Designing Empathic Agents: Adults vs. Kids

Lynne Hall[1], Sarah Woods[2], Kerstin Dautenhahn[2], Daniel Sobral[3], Ana Paiva[3], Dieter Wolke[4,] Lynne Newall[5]
[1] School of Computing & Technology, University of Sunderland, UK, lynne.hall@sunderland.ac.uk
[2] Adaptive Systems Research Group, University of Hertfordshire, UK, s.n.woods, k.dautenhahn@herts.ac.uk
[3] Instituto Superior Technico & INESC-ID, Porto Salvo, Portugal ana.paiva@inesc-id.pt
[4]Jacobs Foundation, Zurich, Switzerland, dieter.wolke@jacobsfoundation.org
[5]Northumbria University, Newcastle, UK, lynne.newell@unn.ac.uk

Abstract. An evaluation study of a Virtual Learning Environment populated by synthetic characters for children to explore issues surrounding bullying behaviour is presented. This 225 participant evaluation was carried out with three stakeholder groups, (children, teachers and experts) to examine their attitudes and empathic styles about the characters and storyline believability. Results revealed that children expressed the most favourable views towards the characters and the highest levels of believability towards the bullying storyline. Children were more likely to have an empathic response than adults and found the synthetic characters more realistic and true-to-life.

Introduction

Virtual Learning Environments (VLEs) populated with animated characters offer children a safe environment where they can explore and learn through experiential activities [5, 8]. Animated characters offer a high level of engagement, through their use of expressive and emotional behaviours [6], making them intuitively applicable for exploring personal and social issues. However, the design and implementation of VLEs populated with animated characters are complex tasks, involving an iterative development process with a range of stakeholders.

The VICTEC (Virtual ICT with Empathic Characters) project uses synthetic characters and Emergent Narrative as an innovative means for children aged 8-12 years to explore issues surrounding bullying behaviour. FearNot (Fun with Empathic Agents to Reach Novel Outcomes in Teaching), the application being developed in VICTEC, is a 3D VLE featuring a school populated by 3-D self-animated agents representing various character roles involved in bullying behaviour through improvised dramas.

The main focus of this paper is to consider the different perspectives and empathic reactions of adult and child populations in order to optimise the design and ultimately usage of a virtual world to tackle bullying problems. The perspective that we have taken is that if children empathise with characters a deeper exploration and understanding of bullying issues is possible [3]. Whilst it is less critical for other stakeholder groups, such as teachers, to exhibit similar empathic reactions to children, the level of empathy and its impact on agent believability [9] has strong implications for

teacher's usage of such applications for classroom-based teaching. As relatively few teachers have exposure to sophisticated, innovative educational environments they may have inappropriately low or high expectations of an unknown technology. To offer an alternative perspective, the views and empathic reactions of discipline-specific experts were also obtained to enable us to gain the view of stakeholders who were "early adopters" of VLEs and synthetic characters.

The main questions we are seeking to answer in this paper are: Are there differences in the views, opinions and attitudes of children and adults? And, if there are differences, what are their design implications? In the first section we discuss development and technical issues for our early prototype. In the second section we discuss our approach to using this prototype. We then present the results and discuss our findings.

FearNot: Technical and Development Issues

FearNot is an interactive 3D environment that allows children to interact and influence the events happening in a story featuring bullying scenarios.

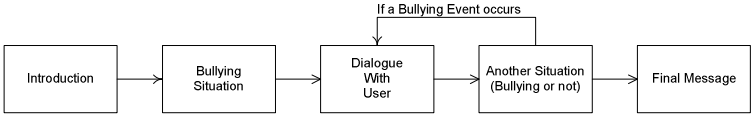

Fig. 1. Interacting with FearNot

Fig. 1. presents a schematic view of the episodes of an interaction with FearNot. After each of these, the victim starts a dialogue probing for user help. This dialogue concludes with the selection of a coping strategy which influences the course of the events in the episodes ahead. The episodes are not pre-scripted, and arise from the actions of the characters in the story that act autonomously, performing their roles in character (as a bully, a victim, a bystander or a bully-victim).

The FearNot Trailer Approach

Fig. 1. identifies how interaction will occur with the final version of FearNot. However, we needed to gain feedback from users and stakeholders at an early stage in the lifecycle when there was no stable version of the final product and where development emerges as a response to research findings. Recognising this as an issue early in the design of FearNot prompted the creation of the trailer approach which is a snapshot vision of the final product, similar to the trailers seen for movies, where the major themes of a film are revealed. Also similar to a movie trailer using real movie clips, our trailer used a technology closely resembling the final application.

The trailer depicts a physical bullying episode containing 3 characters, Luke the bully, John the victim and Martina the narrator. The trailer begins with an introduction to the main characters, Luke and John and subsequently shows Luke knocking John's pencil case off the table and then kicking him to the floor. John then asks the user what he should do to try and stop Luke bullying him and arrives at 3 possible choices: 1) Ignore Luke, 2) Fight back, 3) Tell someone that he trusts such as his teacher or parents.

Developmental constraints of the application did not allow us to include the dialogue phase in the first trailer developed. Nonetheless, the importance of the dialogue phase for the overall success of the application required us to include it. As an advance, we built a dialogue phase between the bullying situation and the final message. We are using the Wizard of OZ technique [1] to iterate on our dialogue system and adjust the user interaction during this stage.

Re-Using the Trailer Technology for FearNot

The re-use of the trailer technology in the final application is possible due to the agent-based approach [14] we adopted for the FearNot application, as depicted in Fig. 2. Several Agents share a virtual symbolic world where they can perform high-level acts. These can be simply communicative acts or can change the symbolic world, which contains domain-specific information, in this case, information regarding bullying situations. A specific agent must manage the translation of such symbolic information and the agents' acts to a particular display system. Such a process is outlined in Fig. 2. (the ellipse outlines the technology used in the trailer).

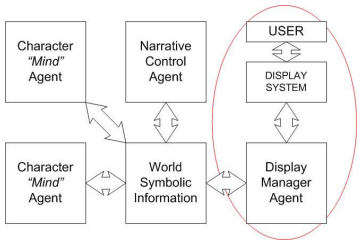

Fig. 2. FearNot Agent-Based Approach

Popular approaches to implementing environments with self-animated characters suffer from being too low-level (e.g. [4]), solely focusing on a realistic display of character behaviour and directly connecting character architecture and display system. Although PAR [2] constitutes an example of a higher-level approach, this is a humanoid-dependent language and too complex for our needs. Flexible Improv [7] systems are becoming the de facto standards in the field, however, current implementations make it impossible to achieve rich high-level character behaviour. Therefore, the approach we have chosen has two different levels: 1) the higher-level act and 2) the lower-level view-action (which then renders to a specific display system).

The modular agent-based approach enables us to work in parallel on components. Whilst defining the act ontology which coordinates agent communication, we were able to focus on the lower-level graphical language definition that was used to implement the trailer. This consists of a scripted sequence of view-actions, depicting the situation and emulating the character acts. For this approach to integrate high-level acts and low level view-actions we assumed a simple trailer-bounded ad-hoc high-level language. Yet, the trailer served equally as a validating tool for our approach.

The trailer was implemented as a Java applet running inside a browser, as demonstrated in Fig. 3. A simple View Manager was developed which emulated character acts and ran a sequence of view actions to a display system, implemented with the use of a proprietary game engine. These provide excellent tools for prototyping, and were sufficiently stable and robust to fully implement the FearNot application. The view action language aims to minimize the effort required to change other displays.

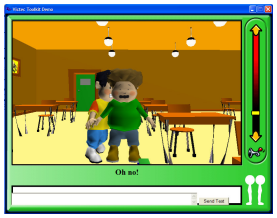

Fig. 3. A screenshot of the FearNot Trailer, Displaying a Physical Bullying Situation**.**

The Trailer Experiment

The trailer was evaluated using a questionnaire applicable for children and adults and focused on character attributes (voice believability, likeableness, conversation content, movement) storyline (believability), character preferences and empathy (sorrow and anger). Measurement was predominantly by a 5 point Likert scale.

225 trailer questionnaires were completed by 128 children from schools in England and Portugal (57%), 54 experts (24%) and 43 teachers / educationalists (19%). Table 1 illustrates the gender and age distribution of the sample.

Table 1: Gender and Age Distribution of the Sample

Gender / Age	Frequency	%
Male child	64	29
Female child	63	28
Male adult	49	22
Female adult	46	21

Teachers in the sample were from a wide range of primary and secondary schools in the South of England. They were predominantly female (90%), aged between 25 to 56. The children, aged from 8-13 (x=9.83, SD=1.04), were from primary schools located in urban and rural areas of Hertfordshire, UK (47%) and Cascais, Portugal

(53%). The experts were attendees at the Intelligent Virtual Agents workshop in Kloster Irsee, Germany and were predominantly male (80%) and under 35 (67%). Table 2 illustrates the procedure used for showing the FearNot trailer and completion of the trailer questionnaire.

Table 2. FearNot Demonstration and Questionnaire Completion

Sample	Procedure
Experts	Trailer shown and questionnaire explained to whole audience. Completed as part of conference workshop.
Teachers	Trailer shown and questionnaire explained to whole audience and completed as part of a teacher workshop, Germany, and a teacher seminar, Bristol, UK.
Children	Trailer shown and questionnaire distributed to all children. The questionnaire was explained to the whole class and the researcher then guided the children through each question ensuring that they understood each question.

3 Results

Frequency distributions were examined using histograms for questions that employed Likert scales to ensure that the data was normally distributed. Chi-square tests in the form of cross-tabulations were calculated to determine relationships between different variables for categorical data. One way analysis of variance (ANOVA) using Scheffe's post-hoc test were carried out to examine mean differences between the 3 stakeholder groups according to questionnaire responses using the Likert scale.

Character Attributes

There were significant differences between the stakeholder groups and views of the believability (F=6.16, (225, df=2), p=0.002), realism F=9.16, (225, df=2), p=0.00) and smoothness (F=12.96, (224, df=2), p=0.00) of character movement with children finding character movement more believable, realistic and smooth compared to adults, see table 3. No significant gender differences were revealed for the believability or smoothness of character movement. An independent samples T-test revealed significant gender differences for the realism of character movement (t=2.91, 225, df=220, p=0.004). Females (m=3.17) found character movement significantly more realistic than males (m=3.63).

Significant differences were found for the believability (F=11.82, (224, df=2), p=0.00) and likeability (F=9.35, (221, df=2), p=0.00) of character voices, with teachers finding voices less believable and likeable. An independent samples T-test revealed significant differences between gender and believability of voices (t=-2.65, 221, df = 219, p=0.01). Females (m=2.53) found the character voices less believable than males (m=2.15).

Table 3. Mean Group Values for Character Movement

Mean values for character movement (1 = positive, 5 = negative)	Child	Teacher	Expert
Believable	3.04	3.77	3.31
Realistic	3.11	3.81	3.76
Smooth	2.82	3.88	3.17

Storyline

No significant differences were found between children, teachers and experts or gender for the believability of character conversation and interest levels of character conversation. Significant differences were found in the views of the storyline believability (F=10.17, (224, df=2), p=0.00) and the true-to-lifeness of both the character conversation (F=6.45, (223, df=2), p=0.002) and the storyline (F=14.08, (225, df=2), p=0.00), with children finding the conversation and storyline more true to life and believable.

There were significant differences between child, expert and teacher views in relation to the match between the school environment and the characters (F=10.40, (220, df=2), p=0.00). Children were significantly more positive towards the match between the school environment and characters compared to teachers (Fig. 4.). Children were also more positive about the School appearance (F=22.08, (224, df=2), p=0.00)

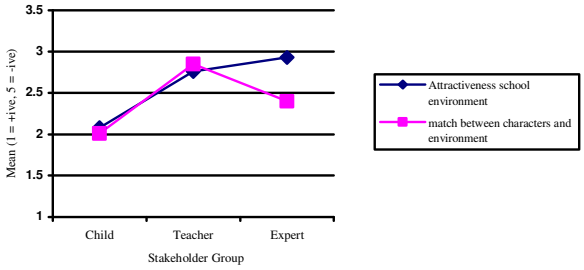

Fig. 4. Mean Group Differences for the Attractiveness of the Virtual School Environment and the Match between Characters and the School Environment.

Character Preferences

Significant gender differences were found for children only when character preference was considered, (x=20.46, N=195, df = 2, p=0.000) indicating no overall gender preferences for John (the victim) but that significantly more female children preferred

Martina (the narrator), and significantly more male children preferred Luke (the bully).

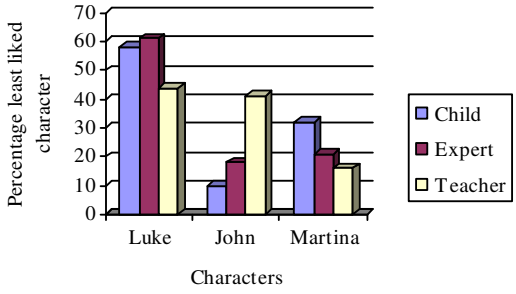

Fig. 5. Percentages for Least Liked Characters According to Children, Experts and Teachers.

Significant differences were revealed between teachers, children and experts for the least liked character (x=18.35, N=201, DF=4, p=0.001) (Fig. 5). Significantly more teachers least liked John (the victim), compared to children and experts. Female adults disliked John (the victim) more than children and experts (37%), and male children disliked Martina the most (52%). 78% of female children disliked Luke the most closely followed by the male adults, 60% of whom disliked Luke the most.

There were no significant differences between children, teachers and experts in which of the characters they would like to be. However, significant differences emerged when gender and age were taken into account. 40% of male children chose to be John and 88% of female children, followed by 73% of female adults chose to be Martina. No female children (n=59) chose to be Luke compared to 44% of male children who chose to be Luke. Male adults did not wish to be John, with 51% wishing to be Martina and 34% wanting to be Luke.

Empathy

Significant differences were found between children, experts and teachers for expressing sorrow (x=10.33, N=216, df=2, p=0.006) and anger (x=26.13, N=213, df=2, p=0.000). Children were most likely to feel sorry or angry, see table 4, however, whilst most children felt sorry for the victim, significantly more experts felt sorry for Luke (the bully) compared to teachers and children (x=13.60, N=175, df = 2, p=0.001). Significant age and gender differences emerged, (x=27.42, N=210, df=3, p=0.000) where more female children expressed anger towards the characters compared to adults. This anger was almost exclusively directed at Luke (90%).

Table 4. Empathic Responses to Characters

	Luke (Bully)	John (Victim)	Martina (Narrator)

% Felt sorry for characters			
Child	5.0	95.0	0.0
Expert	27.0	74.0	0.0
Teacher	7.0	93.0	0.0
% anger towards characters			
Child	85.0	13.0	2.0
Expert	70.0	30.0	0.0
Teacher	73.0	9.0	18.0

4. Discussion

The main aims of this paper were to consider whether there were any differences in the opinions, attitudes and empathic reactions of children and adults towards FearNot, and whether differences uncovered offer important design implications for VLEs addressing complex social issues such as bullying.

A summary of the main results revealed that (1) Children were more favourable towards the appearance of the school environment, character voices, and character movement compared to teachers who viewed these aspects less positively. (2) Children, particularly male children found the conversation and storyline most believable, realistic and true-to-life. (3) No significant differences were revealed between children and adults for most-liked character, although teachers disliked 'John' the victim character the most compared to children and experts. (4) Children preferred same-gender children, with male characters disliking the female narrator character; female children disliking the male bully; and children choosing to be same-gender characters. (5) Children, particularly females, expressed more empathic reactions (feeling sorry and/or angry for the characters) compared to adults.

Throughout the results, a recurrent finding was the more positive attitude and perspective of children towards the FearNot trailer in terms of the school environment, character appearance, character movement, conversation between the characters and engagement with the storyline. Children's views expressed were typically within the positive range under 3 (scale 1 to 5). Children's engagement and high level of empathic reactions to the trailer are encouraging as they indicate the potential for experiential learning with children clearly having a high level of belief and comprehension of a physical virtual bullying scenario.

The opposite trend seems to have emerged from the teacher responses, where teachers clearly have high expectations that are not met or possibly are unable to engage effectively with such a novel system such as FearNot. Experts were positive about the technical issues of FearNot such as the physical representation of the characters. However, they failed to engage with the educational theme of bullying and applied generic criteria ignoring the underlying domain. Thus, whilst character movement and voices were rated highly, limited levels of empathy were seen with experts taking a somewhat voyeuristic approach.

We consider that self-animated characters bring richness to the interaction essential to obtain believable interactions. Nevertheless, danger of unbelievable "schizophrenic" behaviour [10] is real, and enormous technical challenges emerge. To overcome these, constant interaction between agent developers and psychologists is crucial. Furthermore, the use of a higher-level narrative control arises as another technical challenge that is being explored, towards the achievement of story coherence that characters are ineffective, on their own, to attain. The use of a cartoon style offers a technical safety net that hinders some jerkiness natural to experimental software. Furthermore, the cartoon metaphor already provides design decisions that most cartoon-viewing children accept naturally.

Conclusion

The trailer approach described in this paper enabled us to obtain a range of viewpoints and perspectives from different stakeholder groups. Further, the re-use of the technology for the trailer within the final application highlights the benefits of adopting an agent-based approach, allowing the development of a mid-tech prototype that can evolve into the final application. Input from a range of stakeholders is essential for the development of an appropriate application. There must be a balance between true to life and acceptable (by teachers and parents) behaviours and language. The use of stereotypical roles (e.g. typical bully) can bias children's understanding and simple design decisions can influence the children's perception of a character (e.g., Luke looks a lot "cooler" than John). The educational perspective inhibits the applicability of the «game» label to the application, which most of the time children instantly apply to an application like this. Achieving a balance between the expectations of all stakeholders involved may be the hardest goal to achieve over and above technical challenges.

References

1. Anderson, G., Höök, K., Paiva, A., & Costa, M. (2002). Using a Wizard of Oz study to inform the design of SenToy. Paper presented at the Designing Interactive Systems.
2. Badler, N., Philips, C., & Webber, B. (1993). Simulating humans. Paper presented at the Computer graphics animation and control, New York.
3. Dautenhahn, K. (2002). Design spaces and niche spaces of believable social robots. Paper presented at the International Workshop on Robots and Human Interactive Communication.
4. Magnenat-Thalmann, N., & Thalmann, D. (1991). Complex models for animating synthetic actors. Computer Graphics and Applications, 11, 32-44.
5. Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents. Cognition and Instruction, 19(2), 177-213.

6. Nass, C., Isbister, K., & Lee, E. (2001). Truth is beauty: researching embodied conversational agents. Cambridge, MA: MIT Press.
7. Perlin, K., & Goldberg, A. (1996). Improv: A system for scripting interactive actors in virtual worlds. Paper presented at the Computer Graphics, 30 (Annual Conference Series).
8. Pertaub, D.-P., Slater, M., & Barker, C. (2001). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. Presence: Teleoperators and Virtual Environments,, 11(1), 68-78.
9. Prendinger, H., & Ishizuka, M. (2001). Let's talk! Socially intelligent agents for language conversation training. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 31(5), 465-471.
10. Sengers, P. (1998). Anti-Boxology: Agent Design in Cultural Context. PhD Thesis, Technical Report CMU-CS-98-151, Carnegie Mellon University.
11. Wooldridge, M. (2002). An Introduction to Multiagent Systems. London: John Wiley and Sons Ltd.