The Cultural Influence Model: When Accented Natural Language Spoken by Virtual Characters Matters

Peter Khooshabeh*@, Morteza Dehghani*, Angela Nazarian*, Jonathan Gratch*

** Institute for Creative Technologies,*
University of Southern California,
1205 Waterfront Dr., Playa Vista, CA 90094, USA
Khooshabeh@ict.usc.edu
www.ict.usc.edu

@ Research Center for Virtual Environments in Behavior (ReCVEB)
Department of Psychological and Brain Sciences
University of California
Santa Barbara, CA 93106-9660, USA

Abstract:

Advances in Artificial Intelligence (AI) and computer graphics digital technologies have contributed to a relative increase of realism in virtual characters. Preserving virtual characters' communicative realism, in particular, joined the ranks of the improvements in natural language technology and animation algorithms. This paper focuses on culturally relevant paralinguistic cues in nonverbal communication. We model the effects of an English speaking digital character with different *accents* on human interactants (i.e., users). Our *cultural influence model* proposes that paralinguistic realism, in the form of accented speech, is effective in promoting culturally congruent cognition only when it is self-relevant to users. For example, a Chinese or Middle Eastern English accent may be perceived as foreign to individuals who do not share the same ethnic cultural background with members of those cultures. However, for individuals who are familiar and affiliate with those cultures (i.e., in-group members who are bicultural), accent not only serves as a motif of shared social identity, it also primes them to adopt culturally appropriate interpretive frames that influence their decision making.

Keywords: natural language, cultural priming, biculturals

For the general public, the term ‘digital virtual character’ typically evokes images of a crude, pixelated, robotic entity that falls short of the rich complexity of humans. Developments in representational software and AI have advanced virtual human characters from their cartoonish beginnings. Today, virtual characters are present in realms well beyond early video games and recreational simulations.

Recent virtual character developments introduce a world of possibility and unexplored research opportunities for facilitating cultural fluency, i.e., social simulations of ethnic groups from various parts of the world. Subsequently, developers of culturally motivated virtual characters (CMVCs) stand to benefit from learning about the essential features for designing effective characters and exploring their effects on users. The behavioral science literature lends support to the role of social cues in communication to influence learning and other cognitively complex behaviors.

Nass and colleagues have demonstrated that people communicate with computers as social actors, even when there is no visual representation of a graphically rendered virtual character (Nass & Brave, 2007; Nass & Yen, 2010; Reeves & Nass, 1996). In cases where there is a graphical representation of an animated virtual character, Blascovich defines communicative realism as a character’s fidelity to human-like movements, anthropometrics and photographic realism (Blascovich & McCall, 2013; Blascovich, In press, 2002). We extend this definition to also include auditory features of paralinguistic nonverbal communication, e.g., accents. Our work examines whether paralinguistic communication cues

elicit different social responses to culturally distinct virtual characters depending on a user's cultural background.

Keeping cross-cultural literacy and fluency in mind, there are potential benefits in designing CMVCs with accents resembling speech from the region of interest. Consider an example when a North American businessperson must negotiate with clients in China. In order to prepare for the anticipated culture differences and efficiently equip herself for the negotiation, the American partner could interact with a CMVC designed with Eastern cultural and economic customs in a digital virtual simulation. The fidelity of this character is magnified by increasing the communicative realism of the CMVC with a Chinese accent, which could increase the trainees' overall immersion in the experience. To the extent that such immersion occurs, trainees' negotiation skills for appropriately interacting with members of the culture represented by the virtual agent should improve. However, immersion is likely to be compromised if trainees simply disregard the accent as foreign. Subsequently, the user's negotiation skills and cross-cultural tolerance are likewise compromised. To address the challenges in these types of social exchanges, the work presented in this paper computationally models and experimentally evaluates cultural interactions with a focus on psychological outcomes (Brewer, 1991; Chen & Bond, 2007).

General Methodology

Here we report on two behavioral experiments that have manipulated auditory communicative characteristics of CMVCs while

maintaining an identical visual appearance[1]. Both studies recruited individuals from bicultural backgrounds, specifically, English speaking Iranian- and Chinese-Americans. The independent variable in the between-subjects experimental designs has been whether the CMVCs in the studies speak to users with either (1) a foreign English accent (e.g., Chinese, or Middle-Eastern English pronunciations) or (2) a native (California) English accent.

Benefits of using embodied CMVCs versus other media

Virtual agents—particularly when rendered as embodied conversational characters—are capable of providing a compelling

[1] This approach represents a significant contribution to the field of communication science. Communication scholars in the area of psycho-social-linguistics typically use Lambert's (1967) matched guise paradigm. In this paradigm, the same speaker (a bicultural research confederate) displays at least two accents and researchers study the effect on listeners. Our method utilizing advanced AI, natural language, and animation technology overcomes traditional problems with the matched-guise technique because it uses a visually identical speaker, hence matched-guise, while being able to use two different voices. The problem with Lambert's original matched guise paradigm, which is still prevalent (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012), is that the same person speaks different variations of an accent. While it is plausible that bi-lingual people can adeptly speak both native and accented English, there is likely a slight difference between the different accents, e.g., one accent might have a slight accent infused from the other language. There are also potential confounds that could arise, due to the person speaking in both accents, because other behaviors congruent with that culture, such as gesture, might get inadvertently primed. Virtual characters allow us to use different real people speaking English in various accents, while still keeping the character's visual appearances and non-verbal behaviors identical.

multimedia platform that serves as an effective interface for research, education, and entertainment applications. CMVCs deliver enhanced experimental control, ease of manipulation (i.e. variables such as visual appearance, speech type, and contextual graphical environments), and controlled measurements in a consistent and cost-effective form (Loomis, Blascovich, & Beall, 1999). Compared to other training media, such as pre-recorded video and text, virtual characters make it possible to more closely replicate visual and auditory dimensions of the physical world. Thus, they present the ideal platform to isolate unique cultural characteristics and present them via simulation. These features make CMVCs useful and reliable tools for studying cultural cognition.

Theoretical Background and Related Work

Communicative cues interact with other dimensions in virtual worlds to influence user behavior. For example, Blascovich's (2002) Threshold Model of Social Influence (TMSI) is comprised of five factors: agency, response system level, communicative realism, self-relevance, and context. Communicative realism is a latent variable made up of movement, anthropomorphic, and photographic realism. Here we focus on the crucial role of communicative realism, specifically a paralinguistic cue such as accent, and study its interaction with self-relevance.

Self-relevance is an intrapersonal dimension of the TMSI that refers to the importance of a social interaction—within a context—to an individual's sense of self. The studies here manipulated self-relevance of a CMVC's spoken accent. We hypothesize that a spoken linguistic *accent* is more self-relevant to individuals who share the same cultural background

adopted by the CMVC. Our goal is to examine the relationship between cultural context and social influence in virtual environments.

Accent and Virtual Characters

Some studies have investigated how culturally congruent virtual agent characteristics influence interactants' cognition and behavior. In an effort to examine the interaction between culture and a CMVC's design in the domain of education, Rader, Echelbarger, and Cassell (2011), developed virtual peers that matched the dialect of children speaking African-American English and asked the children to complete a bridge building exercise. The children alternated between playing the role of student and teacher while explaining the building process. Rader et al. found that students who tend to speak more dialected English did so less when they played the role of a teacher. This work suggests that the virtual peer and culturally congruent context, coupled with the role switch, influenced students to speak mainstream English, which is shown to be related to higher student achievement.

In another line of work, Yin, Bickmore, and Cortes (2010) report that individuals who process information using peripheral cues were influenced by and trusted an agent tailored to their own culture. Yin et al. had two female agents; one Latina and the other Anglo. While the Latina agent was visibly busty, in a colorful room, and pictured with a Frida Kahlo painting, the Anglo agent was in a sterile, white room, with only a (Harvard) college banner. Yin and colleagues seemed to be biased in their own (American) cultural values and therefore over-marked the Latino culture (Medin, Bennis, & Chandler, 2010). Our work overcomes these

confounds by keeping visual appearance constant across studies and only manipulating the virtual characters' spoken accent.

Frame Switching

Cultural frame-switching refers to the idea that interpretive frames can shift due to situational cues in individuals who have two, internalized cultural identities (Benet-Martinez, Leu, Lee, & Morris, 2002). For example, Hong et al. (2000) primed Chinese-Americans with American, Chinese or neutral iconic images and demonstrated that participants in the Chinese primed condition interpreted the next task with more of a Chinese interpretive lens (focusing more on external attributions) than those in the American or control conditions. On the other hand, participants in the American primed condition projected more American cultural values by focusing on individual attributions for the same task. The work of Hong et al. (2000) suggests that cultural cues can influence biculturals to switch interpretive frames. A plausible explanation for frame-switching is that multi-cultural individuals internalize a host of different norms and culturally specific systems of meaning (D'Andrade, 1984) which can get activated depending on the context or social environment. Hence, activation of different cultural interpretive frames can result in varied constraints on the individual's psychological interpretations. The shift in interpretive frames can be especially notable if the difference in normative behavior between the bicultural individuals' cultures is significant.

Culturally Motivated Middle Eastern Virtual Characters

Here, we examined whether the accent of a CMVC can be used as a marker for culture by evoking cultural frame-switching in bicultural individuals. There is growing evidence in the social sciences that culturally normative behaviors vary across cultures (see Henrich, Heine, & Norenzayan, 2010 for a review). In other words, behaviors that are considered typical in one culture may be considered abnormal in another. This variability in culturally normative behavior and cognition has been observed in various aspects of human behavior. Research demonstrates the presence of cultural differences in morally motivated decision-making through the identification of moral domains present (or salient) in some cultures, but not in others (Haidt, Koller, & Dias, 1993; Shweder, Much, Mahapatra, & Park, 1997).

We conducted cross-cultural evaluations of our CMVCs to study a specific component of communication, namely accent. By carefully controlling for non-verbal behavior and solely manipulating the CMVC's accent, we could directly measure its effect on participants' perception and cognition. Based on previous literature and our *cultural influence model*, we predicted that bicultural individuals interacting with a culturally congruent (i.e., self-relevant) agent would adopt interpretive frames specific to that culture.

Eighty-one individuals (66 Americans) participated in the study. The study employed a 2 X 2 between-subjects design, where the first factor was the virtual character's accent (Iranian- or California-accented spoken English) and the second factor was the cultural background of the

participant (American or Iranian-American). All participants read a vignette that incorporated both Iranian and North American cultural values. The story is about an ostensibly Western student named Anthony who is asked to go to dinner at his classmate Shawn's house, who is the Iranian character, although not explicitly referred to as such. Various incidents occur in the story that play upon common social customs from both American and Iranian culture. The story (see Figure 1) also includes a balanced number of ambiguous American and Iranian cultural referents

> Yesterday was an interesting day. I got out of class at around 4:30. The weather was great and spring was in the air. My classmate Shawn, who I'm not really good friends with and don't know too well, invited me to a bonfire at the beach. I told him that I was hungry and needed to get something to eat before going to the beach. He was going to his parent's house for dinner and invited me over. I don't know him well, so I initially refused his offer. But he kept on insisting that it's the beginning of Spring and I should go and have dinner with his family, he said then we can all go to the bonfire together. On the way to his house, Shawn asked why I was wearing almost all green. I thought it was a strange question as a lot of students were wearing green that day.
>
> I went over to his house and met his dad at the living room. I thought to myself that the apple doesn't fall far from the tree. Upon seeing his father, Shawn introduced me to him, saying "this is my good friend Anthony". He seemed surprise by my presence. Now I wasn't sure if they were expecting a guest for dinner. Then Shawn's mother came to the living room. I had met her before. She used to work in the registrar of the school. I said hi to her and she greeted me back saying that it looked like water had gone under my skin.
>
> They had some apples on the table, they also had some coins right by the apple. He offered some salad to me, I asked if it was okay to chop up the garlic which was conveniently already on the table by the apples, which I think he had forgotten to put in the salad. He hesitated and took a few seconds before saying yes…

Figure 1: A snippet of the vignette of cultural events and products

(e.g. proverbs, values, and events) that could be tied to either culture. None of these idea units were explicitly labeled with their cultural referent (e.g., there was no explicit reference to the cultural event of St. Patrick's Day or Noruz, which is the Persian new year) and the idea units were interleaved so as to minimize memory distortions due to recency or primacy. After reading the story, participants interacted with a virtual confederate (see Figure 2) that spoke English with either an American (Californian) or Iranian Accent. They recounted the story to the

confederate, who simply listened and provided backchannel nonverbal behavior.

Rapport Agent

The agent used in this experiment, Utah (Hartholt, Gratch, Weiss, & Team, 2009) (Figure 2), is designed to establish rapport with human participants by providing contingent non-verbal feedback while the human speaker is speaking. To produce this feedback, the agent first detects and analyzes the human speakers' audiovisual features in real-time. These features include silence, head nod, eye-gaze (looking at the agent or not), and smile. An audio feature detector also extracts vocal intensity from the raw signal every 100ms using Praat[2]. With the intensity information, it outputs a binary feature, speech or silence, every 100ms. A visual feature detector[3] tracks the position of the face, facial feature points, direction of eye gaze, and smile level. Using this information, the system outputs visual features indicating whether the human is nodding, looking away, or smiling. Based on these perceived audiovisual features, the response model (Huang, Morency, & Gratch, 2011) decides in real-time what would be the most appropriate response (i.e. head nod and smile). These animated responses are first converted into Behavior Markup Language (BML) (Kopp et al., 2006), then sent to an action scheduler that keeps track of the duration of each animation. If the current animation has not been completed, the new animation will be ignored. The BMLs are passed to Smartbody (Thiebaux & Marsella, 2007), a virtual human animation system designed to seamlessly blend animations and procedural

[2] Praat, http://www.fon.hum.uva.nl/praat/

behaviors. Finally, the by-products of Smartbody are rendered by a commercial game engine, Gamebryo[4], and displayed to users. Our experiment employed an approach similar to the matched-guise paradigm[1]. The voice of the ECA was pre-recorded using the voice of the first author, who is a bi-cultural Iranian- American. The first author did not participate in recruitment or in conducting the experiments.

Figure 2: The CMVCs in the two experiments

After reading the vignette, participants completed a survey that included questions about the different events and perceived intentions of characters within the story. There were two questions that were designed to tap into culturally relevant social dilemma decisions in the context of the vignette. The first asked participants about the appropriateness of the non-Iranian character's behavior. Specifically, the question asked whether it was appropriate for Anthony to leave his host's house prior to them serving dinner. The dilemma of propriety in this case is refusing the host's hospitality.

---

[3] OKAO Technology, http://www.omron.com/r_d/coretech/vision/okao.html
[4] http://www.gamebryo.com

Using a *t*-test, results of the manipulation check questions showed that participants reliably judged the Iranian accented agent to sound more Middle Eastern. The Iranian accent also biased participants' visual perceptions of the virtual character such that they thought that the virtual character also looked more Middle Eastern.

We conducted an ANOVA statistical test with participant culture (Iranian or American) and virtual confederate accent (Middle-Eastern or California English) as two independent variables. The results showed that Iranian-American participants decided that it was less appropriate for Anthony to leave before dinner when they interacted with the Iranian-accented virtual character. When interacting with the California accented virtual character, the Iranian-Americans decided that it was less inappropriate for Anthony to leave before dinner (see Figure 3).

A similar pattern was observed for a social dilemma regarding the offering of a gift by a host. Iranian-Americans thought that the offering of the gift was more genuine when they interacted with the Iranian-accented character than with the California accented character. Neither of the accents had a significant effect on the American control group. With respect to the *cultural influence model*, these results suggest that the communicative cue of accent (Iranian spoken English) only influenced users' decision making when it was self-relevant to Iranian Americans.

Culturally Motivated Chinese Virtual Characters

The results from the Iranian American study demonstrate that culturally motivated virtual characters can affect user behavior. When the virtual character instructed the participants with Middle-Eastern accented

English, the Iranian-Americans adopted a congruent cultural frame to interpret a vignette and make decisions.

In the second study we examine the role of accent on Chinese-American biculturals. Moreover, we recorded two different speakers for each accent condition. This was intended to avoid any potential confounds related to specifics of one speaker's voice. Analyses did not indicate that there were speaker specific effects on any of the dependent measures so all analyses collapsed across both speakers for each accent condition.

The design of the study was similar to the previous 2 X 2 between subjects experiment. The first factor was the accent of the virtual character (Chinese English or California English) and the second factor was participants' cultural background (Chinese-American or American). Data was collected over the Internet using a combination of Amazon's Mechanical Turk and Qualtrics. Chinese-Americans were from a cultural student organization in Southern California. Recruiting was through email and snowball sampling. Once participants arrived at the URL for the experiment and after they consented to participate voluntarily, either the Chinese or California English accented agent introduced them to the study and gave them instructions on the task.

Participants included eighty-five Chinese-Americans and 87 for our control monocultural group. Participants completed the causal attribution task that consisted of being shown an image of a group of fish where one is in front of the others. The causal attribution task and variations of it have been used widely to measure collectivistic/individualistic tendencies (Benet-Martinez et al., 2002; Chattaraman, Lennon, & Rudd, 2010; Friedman, Liu, Chi, & Hong, 2006; Hong et al., 2000). On a 12-point Likert

scale, participants had to decide why one fish was in front of the others. A response on the higher end of the scale—closer to 12—indicated that participants were, "Very confident that it is because the one fish is being chased by the other fish." Decisions on the opposite side of the scale—on the lower end, closer to 1—suggested that participants were, "Very confident that it is because the one fish is leading the other fish." Higher scores on this task reflect a collectivistic causal attribution indicating that the group of fish causes the behavior of the front fish. Lower scores indicate an individualistic attribution that the fish volitionally leads the others by being in front.

We also asked participants manipulation check questions about how the agent sounded and appeared. Participants listening to the Chinese English accented character rated the character as sounding and appearing more Asian. Participants rated the California English accented character as sounding and appearing more American. There is also a main effect of culture across all conditions such that Chinese-Americans rated the virtual character as appearing and sounding more Asian than the American participants. Finally, there is an interaction between culture and accent such that the Americans rated the California English accent as more Western, and the Chinese-English accent as more Asian compared to Chinese-Americans.

Based on previous literature, we predicted that there would be a main effect of culture such that Chinese-Americans should score on the upper end of the scale on the causal attribution task. We conducted an ANOVA statistical test with participant culture (Chinese or American) and virtual confederate accent (Chinese or California English) as two

independent variables. The results were consistent with this hypothesis. As Figure 3 shows, the aggregate Chinese American group scored higher than the American control group. If accent is an important cultural cue, then the virtual characters with a Chinese English accent will prime the Chinese-Americans to adopt a collectivistic interpretive frame. As Figure 3 shows, the Chinese English speaking virtual character primed Chinese Americans to score higher on the causal attribution task compared to the California English accented character. This result lends support to the *cultural influence model* by demonstrating that accent influences users when it is self-relevant to them.

The Chinese-American sample in our study seems to be generally individualistic because they are significantly below the scale midpoint of 6 (which is approximately the average from the Chinese in Figure 4 from Hong et al., 2000). Nonetheless, the Chinese-English accented CMVCs is robust enough to cause relative differences compared to the Californian CMVCs. This suggests that Chinese Americans are indeed bicultural and that accent is a reliable cue to prime dual cultures.

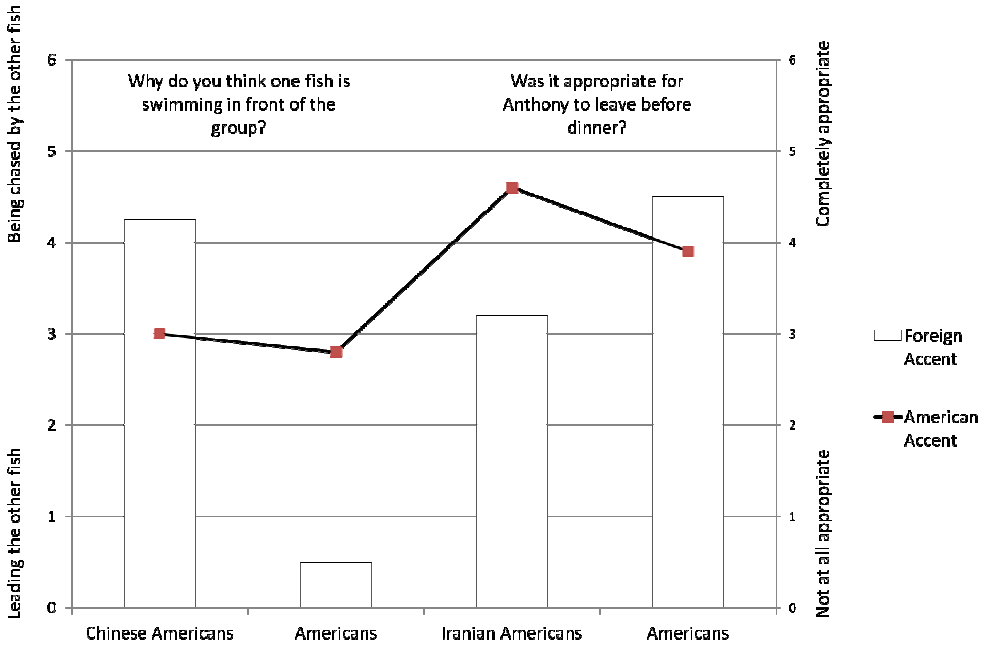

Figure 3: Summary results from both studies

Discussion and Implications

By using advanced digital technology to maintain a visually constant virtual character, these experiments manipulated only accent as a culturally relevant paralinguistic cue and controlled for other confounds. Across two studies with distinctly separate tasks, the results suggest that accent primes biculturals, from two different ethnic backgrounds, to switch their interpretative frames. Iranian-Americans interacting with an Iranian accented, English speaking CMVC were more likely to make decisions congruent with Iranian cultural customs. Similarly, Chinese-Americans listening to a Chinese accented, English speaking CMVC were more likely to make causal attributions congruent with collectivistic, East Asian cultural ideologies. Our work replicates results from Rader and colleagues (2011) by showing that biculturals switch their interpretive frames and related cognitive processing as a function of culturally congruent accent. The

results here are also consistent with the work of Nass and colleagues and show that accent can be a strong cue in how users treat computers as social actors (Dahlbäck, Wang, Nass, & Alwin, 2007; Nass & Brave, 2007; Reeves & Nass, 1996)

These results contribute to our understanding of what cues elicit social responses from users, which can provide design guidelines for building effective CMVCs in appropriate contexts. Moreover, in contrast to previous work in accented voice interfaces, the results here are more pervasive in that they show how interacting with a CMVC carries residual effects on a concurrent, but irrelevant task.

Developers who use their intuitions for design decisions to create CMVCs might assume that if a virtual character is more real, then it will be more successful at serving its purpose (Andre & Wickens, 1995). It could be argued that one reason to use ethnically accented CMVCs is to increase realism for cross-cultural training simulations for which virtual characters are developed. The logic here is that if increased communicative realism increases presence—a psychological state that a "mediated experience is not mediated" (Lombard & Ditton, 1997)—then increased presence will lead to effective CMVC-based cultural awareness training and fluency. For example, in cross-cultural contexts, developers of CMVCs might design virtual characters to simulate individuals from various cultures, such as those from the Middle East or Asia ("Alelo, Inc.," 2012). When creating the CMVCs, developers would probably add paralinguistic cues, such as accents, to the natural language that are congruent with the visual features of the target culture they are simulating (e.g., ethnically specific facial features and dress). Our work provides

experimental evidence that accent can affect individuals differently depending on whether they share the same ethnic cultural background as the target culture.

Extending theoretical models from social psychology, the *cultural influence model* put forth in this paper states that paralinguistic communicative cues, in the form of accented speech, affects users when it is self-relevant. In two studies, CMVCs with accents primed bicultural individuals, thereby leading them to adopt culturally congruent interpretive frames. Because participants in the control conditions for both studies were recruited and assumed to be a heterogeneous sample of monocultural North Americans, we did not predict that accent would affect them.

Results from the second study with Chinese accented CMVCs present an interesting pattern. North American participants in the English accent control condition and Chinese Americans primed by the English accent responded individualistically. The notable pattern in the results is that the Chinese accented CMVC garnered more extreme individualistic responses from the American control condition.

This result requires some theoretical speculation. According to Brewer's (1991) optimal distinctiveness theory, individuals have a need to differentiate themselves more or less from an out-group depending on the context. For example, if a less unique social identity is made salient (e.g., a more general social category such as a college student as opposed to a unique social category such as a computer science major in California), then opposing social psychological processes lead individuals to yearn more uniqueness. Therefore, one possible explanation of why the

heterogeneous control participants in the Chinese condition responded more individualistically is that the foreign accent made them feel less unique. This, in turn, biased them to use an individualistic, in-group interpretive frame. Another possible account for the pattern of results in the American control group is based on the concept of ethnic affirmation (Chen & Bond, 2007).

The *cultural influence model* can help designers of virtual humans decide what cues to add to simulations based on psychological theory and empirical experimental research. A design implication is that virtual human developers should consider using realistically accented, high fidelity spoken natural language modules. Although this work used pre-recorded speech for the virtual accent utterances, a follow-up line of research is to study synthetically accented natural language. This approach to natural language processing is nascent given that the First Workshop on Algorithms and Resources for Modelling of Dialects and Language Varieties just took place in 2011 (Jancsary, Neubarth, & Trost, 2011). Accent has psychological effects on users and the field of AI can help improve auditory aspects of virtual human realism in addition to the prevailing advancements in animation and photographic realism.

Author Note and Acknowledgment:

The authors would like to thank Jim Blascovich and Ari Shapiro for comments on drafts of this manuscript.

A postdoctoral fellowship from the ARL HRED Cognitive Sciences Branch supported the first author.

References

Alelo, Inc. (2012). Retrieved December 12, 2012, from www.alelo.com

Andre, A. D., & Wickens, C. D. (1995). When users want what's NOT best for them. *Ergonomics in Design*, *3*(4), 10–14.

Benet-Martinez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating Biculturalism: Cultural Frame Switching in Biculturals with Oppositional Versus Compatible Cultural Identities. *Journal of Cross-Cultural Psychology*, *33*(5), 492–516. doi:10.1177/0022022102033005005

Blascovich, J. (In press). Challenge, Threat and Social Influence in Digital Immersive Virtual Environments. In J. Gratch & S. Marsella (Eds.), *Social Emotions in Nature and Artifact: Emotions in Human and Human-computer Interaction*. Oxford: Oxford University Press.

Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments (pp. 25–30). ACM Press. doi:10.1145/571878.571883

Blascovich, J., & McCall, C. (2013). Social influence in virtual environments. In K. Dill (Ed.), *Oxford Handbook of Media Psychology* (pp. 305–315). Oxford.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482. doi:10.1177/0146167291175001

Chattaraman, V., Lennon, S. J., & Rudd, N. A. (2010). Social identity salience: Effects on identity-based brand choices of Hispanic consumers. *Psychology and Marketing*, *27*(3), 263–284. doi:10.1002/mar.20331

Chen, S. X., & Bond, M. H. (2007). Explaining Language Priming Effects: Further Evidence for Ethnic Affirmation Among Chinese-English

Bilinguals. *Journal of Language and Social Psychology*, *26*(4), 398–406. doi:10.1177/0261927X07306984

D'Andrade, R. G. (1984). Cultural meaning systems. In R. A. Schweder & R. A. Levine (Eds.), *Culture Theory. Essays on Mind, Self and Emotion* (pp. 88–119). Cambridge: Cambridge University Press.

Dahlbäck, N., Wang, Q., Nass, C., & Alwin, J. (2007). Similarity is more important than expertise (p. 1553). ACM Press. doi:10.1145/1240624.1240859

Friedman, R. A., Liu, W., Chi, S.-C., & Hong, Y. (2006). Western Exposure: the Effects of Foreign Experience on Taiwanese Managers' Cognition. *SSRN Electronic Journal*. doi:10.2139/ssrn.935134

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, *42*(1), 120–133. doi:10.1002/ejsp.862

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*, 613–628.

Hartholt, A., Gratch, J., Weiss, L., & Team, T. G. (2009). At the virtual frontier: Introducing gunslinger, a multi-character, mixed-reality, story-driven experience. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents, Lecture Notes in Computer Science* (pp. 500–501). Springer.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–135.

Hong, Y., Morris, M. W., Chiu, C., & Benet-Martínez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*(7), 709–720. doi:10.1037/0003-066X.55.7.709

Huang, L., Morency, L. P., & Gratch, J. (2011). Virtual Rapport 2.0. In *International Conference on Intelligent Virtual Agents* (pp. 68–79).

Jancsary, J., Neubarth, F., & Trost, H. (2011). First Workshop on Algorithms and Resources for Modelling of Dialects and Language Varieties. Presented at the Dialects 2011, Edinburgh, Scotland: Association for Computational Linguistics.

Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., … Vilhjálmsson, H. (2006). Towards a Common Framework for Multimodal Generation: The Behavior Markup Language. In *International Conference on Intelligent Virtual Agents* (pp. 21–23).

Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social issues*, *23*, 91–109.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2), 0–0. doi:10.1111/j.1083-6101.1997.tb00072.x

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environments as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers*, *31*(4), 557–564.

Medin, D., Bennis, W., & Chandler, M. (2010). Culture and the Home-Field Disadvantage. *Perspectives on Psychological Science*, *5*(6), 708–713. doi:10.1177/1745691610388772

Nass, C. I., & Brave, S. (2007). *Wired for speech : how voice activates and advances the human-computer relationship*. Cambridge, Mass.; London: MIT.

Nass, C. I., & Yen, C. (2010). *The man who lied to his laptop : what machines teach us about human relationships*. New York: Current.

Rader, E., Echelbarger, M., & Cassell, J. (2011). Brick by brick. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11* (p. 2971). Vancouver, BC, Canada. doi:10.1145/1978942.1979382

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York, NY.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). Morality and health. In A. Brandt & P. Rozin (Eds.), (pp. 119–169). New York: Routledge.

Thiebaux, M., & Marsella, S. (2007). Smartbody: Behavior realization for embodied conversational agents. In *In 7th International Conference on Intelligent Virtual Agents (IVA)*.

Yin, L., T. W. Bickmore, & Cortes, D. E. (2010). The impact of linguistic and cultural congruity on persuasion by conversational agents. In *Proceedings of IVA, Lecture Notes in Computer Science* (pp. 343–349). Presented at the Intelligent Virtual Agents, Philadelphia, PA: Springer. Retrieved from http://relationalagents.com/publications/IVA10.culture.pdf