A Cross-Cultural Study on the Perception of Sociability within Human-Computer Interaction

Fang-Wu Tung¹, Keiichi Sato², Yi-Shin Deng³, and Tsai-Yi Lin⁴

¹ Dept of Industrial Design, National United University, 1, Lienda, Miaoli, 36003 Taiwan
² Institute of Design, Illinois Institute of Technology,
350 North La Salle St. 4th Floor, Chicago, Illinois 60654 USA
³ Institute of Applied Arts, National Chiao Tung University
1001 University Road, Hsinchu, 30080 Taiwan
⁴ Department of Fashion Imaging, MingDao University
369, Wen Hwa Road, Pee-Tow, Chang-Hwa, 52345 Taiwan
fwtung@gmail.com, sato@id.iit.edu, ydest@faculty.nctu.edu.tw,
lintsaiyi99@yahoo.com.tw

Abstract. This study tries to use speech and dynamic emoticons as social cues to create a more sociable human-computer interaction. A cross-cultural study was conducted to investigate the influence of cultural backgrounds (Taiwan and America) on children's perceptions of sociability within human-computer interaction and explore how the management of social cues affects their engagement in e-learning environments. A 2x2 (Taiwan/America, speech/dynamic emoticon) quasi experiment was conducted to investigate the effects of the independent variables on children's perception of social presence and intrinsic motivation. Cultural differences in the perception of social presence are observed. American children reported higher perceived social presence than Taiwanese children did. No differences of effects of speech and dynamic emoticons on children's feelings of social presence and motivation are found. It suggests that children's social responses and learning motivations are triggered equally strongly by the two social cues. These findings suggest that designers of educational technology could use speech or dynamic emoticons to build more sociable interfaces that could boost children's motivation in learning.

Keywords: Cultural difference, Sociability, Interaction design, Speech, Dynamic Emoticon, Children.

1 Introduction

1.1 The Social-Emotional Interaction between Human and Computer

People are emotional and social. The role of human emotions is an influential factor in the way people deal with and relate to objects and artifacts [22]. Given this, human emotional needs and social desires have to be taken into consideration in developing products for people. Research also reveals that human-computer interaction is on both a social and emotional level. In light of the findings, new theories that enable or augment socio-emotional interaction between people and computers are explored [18, 24].

Indeed, technologists have aspired to make computer interfaces more human-like and sociable because it is suggested that more humanized interfaces convey a sense of comfort and ease to the user [15, 27]. Besides the use of sophisticated computing technology or artificial intelligence, utilizing social cues in user interface may offer an uncomplicated and inexpensive way to achieve the goal of humanized interface.

Moving beyond theory, several experimental studies have demonstrated that people do not respond to a computer merely as a tool. Instead, a wide range of social rules and learned behaviors lead humans to interact with and attitudes toward computers [23, 26]. The finding that people appear to have social relationships with computers raised attention for its potential to promote the interaction between humans and computers. Computers could act as social partners of humans. From this perspective, Nass et al. [20] have empirically proven that people socially interact with computers and claim that "Computers are social actors (CASA)". CASA claims that computers that exhibit social cues can convey a sense of sociability and intimacy and thereby induce social responses from people, which lead people to treat computers in the way as they treat other people. Just as mentioned by Norman [22], people have evolved to interpret even the most subtle of indicators and are predisposed to anthropomorphize, and so to project human emotions into everything. Thus, anthropomorphic responses can bring great delight and pleasure to the user of a product due to human tendency to interpret and anthropomorphize things.

1.2 Enhancing Social Presence in Computer-Mediated Learning Environments

CASA paradigm is based on a concept of social presence, and involves the social responses of people not to other entities within a medium, but to cues provided by the medium itself [16]. The degree of social presence supported by the media used in learning environments can also affect learning. Gunawardena [6] agues that social presence is necessary to enhance and improve effective instruction in both traditional and technology-based classroom. Studies have suggested that enhancing social presence in an e-learning environment can instill the learner with an impression of a quality learning experience [21, 29, 34]. As claimed by CASA paradigm, computers are social actors. It suggests that a learner can perceive the social presence created by the computer itself. Thus, computers may be perceived on social dimension to improve motivation while a single learner participates in a computer learning activity with no instructor involved. Such experiences may enable an individual to perceive that another social being exists and is interacting with them.

Studies have tried to study the affective and emotional dimension of people's interactions with computers in the hope of proposing a new direction for educational software design [1, 177, 30, 31]. The results suggest that designers may consider developing products that deliberately elicit social responses from learners. Employing social cues derived from social psychology and sociology to interface design can enable users to sense a high degree of social presence from the medium itself and a sense of intimacy accordingly. Given this, human-computer interactions can be improved by incorporating a set of social cues into user interface design by replicating human-human interaction. The current study attempts to use two important social cues in interpersonal communication, speech and facial expressions, as social cues in e-learning environments and examine the effects of the two social cues on children's perceived social presence and learning motivation.

1.3 Speech vs. Dynamic Emoticons

Just as the brain processes voices differently than all other sounds, it processes faces differently than all other objects [19]. To compare the two social cues, dynamic emoticons and speech, involves further issues of nonverbal communication versus verbal communication as well as visual modality versus audio modality. Posner et al. [255] argued that visual stimuli are less likely to automatically engage attention than auditory stimuli, and people have to learn to direct their attention to visual information. Compared with texts and symbols in the visual channel, speech in the auditory channel is less cognitively demanding [73]. Besides, speech is a dynamic process which can trigger people to focus on the content and keep the audience in awareness during its creation [1], leading to exerting a high degree of social presence.

Judging from above, speech seems to be superior to emoticons. The sound of speech is effective in gaining users' attention; nevertheless, Zaidel and Mehrabian [36] claimed that faces are more important than voices in interpersonal interaction. As mentioned above, dynamic emoticons being visual cues may not draw as much attention as speech in terms of sensory aspects. However, the dynamic emoticon as a representation of the facial expression which is most powerful channel of nonverbal communication and expresses more explicit emotions fitting in the situation than speech does. Speech may transmit emotional messages by changing pitch, tempo as well as loudness, but research revealed information about emotions transmitted by the facial expression is more precise than speech [3]. Burns and Beier [1] also pointed out that visual cues are also more influential and accurate than vocal cues in the designation of a mood state. Moreover, Ekman [3] indicated that happiness is an easily recognizable facial expression, while it is more difficult to identify in the voice. These strengths may enable dynamic emoticons to receive more attention and as such yield higher social presence. Thus, it is interesting to investigate whether dynamic emoticons used in computer interfaces have the same effect on children's feeling of social presence as speech does. The two social cues, speech and dynamic emoticons, used to enhance the level of social presence in e-learning environments are derived from interpersonal communication. As communication is involved in the issue, it is inevitable to consider the factor of the cultural differences because the communication process is different in individualistic and collectivistic cultures or in terms of low-context and high-context communication.

1.4 Cultural Differences in the Perception of Social Cues

To understand cultural differences, Hofstede's International cultural dimensions and Hall's high-context and low-context culture provide useful frameworks to approach the issue [7, 9]. Hofstede proposed four international dimensions including power distance, uncertainty avoidance, Individualism, and Masculinity. Although the concept of culture is multidimensional proposed by Hofstede, individualism-collectivism has been empirically developed and is found to have the strongest variation across cultures.

The individualism-collectivism dimension is the preference to act as individuals rather than as a member of a group [1010, 28]. It refers to the tendency to focus on the needs of self as opposed to community and society. Cultures high on the individualism dimension promote individual identity while collectivistic cultures are more

group oriented. In terms of communication, collective cultures depend more on high context messages, while individualistic cultures depend ore on low context messages [11]. Thus, members in collectivistic culture value a more subtle form of communication. In other words, they would be more negatively affected by the less social cues when going from human-human interaction to human-computer interaction in comparison to individualistic ones. As a result, the study expects that children from Taiwan perceive a lower level of social presence in e-learning environments than children from the US do.

Another lens to view cultural differences in communication is Hall's concept of high-context and low-context communications. Hall [8] points out that a high-context individual as someone who engages feelings in a relationship. It suggests that low-context cultures don't differentiate as much as high-context cultures between in- and out-groups. Thus, high-context communication involves emotions and closer relationships while low-context communication relies on logical part of the brain and is less personal. Individuals in high-context culture learn to understand others through non-verbal responses while ones in low-context culture rely on verbal communication more. Also, research suggests that information in the verbal content of East Asian language is relatively low so nonverbal cues provide the missing link in the communication process [1212]. According to the difference in the usage of nonverbal communication, the study assumes that speech and dynamic-emoticons feedbacks have different impacts on the level of social presence perceived by children from different cultural backgrounds.

Concluded from above, the current study attempts to employ speech and dynamic emoticons as social cues in e-learning environments and to compare their effects on children's perceived social presence and motivation. Further, cultural differences in the research issue are examined by recruiting participants from Taiwan and America. Three questions are addressed in this paper. First, do the social cues of speech and dynamic emoticons yield the same impacts on children's perceived social presence and learning motivation? Second, do different cultural backgrounds influence perceived social presence and learning motivation? Third, what is the relationship between perceived social presence and learning motivation?

2 Hypotheses and Research Questions

Concluded from the individualism-collectivism dimension and low-context and high-context communication in terms of cultural differences, it leads to the following hypotheses:

- H1. The effects of speech and dynamic-emoticon feedbacks on Taiwanese and American participants' perceived social presence are different.
- H2. American participants perceived higher level of social presence in e-learning environments than Taiwanese participants do.

The effects of speech and emoticons on children's motivation in e-learning environments are examined and the correlation between perceived social presence and their motivation is investigated in the study. Three research questions are as follow.

R1. Do the two social cues, speech and dynamic emoticons, have same impacts on children's motivation in e-learning environments?

- R2. Do different cultural backgrounds influence children's motivation differently while participating in e-learning activities?
- R3. Does the social presence perceived from the medium itself correlate with learners' motivation?

3 Method

Participants. Seventy-seven subjects participated at the study. Forty 5th and 6th graders (21 girls and 24 boys) were recruited from an elementary school in Hsinchu, Taiwan. Thirty-two 5th and 6th graders (15 girls and 17 boys) were recruited from 2 elementary schools in Chicago, America.

Experiment design. A mixed factorial design of experiment was employed to investigate the effect of two factors. The first factor, cultural background, is a between-subject factor of Taiwan and America. The second factor, social cue, is a within-subject factor of speech or dynamic emoticons. It is a two by two mixed factorial design. The dependent variables were the children's perceived social presence and their intrinsic motivation.

Experiment material. The instructional material was prepared as a math problem-solving practice program designed in Macromedia Flash. The instructional program had seven math problems. The degree of level of the questions was discussed with teachers of the selected schools. The feedbacks provided by the program including presenting a text-based greeting at the beginning, showing the "right" symbol as feedback and automatically guide the participant to the next question when a participant answers correctly, or showing the "wrong" symbol and text-based message which suggested the participant to press "again" or "next" icon for proceeding when a participant fails to answer correctly, presenting the counts of correct and positive comments after all questions have been attempted.

The program then developed into two versions to present the two manipulated conditions by providing speech outputs and dynamic emoticons in the above-mentioned feedbacks. The emotion of speech and facial expression suitable for each feedback were determined by the focus group, which was formed by two teachers and four students. Speech used in the instructional program was created from recordings made by two young ladies; one is from Taiwan and the other one is from America. Dynamic emoticons start with a neutral face, which then continuously evolves into full expression. The speed of the dynamic display is 12 frames per second.

Measurement tools. The dependent variables of social presence and intrinsic motivation were measured using a set of paper-and-pencil questionnaires.

- 1. The first set questions adopted the four items proposed by Short, Williams, and Christie (1976) to measure social presence, sociable/unsociable, personal/impersonal, sensitive/insensitive, and warm/cold, and applied a semantic differential technique. (Cronbach's $\alpha = 0.86$)
- 2. The second set of questions was adapted from the Activity-Feeling Scales (AFS) developed by Reeve and Sickenius (1994) and used to measure the subjects' intrinsic motivation. The 12-item measure made up of separate 3-item scales to assess

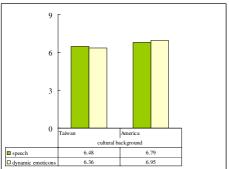
self-determination, competence, relatedness, and tension. The name and individual items for each index are as follows: self-determination-- offered choice what to do, I want to answer the questions, and my participation is voluntary; competence-- capable, competent, and achieving; relatedness-- intimate, involved with friends, part of a team; Tension-- pressured, uptight, and easy (reverse). (Cronbach's $\alpha = 0.74$)

Procedure. The study took place in the computer labs of the selected schools during one of their computer class session or recess. Each participant was assigned to one computer in the lab. The two conditions of speech and dynamic emoticon are counterbalance. The experiment took around 100 minutes. Upon completing the experiment, the subjects were debriefed and thanked with a toy.

4 Results

Social presence. Fig 1. shows the results of the perceived social presence. The analysis of variance of within-subjects effects indicated that the effect of speech and dynamic emoticons on perceived social presence is not significant (F(1,75)=0.05, p=0.83, 2 tails). No significant interaction effect is found (F(1,75)=2.01, p=0.16, 2 tails). The H1 is not supported. The American participants reported stronger feelings of social presence 6.85 (sd 1.47) than Taiwanese participants did 6.42 (sd 1.09). The analysis of variance of between-subjects effects indicates that the American participants perceived significant higher social presence than Taiwanese participants did (F(1,75)=2.89, p<0.05, 1 tail). The H2 is supported.

Intrinsic motivation. Fig 2. shows the results of the intrinsic motivation. The analysis of variance of within-subject effects indicated that the effect of social cues on intrinsic motivation is not significant (F(1, 75)=1.13, p=0.26, 2 tails). No significant interaction effect is found (F(1, 75)=0.88, p=0.35, 2 tails). The analysis of variance of between-subject effects indicates that the effect of cultural backgrounds on intrinsic motivation is not significant (F(1, 75)=0.50, p=0.48, 2 tails). No cultural differences are found in participants' intrinsic motivation. Also, it is observed that no significant difference in effects of social cues on intrinsic motivation.



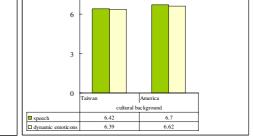


Fig. 1. Results of perceived social presence

Fig. 2. Results of intrinsic motivation

Correlation between social presence and intrinsic motivation. The mean of American participants' social presence and intrinsic motivation is 6.87 (SD. 1.20) and 6.66 (SD. 1.32), respectively. A highly positive correlation (Pearson r = 0.75, p < 0.001) is found between the two variables. The mean of Taiwanese participants' social presence and intrinsic motivation is 6.42 (SD. 1.09) and 6.40 (SD. 1.28), respectively. A positive correlation (Pearson r = 0.69, p < 0.001) is also found between the two variables. The results indicate that feelings of social presence from a computer itself can mediate children's intrinsic motivation.

5 Discussion

The cultural differences in the perception of social presence are observed in the study. The results support that American children perceived a higher degree of social presence than Taiwanese children did during interaction with computers. As mentioned above, collectivistic cultures depend more on high context messages while individualistic cultures depend more on low context messages. Given this, members in individualistic cultures would still feel comfortable with less social cues in human-computer context compared to the real context. As predicted in H2, the finding shows that American children feel stronger social presence than Taiwanese children do. Wegerif [35] has proposed the concept of social threshold. The social dimension of learning networks depended upon the extent to which learners were able to cross a threshold from feeling like outsiders to feeling like insiders. It means that a leaner would feel computer on social dimension while crossing the threshold. Regarding the concept of social threshold, it has been observed that members in low-context cultures like American don't differentiate as much as high-context cultures between in- and out-groups. The current study echoes the social threshold idea proposed by Wegerif.

No cultural differences in children's intrinsic motivation in e-learning environments are found in the study. In other words, cultural backgrounds do not affect children's motivation while participating in e-learning activities. That children both from America and Taiwan are exposed to various digital media and educational technology nowadays as computers are wildly used may account for the results.

As for the effects of speech and dynamic emoticons, the results do not support the expectation that effects of social cues on Taiwanese and American participants' feelings of social presence were different. The nonverbal usage in communication is different between in low-context and high-context cultures. The findings, however, are not consistent with cultural differences in low- vs. high-context communication. It indicates that speech and dynamic emoticons could yield the same level social presence for children regardless of cultural backgrounds. Besides, the effects of speech and dynamic-emoticon feedbacks on children's motivation in e-learning environments are not different either. Therefore, the e-learning environments with the application of speech or dynamic emoticons could equally motivate children to practice learning activities.

The development of a sociable interface has never meant the direct embedding of social cues without prudential consideration. Rather, it is required to investigate users' attitudes and perceptions more closely. Comparing speech with dynamic-emoticon feedbacks, the former is a channel of verbal communication and an audio modality while the latter is one of nonverbal communication and a visual modality. The research provides evidence that dynamic emoticons and speech almost have the same impacts on

inducing social responses from children and engaging them in e-learning environments. The results may provide designers with research based guidelines in consideration of speech or emoticon techniques to design sociable interactions for children.

The study supports that the social presence perceived from the medium itself positively correlate with learners' motivation. It suggests that the social presence provided by a computer itself also can improve children's motivation in computerized learning with no instructor or other colleagues involved. Some research in educational technology has considered computers as being neutral cognitive tools and has emphasized the cognitive and information processing aspects of learning [14]. Turkle[32] described the computer as Rorschach to present the computer as a relatively neutral screen onto which people were able to express their thoughts and feelings. Nevertheless, Turkle [333] indicates that the computational object is no longer effectively neutral. People perceive computers on a social dimension and attribute personality and emotion to them. Thus, the findings suggest that designers of educational technology may move beyond an emphasis on merely cognitive aspects of learning with computers and pay attention to the effects of social traits of computers.

6 Conclusion

People appear to form affective attachments to and affective relationships with objects. Our anthropomorphic perceptions may reflect the social nature of humans. Computers are responsive and interact with people in more engaging ways, with which people might expect to have a social relationship. From this perspective, this study aims at incorporating social cues adopted from human-human interaction into human-computer interaction to intensify the sense of social presence in child-computer interaction and hence motivate them in learning activities. With technology advancing rapidly, computer-mediated learning may come in a variety of applications such as educational products, systems, interactive toys, or even robots. Creating a social interface for those applications can help counterbalance the feeling of isolation when children interact with them individually. In a world that is becoming increasing globalized, a cross-cultural study helps us to obtain a comprehensive understanding of the issue and the findings may benefit design and development for related products.

Acknowledgments. This material is based upon work supported by the National Science Council of the Republic of China under grant NSC 95-2411-H-158-003 and 096-2917-I-009-005. We thank Yumao Feng for help developing the instructional programs. We thank David Groves, Gabrielle Parra, Hua-Min Chang, Jun-Zhi Chen, Nikki Krieger, and Yung-Yuan Liu for their assistance and helpful input in the study.

References

- 1. Burns, K., Beier, E.: Significance of vocal and visual channels in the decoding of emotional meaning. Journal of Communication 23, 118–130 (1973)
- Chalfonte, B.L., Fish, R.S., Kraut, R.E.: Expressive Richness: A Comparison of Speech and Text as Media for Revision. In: Robertson, S.P., Olson, G.M., Olson, J.S. (eds.) Proceedings of the ACM CHI 1991 Human Factors in Computing Systems Conference, New Orleans, Louisiana, pp. 21–26 (1991)

- 3. Ekman, P.: The nature of emotion: fundamental questions. Oxford University Press, New York (1994)
- 4. Ekman, P., Friesen, W.V.: Unmasking the face. A guide to recognizing emotions from facial clues. Prentice-Hall, Englewood Cliffs, New Jersey (1975)
- Ferdig, R., Mishra, P.: Emotional Responses to Computers: Experiences in Unfairness, Anger, and Spite. Journal of Educational Multimedia and Hypermedia 13(2), 143–161 (2004)
- 6. Gunawardena, C.N.: Social presence theory and implications for interaction and collaborative learning in computer conferences. International Journal of Educational Telecommunications 1(2/3), 147–166 (1995)
- 7. Hall, E.T.: Beyond Culture. Anchor Books Editions (1989)
- 8. Hall, E.T.: The power of hidden differences. In: Bennett, M. (ed.) Basic concepts in intercultural communication, pp. 53–67. Intercultural Press, Inc., Maine (1998)
- Hofstede, G.: Dimensions of National Cultures in Fifty Countries and Three Regions. In: Deregowski, J.B., Dziurawiec, S., Annis, R.C. (eds.) Expiscations in Cross-Cultural Psychology, Lisse NL, Swets and Zeitlinger, pp. 335–355 (1983)
- Hofstede, G.: Cultural Differences in Teaching and Learning. International Journal of Intercultural Relations 10(3), 301–320 (1986)
- 11. Hofstede, G.: Culture and Organizations: Software of the Mind. McGraw-Hill International, New York (1991)
- 12. Kitayama, S., Ishii, K.: Word and voice: Spontaneous attention to emotional utterances in two languages. Cognition & Emotion 16, 29–59 (2002)
- 13. Kroll, B.M.: Cognitive egocentrism and the problem of audience awareness in written discourse. Research in the Teaching of English 12, 269–281 (1978)
- 14. Lajoie, S.P., Derry, S.J.: Computers as cognitive tools. Erlbaum, Hillsdale, NJ (1993)
- Laurel, B.: Interface Agents: Metaphors with Character. In: Laurel, B. (ed.) The Art of Human-Computer Interface Design, pp. 355–365. Addison-Wesley, Reading (1990)
- Lombard, M., Ditton, T.B.: At the heart of it all: The concept of presence. Journal of Computer-Mediated-Communication 3(2) (1997), http://www.ascusc.org/jcmc/vol3/issue2/lombard.html
- 17. Mishra, P.: Affective Feedback from Computers and its Effect on Perceived Ability and Affect: A Test of the Computers as Social Actor Hypothesis. Journal of Educational Multimedia and Hypermedia 15(1), 107–131 (2006)
- 18. Muller, M.: Multiple paradigms in affective computing. Interacting with Computers 16(4), 759–768 (2004)
- 19. Nass, C., Brave, S.: Wired for speech: How voice activates and advances the human-computer relationship. MIT Press, Cambridge, MA (2005)
- Nass, C., Steuer, J., Tauber, E.: Computers are social actors. In: Proceedings of the SIG-CHI conference on Human factors in computing systems: celebrating interdependence, Boston, Massachusetts (1994)
- Newberry, B.: Raising student social presence in online classes. In: Fowler, W., Hasebrook, J. (eds.) Proceedings of WebNet 2001 World Conference on the WWW and the Internet, pp. 905–910. Association for the Advancement of Computing in Education, Orlando, FL, Norfolk, VA (2001)
- Norman, D.: Emotional Design: Why we love (or hate) everyday things. Basic Books, NY (2004)
- 23. Picard, R.: Affective Computing. The MIT Press, Cambridge, Saarni and Harris (1997)
- 24. Picard, R., Wexelblat, A., Nass, C.: Future interfaces: social and emotional. In: CHI 2002 extended abstracts on Human factors in computer systems, Minneapolis, Minnesota (2002)

- 25. Posner, M.I., Nissen, M.J., Klein, R.M.: Visual dominance: An information-processing account of its origins and significance. Psychological Review 83, 157–171 (1976)
- 26. Reeves, B., Nass, C.: Media equation: how people treat computer, television, and new media like real people and places. Cambridge University Press, New York (1996)
- 27. Sproull, L., Subramani, M., Kiesler, S., Walker, J.H., Waters, K.: When the interface is a face. Human-Computer Interaction 11, 97–124 (1996)
- 28. Triandis, H.C.: The self and behavior in different cultural contexts. Psychological Review 96, 506–520 (1989)
- 29. Tu, C.H.: How Chinese perceive social presence: An examination of interaction in an online learning environment. Educational Media International 38(1), 45–60 (2001)
- 30. Tung, F.W., Deng, Y.S.: Designing social presence in e-learning environments: Testing the effect of interactivity on children. Interactive learning environments 14(3), 251–264 (2006)
- 31. Tung, F.W., Deng, Y.S.: Increasing Social Presence of Social Actors in E-learning environments: Effects of Dynamic and Static Emoticons on Children. Displays 28(4-5), 174–180 (2007)
- 32. Turkle, S.: Computer as Rorschach. Society 17(2), 15–24 (1980)
- 33. Turkle, S.: Sociable Technologies: Enhancing Human Performance when the computer is not a tool but a companion. In: Roco, M.C., Bainbridge, W.S. (eds.) Converging Technologies for Improving Human Performance, pp. 151–158. Kluwer Academic Publishers, Netherlands (2003)
- 34. Walther, J.B.: Interpersonal effects in computer-mediated interaction: A relational perspective. Communication Research 19, 52–90 (1992)
- 35. Wegerif, R.: The social dimension of asynchronous learning networks. Journal of Asychronous Learning Networks 2(1), 34–49 (1998)
- 36. Zaidel, S.F., Mehrabian, A.: The ability to communicate and infer positive and negative attitudes facially and vocally. Journal of Experimental Research in Personality 3, 233–241 (1969)