Making It Personal: End-User Authoring of Health Narratives Delivered by Virtual Agents

Timothy Bickmore, and Lazlo Ring

College of Computer and Information Science, Northeastern University
360 Huntington Ave - WVH 202, Boston, MA 02115
{bickmore, lring}@ccs.neu.edu

Abstract. We describe a design study in which five different tools are compared for end-user authoring of personal stories to be told by an embodied conversational agent. The tools provide varying degrees of control over the agent's verbal and nonverbal behavior. Results indicate that users are more satisfied when their stories are delivered by a virtual agent compared to plain text, are more satisfied when provided with tools to control the agent's prosody compared to facial display of emotion, and are most satisfied when they have the most control over all aspects of the agent's delivery.

1 Introduction

Storytelling plays a central role in group counseling for health behavior change, and is the cornerstone of many formal and informal counseling and support group methodologies, such as the 12-step program used by Alcoholics Anonymous. Hearing successful stories of change from peers not only conveys specific information that may be instrumental in effecting change, but increases motivation and self-efficacy to change through social learning [1]. The act of writing such stories also has psychological benefits for the author, including better coping with past traumatic events [2] and improvement in self-efficacy [3]. In one study of participants attending a cancer-related storytelling workshop, 85% of respondents indicated that hearing stories of others living with cancer gave them hope, and 97% indicated that the act of storytelling helped them cope with cancer [4].

We plan to leverage the power of personal change narratives in an online system to promote "preconception care" among young African American women. Preconception care involves addressing health issues that may negatively impact the health of a baby before a woman conceives [5]. We are focusing on African American women since they have roughly twice the rate of infant mortality and low infant birth weight compared to Caucasian women in the U.S. Our automated intervention involves first screening women for 53 risk factors during a clinic visit, ranging from vaccinations, diet and chronic disease management, to smoking and drug abuse. Following this, screened women will interaction with a web-based virtual agent

(based on [6]) approximately three times a week for two months, during which time the agent will counsel them on how to address any health issues identified in the screening and motivate them to take action.

In addition to having the agent give advice authored by the clinicians on the design team, we will also elicit stories of health behavior change from the users of the system. These (moderated) stories will then be shared with other users who are struggling with issues addressed in the stories. We face several fundamental design choices in developing this system, including: how these stories should be authored by users; how the stories should be told to other users; how users should be motivated to contribute stories (e.g., using incentive mechanisms as in [7]); how the stories should be indexed (e.g., as in [8]); and how their telling should be integrated into the rest of the online counseling dialogue system [9]. Among these many issues, we have decided to initially focus our design efforts on the first one: what kind of tools should we provide to users for them to author their stories?

The simplest end-user story authoring tools would just let users write their stories in text and display them in a similar fashion. However, since the system will incorporate a conversational agent, users could be provided with the means to author their stories so that they could be told to other users by the agent in dialogue, taking advantage of the additional expressivity afforded by the agent. Given this, how much control over the conversational agent and its delivery of the story would users want? Is it sufficient to let users write the text of their story and have a text-to-embodied speech translation engine (such as BEAT [10]) automatically generate all of the verbal and nonverbal agent behavior used in the story telling? Or, despite evidence that naïve users have difficulty with task-specific programming languages [11], will they want full control over the agent and its delivery of their story, and if so, which aspects of that control are most important to them?

In this article we present the results of an initial design study comparing a range of authoring tools to support users in creating health behavior change narratives.

2 Related Work

Several researchers have investigated the use of embodied conversational agents as aides in storytelling. For example, Umaschi and Cassell developed the SAGE system for the purpose of eliciting stories from children [12]. The system's interface allowed children to author stories using a branching dialogue representation, with the interactive stories told by a conversational rabbit character.

Many others have developed authoring tools to enable non-programmers to create stories or storytelling systems, in which the stories are told using a variety of media. Cassell and Ryokai developed the StoryMat system to elicit stories from preliterate children using a combination of physical location cues (where a child was playing on a play mat), audio prompts and voice recording [13]. Skorupski and Mateas developed Wide Ruled, an authoring system to enable non-programmers to create AI plan-driven interactive storytelling systems [14].

Avatar-based CMC systems typically provide users with some control over the verbal and nonverbal behavior of their avatar. For example, Comic Chat allows users

to communicate with each other online using a comic strip metaphor, in which many aspects of a character's nonverbal behavior, such as its emotional state, can be manipulated.

3 An Initial Design Study of Storytelling Support Tools

In order to determine how much and what kind of control end users would want over an embodied conversational agent telling their stories to others, we conducted an experiment comparing five different versions of story authoring and re-telling tools. In all versions, the story fragment was comprised of a short linear sequence of utterances, and the tools provided different sets of controls over agent verbal and nonverbal behavior. The five tools evaluated were:

- TEXT – The author enters their story in a text edit box, and the story is "told" by simply displaying the text (intended as a baseline control condition).
- AGENT – The author enters their story in a text edit box, and the story is told by an embodied conversational agent speaking the story ("Gladys", Figure 1), with nonverbal behavior (eyebrow raises, head nods, eye gaze motion, posture shifts) automatically generated using BEAT [10], and prosody generated by the default rules in the text-to-speech engine (Loquendo "Susan" voice). The agent maintains a neutral facial display throughout.
- FACE – The same as AGENT, but the author can specify alternative facial displays (happy and sad/concerned displays in addition to neutral) to be used when delivering different parts of the story.
- VOICE – The same as AGENT, but the author can use a variety of prosody commands to affect speech delivery, including speed (increase/decrease), volume (increase/decrease), baseline pitch (increase/decrease), emphasis, and pause at specified locations in the story.
- ALL – Combination of FACE and VOICE (Figure 2).

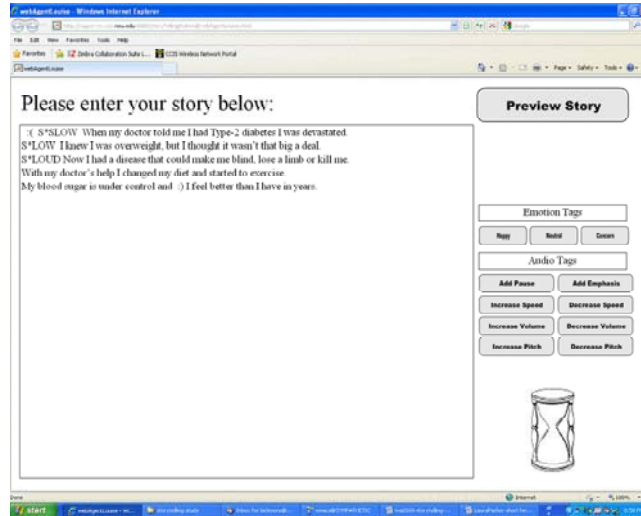**Figure 1.** Conversational agent interface

Figure 2. Story authoring interface

The study had a five treatment counterbalanced within-subjects design. In order to minimize burden on study participants, they were provided with a set of 10 pre-authored health stories and asked to select 5 that they wanted to work with (rather than having to write their own). Following administration of demographic and personality questionnaires, participants conducted a brief training interaction with the agent, then filled out a questionnaire assessing their attitude towards the agent. Next they were asked to input each of the five stories they selected into different versions of the tool and asked to modify each until they were happy with its delivery, filling out evaluation questionnaires after each story.

Measures. After submitting a story with each version of the tool, participants answered the questions shown in Table 1. In addition, we measured the number of controls they used in each story and the amount of editing they did on the text of the story that was provided to them, using Levenshtein distance [15].

Participants. Nine African American females aged 20-24 were recruited via craigslist.org. All had high levels of computer proficiency and were well educated (one had a high school level education, the rest had at least some college).

3.3 Quantitative Results

For all rating measures (with larger values representing more positive evaluations), TEXT was always rated the lowest, ALL the highest, and VOICE second highest (Table 2). Analyses across the five treatments were performed using SPSS Repeated Measures ANOVA, with LSD post hoc tests.

There was a significant difference in ratings of how much the agent helped users tell their stories, $F(4,24)=8.6$, $p<.001$. Post hoc tests indicated significant differences between TEXT and the other conditions, and AGENT and the other conditions, but no significant differences among FACE, VOICE and ALL.

Table 1. Tool Evaluation Questions

Item	Question Text	Anchor 1	Anchor 7
Helpful	How much did Gladys help you tell your story?	Not at all	Very much
Confident	How confident are you that Gladys will re-tell your story the way you want it to be heard by others?	Not at all	Very much
Express	How much did you feel you could express yourself in Gladys' telling of your story?	Not at all	Very much
Easy	How easy was it to author a story using this version of the tool?	Very difficult	Very easy
Satisfied	How satisfied were you with your authoring experience using this version of the tool?	Not at all	Very much

Table 2. Primary Study Outcomes – Mean (SD) - *$p<.05$

Item	TEXT	AGENT	FACE	VOICE	ALL
Helpful*	2.86 (1.57)	4.43 (1.40)	5.14 (1.34)	5.43 (0.79)	5.86 (1.46)
Confident	3.22 (1.86)	4.22 (1.20)	4.00 (2.24)	4.78 (2.28)	5.11 (2.21)
Express	3.67 (1.50)	4.00 (1.32)	4.00 (2.40)	4.89 (2.21)	5.22 (2.22)
Easy	3.44 (2.01)	4.22 (1.56)	4.22 (2.22)	5.00 (2.29)	5.44 (2.19)
Satisfied*	3.00 (1.87)	4.00 (1.41)	3.78 (2.11)	5.00 (2.34)	5.33 (2.24)
ControlsUsed*	N/A	N/A	3.33 (2.40)	7.22 (6.69)	8.00 (6.27)
EditDistance	24 (32)	25 (42)	21 (56)	63 (89)	80 (186)

There was also a significant difference in ratings of overall satisfaction with the authoring experience, $F(4,32)=2.8$, $p<.05$. Post hoc tests indicated significant differences between TEXT and ANIMATION, TEXT and VOICE, and TEXT and ALL, but no other differences.

There were no significant differences among the treatments on reported confidence that the agent would properly re-tell the story, $F(4,32)=1.6$ n.s., or in the amount they felt they could express themselves, $F(4,32)=1.2$ n.s., or in ratings of ease of use, $F(4,32)=1.6$ n.s.

3.4 Qualitative Results

When asked for their overall impressions of the storytelling system, six respondents provided positive responses ("interesting", "fun", "cool" being the most common):

> *"I thought it was really cool, um, I've never worked with inputting text and seeing it read back to me or portrayed to me um virtually so that was a new experience for me and I really really liked it."*

> *"Gladys was able to use emotion and express it the way I wanted her to express it so I felt like I was telling the story."*

Two respondents complained that the tools still did not let them express themselves enough:

> *"I thought it was an interesting concept but sometimes working with the tool was a little difficult. It was kind of hard to get the person online to talk the same way that I wanted it to sound like in my head."*

When asked which of the tools they thought was most effective, seven respondents said they preferred having all of the tools available to them. Given a choice between control of facial display and prosody, two respondents indicated they preferred the speech controls, while two others indicated they preferred control over facial display:

> *"The one where you can add all of the emotions and pauses and increase pitch um I think that one was better because you can make it more human."*

Finally, when asked about suggestions to improve the authoring tool, several mentioned better voice quality, two suggested a broader range of emotion displays for the face and one suggested more emotion controls for the voice. One participant also volunteered that the ability to customize the look of the agent would be important to her.

4 Discussion

We find that, rather than being intimidated by a wide range of controls for verbal and nonverbal behavior of an embodied conversational agent, participants actually preferred having as many controls at their disposal as possible for crafting their stories. They also appeared to prefer speech controls over facial display controls, although this may also simply be an artifact of the number of options available in each (8 prosody controls vs. 3 facial display controls). Participants used the most controls and made the most edits on the pre-written story text when the speech controls were available to them (VOICE and ALL), likely due to a focus on getting their story to sound right given the quirks of the speech synthesizer. Finally, participants preferred all versions of agent storytelling to simply having the text of their story displayed to others.

4.1 Future Work

We plan to explore several additional design alternatives for story authoring tools, including "storyboards" in which users can select from static images of the character telling the story, allowing a much broader range of characters and nonverbal behavior than we can currently support, but without animation. We also plan to explore the space of tools to help users write well-formed and impactful stories, for example by providing them with relevant examples, forms to fill out, prompts, coaching, or guiding them through a process of oral storytelling and transcription.

References

1. Bandura, A.: Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev **84** (1977) 191-215
2. Andersen, P., Guerrero, L.: Principles of Communication and Emotion in Social Interaction. In: Andersen, P., Guerrero, L. (eds.): Handbook of Communication and Emotion. Academic Press, New York (1998)
3. Roberts, L.: Giving and receiving help: Interpersonal transactions in mutual-help meetings and psychosocial adjustment of members. American Journal of Community Psychology **27** (1999) 841–868
4. Chelf, J., Deshler, A., Hillman, S., Durazo-Arvizu, R.: Storytelling: A Strategy for Living and Coping with Cancer. Cancer Nursing **23** (2000) 1-5
5. Mitchell, S., Bickmore, T., Paasche-Orlow, M., Williams, C., Forsythe, S., Atrash, H., Johnson, K.J., B Increasing Access to Preconception Care Using Health Information Technology Salud(i)Ciencia (2010)
6. Bickmore, T., Schulman, D., Shaw, G.: DTask & LiteBody: Open Source, Standards-based Tools for Building Web-deployed Embodied Conversational Agents. Intelligent Virtual Agents, Amsterdam, Netherlands (2009)
7. Farzan, R., DiMicco, J.M., Millen, D.R., Dugan, C., Geyer, W., Brownholtz, E.A.: Results from deploying a participation incentive mechanism within the enterprise CHI'08 (2008) 563-572
8. Domeshek, E.: Do The Right Thing: A Component Theory for Indexing Stories as Social Advice. Institute for the Learning Sciences, Northwestern University (1992)
9. Jefferson, G.: Sequential aspects of storytelling in conversation. In: Schenkein, J. (ed.): Studies in the organization of conversational interaction. Academic Press, New York (1978) 219-248
10. Cassell, J., Vilhjálmsson, H., Bickmore, T.: BEAT: The Behavior Expression Animation Toolkit. SIGGRAPH '01, Los Angeles, CA (2001) 477-486
11. Nardi, B.A.: A Small Matter of Programming: Perspectives on End User Computing. The MIT Press, Cambridge, MA (1993)
12. Umaschi, M., Cassell, J.: Storytelling Systems: Constructing the Innerface of the Interface. Cognitive Technologies Proceedings (1997) 98-108
13. Cassell, J., Ryokai, K.: Making Space for Voice: Technologies to Support Children's Fantasy and Storytelling. Personal Technologies **5** (2001) 203-224
14. Skorupski, J., Jayapalan, L., Marquez, S., Mateas, M.: Wide Ruled: A Friendly Interface to Author-Goal Based Story Generation. Proceedings ICVS 2007.
15. Levenshtein, V.: Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady **10** (1966) 707-710