger and chat services are frequently used in our daily communication beyond nationality and languages, emoticons and expressive avatars are widely used to provide nonverbal cues to text-only messages [1, 2, 3]. Recent growth of Second Life [4] attracts worldwide attention to avatar mediated communication both from entertainment and businesses. Studies on emoticons and avatars report positive effects on computer-mediated communication. Those studies indicate that emoticons and avatars improve user experiences and interactions among participants [5, 6, 7] and build enthusiasm toward participation and friendliness in intercultural communication [8, 9].

However, these avatars are used based on an implicit assumption that avatar expressions are interpreted universally across cultures. Since avatars work as graphical representations of our underlying emotions in online communication, those expressions should be carefully designed so that they are recognized universally. We need to closely examine cultural differences in the interpretation of expressive avatars to avoid misunderstandings in using them.

In order to examine cultural differences in avatar facial expression recognition, we apply findings from psychological studies on human facial expressions, since there have been a much wider variety of studies in psychology on human expressions than

on avatar expressions. Recent psychological research found evidence for an “in-group advantage” in emotion recognition. That is, recognition accuracy is higher for emotions both expressed and recognized by members of the same cultural group [10]. Elfenbein et al. state, “This in-group advantage, defined as extent to which emotions are recognized less accurately across cultural boundaries, was smaller for cultural groups with greater exposure to one another, for example with greater physical proximity to each other [10].” Also, the decoding rule [11] implies that we concentrate on recognition of negative expressions, since misinterpretation of negative expressions leads to more serious social problems than misinterpretation of positive expressions would cause.

We conducted a web experiment to compare interpretations of Japanese designed avatars' facial expressions among 8 countries, namely, Japan, South Korea, China, United States, United Kingdom, France, Germany, and Mexico in our previous study [12, 13]. The results showed the following:

- 1) Cultural differences do exist in interpretation of avatar facial expressions, which confirms the psychological findings that physical proximity affects recognition accuracy. The in-group advantage was found within Japan and between Korea and Japan.
- 2) There are wide differences among cultures in interpreting positive expressions, while negative expressions had higher recognition accuracy. This result indicates that the decoding rule is found in avatar expression interpretation.

This paper reports the results of our successive experiment that uses western designed avatars. The objective of the current experiment is to examine whether in-group advantage and decoding rule are also applicable to avatars' facial expressions drawn by western designers.

2 Experiment Overview

We are currently conducting a web based experiment since February, 2008. The experimental procedure is the same as the previous experiment except that the avatars used in the current experiment are drawn by French, British, and American designers and one made digitally, while the avatar designs used in the previous experiment were by Japanese artists [12, 13].

Participation in the experiment was by invitation only. We have collected total of 293 answers from United States (n=98), France (n=23), Germany (n=75), and Japan (n=97). The answers of which the common language of a participant's country and his/her native language match are selected for later analyses. The participants' gender ratio is 65% of male and 35% of female, and their age ranges are 80% in the 20's and 10% in the 30's. 53% of the participants rated themselves as expert computer users, 43% as intermediate.

2.1 Experimental Procedure

The experiment was developed using the application of Macromedia Flash. Participants first answer a brief questionnaire on their background profile such as their

nationality and mother tongue. The main experiment starts after the questionnaire, which is presented as a matching puzzle game as shown in Fig. 1. Participants are requested to match 12 facial expressions to 12 adjectives. The 12 facial expressions are displayed in a 4 x 3 matrix and the 12 adjectives as buttons below the matrix. As shown in Fig. 1, participants can drag/drop the adjective buttons to/on the 12 expressions and continue changing the location of each button until they are satisfied with their answer. One avatar representation is chosen randomly from 7 avatars, and facial expression images are randomly placed in the 4 x 3 matrix. The adjective buttons are always displayed in the same order, and the 12 adjectives are always the same (see the next section for the adjectives used in the experiment).

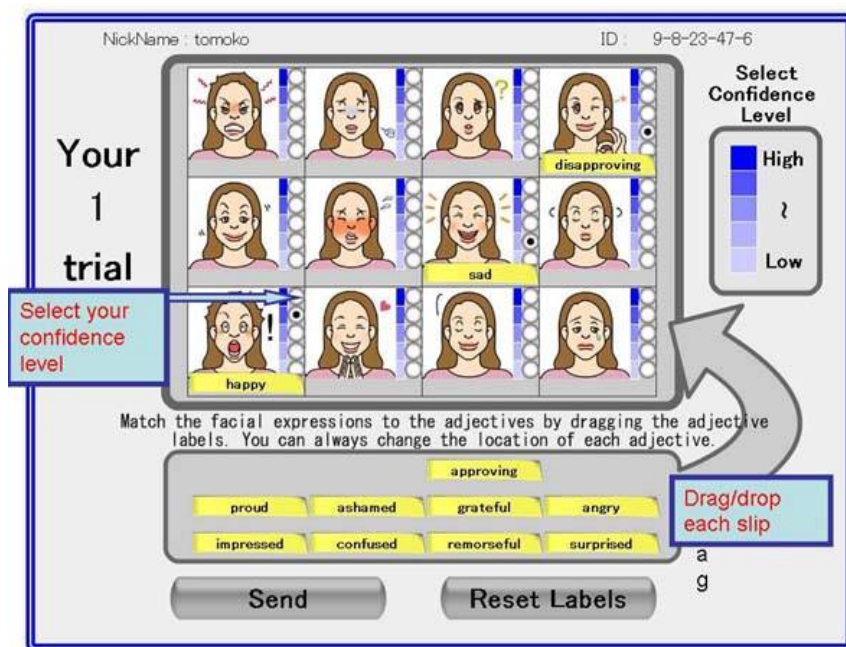


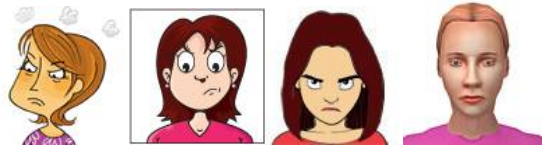
Fig. 1. Experiment screen shown as a matching puzzle game (Example screen for English speaking participants)

Participants' answers to the puzzle game as well as their background profile including gender, age, country of origin, native language, and computer expertise are logged in the server for later analyses. Participants are required to continue the experiment until they finish evaluating all the 7 avatar designs. The adjectives can be shown in English, German, French, and Japanese (all validated by native speakers). Participants from countries where the above languages are primarily spoken can see the adjective selections in their native language according to the background profile.

2.2 Avatar and Expression Design

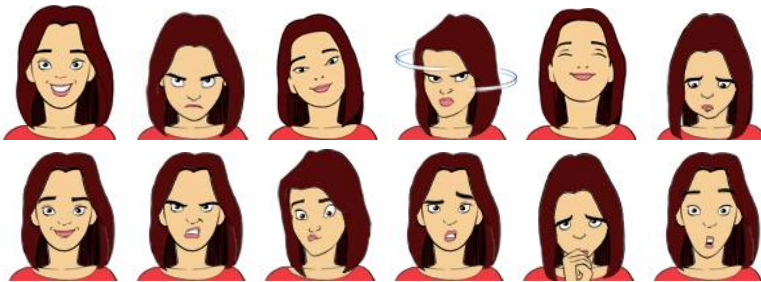
Commercially used avatars are represented not by photo-realistic images but as caricatures or comic figures. We prepared 7 avatar representations drawn by French

(2 avatars), British (2 avatars), and American (1 avatar) professional designers using their countries' comic/anime drawing styles, and 2 avatars created by a 3D facial expression modeling tool by a German researcher (the last avatars are referred to as German designs). The designers grew up and speak the first common language in the referred countries. By using avatars drawn with techniques from one culture, we can use those avatars as “expressers” and subjects as “recognizers” as in [11]. Accordingly, comparing the answers between users of the designer's country and those of other countries made it easier to validate the in-group advantage. Fig. 2 shows examples of “displeased” expression drawn by French, British, American designers and German design.



Note: from left to right, drawn by French, British, American designer, and made by a 3D modeling tool (German design).

Fig. 2. Examples of avatar expression



Note: From top left, pleased, displeased, approving, disapproving, proud, ashamed, grateful, angry, impressed, confused, remorseful, and surprised.

Fig. 3. Twelve facial expressions of one of the avatars designed by an American designer

The 12 expressions used in the experiment are “pleased,” “displeased,” “approving,” “disapproving,” “proud,” “ashamed,” “grateful,” “angry,” “impressed,” “confused,” “remorseful,” and “surprised” as shown in Fig. 3. Those expressions are selected from Ortony, Clore and Collins’ global structure of emotion types, known as the OCC model [14]. These are commonly used expressions in chat and instant messenger systems [1, 2, 3], and they reflect those emotions desired by the subjects for intercultural communication in [8].

These 12 expressions are paired as valanced expressions as defined in the OCC model, that is, negative/positive emotions that arise in reacting to an event or person. “pleased,” “approving,” “proud,” “grateful,” and “impressed” are positive expressions, while “displeased,” “disapproving,” “ashamed,” “angry,” “confused,” and “remorseful” are negative expressions, leaving “surprised” as a neutral expression.

3 Results

The participants' answers to the puzzle game are analyzed by calculating matching rates between expressions and adjectives. There is no correct answer to the matching puzzle, but the avatar designers' original intention can be used as an expresser's "standard" answer. Each expression and adjective is assigned a number (1-12) within the system. The designer's intended pairs are described as (1,1), (2,2), (3,3), (4,4) reflecting (expression number, adjective number). We calculated each country's number of "expression-adjective" pairs that are the same as the designers' pairs. Consequently, here, "matching rate" means the percentage of pairs of expressions and adjectives that match the avatar designer's intentional pairs. We use this expression-adjective matching rate in comparing the answers to the 12 facial expressions.

3.1 Analysis of Average Matching Rates by Designs

Table 1 shows the matching rates of the 12 expressions of German, UK, French, and US designs. The average matching rates of negative expressions are always higher than the ones of positive expressions in any designs (average matching rates of positive expressions: negative expressions (%) German 34.5:39.6, UK 46.6:73.9, French 65.6:79.6, US 40.9:67.0).

When we compare the matching rates by country, the average matching rate of the French design by French participants, and the matching rate of the American design by American participants are the highest among other countries. In addition, the matching rates of Japan are always the lowest in any avatar designs made by western designers.

3.2 Analysis of Expressions by Recognition Accuracy

In this section, we categorize the expressions into three groups according to their recognition accuracy. Those categories are as follows:

- 1) Highly recognized expressions (Expressions that have higher than 70% matching rates in all four countries): The expressions belong to this category are German Angry, British Ashamed, British Angry, British surprised, French disapproving, French proud, French ashamed, French angry, and American surprised.
- 2) Poorly recognized expressions (Expressions that have lower than 30% matching rates in all four countries): The expressions belong to this category are German disapproving, German impressed, German remorseful, British impressed, and American grateful.
- 3) Culture dependent expressions (Expressions that have wider than 50% distance among the matching rates of four countries): The expressions belong to this category are British Pleased, British proud, French grateful, French confused, American pleased, American disapproving, American proud and American impressed.

Highly recognized expressions are consistently highly recognized across four participating country, and between western countries and Japan. Thus we can assume those expression designs would cause less misinterpretations than other designs when used

across countries. Poorly recognized expressions are consistently recognized as different expressions from the designer’s intended expressions. The poorly recognized expressions are more subtle expressions than the highly recognized expressions (i.e., impressed vs. angry). Culture-dependent expressions have wider variances in their matching rates, and Japanese matching rates are the lowest among them.

Table 1. Matching rates shown by designers’ and participants’ countries

Note: Bold font: the matching rates of the designer’s country, *: highest rates, **: lowest rates

| German 3D design | | | | | | UK design | | | | | |
|---------------------|-------|--------|-------------|--------|---------|---------------------|------|--------|---------|--------|---------|
| | US | France | Germany | Japan | Average | | US | France | Germany | Japan | Average |
| Pleased | 42 | 50 | 36 | 48 | 35.2 | Pleased | 55 | 85 | 42 | 28 | 52 |
| displeased | 14 | 50 | 20 | 15 | 19.8 | displeased | 41 | 57 | 53 | 62 | 52.6 |
| approving | 32 | 50 | 40 | 33 | 31 | approving | 35 | 42 | 50 | 31 | 41.6 |
| Disapproving | 25 | 0 | 12 | 30 | 13.4 | Disapproving | 52 | 57 | 61 | 31 | 50.2 |
| Proud | 53 | 66 | 64 | 51 | 56.8 | Proud | 61 | 85 | 76 | 25 | 69.4 |
| ashamed | 50 | 66 | 44 | 36 | 39.2 | ashamed | 97 | 100 | 92 | 100 | 97.8 |
| Grateful | 35 | 16 | 44 | 15 | 32 | Grateful | 52 | 28 | 65 | 40 | 47 |
| Angry | 96 | 83 | 92 | 87 | 81.6 | Angry | 100 | 100 | 92 | 100 | 98.4 |
| Impressed | 32 | 16 | 28 | 12 | 17.6 | Impressed | 26 | 14 | 11 | 15 | 23.2 |
| confused | 64 | 16 | 24 | 15 | 43.8 | confused | 70 | 71 | 46 | 65 | 70.4 |
| Remorseful | 21 | 16 | 24 | 12 | 14.6 | Remorseful | 67 | 42 | 53 | 28 | 58 |
| surprised | 39 | 33 | 44 | 39 | 31 | surprised | 79 | 71 | 80 | 93 | 84.6 |
| Average | 41.9* | 38.5 | 39.3 | 32.8** | 34.7 | Average | 61.3 | 62.7* | 60.1 | 51.5** | 60.1 |
| Positive expression | | | | | 34.5 | Positive expression | | | | | 46.6 |
| Negative expression | | | | | 39.6 | Negative expression | | | | | 73.9 |

| French design | | | | | | American design | | | | | |
|---------------------|------|--------------|---------|--------|---------|---------------------|--------------|--------|---------|--------|---------|
| | US | France | Germany | Japan | Average | | US | France | Germany | Japan | Average |
| Pleased | 46 | 75 | 60 | 30 | 52.2 | Pleased | 37 | 75 | 25 | 12 | 49.8 |
| displeased | 75 | 62 | 68 | 72 | 75.4 | displeased | 43 | 75 | 41 | 50 | 61.8 |
| approving | 90 | 87 | 84 | 72 | 86.6 | approving | 37 | 0 | 16 | 0 | 10.6 |
| Disapproving | 96 | 87 | 92 | 72 | 89.4 | Disapproving | 62 | 75 | 33 | 100 | 74 |
| Proud | 78 | 100 | 76 | 84 | 87.6 | Proud | 75 | 75 | 100 | 12 | 72.4 |
| ashamed | 78 | 87 | 76 | 78 | 73.8 | ashamed | 50 | 50 | 50 | 31 | 56.2 |
| Grateful | 62 | 75 | 28 | 39 | 60.8 | Grateful | 25 | 25 | 16 | 18 | 16.8 |
| Angry | 96 | 100 | 96 | 96 | 97.6 | Angry | 37 | 50 | 41 | 62 | 58 |
| Impressed | 50 | 50 | 24 | 30 | 40.8 | Impressed | 81 | 50 | 25 | 18 | 54.8 |
| confused | 62 | 87 | 32 | 27 | 61.6 | confused | 100 | 75 | 58 | 93 | 85.2 |
| Remorseful | 75 | 75 | 52 | 57 | 61.8 | Remorseful | 50 | 25 | 50 | 31 | 51.2 |
| surprised | 78 | 100 | 64 | 63 | 81 | surprised | 100 | 100 | 91 | 100 | 98.2 |
| Average | 73.8 | 82.1* | 62.7 | 60.0** | 72.4 | Average | 58.1* | 56.3 | 45.5 | 43.9** | 57.4 |
| Positive expression | | | | | 65.6 | Positive expression | | | | | 40.9 |
| Negative expression | | | | | 79.6 | Negative expression | | | | | 67.0 |

4 Discussion

The analysis of the average matching rates by designs showed that negative expressions have higher recognition accuracy than positive expressions in western designed avatars. This may imply another decoding rule in western designs. The result of Japan’s lowest average matching rates in any western designs suggests there is a tendency of in-group advantage among western countries.

When we categorized avatar designs by their recognition accuracy, the result showed that Japanese matching rates of the culture dependent designs are the lowest among the four countries. Thus again, there is a tendency that Japanese recognition accuracy is lower than western countries in culture-dependent designs, which suggests in-group advantage among western countries.

Further issues to be examined are as follows:

- 1) More participants: Further statistical analyses should be made with more participants. We analyzed 293 answers in the current experiment, while there were more than 1200 answers in the previous experiment.

- 2) Variations in avatar designs: There should be more variations in avatar designs in order to avoid one designer's judgment and drawing style. We used total of seven avatar designs drawn by four western designers (one designer per country) in the current experiment, while there were 40 avatar designs made by three Japanese designers in the previous experiment. Although the four western designers are carefully selected to have equal professional skills, we could guarantee the quality of avatar designs by having more variety in designers and avatar designs.
- 3) Translation of adjectives: Translations of adjectives are not completely equal as some languages do not have the exact meaning of the adjectives in other languages. One solution is to use scenario based interpretation for each avatar expression.
- 4) Definition of culture: We simply defined culture by participants' country and first language. However, the definition of culture is more complicated, e.g., religion. Further study should consider cultural models such as Hofstede's rankings [15], and cultural uncertainty avoidance models [16, 17].

5 Conclusion

The goal of the study is to investigate cultural differences in avatar expression interpretations and apply findings from psychological study in human facial expression recognition. Our previous study using Japanese designed avatars showed there are cultural differences in interpreting avatar facial expressions, and the psychological theory that suggests physical proximity affects facial expression recognition accuracy (in-group advantage) and decoding rule are also applicable to avatar facial expressions. We again observed tendencies of decoding rule and in-group advantage among western countries in the current experiment using western designed avatars.

Acknowledgements

This research is supported by a Grant-in-Aid for Young Scientists (Start-up: 18800063) from the Japan Society for the Promotion of Science (JSPS). The Universal Character Experiment 2008 is jointly conducted with University of Augsburg under DFG research grant RE 2619/2-1 (CUBE-G).

References

1. MSN Messenger: <http://www.msn.com>
2. Yahoo! Messenger: <http://messenger.yahoo.com/>
3. Kurlander, D., Skelly, T., Salesin, D.: Comic Chat. In: Proceedings of Computer Graphics and Interactive Techniques, pp. 225–236. ACM Press, New York (1996)
4. Second Life: <http://secondlife.com/>
5. Damer, B.: Avatars: Exploring and Building Virtual Worlds on the Internet. Peachpit Press, Berkeley (1997)
6. Smith, M.A., Farnham, S.D., Drucker, S.M.: The Social Life of Small Graphical Chat Spaces. In: Proceedings of CHI, pp. 462–469. ACM Press, New York (2000)

7. Pesson, P.: ExMS: an Animated and Avatar-based Messaging System for Expressive peer Communication. In: Proceedings of GROUP, pp. 31–39. ACM Press, New York (2003)
8. Koda, T.: Interpretation of Expressive Characters in an Intercultural Communication. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3214, pp. 862–868. Springer, Heidelberg (2004)
9. Isbister, K., Nakanishi, H., Ishida, T.: Helper Agent: Designing and Assistant for Human-Human Interaction in a Virtual Meeting Space. In: Proceedings of Human Factors in Computing Systems (CHI2000), pp. 57–64. ACM Press, New York (2000)
10. Elfenbein, H.A., Ambady, N.A.: Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology* 34, 92–110 (2003)
11. Elfenbein, H.A., Ambady, N.: On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis. *Psychological Bulletin*, American Psychological Association, Inc. 128(2), 203–235 (2002)
12. Koda, T., Ishida, T.: Cross-cultural study of avatar expression interpretation. *International Symposium on Applications and the Internet (SAINT)* (2006)
13. T. Koda. Cross-cultural study of avatars' facial expressions and design considerations within Asian countries. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.): *Intercultural Collaboration I*. LNCS, Springer-Verlag, pp.207-220 (2007).
14. Ortony, A., Clore, G.L., Collins, A.: *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge (1998)
15. Hofstede, G.: *Cultures Consequences: International Differences in Work-Related Values*. Sage Publications, Thousand Oaks (1984)
16. Berger, C.R., Calabrese, R.: *Some explorations in initial interactions and beyond* (1975)
17. Gudykunst, W.B., Mody, B.: *Handbook of International and Intercultural Communication* SAGE 2002. In: Hall, E., Hall, M.R. (eds.) *Understanding Cultural Differences*, 2nd edn. Intercultural Press, Yarmouth, Maine (1990)