

Designing Empathic Agents: Adults vs. Kids

Lynne Hall¹, Sarah Woods², Kerstin Dautenhahn², Daniel Sobral³, Ana Paiva³,
Dieter Wolke⁴, Lynne Newall⁵

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, Porto Salvo, Portugal ana.paiva@inesc-id.pt

⁴ Jacobs Foundation, Zurich, Switzerland, dieter.wolke@jacobsfoundation.org

⁵ School of Informatics, Northumbria University, Newcastle, UK,
lynne.newall@unn.ac.uk

Abstract. An evaluation study using a trailer approach of a Virtual Learning Environment populated with animated characters focusing on physical bullying 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 from 127 children and 95 adults revealed that children expressed the most favourable views towards the characters and the highest levels of believability towards the bullying storyline. Results are discussed in terms of the importance of child-informant processes in leading the design of child-based systems and the use of animated cartoon characters alongside storyline narratives to incite engaging interactions.

Introduction

Virtual Learning Environments (VLEs) populated with animated characters offer children a safe environment where they can explore and learn through experiential activities [8,11]. Animated characters offer a high level of engagement, through their use of expressive and emotional behaviours [9], 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 aims to use synthetic characters and Emergent Narrative as an innovative means for children aged 8-12 years to explore the different 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. FearNot will potentially be used within Personal and Social Health Education (PHSE) curriculum work to deal with the difficult and sensitive social issues involved with bullying.

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 [4]. 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 [12] has strong implications for teacher's usage of such applications for classroom-based teaching. However, the majority of teachers have had limited contact with sophisticated, innovative educational environments and due to this lack of exposure and awareness 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 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 do these have any important design implications for embodied agents? In the first section we discuss the development and technical issues related to the approach we have used to enable us to gain stakeholder feedback with an early prototype. In the second section we discuss results gained using this approach with three different stakeholder groups. In the third section we discuss the findings and their design implications for developing empathic agents for VLEs for an educational context for children. Finally, some ideas for future work are presented and the principal conclusions.

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. Figure 1 presents a schematic view of an interaction with FearNot episodes, where after each episode, 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. If the victim used a successful coping response then a final educational message about bullying prevention is displayed. Conversely, if an unsuccessful coping mechanism is selected, further bullying situations are depicted, corresponding with the victim asking the user for help. 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).

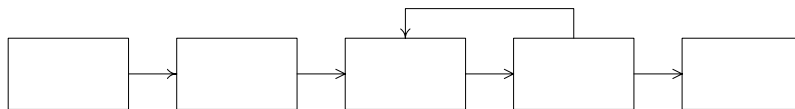


Fig. 1. Interacting with FearNot

The FearNot Trailer Approach

The trailer approach used provides a snapshot vision of the final product, similar to the trailers seen for movies, where the major themes of a film are revealed. Similar to a movie trailer using real movie clips, our trailer used a technology closely resembling the final application.

The trailer depicts one physical bullying episode based on a 2D storyboard (developed using storyboarding software [6] developed by experts in bullying research in conjunction with teachers and pupils. All characters and animations, places and objects were transposed to 3D by a team of designers. The dialogues were obtained through the recording of real voices.

The physical bullying episode contains 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 became 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).

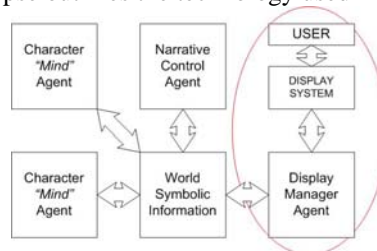


Fig. 2. FearNot Agent-Based Approach

Popular approaches to implementing environments with self-animated characters suffer from being too low-level (e.g. [7]), solely focusing on a realistic display of

character behavior and directly connecting character architecture and display system. PAR [2] constitutes an example of a higher-level approach. However, this is a humanoid-dependent language which we considered too complex for our needs. Flexible Improv [10] systems are becoming the de facto standards in the field, although we believe that current implementations make it impossible to achieve a rich high-level behavior from the characters. 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).

As previously mentioned, this modular agent-based approach enables us to work in parallel on different components. Although we were defining the act ontology which coordinates agent communication, we were able to focus on the definition of the lower-level graphical language which was extensively used to implement the trailer. This consists of a scripted sequence of view-actions, depicting the situation and emulating the character acts. For our first 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. However, the view action language was done to minimize the effort required to change other displays.



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

Using the Trailer

The overriding research question that we are seeking to study is does the population of a VLE with synthetic agents provide a suitable mechanism to permit the exploration of bullying within the context of a formal educational environment. This question was considered through using a single episode trailer depicting the exploration of a physical bullying scenario within FearNot and obtaining perspectives and empathic reactions from three different stakeholder groups.

Approach

A questionnaire applicable for children and adults was designed in order to evaluate aspects of FearNot. These were predominantly measured according to a 5 point Likert scale. The questions focused on a number of topics including synthetic character attributes, the storyline, character preferences and empathy for the characters.

Table 1. FearNot Demonstration and Questionnaire Completion

Sample	Procedure
Experts	FearNot trailer shown and trailer questionnaire explained to whole audience and completed as part of conference workshop.
Teachers	FearNot trailer shown and trailer questionnaire explained to whole audience and completed as part of a teacher workshop, Germany, and a teacher seminar, Bristol, UK.
Children	FearNot trailer shown and trailer 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.

225 trailer questionnaires were completed by 128 children from schools in England and Portugal (57%), 54 experts (24%) and 43 teachers / educationalists (19%), using the procedure outlined in Table 1. Table 2 illustrates the gender and age distribution of the sample.

Table 2: 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. The teachers were predominantly female (90%) and aged between 25 and 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%).

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 of character movement ($F=6.16$, (225, $df=2$), $p=0.002$), the realism of character movement ($F=9.16$, (225, $df=2$), $p=0.00$) where children reported that the character movement was significantly more believable and realistic compared to teachers and experts.

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 between groups for the smoothness of character movement ($F=12.96$, (224, $df=2$), $p=0.00$). Teachers thought that the character movement was significantly more jerky than children ($p=0.00$) and experts ($p=0.01$).

Significant differences were found for the believability of character voices ($F=11.82$, (224, $df=2$), $p=0.00$) where teachers ($m=3.00$) reported significantly lower believability for the character voices compared to children ($m=2.24$, $p=0.00$), and experts ($m=2.02$, $p=0.00$). An independent samples T-test revealed significant differences between gender and believability of character voices ($t=-2.65$, 221, $df = 219$, $p=0.01$). Females ($m=2.53$) found the character voices less believable than males ($m=2.15$).

There were significant differences for the likeability of character voices ($F=9.35$, (221, $df=2$), $p=0.00$) where teachers ($m=2.88$) reported significantly less likeability for the character voices compared to children ($m=2.05$, $p=0.00$) and experts ($m=2.09$, $p=0.003$). No significant gender differences.

Storyline and Conversations

Significant differences were found between groups for views of the believability of the storyline ($F=10.17$, (224, $df=2$), $p=0.00$) where children found the storyline significantly more believable than teachers.

Significant differences emerged in the views of the true-to-lifeness of the storyline ($F=14.08$, (225, $df=2$), $p=0.00$) where teachers found the storyline significantly less believable compared to children ($p=0.00$), and experts ($p=0.006$) who found it very believable. Overall, there were no significant differences for gender and true-to-lifeness of the 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 true-to-lifeness of character conversation ($F=6.45$, (223, $df=2$), $p=0.002$) where teachers ($m=2.67$) found the character conversation significantly more false and less true to life compared to children ($m=2.02$, $p=0.005$) and experts ($m=1.94$, $p=0.007$).

Character Preferences

Significant gender differences were found for children only when character preference was considered, ($\chi=20.46$, $N=195$, $df = 2$, $p=0.000$) 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).

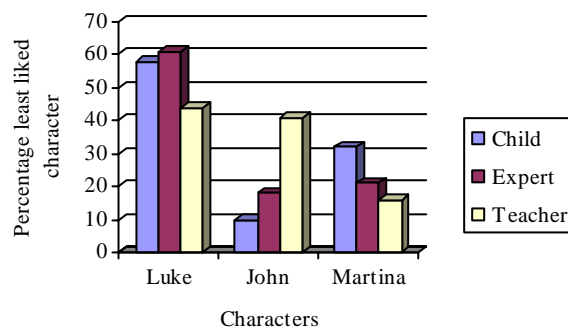


Fig. 4. Percentages for Least Liked Characters According to Children, Experts and Teachers.

Significant differences were revealed for the least liked character and teachers, children and experts ($\chi=18.35$, $N=201$, $DF=4$, $p=0.001$) see figure 4. 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 affective empathy ($x=10.33$, $N=216$, $df=2$, $p=0.006$). More children (80%) expressed feeling sorry for the characters compared to teachers and experts (70%). Affective empathy was only expressed for Luke and John, not for Martina. 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 differences were found between children, experts and teachers for cognitive empathy ($x=26.13$, $N=213$, $df=2$, $p=0.000$) where more children (71%) expressed cognitive empathy towards characters compared to experts (47%) and teachers (28%). Significantly more experts expressed anger towards John compared to children and teachers and significantly more teachers expressed anger towards Martina compared to experts and children. 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%).

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 (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) All groups reported that they found the character conversation believable although teachers viewed the conversation as the least realistic compared to children and experts, (3) Children and in particular male children viewed the storyline as the most believable and true to life. Teachers were the least favourable, (5) No significant differences were revealed between children and adults for most-liked character, (6) Teachers disliked 'John' the victim character the most compared to children and experts. A child gender specific result was revealed for least liked character where male children disliked the female narrator character Martina, and female children disliked the Luke, the male bully, (7) A child gender specific result was revealed for prime character where no female children wanted to be Luke the bully and male children chose to be John the victim or Luke, and (8) Children expressed more affective empathy (feeling sorry for the characters) and cognitive empathy (anger) compared to adults. Female children felt the most sorry and the most anger for characters.

An encouraging result for FearNot was that children viewed the school appearance, character voices and character movement favourably. The finding that male children reported the highest believability for the character voices would be expected as the focus of the bullying scenario was a physical bullying incident that occurred between John the victim and Luke the bully and therefore focused on male voices as

opposed to female voices. Less encouraging was that teachers in particular were not overly positive towards character and school appearance.

A positive result related to the believability of the storyline and character conversation across children, experts and teachers. Male children in particular were receptive to the content of the bullying storyline and this can be attributed to the fact that the story was a physical bullying scenario occurring between two male characters. Future research must consider a relational bullying scenario which is more gender-specific for females in order to determine whether females would express higher engagement and believability towards this type of bullying behaviour. Teachers expressed less storyline believability and this could be ascribed to the fact that some teachers may have found the storyline too stereotypical and not realistic within a real school environment. It could also indicate that adult teachers are not primed towards the nature of physical bullying as they are no longer subjected to this kind of bullying as an adult and may identify more with relational bullying.

Gender specific results were revealed for least liked characters and prime character in the scenario. This result was most clear-cut for the children compared to adults, where girls least liked and did not want to be the male bully character, Luke, and boys did not like or want to be Martina, the female narrator. These results support previous findings related to identification with ones own gender and suggest the need for focused scenarios offering children same gendered synthetic characters to identify with. The above results could also be explained by the nature of the bullying scenario as more males are involved in physical bullying situations compared to females where relational and verbal bullying is more prevalent [3].

A limitation of the FearNot trailer related to the feedback received by teachers for the least liked character. 41% of teachers least preferred John the victim compared to 10% of children and 18% of experts. This finding could be explained in light of the perceived stereotypical nature of John's behaviour by the teachers and they may believe that characters such as John are not realistic of the types of children who are victims of physical bullying at school. Alternatively, teachers may have least liked John as they thought he should have defended himself more and tried to prevent the bullying happening by being proactive. Typical comments expressed by teachers about John were: 'John was rather unbelievable as he didn't seem introverted enough', 'not believable enough as a victim', 'too much of a stereotype', 'he was overbearing', 'John was a bit annoying'

Children were more likely to feel empathy, both affective (sorrow) and cognitive (anger), for the characters in the scenario compared to adults. Female children expressed the most affective and cognitive empathy compared to male children, experts and teachers. An explanation for this could be that female children are shocked by the severe nature of physical bullying as this is not part of their usual social experience at such an explicit level. Teachers expressed the lowest levels of both affective and cognitive empathy. This could be explained by their overall lack of willingness to suspend disbelief and engage with the application.

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 [13] 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. Crick, N. R., & Grotper, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, 66, 710-722.
4. Dautenhahn, K. (2002). Design spaces and niche spaces of believable social robots. Paper presented at the International Workshop on Robots and Human Interactive Communication.
5. Dautenhahn, K., & Bond, A. H. (Eds.). (2002). *Socially intelligent agents: Creating relationships with computers and robots*. Massachusetts, USA: Kluwer Academic Publishers.
6. Immersive Education. (2001). Kar2ouche. Oxford.
7. Magnenat-Thalmann, N., & Thalmann, D. (1991). Complex models for animating synthetic actors. *Computer Graphics and Applications*, 11, 32-44.
8. 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.
9. Nass, C., Isbister, K., & Lee, E. (2001). *Truth is beauty: researching embodied conversational agents*. Cambridge, MA: MIT Press.
10. 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).
11. 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.
12. 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.
13. Sengers, P. (1998). *Anti-Boxology: Agent Design in Cultural Context*. PhD Thesis, Technical Report CMU-CS-98-151, Carnegie Mellon University.
14. Wooldridge, M. (2002). *An Introduction to Multiagent Systems*. London: John Wiley and Sons Ltd.