Expectancy Violations Related to a Virtual Human's Joint Gaze Behavior in Real-Virtual Human Interactions

Kangsoo Kim[1], Arjun Nagendran[1], Jeremy Bailenson[2], and Greg Welch[1]

[1]University of Central Florida
[2]Stanford University

Abstract

Joint gaze—the shared gaze by the individuals toward a common object/point of interest—offers important non-verbal cues that allow interlocutors to establish common ground in communication between collocated humans. Joint gaze is related to but distinct from *mutual gaze*—the gaze by the interlocutors towards each other such as during eye contact, which is also a critical communication cue. We conducted a user study to compare real human perceptions of a virtual human (VH) with their *expectancy* of the VH's gaze behavior. Each participant experienced and evaluated two conditions: (i) VH with mutual gaze only and (ii) VH with mutual and joint gaze. We found evidence of positive responses when the VH exhibited joint gaze, and preliminary evidence supporting the effect of *expectancy violation*, i.e., more positive perceptions when participants were presented with VH's gaze capabilities that exceeded what was expected.

Keywords: joint gaze, expectancy violation, human perception, virtual human, avatar

1 Introduction

Expectancy violation (EV) is a well-known phenomenon in human communications and psychology [1]. The phenomenon of EV arises when one encounters an unexpected behavior, and as a result experiences either positive or negative feelings. For example, if a child is given a gift, she will likely be happier if the gift was unexpected (e.g., "out of the blue") than if the gift was expected (e.g., it was her birthday). This would be a *positive violation*. Conversely, not receiving a gift on her birthday, when she clearly expected it, would be a *negative violation* that could cause her to have an unfavorable response.

Figure 1: Virtual human exhibiting mutual gaze (left) and joint gaze (right).

In this paper, we conduct a user study that independently varies a virtual human's (VH) joint gaze behavior (Fig. 1), and investigate the effects of a mismatch between user expectations of the VH's gaze behavior and the VH's actual gaze behavior, with respect to the user's perceptions of the VH. *Joint gaze* is the shared gaze that interacting interlocutors typically exhibit when attending to a common object of interest. Joint gaze is an important aspect of establishing common ground, so interlocutors generally expect joint gaze when attempting to establish *joint attention* to a shared object. For example, if you explain directions to a partner while pointing toward features on a map, you would expect your partner to look at the map. If your partner does *not* look at the map, you might be puzzled and wonder whether your partner is paying attention. A positive (or negative) violation corresponds to when a subject initially having a low (or high) expectation of a VH's joint gaze later evaluates the VH more positively (or negatively) after they actually meet a VH with (or without) joint gaze. We hypothesize that an expectation violation related to the VH's joint gaze will influence one's perceptions of the VH.

Among the prior research looking at the importance of gaze behavior in VH systems, some work has looked at the gaze behavior between VHs in a *virtual* environment [2, 3], while other work has looked at gaze between real humans and VHs in *real/mixed* environments [4, 5, 6, 7, 8]. Our interest is in the latter due to the involvement of real objects in the interaction, as opposed to interactions in a purely virtual environment. Previous research supported the importance of VH eye gaze (mostly mutual gaze, i.e. eye contact) in human perception (e.g., social-presence) or task performance. However, there is relatively little research narrowing the focus down to joint gaze and one's expectancy. This paper presents preliminary results about the effects of VH's joint gaze and its expectancy violation in one's perceptions of the VH.

2 Virtual Human

A remote human controller manipulated our VH from a separate room (Wizard of Oz). We provided the controller with a video feed of the experimental space so he could see the experimental environment and affect the VH gaze either toward the subjects face or toward the map, depending on the current trial/subject (Fig. 1). The controller used an interface with an infrared camera (TrackIR) and a magnetic tracking device (Razer Hydra) to perform the VH's facial expressions, mouth movement, and change of gaze direction effectively via our system developed previously [9]. The upper torso of our VH was displayed in near human-size on a 55" 2D flat screen, and a table with black curtains blocked the place where the lower torso should have been, so subjects could feel that the VH was behind the table (Fig. 2). In our scenario, the VH expressed a normal or slightly pleasant facial expression during the interaction, so that the subjects could feel the VH's emotional state was consistent. The VH generally initiated the conversation unless the subject started talking first, but did not say anything proactively during the interaction. In other words, the VH only made positive reciprocal answers (e.g. "Yes, I understand.") to the subject's affirmative question "Do you understand?".

Figure 2: Virtual human setup (left) and facial expressions (right).

3 Experiment

3.1 Scenario and Manipulation

Our human subjects were introduced to a VH and told his name ("Michael"). They were then told that the VH was a new student at the university who was currently in a building off campus, but needed to return for a lecture on campus, and that he was late. The subjects were then asked to staff a "help desk" and to provide the VH with directions using a Campus Map and a pen. We had two conditions of the VH's gaze behavior: (i) *mutual gaze only* and (ii) *mutual gaze with joint gaze* (Fig. 1). While the VH always looked at the subject's face without looking down to the map in "mutual-only" condition, he looked at the map occasionally in the "mutual+joint" condition. In both conditions, the VH exhibited small natural upper-torso movement and eye blinks. Subjects experienced both conditions and evaluated the two VHs. The overall procedure is illustrated in Fig. 3. First, subjects saw the VH verbally explaining the situation that he would look for directions to the campus, and completed a demographic pre-questionnaire. They experienced both interaction 1 and 2 explaining the map, but the VH performed a different gaze behavior in each interaction—either "mutual-only" or "mutual+joint." After each interaction, subjects were asked to complete a questionnaire about their perceptions of the VH and sense of their EV with respect to the VH's gaze behavior (5-scale Likert). Finally, they compared two conditions and reported their preference in questionnaire 3. To prevent the subjects from familiarizing themselves with the same set of directions, we counter-balanced a different destination on the map as well as the VH's gaze behavior for each interaction.

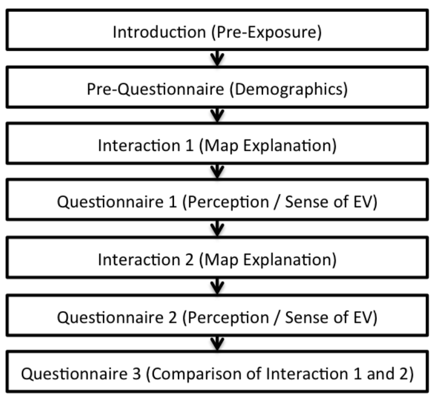

Figure 3: Overall procedure.

3.2 *Human Subjects*

A total of 28 subjects were recruited from The University of Central Florida, and received \$15 of monetary compensation for the experiment. Subjects were 75% male (n = 21, mean age = 19.95, range = 18–26) and 25% female (n = 7, mean age = 20.71, range = 18–24). Most of them (n = 26) had previous experience of virtual characters through video games or virtual reality applications. All the subjects were aware that what they interacted with was a VH.

4 Results and Discussion

Joint Gaze: We evaluated subjects' responses from (comparison) quesionnaire 3 to check which gaze condition of VH subjects preferred. More than 50% of subjects chose "mutual+joint" VH as their preference in all the questions, which indicates the importance of joint gaze feature in VH system (Table 1). However, there were still considerable number of people who did not feel any difference between two conditions. According to informal discussion with subjects after the study, a majority of the subjects addressed that the VH's verbal response capability far exceeded what they had previously experienced although our VH merely responded their affirmative questions. We guess that highly engaging verbal communication might overwhelm the effect of joint gaze, so subjects could not feel any difference between two conditions.

Expectancy Violation (EV): Although we used the same questions for questionnaire 1 and 2 (perception / sense of EV) for experiment consistency, we only analyzed subjects' responses from questionnaire 1 to evaluate EV effects, because their expectation might be biased for the multiple interactions. After the first interaction, we asked for their sense of EV with respect to the VH's gaze behavior, e.g., "How would you rate the capability of virtual human's gaze compared to what you expected?" (5-scale, 1: more negative than what I expected, 3: same as what I expected, 5: more positive than what I expected). We expected both negative and positive responses, but the responses were mostly positive, so we focused on *positive* violations. The results indicated that "mutual+joint" VH encouraged more positive violation than "mutual-only" VH. In other words, subjects with "mutual+joint" VH tended to evaluate the VH's gaze behavior more positively, compared to what they expected before, than with "mutual-only" VH (Fig. 4). T-tests showed a significant difference in subject-reported EV of VH's gaze behavior for "mutual-only" (M = 3.643, SD = 0.842) and "mutual+joint" (M = 4.500, SD = 0.650) conditions; $t(24) = -3.01$, $p = 0.006$.

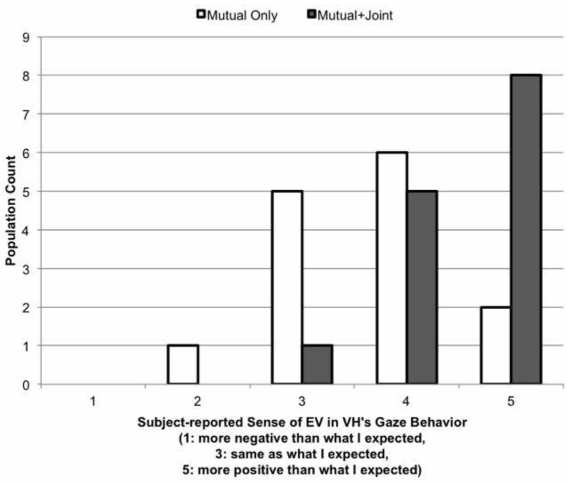

Figure 4: Population by subject-reported sense of EV in VH's gaze behavior. Population with "mutual+joint" VH tends towards the highest (5) in the sense of EV while population with "mutual-only" is more towards (4) in the sense of EV, which can be interpreted that VH's joint gaze encourages more positive violation.

When we analyzed the relationship between subject's perceptions and their sense of EV, we observed high-reliability between the responses from 9 questions in Table 2 (Cronbach's alpha > 0.80), so we averaged their responses into a single value and used it as their perception re-

Table 1: Subject's responses from comparison questionnaire 3. The value indicates the number of people who preferred the condition, and its percentage out of total 28 subjects in parentheses.

Question	Mutual+Joint	Mutual-Only	No Difference
Which virtual human did you like more?	17 (61%)	2 (7%)	9 (32%)
Which interaction did you enjoy more?	16 (57%)	8 (29%)	4 (14%)
Which interaction were you more engaged with?	14 (50%)	3 (11%)	11 (39%)
Which virtual human did you think more pay attention to what you were explaining?	21 (75%)	2 (7%)	5 (18%)
Which virtual human did you feel that more understood what you were explaining?	17 (61%)	6 (21%)	5 (18%)
Which virtual human did you feel more as if it was a real human?	16 (57%)	3 (11%)	9 (32%)
Which virtual human gave you more sense of physical presence?	14 (50%)	2 (7%)	12 (43%)
Which virtual human was more natural (human-like)?	18 (64%)	3 (11%)	7 (25%)

sponse. In Fig. 5, we found evidence of a tendency that a subjects' perception became more positive (i.e. larger values in y-axis) as their expectancy of gaze was more positively violated (i.e. more towards 5 in x-axis) when we compared 3, 4, and 5 columns. Although the sample size (N) was small and varied, the tendency could be interpreted that subject's perception was influenced by their expectancy, which was positively violated after the interaction.

Table 2: Nine questions for subject's perception responses from questionnaire 1 (5-scale, 1: strongly disagree, 5: strongly agree). Subject's responses from these questions were correlated (Cronbach's alpha > 0.80).

1. You liked the virtual human.
2. You enjoyed the interaction with the virtual human.
3. You were engaged in the interaction with the virtual human.
4. You had the feeling that the virtual human was paying attention to what you explained.
5. The virtual human seemed to understand what you explained.
6. You had the feeling that the virtual human was a real human.
7. You had the feeling that the virtual human was physically present in real space.
8. The interaction with the virtual human was natural.
9. You had the feeling that the virtual human looked at the map.

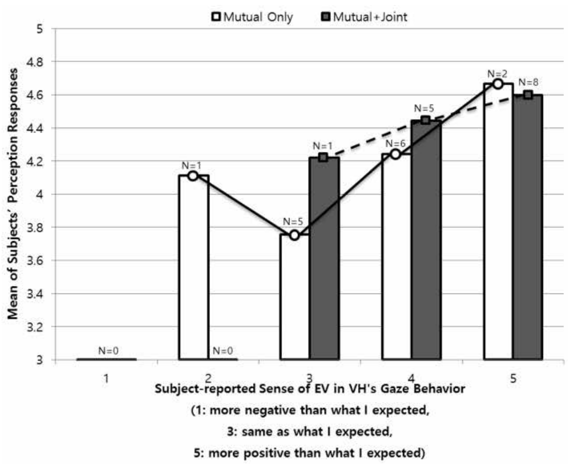

Figure 5: Mean of subjects' perception responses over self-reported sense of EV. In both "mutual-only" and "mutual+joint" conditions, a higher positive EV (x-axis) resulted in a higher value of perception (y-axis).

5 Conclusions

We have presented a user study aimed at understanding the effects of a VH's joint gaze behavior and the phenomenon of expectancy violation (EV) with respect to a human's perception of the joint gaze behavior of a VH. As expected, joint gaze was found to be an important characteristic for subjects to build positive responses to the VH during a map explanation scenario. We also discovered preliminary evidence of a positive EV effect—subjects evaluated the VH more positively corresponding to how much the VH's gaze behavior exceeded their expectation (positively) regardless of the presence of joint gaze. In the future, we will consider a large-sample study investigating the effects of a user's previous experience and expectations related to various features of virtual or robotic humans. If we find a certain feature that causes a negative violation in general, which means people normally have high expectations about the feature, it would indicate that the feature should be carefully considered for future VHs.

Acknowledgements

This work was supported in part by the U.S. Office of Naval Research under Dr. Peter Squire, Code 30, awards N00014-14-1-0248 and N00014-12-1-1003.

References

[1] Judee K. Burgoon, Deborah A. Newton, Joseph B. Walther, and E. James Baesler. Nonverbal expectancy violations and conversational involvement. *Journal of Nonverbal Behavior*, 13(2):97–119, 1989.

[2] Jeremy N. Bailenson, Andrew C. Beall, and Jim Blascovich. Gaze and task performance in shared virtual environments. *The Journal of Visualization and Computer Animation*, 13(5):313–320, 2002.

[3] Rutger Rienks, Ronald Poppe, and Dirk Heylen. Differences in head orientation behavior for speakers and listeners: an experiment in a virtual environment. *ACM Transactions on Applied Perception*, 7(1):1–13, 2010.

[4] Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 301–308, 2001.

[5] Alex Colburn, Michael F. Cohen, and Steven Drucker. The Role of Eye Gaze in Avatar Mediated Conversational Interfaces. In *Technical Report MSR-TR-2000-81, Microsoft Research*, 2000.

[6] William Steptoe, Robin Wolff, Alessio Murgia, Estefania Guimaraes, John Rae, Paul Sharkey, David Roberts, and Anthony Steed. Eye-tracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments. In *ACM Conference on Computer Supported Cooperative Work*, pages 197–200, 2008.

[7] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 529–536, 2003.

[8] Gary Bente, Felix Eschenburg, and Nicole C. Krämer. Virtual gaze. A pilot study on the effects of computer simulated gaze in avatar-based conversations. *Virtual Reality (LNCS)*, 4563:185–194, 2007.

[9] Arjun Nagendran, Remo Pillat, Adam Kavanaugh, Greg Welch, and Charles Hughes. AMITIES: Avatar-mediated Interactive Training and Individualized Experience System. In *ACM Symposium on Virtual Reality Software and Technology*, pages 143–152, 2013.