

# Using Theory of Mind to investigate empathic engagement with synthetic characters

Lynne Hall<sup>1</sup>, Sarah Woods<sup>2</sup>, Ruth Aylett<sup>3</sup> and Ana Paiva<sup>4</sup>

<sup>1</sup>School of Computing and Technology, University of Sunderland, Sunderland, SR6 0DD, UK, lynne.hall@sunderland.ac.uk

<sup>2</sup>Adaptive Systems Research Group, University of Hertfordshire, Hatfield, Herts, AL10 9AB, UK. s.n.woods@herts.ac.uk

<sup>3</sup>Mathematics and Computer Science, Heriot-Watt University, Edinburgh, EH14 4AS, UK, ruth@macs.hw.ac.uk

<sup>4</sup>INESC-ID, Av. Prof. Cavaco Silva, Tagus Park, 2780-990 Porto Salvo, Portugal, ana.paiva@inesc-id.pt

This paper is concerned with the simulation of human-like capabilities in synthetic characters within the domain of Personal and Social Education. Our aim was to achieve socially meaningful and engaging interactions with children in the 8-12 age group to enable an exploration of bullying and coping strategies. In this paper, we consider the engagement between the interacting partners, focusing particularly on the affective and empathic aspects of this relationship. We have used Theory of Mind to enable us to evaluate children's understanding of social scenarios and the thinking of others. The results from this 345 children study highlight that children are able to recognise and interpret affect in synthetic characters and are empathically engaged with the characters in the scenarios.

## 1 Introduction

Successful virtual environments provide an interface to the user that convinces them to be “in” a different world. Synthetic characters appear as part of these virtual environments as the entities that act, and by their actions the world changes, the story evolves and provides the user with the illusion of life. Empathic engagement is the fostering of emotional involvement intending to create a coherent cognitive and emotional experience resulting in the development of empathic relations between the user and the character. Through empathic engagement the user perceives and models the emotion of the character experiencing an appropriate emotion as a consequence. Empathic engagement is realized through affective interaction, that is, interaction that relates to, arises from, or influences feelings and emotion [31]. Through empathically engaging the user, we can enable them to experience the character's emotions and problems in a distanced way, while at the same time being engaged in what happens to the characters [3, 11, 22, 28].

Empathising with characters permits a deeper exploration and understanding of sensitive social and personal issues [7]. Results highlight the potential of synthetic characters for empathic engagement in theatre [5], educational drama and story telling applications e.g. Ghostwriter [34], Virtual Puppet Theatre [1, 18]; and Personal, Health and Social Education [5, 11, 22]. Empathic engagement with synthetic characters provides users with a safe environment where they can explore and learn through experiential social and emotional activities [23, 30]. It can lead to real-life impacts such as the development of constructive solutions, e.g. Carmen's Bright Ideas [21, 22].

Evaluating affective interaction and empathic engagement with synthetic characters provides significant challenges with no clear consensus on how to accurately measure and interpret affect. A number of alternative techniques have been used, including questionnaires [20], discussion groups [15], innovative input devices [27] and bio-measurements [33] with varying degrees of success. Here, we propose the development of a Theory of Mind assessment, to investigate children's affect and empathic engagement with synthetic characters.

Theory of mind refers to the ability to predict and explain the behaviour and feelings of others based on reference to mental states, beliefs, desires and percepts [2, 40]. Theory of Mind concepts are employed as a strategy to assist in the understanding of social scenarios, the thinking and intentions of others [36]. It provides a potential framework to investigate a child user's understanding of characters' actions and

behaviours, and permits an exploration of the child user's empathic engagement with the synthetic characters.

This paper discusses FearNot! (Fun with Empathic Agents to Reach Novel Outcomes in Teaching), a school-based VLE populated by synthetic characters representing the various characters in a bullying scenario. FearNot! was developed in the VICTEC (Virtual ICT with Empathic Characters) project [3], that applied synthetic characters and emergent narrative to Personal and Health Social Education (PHSE) for children aged 8-12, in the UK, Portugal and Germany, using 3D self-animating characters to create improvised dramas. FearNot! aims to enable children to explore bullying issues, and coping strategies, through providing synthetic characters that affectively engaged with the user and that through their appearance, behaviours, affect and features were able to trigger empathic relations with the user.

In this paper, we present the results from the development of a Theory of Mind (ToM) assessment to evaluate a scripted version of the FearNot! prototype, further discussed in [3, 12]. This occurred at an evaluation event called "Virtually Friends" held at the University of Hertfordshire, UK, in June 2004, involving 345 children aged 9-11 years. As FearNot! has been developed to explore bullying and coping strategies, a key aim of the evaluation was to determine whether user characteristics are reflected in user choices regarding two bullying scenarios, in terms of mental representations as provided through assessing children's ToM abilities.

There are abundant studies within the psychological literature that have used ToM assessments to provide evidence that autistic children have under-developed Theories of Mind [4]. For this study, we were interested in developing a Theory of Mind assessment to investigate particular child user characteristics related to bullying. Research studies about bullying roles traditionally focused on the simple dichotomy of the bully (perpetrator) and victim roles. However, more recent research findings have indicated that this is an oversimplification of human behaviour, and there is growing awareness of the need to consider other bullying roles, including bully/victims (those children who both bully other children and become victimized) [6] [38], bystanders / neutrals (onlookers who do not intervene in situations), defenders (individuals who defend the victim) and bully assistants (individuals who help the bully) [35]. A further advancement within the field of bullying research pertains to the need to distinguish between physical (direct) bullying incidents (e.g. being hit, kicked, beaten up, threatened) and relational (indirect) bullying behaviour (e.g. social exclusion, malicious rumour spreading) [41].

Section 2 introduces FearNot! and briefly discusses the simulation of human-like capabilities in the synthetic characters inhabiting FearNot!. Section 3 discusses the approach taken to evaluating socially meaningful and engaging interactions with FearNot!, outlining the Theory of Mind assessment. Section 4 briefly outlines the study methodology, and the main results are presented in section 5. In section 6 the results are discussed and some brief conclusions are presented.

## **2 FearNot!**

FearNot! depicts bullying incidents in the form of an episodic virtual drama, see figure 1. Two different scenarios were developed for the FearNot! prototype, a physical bullying scenario and relational bullying scenario each comprised of several episodes to create a believable and engaging story. In this paper, we are focusing on the child's interactions with the physical scenario. The child user views the bullying incidents that take place between autonomous agents in a virtual school and acts as the 'invisible friend' to the victimised character in between episodes, providing help and advice about how to deal with the bullying incidents. Each episode is framed by an introduction segment at the start of the episode, and a reflective interactive segment at the end.



**Figure 1: Screenshots from FearNot!**

Creating believable characters that give the illusion of life allowing the user's suspension of disbelief is essential for FearNot! because not only is believability a precondition for an empathic relationship between characters and users, but also, because we are dealing with a dramatic environment, and to tell a story our characters must be believable. With the FearNot! characters, we have attempted to provide coherent and compelling behaviours, conversations and storylines, providing characters with sufficient *depth* to encourage children to believe the drama.

Interaction with FearNot! is achieved through scenarios in which the main purpose of the communication was to engage in social interaction as opposed to accomplishing a task as efficiently as possible. During a FearNot! interaction the child user is asked to take responsibility for the victim character, by providing support and advice about how to deal with different bullying scenarios as they would to another child. This approach enables the child to become engaged in and affected by what happens to the characters [3, 11, 22, 28], and can lead to empathic engagement.

FearNot! characters each have a role in the episode, such as the bully, the victim, bully-victim, defender and bystanders. Each character has an internal architecture, see [29], that is a subset of the taxonomy of emotions of Ortony, Clore and Collins [26], known as the OCC. This has been implemented in order to give characters the appropriate emotional responses and expressive behaviour.' Using the OCC, we modelled the different bullying roles by specifying different goals and emotional reactions to events. Using these parameters, the OCC appraisal process leads to different emotions being triggered and thus to different behaviours and expressions emerging. A conceptual planning structure was used, which takes on the goals of the synthetic character and is influenced by the character's emotions at a certain time. This process leads to a particular behaviour, characterizing the role by producing adequate actions. FearNot! characters were designed to be reactive to user interactions, responding either immediately or by changing the mode of behaviour. However, they were also designed to be autonomous, thus resulting in behaviour that is not completely predictable and that could only be influenced by the user to a certain degree. For example, the user may advocate that the victim should tell someone about the bullying, however, before the victim has a chance to do this they may be directly bullied again and may choose to hit back at the bully.

A key issue for FearNot! was the physical appearance of the characters, their perceivable actions and their expressions. The child as informant approach has been extensively used in the development of FearNot! This had a number of significant design impacts, including the development of synthetic characters with a cartoon style appearance. Not only were children most in favour of cartoon characters [42], but additionally this offers a technical safety net in that highly naturalistic behaviour is not expected in cartoons making the element of jerkiness natural to experimental software less of an issue for children. Furthermore, the cartoon metaphor already provides design decisions that most cartoon-viewing children accept naturally.

Of particular importance for the FearNot! characters was the expression of emotions and the need to ensure that these made characters believable and that they appeared life-like. According to Thomas and Johnston [39], animators from Disney, there are three important points when expressing emotions:

- the emotional state of the character must be clearly defined, in such a way that is undoubtedly perceived by the viewer
- the emotional state affects the reasoning process and consequences must be perceivably reflected in the actions of the characters

- emotions can be accentuated or exaggerated, to clearly communicate to the viewer, the emotional state of the character.

For the FearNot! characters, clear definition of the emotional state of the character has been realized mainly through facial expressions of emotions, with the faces of the characters reflecting their emotional reaction. A design decision was made to provide FearNot! characters with only crude facial expressions (depicting anger, happiness, sadness, fear and a neutral state) and limited gestures [3] and this was evaluated using a Classroom Discussion Forum [14]. We found that children in the 8-12 age group had relatively little to say about emotions, either their own or those of the characters. However, children clearly understood the characters' expressions having no problems in identifying the emotional state of the characters and found the appearance appropriate. The simple cartoonish approach supported mediation by expression with children empathizing with the characters, feeling both sympathy and anger.

### **3 Evaluating FearNot! using a Theory of Mind Assessment**

Synthetic characters offer a high level of engagement, through their use of expressive and emotional behaviors [24]. However, it is not clear how to accurately measure this engagement, and a number of different techniques have emerged. For VICTEC, a number of evaluation approaches have been developed including Classroom Discussion Forums [14], Character Evaluation Questionnaires [16], Interaction Analysis [13], etc. Here, we focus on a new Theory of Mind (ToM) assessment specifically developed for this study, to assist in gaining a deeper understanding of children's empathic engagement with FearNot!

Human social dynamics and interactions rely upon the ability to correctly attribute beliefs, desires, goals and percepts to others. These metarepresentational abilities enable us to understand the actions and behaviours of intentional others within a goal-directed situation. ToM develops in early childhood [10] and is a crucial step in development, with the awareness that other people have different knowledge, beliefs and goals to ourselves.

ToM skills allow children and adults to form representations of mental states such as pretending and knowing, and secondly to understand the relationships between states and actions. ToM tasks can be used to assess whether children can attribute mental states to themselves and others in order to explain and predict behaviour. The ability to recognise that others can have false as well as true beliefs is a central tenet to the development of Theory of Mind skills (e.g. Sally-Anne task, [4, 32])

ToM is a broad cognitive concept where the understanding of emotions and empathy play a role. For example, ToM can be used to explain and predict a great deal about human talk and action. The theory encapsulates the mental states of thinking, knowing, guessing, remembering, hoping, fearing, perceptions, intentions and emotions. These are all important elements for the development, understanding and display of empathy. Through using a ToM assessment within FearNot! we evaluated a child's perception of the synthetic characters and their behaviours. ToM allowed us to investigate whether children could recognise the synthetic characters' behaviours, appearance and affect, what children understood and interpreted from the characters, and what goals and intentions they ascribed to characters.

In addition to evaluating children's perceptions of the characters' behaviour, the ToM assessment allowed us to concentrate on some of the characteristics of the child user. Recent research has suggested that there may be individual differences in the emotional and social competence of children involved in different bullying roles, which may contribute to the stability of these roles. For example, some research has suggested that 'pure' bullies (perpetrators) may have superior ToM abilities, which allows them to carefully plan and execute successful bullying incidents. In contrast, it is believed that victims may have some deficits in ToM abilities resulting in the repeated cycle of victimisation behaviour [37, 38]. However, this research area has received little empirical research attention. Based on previous research, we hypothesise that bullies and neutrals, in particular relational bullies will have superior theory of mind abilities compared to children who were classified as victims. Here, we focus on whether there are any differences in ToM responses for children classified as 'pure' bullies, 'pure' victims, bully/victims or neutral children for direct bullying behaviour. Additionally we investigate whether there are any differences in how children empathise with the characters in the scenarios according to bullying role.

### 3.1 Developing the Theory of Mind Assessment for FearNot!

The ToM questions were devised by experts in the field and are related to previous first order and second order false belief questions used by Happe and Frith [17]. They were extensively piloted in the U.K. in terms of child comprehension and validity with other measures.

The ToM questions were presented electronically to the child, see figure 2, immediately after they had interacted with FearNot! during the evaluation session at the University. They included questions about inferring the emotions, mental states and intentions of the main characters in the story. The ToM assessment included storyboard shots from FearNot!, providing the child with a still from various events, such as Luke (the bully) physically bullying John (the victim). The still shots were used to outline the most important events that happened during the FearNot! interaction with the bullying scenarios, and acted as reminder anchors for the child.



**Figure 2 Child completing the ToM assessment at Virtually Friends**

The ToM assessment was comprised of two response formats: categorical responses where the child was instructed to select the correct response box, and text responses where the child was instructed to write brief sentences.. Questions that made reference to emotions were based on four of the Ekman emotions [9] happiness “Happy,” sadness “Sad”, anger “Angry” and fear “Fearful” with the additional category of “Neutral.” Surprise and disgust were removed after a pilot study indicated that children had difficulty in clearly identifying these emotions. Furthermore, these emotions were not modelled in the synthetic characters. The child was instructed to click on the button that they thought represented the emotions of the character they are being asked about.

The majority of the emotion questions with the exception of the first question permitted the generation of frequency and percentage data. This allowed analysis between choices made by children based on their bullying role assignment (bully, victim, bully/victim and neutral) and discrepancies between selected emotions (radio button pressed), and interpretation of the emotions portrayed by the characters in the physical bullying scenario. The assessment included questions about:

*Comprehension:* the first question related to story comprehension, does the child recognise this as a bullying event (provided as a free text field).

- What do you think is happening in this scene?

*Initial emotion questions:* these relate to character emotions at the beginning of the interaction with FearNot!

- How does Luke (bully) feel at the beginning of the story?
- How does John (victim) feel at the beginning of the story?

The initial emotion questions also refer to character emotions directly after the main bullying incident(s).

- How does Luke (bully) feel after he has hit and pushed John over?
- How does John (victim) feel after Luke has hit and pushed him over?

*Bullying Event Questions:* These questions follow the initial emotion questions and refer to the main bullying incident(s). The questions are formatted as follows and the child types a response.

- What does Luke (bully) think about John (victim)?
- What does John (victim) think about Luke (bully)?
- If you were John (victim), why do you think that Luke (bully) is doing this?
- If you were Luke (bully), why is he doing this to John (victim)?

*End Emotion Questions:* These questions follow the Bullying Event questions and ask the child about the characters' feelings at the end of the scenario (once coping styles have been tried etc).

- How does John (victim) feel at the end of the story?
- How does Luke (bully) feel at the end of the story?

The final end emotion question asks the child how they felt at the end of the scenario.

## 4 Method

The evaluation study ('Virtually Friends' event) took place at the University of Hertfordshire in a large IT suite for a period of two weeks (14-23 June 2004) with up to 65 children from two different schools participating each day. All children individually interacted with FearNot! on standard PCs. Each child user initially provided their personal information (name, gender and age) and a unique personal code that matched their off-line questionnaires. FearNot! began with a physical bullying scenario comprised of three episodes and children had the role of an advisor to help provide the victim character with coping strategies, for example tell the teacher, hit back, run away, tell a friend, etc., to try and stop the bullying behaviour. After the physical scenario, children had the opportunity to interact with the relational scenario showing the drama of bullying among four girls. Immediately after the interaction children completed the ToM questionnaire.

### 4.1 Interacting with FearNot!

The child user interacted with one physical bullying scenario and one relational scenario. After the introduction of the characters, school and situation, users view the first bullying episode, followed by the victimised character seeking rescue in the school library, where it starts to communicate with the child user. Within the initiated dialogue the user selects an advice from a list of coping strategies (shown as a drop down menu). The user also explains his/her selection and what he/she thinks will happen after having implemented the selected strategy, by typing it in (see figure 3).



Figure 3 Interaction with victim

The next episode then starts. The content of the final episode depends on the choices made by the user concerning the coping strategies: Paul, the bystander in the physical bullying scenario, might act as a defender for John (the victim), in case the user has selected a successful strategy, i.e. “telling someone”; or Martina (the bystander) might offer Frances (the victim) help. However, if the user has selected an unsuccessful strategy, i.e. “run away”, the victim rejects the help in the final episode. At the end of the scenario, a universal educational message is displayed pointing out that “telling someone you trust” is usually a good choice. This universal message had to be incorporated as all teachers had strong preferences for children to finish the interaction with a positive feedback message.

Figure 4 shows a flow chart of one scenario (physical or relational) for the evaluation version of FearNot!. The symbols indicate the following:

- Introduction (I): Type in of code, name, age and gender, introduction of characters and school
- Bullying episode (1-3)
- In between episodes: interaction with victim character in resource room (cope)
- Educational message (F): after end of episode 3.

Phases I to F appear twice, once for the physical scenario and once for the relational bullying scenario.

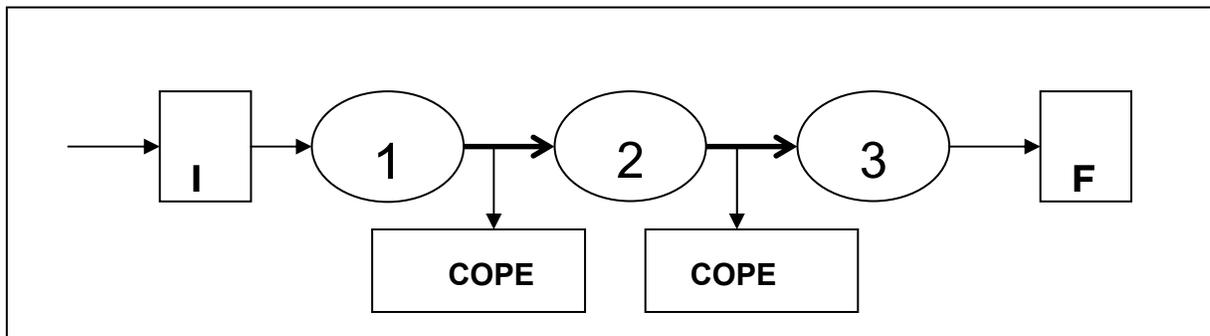


Figure 4 Flow chart of interaction with FearNot!

## 4.2 Identifying Bullying Roles

Children’s bullying roles were determined through the use of a Bullying Questionnaire, based on the Olweus [25] bullying questionnaire which has been widely used to assess bullying behaviour in a wide range of countries. The items on the questionnaire were devised by experts in the bullying field indicating high content validity. Children were asked to identify (using code numbers) the three children who they liked the most in their class, and the three children they disliked the most in their class. Examples were given about physical bullying behaviour over the past 6 months which included being hit or beaten up, having their belongings stolen, being threatened or blackmailed, and having nasty tricks played on them were asked. Children were asked to name up to six children in their class they believed physically bullied other children, and up to six children they thought were physically victimised. The next section of the questionnaire posed questions about relational bullying behaviour over the past 6 months and included getting called nasty names, being deliberately left out of games, withdrawal of friendship, and nasty rumour spreading. Children were asked to name up to six children in their class they believed relationally bullied other children, and up to six children they thought were relationally victimised. The bullying questionnaire was administered to the children prior to their interaction with FearNot!

## 5 Results and Interpretation

373 children aged between 8-11 year (Mean 9.95, SD: 0.5) (172 male (49.9%) and 173 female (50.1%)) participated in the event,. The sample comprised of children from a wide range of primary schools from Hertfordshire in the UK. The preliminary results reported here are for the initial and final emotion questions related to the FearNot! physical bullying scenario.

## 5.1 Incidence of children's bullying roles

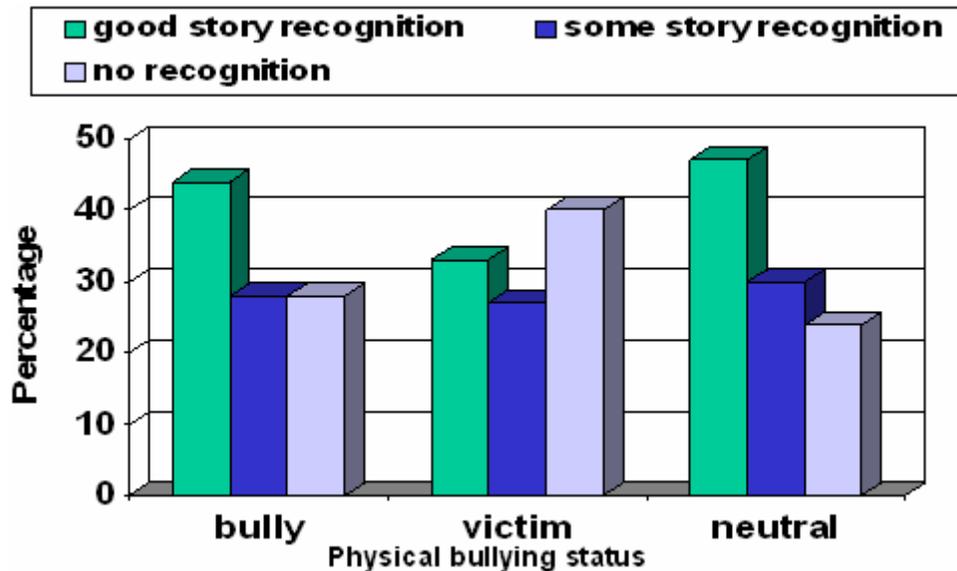
Peer-nominations of physical and relational bullying roles are detailed in table 2. As there were only 2% of children classified as physical bully/victims and 4% as relational bully/victims, the decision was taken to proceed with all further analysis omitting bully/victims as a category.

	Bully (%)	Victim (%)	Bully/victim (%)	Neutral (%)
Physical	53 (14.2)	61 (16.4)	8 (2.2)	250 (67.2)
Relational	51 (13.7)	52 (14.0)	14 (3.8)	255 (68.5)

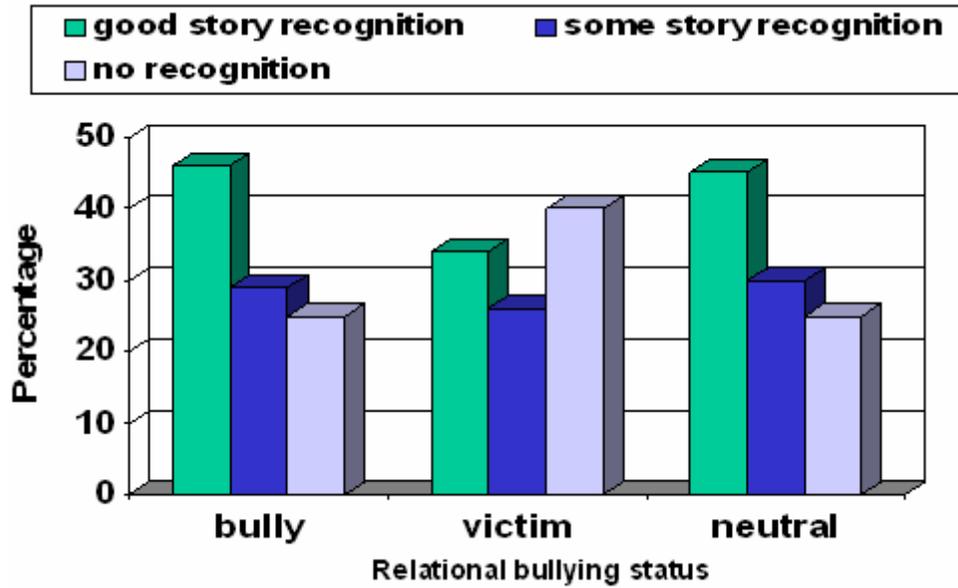
**Table 1** Incidence of peer-nominated bullying roles

## 5.2 Scenario comprehension and initial emotions

In general, children expressed good story comprehension, see figures 5 and 6, however, chi-square analysis in the form of cross-tabulations revealed a significant trend for both physical and relational victims to have poorer overall story comprehension for the physical bullying scenario compared to 'pure' bullies and neutral children ( $p = 0.1$ ).

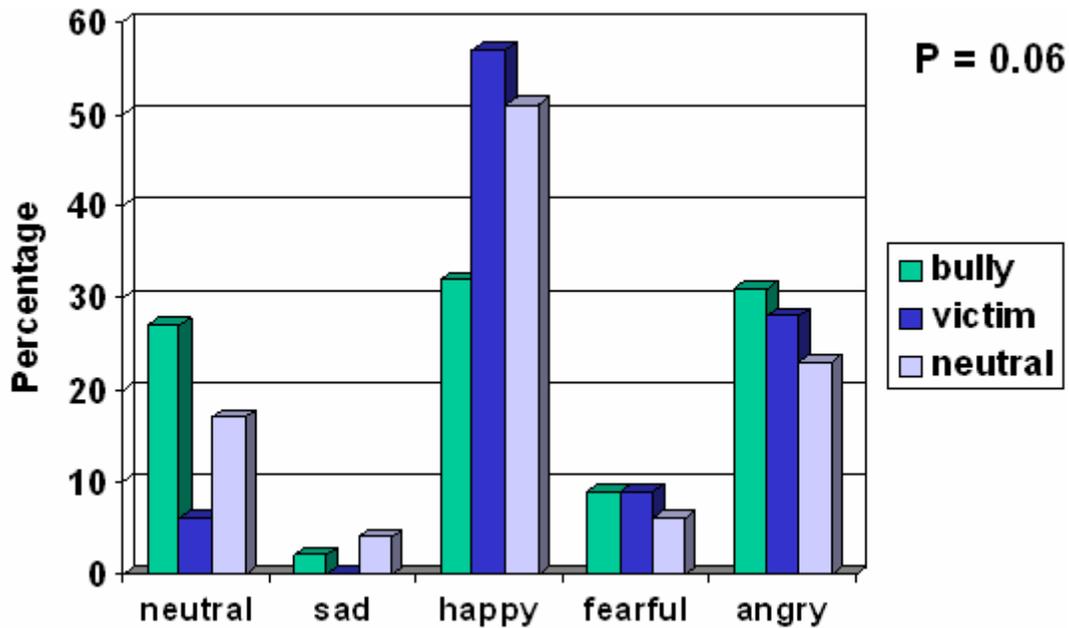


**Figure 5** Recognition of physical bullying scenario & relation with child's physical bullying status



**Figure 6 Recognition of physical bullying scenario & relation with child's relational bullying status**

When children were asked 'how does Luke (bully) feel at the start of the scenario?', a small significant difference ( $p = 0.06$ ) emerged for children's responses. Relational bullies were significantly less likely to say that Luke (bully) felt 'happy' at the beginning of the story compared to victims and neutral children. Bullies stated that Luke felt 'neutral' more frequently.



**Figure 7 Responses to: How does Luke (bully) feel at start of scenario?**

No significant differences were found between children's bullying status and responses to the question 'How does John (victim) feel at the beginning of the story?'. Overall, children demonstrated good emotion recognition for John's (victim) feelings, with 76% of children stating that he felt 'sad' and 20% of children stating that he felt 'fearful'.

### 5.3 Empathic engagement

Children demonstrated good emotion comprehension in relation to John's (victim's) feelings, 55-70% said that he felt 'sad', 15-18% 'angry' and 5-20% said he felt 'fearful'. – Is this for after the bullying incident has occurred or at the beginning of the scenario after the intro?

Questions that enquired about the children's interpretation of characters' emotions after the bullying incident revealed a slight trend for 'how Luke (bully) felt after he had hit/pushed John' but no differences at all for 'how John (victim) felt after being hit/kicked' according to children's bullying status. Children who were classified as physical bullies again stated that 'Luke' (bully) was less likely to feel happy after the incident compared to victims and neutral children ( $p = 0.2$ ).

Table 5.2.6 illustrates that most children stated that Luke (bully) thought John (victim) was a wimp (53.5%) and no significant differences were revealed between children's bullying status and responses to this question. When children were asked what they thought John (victim) thought about Luke (bully), 40% said that he would think he was mean/cruel. 16% of children however did not complete this question or did not understand it.

Response	(%)
Weak/Wimp	53.5
Stupid/idiot	14.6
Disliked John	15.5
No response/did not understand question	16.3

**Table 2 Children's responses to the question 'What does Luke (bully) think about John (victim)?'**

Response	(%)
Mean/cruel	40.0
No response/not understood	20.0
Bully	19.0
Scary/frightening	11.0
Doesn't like him	7.0
Coward/wimp	4.0

**Table 3 Children's responses to the question 'What does John (victim) think about Luke (bully)?'**

### 5.4 End Emotions

No significant relationships were uncovered between children's bullying status and their responses for how they thought 'Luke' (bully) and 'John' (victim) felt at the end of the scenario. 58% of children stated that they thought 'John' was happy at the end of the story followed by 24% who said he felt sad. 35% of children stated that Luke felt angry at the end of the scenario followed by 29% who said he felt happy and 20% who said he felt sad. When children were asked to state how they felt at the end of interacting with the physical bullying scenario, a trend emerged ( $p = 0.1$ ) for victims to state that they felt happier compared to bullies and neutral children. Children who were bullies were more likely to state that they felt neutral at the end of the story.

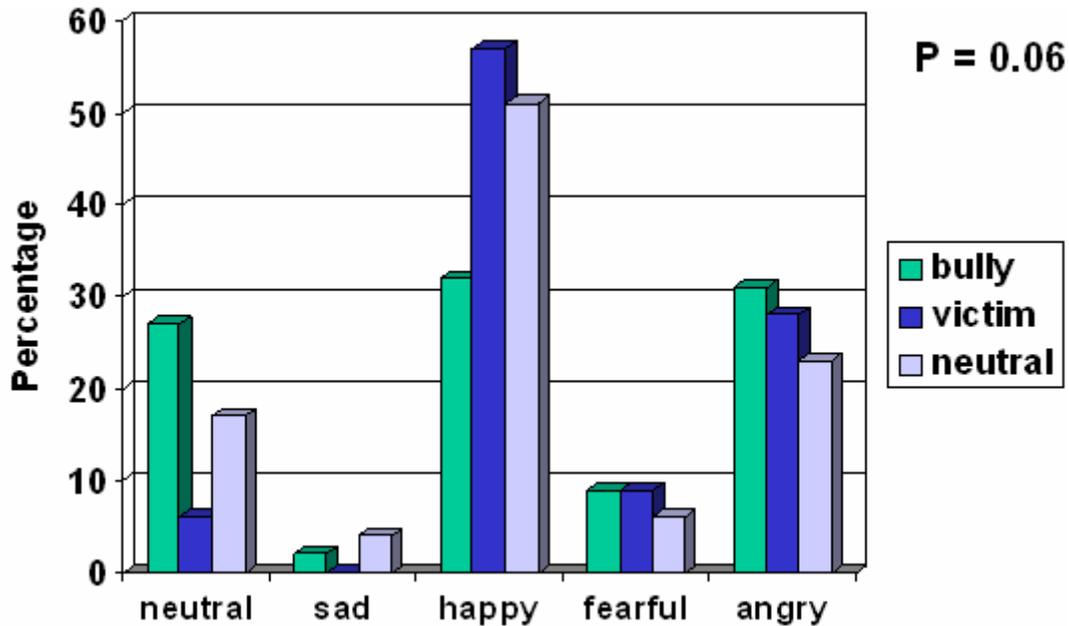


Figure 8 Physical bullying status and end emotion

## 5.5 Gender and ToM abilities in relation to the physical scenario comprehension

Chi-square analysis revealed a significant relationship for overall story comprehension between boys and girls ( $p = 0.04$ ), with girls having significantly better story comprehension compared to boys. When asked how the bully and victim characters felt at the beginning of the physical bullying story, more females thought that John (victim) felt angry compared to males, who stated that John (victim) felt happy. No significant associations were revealed between boys and girls for how Luke (bully) felt at the beginning of the story. Most children stated that he felt happy or angry.

A significant association was revealed between boys and girls responses to the question ‘How does Luke (bully) feel after hitting John (victim)?’ ( $p = 0.007$ ). Females were significantly more likely to state that Luke felt neutral or fearful after hitting John compared to males. A significant trend (0.1) was found between males and females responses to ‘how does John (victim) feel after being hit by Luke (bully) with males more likely to state that John felt sad and females were more likely to state that he was angry.

A significant relationship was uncovered between males and females responses to the question ‘What does John (victim) think about Luke (bully)?’ ( $p = 0.004$ ). Significantly more females stated that John thought Luke (bully) was mean compared to males who stated that John thought Luke did not like him, more than females.

When children were asked how they felt at the end of the physical bullying scenario, a significant association was found for gender ( $p = 0.001$ ). Males were significantly more likely to state that they felt neutral or angry compared to females, and females were more likely to state that they felt sad or happy.

## 6 Discussion

The results clearly identify that children empathically engaged with FearNot! The majority of children understood the story and empathised with the characters identifying appropriate emotional responses for the various events in the interaction. Overall, children appeared to have good comprehension of the mental and emotional states of the two characters (the bully & victim) in the physical bullying scenario.

At the start of the story, there was considerable variation in what children felt that Luke would feel. Just under half of the children felt that he would be feeling happy, with other children identifying a neutral or

angry state. This lack of consensus for Luke's emotional stance at the beginning of the scenario can be contrasted with the children's agreement about John, with 76% identifying that he was sad and the remainder mainly that he was fearful. This is the emotional stance that John had at the beginning of the scenario and children correctly interpreted it. This result can be explained by most children having experience of being sad and afraid, and thus can readily attribute this to John, although they may not have themselves been a direct victim of bullying. Fewer children have experience of perpetuating aggressive behaviour and thus the emotional state of the bully is less apparent to them.

The children were engaged with the bullying incident, with many children attributing a change in the emotion of the bully character. 76% of the children identified that they thought that Luke (the bully) would be happy after the bullying incident. This is in contrast to the lack of consensus about Luke's state at the beginning of the story. Children also demonstrated good recognition of John's feelings after the bullying incident and could empathise with him, attributing sadness and fear.

Children were readily able to empathise with the characters. Although most children were able to understand and determine a mental representation for the characters' states suggesting what the characters thought of one another, 20% of children failed to respond to this question when asked how John (the victim) felt about Luke (the bully) and 16% failed to respond to how Luke felt about John. This result could suggest that some children find it too complex to understand how a character would perceive of another character.

In this bullying scenario, there is a high likelihood for John (the victim) to successfully resolve the bullying situation, through the support of Paul (the defender). Even where the bullying situation is unresolved John (the victim) still thanks the child for their support and is relatively positive. To some extent this helps to explain the 58% happy rating given to John at the end of the scenario. With Luke, similar to the initial emotion state, children are less consensual, however, what is clear is that they no longer believe that Luke is happy, the state he was in directly after bullying John. This changing view of Luke's emotional state supports the view that the child is empathically engaging with the characters, comprehending the story and recognising and interpreting affect.

Although we had hypothesised that bullying roles would have an impact on Theory of Mind, few significant differences emerged when they were taken into account. Those differences uncovered showed that both physical and relational victims tended to have overall poor story comprehension for the physical bullying scenario. Bullies were more likely to remain emotionally neutral when asked to put themselves in the position of the victim character, and when asked about their own emotional states. Although these results were not robust, they do provide some support for the assertion that bullies are 'cool' manipulators of their victims and are able to control their empathic responses [8, 19, 37].

Some gender differences emerged regarding the emotional states children felt that the characters felt. Boys were poorer at recognising what the physical bullying scenario was about compared to girls. Males were more likely to attribute 'angry' emotions to the victims emotional state compared to females and cited the fact that they felt the bully didn't like the victim more compared to females. More females stated that the victim thought the bully was mean compared to males. Boys own emotional states were different to girls as they were more likely to state feeling neutral or angry at the end of the scenario.

The Theory of Mind assessment provided a useful approach to evaluate affect and empathy in interactions with synthetic characters. The responses to the ToM questions suggest that children found the synthetic characters of FearNot! believable and saw the characters as having meaningful social interactions. Children were able to recognise, interpret and attribute affect in synthetic characters and appeared to empathise with the characters in the scenarios. ToM offers an alternative approach to evaluating affective interaction and empathic engagement for the users of synthetic characters.

## 7 Conclusions

The synthetic characters in FearNot! are able to evoke an emotional response in children, with children able to understand and interpret their emotional states. Through using a Theory of Mind assessment, it can be seen that children attribute life-like qualities to the characters and can readily mentalize the characters' states. Children empathically engaged with the characters attributing a range of emotions to the character depending on the events within the scenario.

## 8 References

- [1] André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T., "Exploiting models of personality and emotions to control the behaviour of animated interface agents," presented at Workshop on Achieving Human-Like Behavior in Interactive Animated Agents, Barcelona, Spain, 2000.
- [2] Astington, J., *The child's discovery of mind*. Cambridge, MA: Harvard University Press, 1993.
- [3] Aylett, R. S., Paiva, A., Woods, S., Hall, L., and Zoll, C., "Expressive Characters in Anti-Bullying Education," in *Animating Expressive Characters for Social Interaction*, L. Canamero and R. Aylett, Eds.: John Benjamins,(in print)
- [4] Baron-Cohen, S., Leslie, A. M., and Frith, U., "Does the autistic child have a 'theory of mind'?", *Cognition.*, vol. 21, pp. 37-46, 1985.
- [5] Bates, J., "The Role of Emotion in Believable Agents," *Communications of the ACM*, vol. 37(7), pp. 122-125., 1994.
- [6] Boulton, M. J. and Smith, P. K., "Bully/victim problems in middle-school children: stability, self-perceived competence, peer perceptions and peer acceptance," *British Journal of Developmental Psychology*, vol. 12, pp. 315-329, 1994.
- [7] Dautenhahn, K., Bond, A. H., Canamero, L., and Edmonds, B., *Socially intelligent agents: Creating relationships with computers and robots*. Massachusetts, USA: Kluwer Academic Publishers, 2002.
- [8] Dautenhahn, K. and Woods, S., "Possible connections between bullying behaviour, empathy and imitation," presented at AISB '03, The University of Wales, Aberystwyth, 2003.
- [9] Ekman, P. and Friesen, W. V., "A new pan cultural facial expression of emotion," *Motivation and Emotion*, vol. 10, pp. 159-168, 1986.
- [10] Fodor, J., "A theory of the child's theory of mind," *Cognition*, vol. 43, pp. 283-296, 1992.
- [11] Gratch, J. and Marsella, S., "Tears and fears: Modeling emotions and emotional behaviors in synthetic agents," presented at Fifth International Conference on Autonomous Agents, 2001.
- [12] Hall, L. and Woods, S., "Empathic interaction with synthetic characters: the importance of similarity," in *Encyclopaedia of Human Computer Interaction*, C. Ghaoui, Ed.: Idea Group, 2005.
- [13] Hall, L., Woods, S., Aylett, R., Paiva, A., and Newall, L., "Achieving empathic engagement through affective interaction with synthetic characters," presented at 1st International Conference on Affective Computing & Intelligent Interaction (ACII 05), Beijing, China, 2005.
- [14] Hall, L., Woods, S., and Dautenhahn, K., "FearNot! Designing in the Classroom," presented at British HCI, Leeds, UK, 2004.
- [15] Hall, L., Woods, S., and Dautenhahn, K., "Research findings from synthetic character research: Implications for Interactive Communication with Robots," presented at Ro-Man 2004, Kurashiki, Japan, 2004.
- [16] Hall, L., Woods, S., Sobral, D., Paiva, A., Dautenhahn, K., Wolke, D., and Newall, L., "Designing Empathic Agents: Adults vs. Kids," presented at Intelligent Tutoring Systems 7th International Conference, ITS 2004, Maceio, Brazil, 2004.
- [17] Happe, F. and Frith, U., "Theory of mind and social impairment in children with conduct disorder," *British Journal of Developmental Psychology*, vol. 14, pp. 385-398, 1996.
- [18] Klesen, M., Szatkowski, J., and Lehmann, N., "The Black Sheep:Interactive improvisation in a 3D Virtual World," presented at i3 Annual Conference 2000, Jönköping, Sweden, 2000.
- [19] Lemerise, E. A. and Arsenio, W. F., "An integrated model of emotion processes and cognition in social information processing," *Child Development*, vol. 71, pp. 107-118, 2000.

- [20] Maldonado, H. and Hayes-Roth, B., "Towards Cross-Cultural Believability in Character Design," in *Agent Culture: Designing Human-Agent Interaction in a Multicultural World*, S. Payr and R. Trapp, Eds. Mahweh, NJ: Lawrence Erlbaum Associates Inc, 2004.
- [21] Marsella, S., Johnson, W. L., and LaBore, C., "Interactive Pedagogical Drama," presented at 4th International Conference on Autonomous Agents, 2000.
- [22] Marsella, S., Johnson, W. L., and LaBore, C., "Interactive Pedagogical Drama for Health Interventions," presented at 11th International Conference on Artificial Intelligence in Education, Sydney, Australia, 2003.
- [23] Moreno, R., Mayer, R. E., Spires, H. A., and Lester, J. C., "The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents," *Cognition and Instruction*, vol. 19, pp. 177-213, 2001.
- [24] Nass, C., Isbister, K., and Lee, E., *Truth is beauty: researching embodied conversational agents*. Cambridge, MA: MIT Press, 2001.
- [25] Olweus, D., "Bullying at school: What we know and what we can do." Oxford: Blackwell Publishers, 1993.
- [26] Ortony, A., Clore, G. L., and Collins, A., *The Cognitive Structure of Emotions*. Cambridge, UK: Cambridge University Press, 1988.
- [27] Paiva, A., Costa, M., Chaves, R., Piedade, M., Mourao, D., Sobral, D., Hook, K., Andersson, G., and Bullock, A., "SenToy: an affective sympathetic interface," *International Journal of Human-Computer Studies*, vol. 59, pp. 227-235, 2003.
- [28] Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobrepez, P., Woods, S., Zoll, C., and Hall, L., "Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents," presented at AAMAS 2004, New York, 2004.
- [29] Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., and Zoll, C., "Learning by feeling: Evoking Empathy with Synthetic Characters," *Journal of Applied AI*, in print.
- [30] Pertaub, D.-P., Slater, M., and Barker, C., "An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience," *Presence: Teleoperators and Virtual Environments*, vol. 11, pp. 68-78, 2001.
- [31] Picard, R., *Affective Computing*. Cambridge, MA: MIT Press, 1997.
- [32] Premack, D. and Woodruff, G., "Does the Chimpanzee have a theory of mind?," *Behavioural and Brain Sciences*, vol. 1, pp. 515-526, 1978.
- [33] Prendinger, H. and Ishizuka, M., "The Empathic Companion: A character-based interface that addresses users' affective states," *International Journal of Applied Artificial Intelligence*, vol. 19, pp. 267-285, 2005.
- [34] Robertson, J. and Oberlander, J., "Ghostwriter: Educational Drama and Presence in a Virtual Environment," *Journal of Computer Mediated Communication*, vol. 8, 2002.
- [35] Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., and Kaukiainen, A., "Bullying as a group process: participant roles and their relations to social status within the group," *Aggressive Behavior*, vol. 22, pp. 1-15, 1996.
- [36] Scholl, B. J. and Leslie, A. M., "Modularity, Development and 'Theory of Mind'," *Mind & Language*, vol. 14, pp. 131-153, 1999.
- [37] Sutton, J., Smith, P. K., and Swettenham, J., "'Cold Cognition" in Bullying," presented at British Psychological Society Developmental Section Annual Conference, Loughborough University, 1997.
- [38] Sutton, J., Smith, P. K., and Swettenham, J., "Social cognition and bullying: Social inadequacy or skilled manipulation?," *British Journal of Developmental Psychology*, vol. 17, pp. 435-450, 1999.

- [39] Thomas, F. and Johnston, O., "The Illusion of Life: Disney Animations." Ney York: Walt Disney Productions, 1981.
- [40] Wellman, H., *The child's theory of mind*. Cambridge, MA: MIT Press, 1990.
- [41] Wolke, D. and Stanford, K., "Bullying in school children.," in *Developmental Psychology*, D. Messer and S. Millar, Eds. London: Arnold Publisher, 1999.
- [42] Woods, S., Hall, L., Sobral, D., Dautenhahn, K., and Wolke, D., "A study into the believability of animated characters in the context of bullying intervention," presented at IVA '03, Kloster Irsee, Germany, 2003.