 ISBN: 978-989-95618-4-7

Empathic Virtual Buddy: Setting Up Informed Empathic Responses

Janneke van der Zwaan and Virginia Dignum and Joost Broekens and Catholijn M. Jonker

Delft University of Technology, The Netherlands
{j.m.vanderzwaan,m.v.dignum}@tudelft.nl, joost.broekens@gmail.com, c.m.jonker@tudelft.nl

Abstract. In order to address the user's emotions and give emotional support, affective agents need to know the emotional state of its users. For complex situations, it is not clear what these emotional states might be. However, this information is needed to design informed and suitable empathic responses for an affective agent. This paper presents normative affective ratings for cyberbullying situations. These normative ratings will be used in the reasoning engine of an empathic agent that provides emotional support and practical advice to victims of cyberbullying. In the experiment conducted to gather the normative ratings, we also determined the AffectButton is an adequate interface for gathering affective feedback for these situations; the AffectButton is valid (i.e. measures what it is supposed to measure) and the usability is acceptable.

1 Introduction

Empathy is the capacity to recognize and, to some extent, experience the emotions of another person (for example sadness or happiness). There has been a growing interest in empathic agents addressing the emotional state of users. Our research concerns the development and evaluation of an empathic agent that provides victims of cyberbullying with emotional support and practical advice [31]. Cyberbullying refers to bullying through electronic communication devices. With victimization rates ranging from 20 to 40% [30], it is a common risk for children and adolescents. In addition, recent findings from the EU Kids Online II survey indicate that cyberbullying has a high impact on victims [18]. The Buddy agent is an Embodied Conversational Agent (ECA) that 'lives' on the computer screen of potential victims of cyberbullying. When a victim feels uncomfortable because of a cyberbullying incident, the buddy responds empathically to lower the negative emotions evoked by the incident.

In order to be able to provide suitable and grounded empathic responses to victims of cyberbullying, the buddy needs to understand how they feel about being bullied. Our approach is to enable users to explicitly input their emotional state through the buddy interface. However, it is not clear what the emotional states of victims in cyberbullying situations are and, therefore it is hard to anticipate appropriate responses. In order to facilitate the empathic reasoning of

the anti-cyberbullying buddy, we performed an experiment to gather normative affective ratings for 6 frequently occurring cyberbullying situations.

In the same experiment, we assessed the adequacy of the AffectButton [4] for obtaining affective feedback about cyberbullying situations. The AffectButton is a candidate component that could enable the buddy's users use to communicate their emotional state. The AffectButton is a button with a rudimentary and gender-neutral face that changes its expression based on the position of the mouse cursor. By clicking the button when it shows the emotional expression the user wants to communicate, a Pleasure-, Arousal-, Dominance-based representation of the expression is send to the reasoning engine of the buddy. For an appropriate empathic response, this representation must be accurate. Even though the AffectButton has been validated in different situations [4,5], it has not been validated for gathering affective feedback about more complex stimuli such as (cyber)bullying situations. Therefore, we need to determine whether the AffectButton produces valid (i.e. measures what it is supposed to measure) affective feedback about such situations. Finally, the usability of the AffectButton was also investigated.

This paper is organized as follows. In the next section, we briefly present the Cyberbullying project. In section 3, we discuss background research in the area of emotion recognition. Section 4 describes the experiment and in section 5 its results are discussed. Finally, in section 6 conclusions and directions for future work are presented.

2 The Cyberbullying Project

According to a recent study, US children aged 8 to 18 on average spend 1.5 hours a day using the computer for recreational purposes [11]. Most of these activities take place on the Internet. The study also found 84% of children has access to the Internet at home. So, many children spend a lot of time online. They use the Internet not only as an educational tool, but also for fun, games and to develop and maintain social contacts. One of the risks they run online is to become a victim of cyberbullying. Cyberbullying can be defined as 'any behavior performed through electronic or digital media by individuals or groups that repeatedly communicates hostile or aggressive messages intended to inflict harm or discomfort on others' [30]. Cyberbullying takes place via e-mail, instant-messaging programs, Internet chat rooms, multi-player online games, or social websites or blogs. Our work is part of the multi-disciplinary Cyberbullying project[1] that aims at designing social, legal and technological measures to protect and empower children and adolescents against bullying in virtual environments.

It is widely recognized that cyberbullying is a complex issue and a 'quick fix' does not exist [7,20,27,29]. Education, both of children and adults, and increasing awareness are suggested to tackle the problem [7,9,32]. In particular,

[1] The project 'Empowering and Protecting Children and Adolescents Against Cyberbullying' is funded by Netherlands Organization for Scientific Research (NWO) under the Responsible Innovation program.

there is a growing recognition of the need for children to receive educational support for social and emotional learning, with awareness of its importance for both non-academic outcomes and improved academic performance. The eCircus project[2] developed innovative technology to support social and emotional learning through role-play and affective engagement for Personal and Social Education involving complex social situations. The project focused on enhanced learning through the use of an interactive 3D environment that explores virtual play and improvisational drama with synthetic characters that evoke empathy.

Even though awareness of the problem of cyberbullying is currently increasing at schools, teachers and parents often lack the knowledge and technical skills to truly help victims of cyberbullying [10]. Peer support has proven to be effective against traditional bullying [8]. However, peer supporters need excellent communication skills, such as active listening, adopting a problem-solving approach, being empathic and the ability to build up trust [8].

In previous work, we proposed a design for an Embodied Conversational Agent (ECA) that empowers victims of cyberbullying by acting as a supportive friend (peer) [31]. The short-term goal of this buddy agent is to lower the victim's negative emotions (coping). On the long(er) term, the buddy aims to teach the victim how to deal with cyberbullying.

3 Background

Within the field of psychology two major strands of emotion theories can be distinguished. Cognitive emotion theories focus on the cognitive appraisal processes necessary to elicit the full range of emotions in adult humans [22]. Dimensional emotion theories [12] are based on the idea of classifying emotions along an arbitrary amount of dimensions of connotative meaning.

The emotion model proposed by Ortony, Clore and Collins (OCC) [22] is based on cognitive emotion theories and has often been the basis for the integration of emotions into cognitive architectures of embodied characters, e.g. [24], as it was designed to be computationally friendly. A well known dimensional emotion model is the Pleasure, Arousal, Dominance (PAD) model of emotions, which assumes an emotion (more precisely: affect) can be defined as a coincidence of values on different dimensions [19,23]. Pleasure (valence) determines how positive or negative the emotion, mood or attitude is. Arousal describes whether the emotion, mood or attitude involves activation or deactivation. Dominance refers to the degree to which the individual feels in control versus feels submissive with regards to the stimulus or situation. For example, anger would be low pleasure, high arousal, high dominance emotion, while sadness would be a low pleasure, low arousal, low dominance emotion.

Many empathic agents use affect detection to determine the emotional state of the user: the empathic chess companion of Leite et al. uses facial and body expression recognition systems and contextual features of the chess game to recognize affective states such as interest, boredom and frustration [17]; Bee et al.

[2] http://www.macs.hw.ac.uk/EcircusWeb/

proposed an empathic listening agent extracts affective information from speech sounds [1]; The Sensitive Artificial Listeners (SAL) system uses a multimodal approach to affect detection, it determines the current user emotion by analyzing facial expressions and speech sound [26]; Prendinger and Ishizuka's Empathic Companion interprets physiological data (skin conductance and muscle tension) as emotions [25]. The interactive caring agent Maggie of Lee et al. uses a combination of two implicit and an explicit method to acquire knowledge about the user's emotional state [16]. Klein et al. and Hone used similar agents in their experiments, which required the user to provide a self-report of their frustration levels [13,14].

Implicit methods for affective feedback are unobtrusive and therefore seem most appropriate for many applications. However, performance of implicit methods is still far than optimal [6], especially in real-world non-controlled environments. Furthermore, for applications such as the anti-cyberbullying agent, where the user's current emotional state explicitly is being addressed during the interaction, explicit communication of affective information is arguably more beneficial as the user will be required to consciously think about her own emotions. By providing emotional-content feedback, which involves letting the emotionally upset person know that his or her emotional state has been effectively communicated [21], the buddy will confirm its understanding of the emotional state of the user.

Explicit methods for affective feedback, also known as self-report methods, have been used for a long time in psychology. The Self-Assessment Manikins (SAM) method [2] is a well-validated method to acquire emotional responses to different stimuli (e.g. text [3] and images [15]). The method is based on the PAD model.The method consists of three rows of pictures, one for each PAD dimension. Each row contains five pictures, resulting in a 9-point scale (five pictures and four intermediary spaces). Participants select the point on the scale that best describes their emotion for each dimension separately. A rating consists of a PAD triplet with values ranging from 1-9.

As mentioned in the introduction, the AffectButton [4] is a tool for explicit affective feedback (see Fig. 1). Affective values are represented by the rendered facial expressions, so the user selects an emotional expression by clicking the button. Based on the PAD coordinates, the face displayed is interpolated between nine prototypical expressions. Therefore, a user can enter mixed emotions (e.g., confused) as well as low and high intensity prototypical ones (e.g., little happy, elated).

4 Empirical Methodology

The main goal of the experiment is to gather normative ratings for common cyberbullying situations. However, as discussed in the introduction, we are also interested in evaluating the adequacy of the AffectButton for gathering the affective attribution to more complex cyberbulling situations.

The main hypothesis for the evaluation of the AffectButton is: *The AffectButton interface provides a suitable way for explicit affective feedback, comparable*

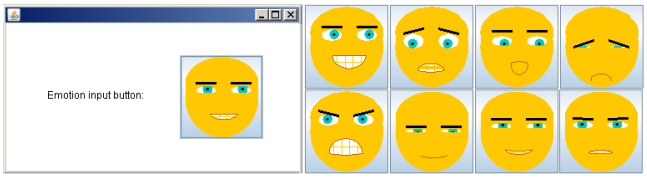

Fig. 1. The AffectButton and its extreme affective states: elated (PAD=1,1,1), afraid (-1,1,-1), surprised (1,1,-1), sad (-1,- 1,-1), angry (-1,1,1), relaxed (1,-1,-1), content (1,-1,1), frustrated (-1,-1,1). Labels are exemplary. Note that the AffectButton allows for continuous input in the PAD space.

to SAM, but simpler in its usage. In order to perform a more detailed analysis of this hypothesis, we investigated 2 sub hypotheses: 1) PAD triplets generated with the AffectButton are valid; 2) the AffectButton is user-friendly.

SAM is considered as a baseline in this paper and usability results obtained for the AffectButton will be compared to the usability results obtained for SAM. Further, translated ANET texts [3], of which the PAD values are known, were used to the validity of the AffectButton feedback by allowing people to rate these texts with the AffectButton so that we could compare these values with the existing ones.

4.1 Experiment design

A between subject design was used to gather affective ratings from participants. Each participant rated 12 stimuli: 2 randomly selected cyberbullying situations and 10 randomly selected texts from ANET, a standard collection of brief texts with associated normative PAD ratings [3]. A subject rated all stimuli with either SAM or the AffectButton (each subject was assigned the same self-report method for all stimuli). The self-report interface used by each participant was selected randomly. The descriptions of cyberbullying situations were created by the authors. Recent research shows common forms of cyberbullying are verbal abuse (35%), sending nasty messages (e.g. through instant messaging; 27%), spreading gossip (20%) and social exclusion (16%) [28]. Based on this information, 6 Dutch sentences describing corresponding situations were constructed. Table 1 lists English translations of the cyberbullying situation descriptions. The 60 ANET texts from ANET set A were all translated to Dutch. A balancing mechanism ensured that all texts (ANET and bullying) were randomly selected approximately the same number of times.

Usability of the affective feedback interface was measured by four items on a 7-point Likert scale: 1) the interface is easy to use, 2) the interface is pleasant to use, 3) the interface is easy to understand, 4) it takes little effort to rate emotions with the interface. In addition, the participants were asked to report demographic information; age, gender, and occupation. Finally, the time needed by each participant to complete the ratings was recorded.

Id	Text
1	All your classmates are online, but nobody responds to your chat requests.
2	A girl in your class doesn't want to be your friend on Facebook, because she only uses Facebook for real friends.
3	You find out that you are the subject of a hate profile on Hyves and multiple classmates posted nasty messages.
4	For the third time this week you receive an e-mail stating 'I HATE YOU AND I'M GOING TO KILL YOU', you don't know the sender.
5	On Facebook people added comments to your photos such as 'You are so ugly' and 'Everybody hates you'.
6	You add a new contact person on MSN who immediately changes his name to 'YOUR WORST NIGHTMARE' and starts to verbally abuse you.

Table 1. The collection of short texts describing frequently occurring cyberbullying situations used in the experiment.

4.2 Experiment setup

The cyberbullying buddy project focuses on children aged 8-18. However, before running our experiment with children and/or adolescents, we want to have a generic evaluation of its functioning and, if needed, adapt it. We have therefore run a first experiment in which 202 subjects participated. Participants were predominantly female (97.5%). 25.5% of the subjects were younger than 25, 11.5% were between 26-35 years old, 25.5% between 36-45, 23.5% between 46-55, and 14% were over 65 years old. The majority of the participants were employed (22.3% working full time and 43.0% working part time), 13.4% were students and 21.3% were either unemployed or had some other occupation. On average, it took participants 4 minutes to complete the experiment.

Participants performed their ratings on a laptop with a mouse. Before starting to rate texts, the purpose of the experiment and the interface to be used (either SAM or the AffectButton) was explained to the participant after which the participant had the opportunity to practice using the affective feedback interface by rating an example text. Figure 2 shows screen shots of the interface used for data collection.

Immediately after completing the rating for one stimulus (three clicks for SAM and one for the AffectButton), the next text appeared on the screen. Just as during the collection of the original ANET ratings [3], participants were instructed not to think too much about the text before rating it. In addition, the rating interface allowed no opportunity to correct ratings. When rating a cyberbullying situation, the participant was reminded by a message on the screen to take the perspective of a 14-year-old person.

5 Results

Of the 202 participants, 102 used SAM to rate the stimuli and 100 used the AffectButton, resulting in 31-36 ratings per interface for the cyberbullying situations and an average of 17 ratings per interface for the ANET texts.

 ISBN: 978-989-95618-4-7

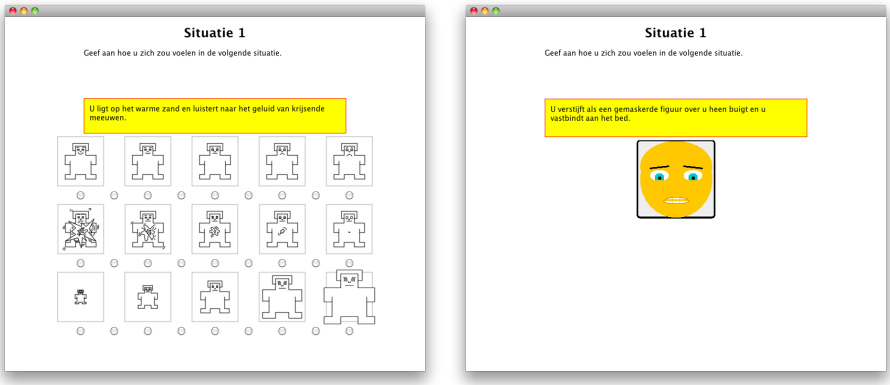

Fig. 2. The interface for data collection with SAM (left) and the AffectButton (right).

5.1 Validity and Usability of the AffectButton

The validity, reliability and usability of the AffectButton has been shown earlier [4]. It has also been shown that high-school students (average age 16.7) can rate affect in music [5] using the AffectButton. Here we focus on two things: first, affective feedback gathered with the AffectButton should correlate with original ANET values, and second, correlate with SAM feedback. First we calculated the average (based on approximately 17 ratings per text) P, A and D values for each of the ANET texts for the AffectButton and SAM feedback. A correlation analysis resulted in the following. AffectButton ratings obtained during the experiment and existing SAM ratings correlated strongly for P ($r(60)=0.94$, $p<0.000$) and D ($r(60)=0.86$, $p<0.000$), and moderately for A ($r(60)=0.55$, $p<0.000$). This indicates the AffectButton results in valid measurements as compared to the existing values for the ANET texts. Further, ratings obtained with the AffectButton and ratings obtained with SAM during the experiment also correlated strongly for P ($r(60)=0.93$, $p<0.000$), D ($r(60)=0.90$, $p<0.000$) and moderately for A ($r(60)=0.55$, $p<0.000$). These results indicate that the AffectButton shows convergent validity, and can therefore be used to obtain PAD values in approximately the same way as intended with SAM in this particular setting.

After having used either SAM or the AffectButton 12 times, participants were asked to express their opinion on the usability of the interface for explicit affective feedback they used. Table 2 lists scores obtained for each item. All scores were on a 7-point Likert scale from 1 (completely disagree) to 7 (completely agree). The differences are small, but significant for items 1 ($t(200)=2.77$, $p=0.006$) and 4 ($t(200)=3.25$, $p=0.001$). The lower score for item 4 indicates that the participants think rating an emotion with the AffectButton takes more effort than rating an emotion with SAM. This is remarkable, because the average text rating time needed per subject with SAM took on average 16 seconds, which is significantly longer than the average of 12 seconds for rating with the AffectButton ($t(200)=4.93$, $p=0.000$).

The usability scores for the AffectButton are lower than the scores for SAM on two of the four measures. On the other two measures, the ratings are not significantly different. Overall, usability scores are acceptable (> 4) and indicate users are able to use the AffectButton. Even though rating emotions with the AffectButton is perceived to take more effort, objectively it takes less time and this is an advantage in situations where users have to rate emotions multiple times. Since the anti-cyberbullying buddy is envisioned to operate in such a situation, the AffectButton is considered suitable.

Item	SAM	AffectButton	Sign.
1. The interface is easy to use	5.44	4.85	+
2. The interface is pleasant to use	5.42	5.13	−
3. The interface is easy to understand	5.60	5.52	−
4. It takes little effort to rate emotions with the interface	5.21	4.50	+

Table 2. Average usability scores (1=Completely disagree, 7=Completely agree).

5.2 Normative Ratings of Common Cyberbullying Situations

Unlike the ANET texts, the situation descriptions of frequently occurring cyberbullying situations have not been designed to evoke certain emotions. Instead, it is the other way around, we want to determine what victims of cyberbullying feel. As we found that the AffectButton is an adequate interface for explicit affective feedback, we use both the AffectButton and the SAM ratings to evaluate the cyberbullying situations.

Figure 3-14 show the histograms of the P, A, and D values obtained with SAM and the AffectButton (AB). The pattern of the frequencies for P, A, and D values obtained with SAM is similar for all 6 texts: pleasure is low (< 5), arousal is high (> 5) and dominance is low (> 5)), except for text 2 which will be discussed below. The histograms for scores obtained with the AffectButton are slightly more complex. To begin with, P values are generally low (< 5), similar to the SAM ratings. These low P values indicate that the cyberbullying situations mostly evoked negative emotions. This is as expected, because being (cyber)bullied is a negative experience. For arousal, there seems to be a peak in A values of 1-2. However, a known issue with the AffectButton is that the center point in the button has an arousal value of 1, meaning that there is a bias towards arousal values of 1 when people do not explore the button enough. This influences the arousal distribution of individual ratings. This issue is being worked on currently. When ignoring the peak in low arousal, The distribution of A values becomes similar to the SAM ratings. The D values of text 1, 2, and 4 are also similar to the SAM D values, i.e. generally low. For text 3, 5, and 6, both low and high dominance value clusters can be distinguished. This will be discussed below. Overall, the common pattern in all PAD values obtained using

SAM and the AffectButton can be characterized as low pleasure, high arousal and low dominance, which indicates an affective state of fear or worry.

The PAD values for text 2 obtained with the AffectButton (and to a lesser extent also for the SAM values, especially the arousal and dominance scales) are spread over the entire PAD space (see Figures 5 and 6). A possible explanation for this wide range of scores is that adults think children do not mind being rejected online. This is probably a misjudgment, possibly caused by their failure to understand the importance of online contacts for children and adolescents. It is expected that children and adolescents, who are the target audience of the anti-cyberbullying buddy, will predominantly rate this situation as negative (low pleasure). This underscores the importance of experimenting with the actual target group in these kind of applications. Text 2 is excluded from further analysis here.

In the three-dimensional PAD space, four negative emotions can be distinguished (which are all characterized by low pleasure): sadness (low pleasure, low arousal and low dominance), fear (low pleasure, high arousal and low dominance), frustration (low pleasure, low arousal and high dominance), and anger (low pleasure, high arousal and high dominance).

Since we know pleasure is low for all of the texts for SAM and AffectButton feedback, we further analyze the negative affective feedback using scatter plots of A and D values. Figures 15, 17, and 19 show scatter plots of the A and D values obtained with SAM for text 1, 3, and 4 respectively. To visualize multiple identical ratings, the data points have been plotted with small random displacements. The scatter plots for texts 5 and 6 are similar to the one of text 3 and have not been depicted for reasons of space. These scatter plots confirm that most ratings map to the fear quadrant of the AD plane (lower right), however, some ratings in the sadness quadrant (lower left) and in the anger quadrant (upper right) can be also found.

As mentioned before, the D values of the ratings obtained with the AffectButton vary for text 3, 5, and 6. From the histograms, it is unclear which negative emotions occur in the ratings. In Figs. 16, 18, and 20 scatter plots of the A and D values obtained with the AffectButton for text 1, 3, and 4 have been depicted. Again, the scatter plots for texts 5 and 6 are similar to the one of text 3 and have been omitted. Note that the arousal bias is clearly visible in the scatter plots; many ratings are spread along the A value of 1.

For text 1, the scatter plot of ratings from the AffectButton is similar to the scatter plot of the SAM results (see Fig. 15 and 16 respectively). Both show ratings in the sadness and fear quadrants (lower left and lower right). For text 3 (and 5 and 6) Fig. 18, three different clusters mapping to three different emotions can be distinguished: there is a cluster in the anger quadrant (upper right), one in the sadness quadrant (lower left) and another one in the fear quadrant (lower right). For text 4, the results for the AffectButton are again similar to the SAM ratings (see Fig. 19 and 20): the ratings mainly map to points in the fear quadrant.

A multivariate ANOVA with age group (<=27 vs. >27) and feedback mechanism as independent factors and average P, A and D values per subject as dependent variable (n=202), did not show a relation between age and rating bully situations ($F(3, 198)=0.29$, p=ns). It did show an interaction effect between age group and rating mechanism ($F(3, 198)=2.61$, $p=0.053$), only significant for the dominance scale ($F(1, 200)=6.35$, $p=0.013$). This indicates that age does influence how subjects used the feedback methods. Younger subjects rated dominance significantly higher ($t(43)=3.38$, $p<0.01$) using the AffectButton (mean D=4.61) than using SAM (mean D=2.38), while this difference was far less pronounced for older subjects (mean D=3,20 and mean D= 3.77 respectively), though still significant ($t(155)=2.0$, $p=0.049$) . We currently do not know how to interpret this difference.

To summarize these findings, the cyberbullying situations elicit the correct affective response, as confirmed by rating obtained with two different feedback mechanisms. They are negative, arousing and involve low levels of control. To some of these situations (3, 5 and 6) different affective responses seem to be possible, i.e., sadness, anger and fear, as shown by the ratings obtained with the AffectButton.

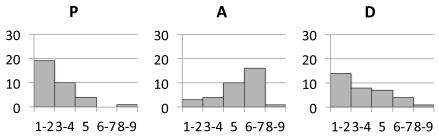

Fig. 3. Text 1: SAM PAD histograms.

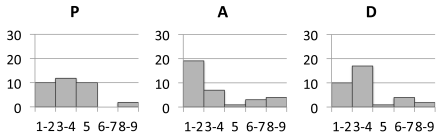

Fig. 4. Text 1: AB PAD histograms.

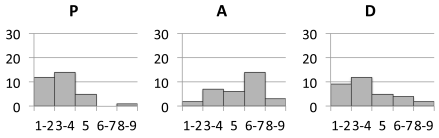

Fig. 5. Text 2: SAM PAD histograms.

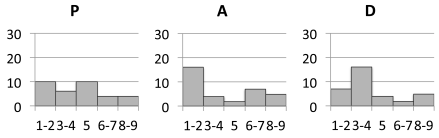

Fig. 6. Text 2: AB PAD histograms.

6 Conclusion

The main goal of this paper was to gather normative affective ratings for emotionally complex situations. These normative ratings are going to be used to provide the reasoning engine of an affective embodied conversational agent that gives emotional support to victims of cyberbullying with suitable and grounded

 ISBN: 978-989-95618-4-7

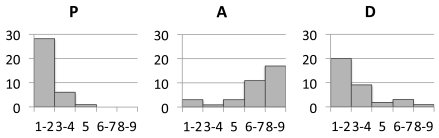

Fig. 7. Text 3: SAM PAD histograms.

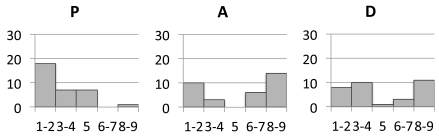

Fig. 8. Text 3: AB PAD histograms.

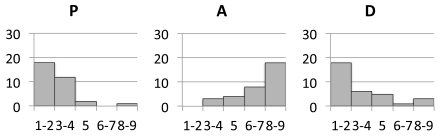

Fig. 9. Text 4: SAM PAD histograms.

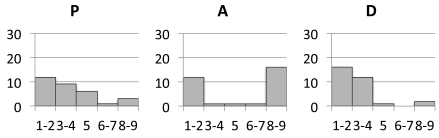

Fig. 10. Text 4: AB PAD histograms.

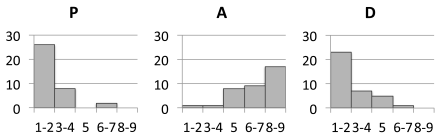

Fig. 11. Text 5: SAM PAD histograms.

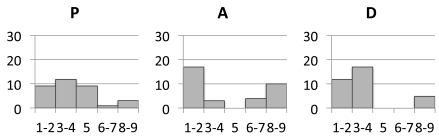

Fig. 12. Text 5: AB PAD histograms.

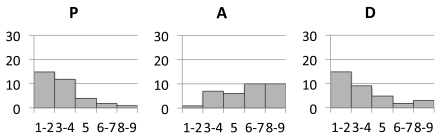

Fig. 13. Text 6: SAM PAD histograms.

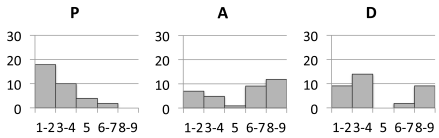

Fig. 14. Text 6: AB PAD histograms.

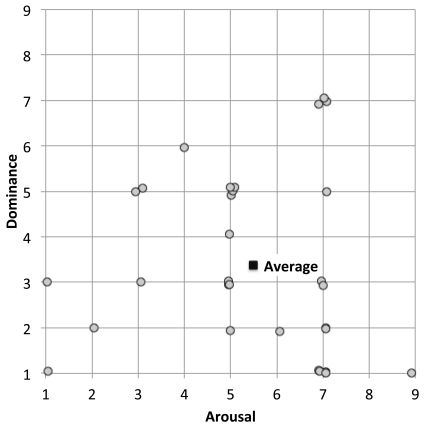

Fig. 15. Text 1: SAM scatter plot of A and D values.

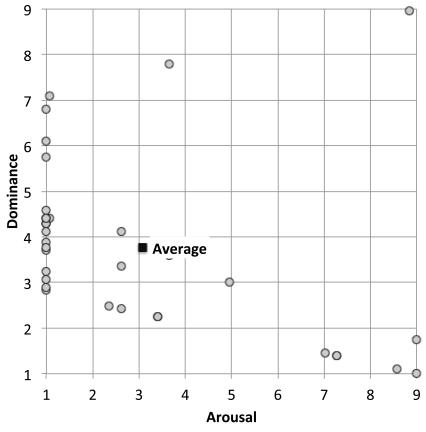

Fig. 16. Text 1: AB scatter plot of A and D values.

 ISBN: 978-989-95618-4-7

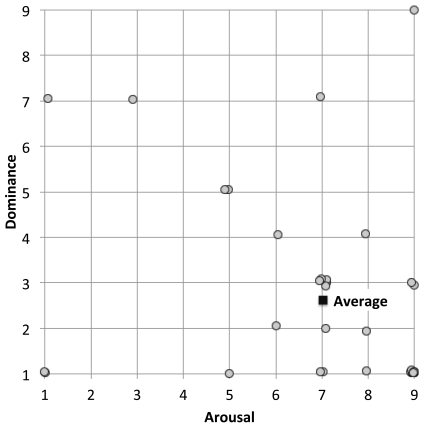

Fig. 17. Text 3: SAM scatter plot of A and D values.

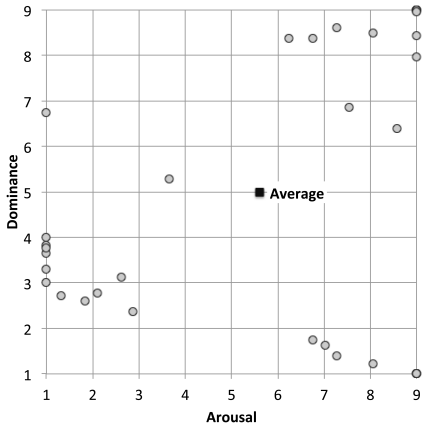

Fig. 18. Text 3: AB scatter plot of A and D values.

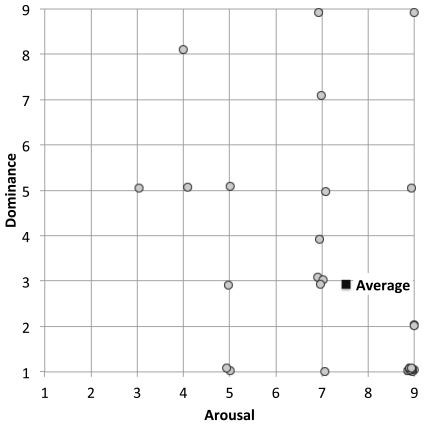

Fig. 19. Text 4: SAM scatter plot of A and D values.

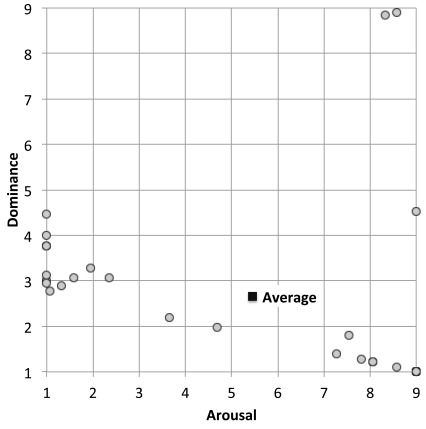

Fig. 20. Text 4: AB scatter plot of A and D values.

empathic responses. In the experiment conducted to gather these ratings, we also evaluated the adequacy of the AffectButton as an interface for communicating complex emotional states. Adequacy was defined as valid (i.e. measures what it is supposed to measure) and easy and pleasant to use. By comparing ratings obtained for sixty standard (ANET) texts using SAM to ratings obtained using the AffectButton, we demonstrated the validity. Results of a post-experiment questionnaire indicate that the usability of the AffectButton is acceptable.

To obtain normative ratings for cyberbullying situations, participants were asked to rate two descriptions of common cyberbullying situations that were randomly selected from a collection of six situation descriptions. The texts in this collection were created by the authors based on recent findings from the cyberbullying literature. Inspection of the results for individual texts revealed that all texts (except one that was probably misinterpreted by the subjects) are perceived to be negative. SAM and AffectButton ratings are predominantly found in the fear quadrant of the arousal/dominance (AD) plane, whereas for three texts different clusters can be found in the ratings obtained with the AffectButton, i.e., sadness, anger, and fear. The results indicate the emotional responses to complex situations (in particular cyberbullying situations) that can be expected by our affective agent. The empathic reasoning engine of this agent should be able to deal with multiple emotional responses (mainly sadness, anger, and fear) to the same situation, and in general respond to the feeling of fear or worry.

An important limitation of this study is that the participants of the experiment (adults) are not the target audience of the cyberbullying project (children/adolescents aged 8-18). Nevertheless, as mentioned before, we wanted to have a generic evaluation of its functioning and, if needed, adapt it. In addition, when rating cyberbullying situations, the participants were explicitly asked to take the perspective of someone in the target audience. However, the results for cyberbullying text two, where the ratings obtained with the AffectButton were spread out over the entire PAD space, possibly indicate that adults sometimes misjudge cyberbullying situations. Further, the main trend of the affective attribution to these situations was one of fear/worry, while one would expect sadness to be at least as dominant. This could also be an indication of the fact that the raters were predominantly mothers, who would experience fear and worry instead of sadness when their child is being cyberbullied. Our results underscore the need to evaluate this kind of application using the actual target group.

For future work, we plan to repeat the experiment with the target audience of the buddy (children/adolescents aged 8-18). Some minor adjustments will be made, such as increasing the number of cyberbullying situations and decreasing the number of ANET texts and the total number of stimuli to rate. Another line of future work concerns embedding the normative cyberbullying ratings in the reasoning engine of the buddy and to empirically link them to suitable embodied empathic responses.

References

1. N. Bee, E. André, and T. Vogd. First ideas on the use of affective cues in an empathic computer-based companion. In *Proceedings of the 8th International Conference on Autonomous agents and Multiagent Systems (AAMAS 2009)*, pages 13–17, 2009.
2. M.M. Bradley and P.J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.
3. M.M. Bradley and P.J. Lang. Affective norms for english text (anet): Affective ratings of text and instruction manual. Technical report, 2007.
4. J. Broekens and W.P. Brinkman. Affectbutton: Towards a standard for dynamic affective user feedback. In *Affective Computing and Intelligent Interaction (ACII) 2009*, 2009.
5. J. Broekens, A. Pronker, and M. Neuteboom. Real time labeling of affect in music using the affectbutton. In *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, AFFINE '10, pages 21–26, New York, NY, USA, 2010. Acm.
6. R.A. Calvo and S. D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1):18–37, 2010.
7. W. Cassidy, M. Jackson, and K.N. Brown. Sticks and stones can break my bones, but how can pixels hurt me?: Students' experiences with cyber-bullying. *School Psychology International*, 30(4):383–402, 2009.
8. H. Cowie, P. Naylor, L. Talamelli, P. Chauhan, and P.K. Smith. Knowledge, use of and attitudes towards peer support: a 2-year follow-up to the prince's trust survey. *Journal of Adolescence*, 25(5):453–467, 2002.
9. F. Dehue, C. Bolman, and T. Völlink. Cyberbullying: youngsters' experiences and parental perception. *Cyberpsychology & Behavior : the Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, 11(2):217–223, 2008.
10. D. Finkelhor, K.J. Mitchell, and J. Wolak. Online Victimization: A Report on the Nation's Youth. `http://www.missingkids.com/missingkids/servlet/ResourceServlet?LanguageCountry=en_US&PageId=869`, 2000.
11. Kaiser Family Foundation. Generation m2: Media in the lives of 8- to 18-year-olds. Technical report, 2010.
12. T.L. Gehm and K.R. Scherer. *Facets of emotion: Recent research.*, chapter Factors determining the dimensions of subjective emotional space, pages 99–113. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc, 1988.
13. K. Hone. Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interact. Comput.*, 18(2):227–245, 2006.
14. J. Klein, Y. Moon, and R.W. Picard. This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, 14(2):119–140, 2002.
15. P.J. Lang, M.M. Bradley, and B.N. Cuthbert. International affective picture system (iaps): Affective ratings of pictures and instruction manual. Technical report, 2008.
16. T.-Y. Lee, C.-W. Chang, and G.-D. Chen. Building an interactive caring agent for students in computer-based learning environments. In *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on*, pages 300–304, 2007.
17. I. Leite, A. Pereira, C. Martinho, A. Paiva, and G. Castellano. Towards an empathic chess companion. In *Proceedings of the 8th International Conference on Autonomous agents and Multiagent Systems (AAMAS 2009)*, pages 33–36, 2009.

18. S. Livingstone, L. Haddon, A. Görzig, and K. Ólafsson. Risks and safety on the internet: The perspective of european children. initial findings. `http://www2.lse.ac.uk/media@lse/research/EUKidsOnline/EUKidsII\%20\%282009-11\%29/home.aspx`, 2010.
19. A. Mehrabian. *Basic Dimensions for a General Psychological Theory*. Oelgeschlager, Gunn & Hain, Cambridge, 1980.
20. F. Mishna, C. Cook, M. Saini, M.J. Wu, and R. MacFadden. Interventions to prevent and reduce cyber abuse of youth: A systematic review. *Research on Social Work Practice*, 2010.
21. W.R. Nugent and H. Halvorson. Testing the effects of active listening. *Research on Social Work Practice*, 5(2):152–175, 1995.
22. A. Ortony, G.L. Clore, and A. Collins. *The cognitive structure of emotions*. Cambridge Univ. Press, 1988.
23. C.E. Osgood. Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, 7(1):1–30, 1966.
24. A. Paiva, J. Dias, D. Sobral, R. Aylett, S. Woods, L. Hall, and C. Zoll. Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence: An International Journal*, 19(3):235–266, 2005.
25. H. Prendinger and M. Ishizuka. The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence: An International Journal*, 19(3):267–285, 2005.
26. M. Schroder, E. Bevacqua, F. Eyben, H. Gunes, D. Heylen, M. ter Maat, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M. Valstar, and M. Wollmer. A demonstration of audiovisual sensitive artificial listeners. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1 – 2, 2009.
27. P.K. Smith, J. Mahdavi, M. Carvalho, S. Fisher, S. Russell, and N. Tippett. Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, 49(4):376–385, 2008.
28. S.R. Sumter and P.M. Valkenburg. Digitaal pesten: de nieuwste feiten. In *CcaM Kennis*, number 1, pages 1–3, 2011.
29. A.D. Thierer. Five online safety task forces agree: Education, empowerment & self-regulation are the answer. *Progress & Freedom Foundation Progress on Point Paper*, 16(13), 2009.
30. R.S. Tokunaga. Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26(3):277–287, 2010.
31. J.M. van der Zwaan, M.V. Dignum, and C.M. Jonker. Simulating peer support for victims of cyberbullying. In *Proceedings of the 22st Benelux Conference on Artificial Intelligence (BNAIC 2010)*, 2010.
32. M.L. Ybarra, K.J. Mitchell, J. Wolak, and D. Finkelhor. Examining characteristics and associated distress related to internet harassment: Findings from the second youth internet safety survey. *Pediatrics*, 118(4):1169–1177, 2006.