

Integration and Evaluation of Prototypical Culture-related Differences

Birgit Endrass¹, Nick Degens², Gert Jan Hofstede^{2,3}, Elisabeth André¹, John Hodgson⁴, Samuel Mascarenhas⁵, Gregor Mehlmann¹, Ana Paiva⁵, Christopher Ritter⁶, and Aleksandra Swiderska⁷

¹ Human Centered Multimedia, Augsburg University,
Universitätsstr. 6a, D-86159 Augsburg, Germany
endrass@hcm-lab.de

² Logistics, Decision and Information Sciences,
Hollandseweg 1, 6706 KN Wageningen, The Netherlands

³ Man-Machine Interaction group, Delft University of Technology, The Netherlands

⁴ The University of Sunderland
City Campus, Chester Road, Sunderland, UK

⁵ INESC-ID, Av. Prof. Cavaco Silva,
2780-990 Porto Salvo, TagusPark, Portugal

⁶ Mathematical and Computer Sciences (MACS), Heriot-Watt University,
EH14 4AS, Edinburgh, UK

⁷ School of Humanities and Social Sciences, Jacobs University Bremen gGmbH,
Campus Ring 1, 28759 Bremen, Germany

Abstract. The integration of culture into the behavioral models of virtual characters requires knowledge from very different disciplines such as sociology and computer science. If culture-related behavioral differences are integrated into a virtual character system, users do not necessarily understand the intent of such a system. In this paper, we present a prototype that tries to integrate the masculinity dimension of culture with prototypical differences in verbal and nonverbal behavior. In a preliminary evaluation study, we investigated how these differences are judged by human observers with different cultural backgrounds.

1 Motivation

Integrating culture as a social factor into behavioral models of virtual characters can serve several purposes: to use virtual characters as training partners to increase cultural awareness, to enhance characters' naturalness by adding a wider variability, to study the user's perception of the characters' behavior or to investigate different interaction modalities.

This integration has been an increasing research area in recent years, while for the simulation of cultural background different aspects of behavior have been investigated. While some systems focus on verbal behavior [1], [2], other systems focus on integrating differences in nonverbal behavior (such as facial expressions, gesture selection, expressivity, spatial behavior or gaze) [3], [4], [5]. Other systems

integrated culture by a variety of behavioral routines such as rituals or different politeness or negotiation strategies [6], [7], [8].

While some of these systems used existing national cultures or cultural groups as a basis for building cultural models for their virtual characters [1], [2], [7], other systems used fantasy cultures to explain the concept without the risk of offending members of an existing culture [6], [9].

For the integration of culture, different approaches might be taken. While some systems start from a broad literature survey to extract culture-specific behaviors [6], [3], other systems are based on empirical data where behavioral tendencies were observed [5].

Whatever approach is taken to integrate culture into the behavior of virtual characters, they have one thing in common: culture is hard to grasp and difficult to implement in a computer-based system. It is thus a time and resource consuming task. In addition, the user does not necessarily need to understand the intent of the designer of such a system and thus cultural differences might even stay unnoticed.

For these reasons, we decided to build a prototype that provides us with a deeper insight on how to simulate different cultures in virtual characters's behavior and on how human observers judge these differences. We therefore focus on only one possible distinction of cultures and build examples that show prototypical differences in behavior. In an evaluation study, we investigated how the virtual characters' behavior is judged by users from different cultural backgrounds.

2 Cultural Dimensions and Synthetic Cultures

Although culture is seen as an abstract concept by many researchers, there are approaches that define culture in a dimensional manner and thus try to categorize groups. A well established model is described by Hofstede and colleagues [10]. This model has already been used with success in a number of computer-based applications featuring virtual characters [9], [6], [5]. At the moment, Hofstede's model exists out of the following six cultural dimensions:

- *Power Distance* (PDI) describes the extent to which a different distribution of power is accepted by less powerful members of a group. Scoring high on this dimension indicates a high level of inequality of power and wealth within a society. A low score means that there is greater equality between social levels, including government, organizations, and families.
- *Individualism* (IDV) describes the degree to which individuals are integrated into a group. On the individualist side, ties between individuals are loose, and everybody is expected to take care of him or herself. On the collectivist side, people are integrated into strong, cohesive groups.
- *Masculinity* (MAS) describes the distribution of roles between males and females. The two extreme sides are 'masculinity', which has values that focus on being assertive or competitive, and 'femininity', which has values that focus on caring and modesty.

- *Uncertainty Avoidance* (UAI) explains a society’s tolerance for uncertainty and ambiguity; the extent to which a member of the culture feels uncomfortable or comfortable in an unknown situation.
- *Long Term Orientation* (LTO) has the two sides: ‘Long Term Orientation’, which is associated with thrift and perseverance, and its opposite ‘Short Term Orientation’, which is about showing respect for tradition, fulfilling social obligations, and protecting one’s face.
- *Indulgence* (IVR) describes the subjective well-being that members of a culture experience. Cultures that score high on this dimension have a high percentage of people that consider themselves as very happy and having high control over their lives. Restrained societies on the other hand are more pessimistic with duty being important.

In total over 70 countries were categorized using this model by assigning a value on each dimension. In [11], Hofstede and colleagues introduce so-called synthetic cultures, which aim to describe the influence of the dimensions of culture on behavior. At the moment there are twelve synthetic culture, that describe the extreme sides of each of the dimensions of culture. The descriptions can be used for cross-cultural learning environments to create extreme examples of the culture-clashes that can occur in real life. They also serve as a good basis of virtual characters’ behavior, since clear behavioral trends are provided for synthetic cultures and no ethical issues have to be considered (as might be the case for the implementation of existing cultures).

3 Prototype

Within our prototype, we focus on only one of the cultural dimensions described in section 2: the Masculinity dimension. It is important to distinguish between prototypical behavior for extreme masculine and feminine cultures. The following ideas and descriptions are taken from Hofstede and colleagues [11].

The authors of [11] defined a profile for each synthetic culture that contains the culture’s values, core distinctions, key elements as well as words with positive or negative connotations.

The extreme masculine synthetic culture has as core value “winning” and as core distinction the distinction between “men and women”. Key elements are statements such as “Material success and progress are dominant values.”, “Bigger and faster are better.” or “Failing (at school, at work, in sports, or wherever) is a disaster”. These key-elements are rules for appropriate behavior in this synthetic culture and explain the way in which members of that culture are thinking. Stereotypical behavior is defined for synthetic cultures as well. Extreme masculine cultures are described as loud and verbal, with a tendency to argue with others. Nonverbally, members of the extreme masculine culture like physical contact, direct eye contact and animated gestures. Words with a positive connotation are: career, competition, fight, aggressive, success, winner, force, fast, big, power or action.

The extreme feminine synthetic culture, which is located on the opposite side of the same dimension, has as core value “caring for others, especially the

weak” and as core distinction the distinction between “caring/needing care”. Key elements are statements such as “Small and slow are beautiful.”, “Everybody is supposed to be modest, soft-spoken and empathetic - men and women alike.” or “Conflicts are resolved through compromise and negotiation”. Typically, members of the extreme feminine culture do not raise their voice and like small talk and agreement. Nonverbally they do not take much space and are warm and friendly in conversation. Words with a positive connotation are: caring, solidarity, modesty, compromise, help, love, soft, slow, tender or touch.

3.1 Cultural Scripts

Based on the above descriptions of masculine and feminine cultures, we designed a scenario that on the one hand demonstrates the differences between masculine and feminine cultures and on the other hand leads to a conflict when members of opposite cultures meet. As a showcase we chose a conversation between a professor and a student in which the student asks for a deadline extension. Each character within the scenario either has a prototypical masculine or feminine behavior. Thus, four different scenarios were scripted (see table 1).

There are two major differences in the four scenarios: the reason for needing an extension, and whether or not the professor agrees to give the student an extension. The agent with the masculine script focuses on ‘performance’, and thus need an extension to improve the assignment. In contrast, the agent with the feminine script focuses on ‘caring’, and thus needs an extension because of attending an important family event before the deadline. A masculine professor would accept the way of argumentation of a masculine student and agree to an extension. The feminine way of argumentation, vice versa, is considered as a weak excuse and leads to a conflict with no extension of the deadline. A professor from a feminine culture would understand the way of argumentation of a feminine student and extend the deadline. The conversation with a masculine student would not necessarily lead to an outwardly visible conflict but to a misunderstanding at least, since the feminine professor would not understand the student’s need to improve the work.

Besides differences in the flow of the scenario, members of masculine and feminine cultures also differ in their choice of wordings and nonverbal behaviors. While members of a masculine culture use direct speech and expressive gestures, members of a feminine culture use more moderate words and soft nonverbal behavior. The integration of these culture-specific differences was based on our previous findings on the integration of cultural backgrounds to the behavior of virtual characters [2] [5].

Table 1. Different cultural constellations for the evaluation study.

| | feminine student | masculine student |
|---------------------|------------------|-------------------|
| masculine professor | scenario 1 | scenario 2 |
| feminine professor | scenario 3 | scenario 4 |

3.2 Virtual Scenario

To implement the scenarios described above, we used the Virtual Beergarden scenario [12]. In the application, an arbitrary number of agents can be loaded that are able to move around in the scenario freely, exhibit gestures and communicate with each other. Verbal behavior is realized by a text-to-speech component. In the current version of the system male and female voices and about 40 different gestures and postures are provided. Animations can be customized in order to match different levels of expressivity.

For the two characters participating in our scenarios, initial values such as positioning in the location, orientation towards each other and voices for the text-to-speech component are set. In order to avoid side effects evoked by the gender of the characters, we left the genders of the virtual characters constant. Thus, only the gender-dimension of the characters' *culture* varied, but not the gender of the characters itself. Figure 1 shows the virtual scenario including our professor-student setup with a female student and a male professor (gender), showing prototypical feminine or masculine nonverbal behavior (culture). To model the virtual characters' behavior, we used the SceneMaker tool [13], while each conversation lasted for about half a minute (between 23 and 32 seconds) and contained from 6 to 10 dialog turns.



Fig. 1. Professor-student scenario with prototypical feminine (left) and masculine (right) behavior for both characters.

4 Evaluation

In an online evaluation study, we investigated how the culture-related differences that we integrated in our scenario are perceived by human observers during agent interaction in different countries.

4.1 Study Setup

For our evaluation study four videos were recorded showing the professor-student scenario in all possible combinations (see table 1).

After answering demographical questions, such as age, gender, and nationality, participants were able to view the videos, while they were given the opportunity to watch the videos several times. Afterwards, participants had to answer the following questions:

- Do you think the student acted appropriately?
- Do you think the professor reacted appropriately?
- Do you think the professor’s decision was fair towards the student who asked for the extension?
- Do you think the professor’s decision was fair towards the other students who did not ask for an extension?
- Would you like to have this professor as a teacher?
- Would you like to have this student as a friend?

Participants were able to rate on a 7 graded Likert scale, rating their agreement with ”yes, absolutely”, ”yes”, ”somewhat yes”, ”neither yes or no”, ”somewhat no”, ”no” or ”no, not at all”. In addition, a comment box was provided where we asked participants to state additional comments about their choices.

For further information on the study setup, introduction, dialogs and videos, please visit the online study ⁸. With it, we wanted to find out how participants judge the virtual characters’ behavior that reflects different cultural backgrounds.

4.2 Results

In total, 69 participants with nine different nationalities took part in our study. Since we only collected enough data from four countries for statistical analysis, we only considered these four in our tests. In that manner, 14 people from Germany (six females), 7 people from Japan (two females), 19 people from the Netherlands (seven females) and 20 people from Thailand (eleven females) were included for analysis, while 9 participants from 5 different countries were excluded from the analysis. The scores for the four participating cultures on Hofstede’s dimensions are provided in table 2.

Table 2. Hofstede’s scores for the four cultures that participated in our study [10].

| Dimension | Germany | Netherlands | Japan | Thailand |
|-----------|---------|-------------|-------|----------|
| PDI | 35 | 38 | 54 | 64 |
| IDV | 67 | 80 | 46 | 20 |
| MAS | 66 | 14 | 95 | 34 |
| UAI | 65 | 53 | 92 | 64 |
| LTO | 31 | 44 | 80 | 56 |
| IVR | 40 | 68 | 42 | 45 |

Our analysis indicates that members of the different cultures did not perceive the scenarios differently with regard to the different ways of behavior. However, we have gained some interesting insights into how people perceive virtual agent scenarios across cultures.

Having a look at the participants’ judgment of fairness towards other students, interestingly all 4 nationalities agreed that only in scenario 1 (masculine

⁸ <http://mm-werkstatt.informatik.uni-augsburg.de/survey/index.php?sid=21954&lang=en>

professor, feminine student), the teacher was perceived as being fair towards his other students. Please note, that due to our scenario construction, this scenario was the only one where a visibly outward conflict unfolded and the deadline extension was not granted, while the other three scenarios where the extension was granted were perceived as being not fair towards other students by all cultures.

Table 3 shows the mean values for the four cultures for this question for all four scenarios as well as the average mean value across the four cultures.

The results from table 3 suggest that the actual outcome of the scenario, e.g. whether a deadline is granted or not, influences the overall perception of human observers more than the behavior of the agents. This perception seems to be similar across cultures, even though the cultures still respond differently.

Table 3. Mean values on the perception of fairness towards others students.

| fairness towards others | extension granted | cross-cultures | Germany n=14 | Netherlands n=19 | Japan n=7 | Thailand n=20 |
|--------------------------------|--------------------------|-----------------------|-------------------------|-----------------------------|----------------------|--------------------------|
| scenario 1 | no | 5.96 | 6 | 6.26 | 5.42 | 6.05 |
| scenario 2 | yes | 3.59 | 3.64 | 3.11 | 3.71 | 3.85 |
| scenario 3 | yes | 2.74 | 2.14 | 2.67 | 3.43 | 3.45 |
| scenario 4 | yes | 3.07 | 2.86 | 3.05 | 3.71 | 2.95 |

In addition, we found a difference in the perception of the appropriateness of the student's behavior between individualistic and collectivistic cultures. According to Hofstede's dimensions, Germany and the Netherlands count as individualistic cultures, while Japan and Thailand are collectivistic cultures (see table 4 for the scores on the Individualism (IDV) dimension). Mean values for the rated appropriateness in the four cultures and the individualistic and collectivistic groups, as well as p-values using the two-sided t-test comparing individualistic and collectivistic cultures are provided in table 4.

For three out of the four scenarios, the individualistic and collectivistic cultures differ significantly in their perception on the student's appropriateness. Participants from the individualistic group found the masculine student's behavior significantly more appropriate (with either a masculine or feminine professor), while participants from the collectivistic group found the feminine student's behavior significantly more appropriate (see table 4). Again, the only exclusion constitutes scenario 1 (feminine student and a masculine professor), where the deadline was not granted.

Another interesting result was found in scenario 3 (feminine professor and feminine student), in which the deadline is granted because of an important family event (marriage of the student's sister). Please note, that this scenario is the only one where no cultural clash (scenarios 1 and 4) and no conflict (scenario 2) occurred between student and professor. In this scenario, cultures differ significantly regarding the teacher's reaction comparing the individualistic and collectivistic group. Members of collectivistic cultures judged the professor's reaction as significantly more appropriate than participants of individualistic cultures (see table 4).

Table 4. Mean values on the perception of appropriateness of behavior with p-values comparing individualistic and collectivistic cultures.

| appropriateness | Germany IDV=67 | Netherlands IDV=80 | Indiv. | Japan IDV=46 | Thailand IDV=20 | Collec. | p |
|-----------------|-------------------|-----------------------|-------------|-----------------|--------------------|-------------|--------|
| student | | | | | | | |
| scenario 1 | 4.29 | 3.78 | 4 | 4.14 | 3.95 | 4 | 1 |
| scenario 2 | 4.64 | 4.95 | 5.24 | 4.29 | 4.35 | 4.33 | 0.02 |
| scenario 3 | 3.79 | 4.53 | 4.21 | 5.43 | 5.1 | 5.19 | 0.01 |
| scenario 4 | 4.57 | 5.11 | 4.88 | 4.14 | 3.95 | 4 | 0.03 |
| teacher | | | | | | | |
| scenario 3 | 3.36 | 3.37 | 3.36 | 5.29 | 4.75 | 4.89 | 0.0002 |

A look at the participants’ comments indicates that members of different cultures perceived the scenarios in a similar manner as we expected. Some participants from masculine cultures (Germany and Japan) stated that the masculine teacher “acted according to the rules” and is “fair”, while participants from feminine cultures (Netherlands and Thailand) stated that the masculine teacher “acted rude” and “judge too soon”. In comparison, the feminine professor was judged “kind” or a “nice man” by members of feminine cultures, while he was judged “too soft” or “not fair” from members of masculine cultures.

5 Conclusions

In this paper, we considered culture-related differences in behavior to build different virtual character scenarios that should resemble prototypical feminine and masculine cultures. Showing these scenarios to human observers from different cultural backgrounds, we investigated the impact on the perception of the observers. Results indicate that participants did judge the scenarios differently, but not necessarily due to the behavioral style of the characters. In particular, the outcome of the scenario seemed to influence the judgment of the observers more than the behavior itself.

We found some interesting insights distinguishing participants from individualistic and collectivistic cultures, who judged the virtual character’s appropriateness of behavior significantly different in some scenarios. In addition, a qualitative analysis of the participant’s comments was promising in terms of the judgment of the agents’ behavior.

We think that other dimensions of culture could have influenced the user’s perceptions. On the one hand, other dimensions of the cultural background of the *human participants* could have influenced their judgment. On the other hand, it is hard to exclude the other dimensions of culture from the *virtual characters’* behavior. The fact that we had a conversation between a student and a professor could, for example, have indicated power distance, as could the fact that student and teacher were of opposite gender.

According to our results, the scenario taps into the Individualism dimension, although differences in the masculinity dimension were intended. An explanation for this could be that the reason for the extension for the feminine student was

a family event, which is an in-group event that collectivistic respondents from Thailand and Japan could have found legitimate, whereas German and Dutch respondents could have assumed equal rights for all students with no regard for family considerations. Thus, as a future work, we intend to repeat the same study including a different reason for the feminine student, such as a medical reason.

Acknowledgments. This work was funded by the European Commission under grant agreement eCute (FP7-ICT-2009-5).

References

1. Johnson, W.L., Valente, A.: Tactical Language and Culture Training Systems: Using Artificial Intelligence to Teach Foreign Languages and Cultures. In: Innovative Applications of Artificial Intelligence (IAAI 2008), AAAI (2008) 1632–1639
2. Endrass, B., Rehm, M., André, E.: Planning Small Talk Behavior with Cultural Influences for Multiagent Systems. *Computer Speech and Language* **25**(2) (2011) 158–174
3. Jan, D., Herrera, D., Martinovski, B., Novick, D., Traum, D.: A Computational Model of Culture-Specific Conversational Behavior. In C. Pelachaud et al., ed.: *Intelligent Virtual Agents 2007*, Springer (2007) 45–56
4. Koda, T., Ruttkey, Z., Nakagawa, Y., Tabuchi, K.: Cross-cultural study on facial regions as cues to recognize emotions of virtual agents. In Ishida, T., ed.: *Culture and Computing*, Springer (2010) 16–27
5. Endrass, B., Rehm, M., Lipi, A.A., Nakano, Y., André, E.: Culture-related Differences in Aspects of Behavior for Virtual Characters across Germany and Japan. In: *Proceedings of AAMAS 2011*. (2011)
6. Mascarenhas, S., Dias, J., Afonso, N., Enz, S., Paiva, A.: Using rituals to express cultural differences in synthetic characters. In Decker, Sichman, Sierra, Castelfranchi, eds.: *Proceedings of AAMAS 2009*. (2009)
7. Wu, P., Miller, C.: Interactive phrasebook: Conveying culture through etiquette. In Blanchard, E.G., Johnson, W.L., Ogan, A., Allard, D., eds.: *3rd International Workshop on Culturally-Aware Tutoring Systems (CATS2010) held on ITS2010*, Pittsburg, USA (2010) 47–55
8. Kim, J., Hill, R.W., Durlach, P., Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D., Hart, J.: BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education* **19** (2009) 289–308 IOS Press.
9. Aylett, R., Paiva, A., Vannini, N., Enz, S., André, E., Hall, L.: But that was in another country: agents and intercultural empathy. In Decker, Sichman, Sierra, Castelfranchi, eds.: *Proceedings of AAMAS 2009*. (2009)
10. Hofstede, G., Hofstede, G.J., Minkov, M.: *Cultures and Organisations. SOFTWARE OF THE MIND. Intercultural Cooperation and its Importance for Survival*. McGraw Hill (2010)
11. Hofstede, G.J., Pedersen, P.B., Hofstede, G.: *Exploring Culture - Exercises, Stories and Synthetic Cultures*. Intercultural Press, Yarmouth, United States (2002)
12. Augsburg University: <http://mm-werkstatt.informatik.uni-augsburg.de/projects/aaa> (2011)
13. Mehlmann, G.U.: *SceneMaker: An Interpreter for Parallel Processes Modeling the Behavior of Virtual Characters*. Master's thesis, German Research Center for Artificial Intelligence (2009)