Affective Feedback for a Virtual Robot in a Real-World Treasure Hunt

Mary Ellen Foster Mei Yii Lim Amol Deshmukh
Srinivasan Janarthanam Helen Hastie Ruth Aylett
School of Mathematical and Computer Sciences
Heriot-Watt University, Edinburgh, Scotland EH14 4AS
M.E.Foster@hw.ac.uk

ABSTRACT

We explore the effect of the behaviour of a virtual robot agent in the context of a real-world treasure-hunt activity carried out by children aged 11–12. We compare three conditions: a traditional paper-based treasure hunt, along with a virtual robot on a tablet which provides either neutral or affective feedback during the treasure hunt. The initial results of the study suggest that the use of the virtual robot increased the perceived difficulty of the instruction-following task, while the affective robot feedback in particular made the questions seem more difficult to answer.

Categories and Subject Descriptors: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – Evaluation/methodology; I.2.9 [Artificial intelligence]: Robotics – Operator interfaces

Keywords: Intelligent tutoring systems; Affective computing

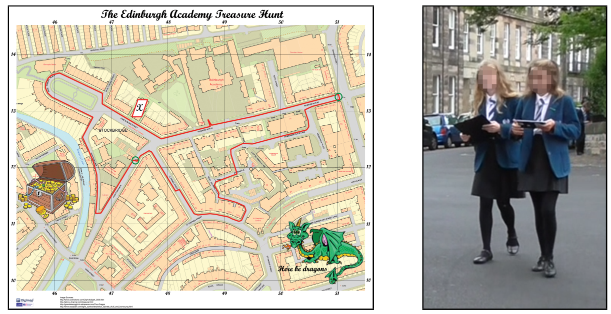
1. THE TREASURE HUNT STUDY

The overall goal of the EMOTE project[1] is to develop an empathic robot tutor for use with 11–14 year olds in a classroom setting. Previous studies on robotic companions in real-world classroom environments [3] have shown that robotic platforms are promising tools for experimental learning. We hypothesise that a robot tutor that is able to detect the user's affective state and respond appropriately will result in increased motivation and better learning outcomes.

A crucial aspect of this overall goal is the specification of appropriate robot behaviour. As part of addressing this issue, we carried out an experiment to investigate how the presence and nature of feedback from a virtual robot affects a child's perception, experience and performance in the context of a real-world treasure hunt activity (Figure 1). This activity requires a child to apply his/her map reading skills in the outside world, and is aimed at children aged 11–12. The students carry out a series of navigation steps in the real world, in groups of two. Each step first requires the students to walk a few yards while making use of their map-reading skills, and then to answer a series of questions regarding their new location: for example, they might need to identify the colour of a nearby door.

[1] `http://www.emote-project.eu/`

MMRWHRI'14, November 16, 2014, Istanbul, Turkey.
ACM 978-1-4503-0551-8/14/11
http://dx.doi.org/10.1145/2666499.2669641.



(a) The treasure map (b) Hunting

Figure 1: The treasure hunt activity

Figure 2: EMYS introducing the activity

To evaluate the impact of affective robot feedback in the context of this treasure hunt activity, we carried out a user study involving 37 students aged 11–12, with three experimental conditions. In all cases, the study began with a physical robot—the EMYS head [2]—introducing the activity to the group of students (Figure 2). One-third of the students then used the paper-based map and questionnaire that have been used in the school treasure hunt activity in previous years.The other students instead used an Android-based tablet application [4] which displayed a digital version of the paper map, along with a virtual EMYS head which presented the navigation instructions and posed the questions. The virtual robot also provided the students with feedback on the correctness of their answers to the questions posed during the treasure hunt; depending on the experimental condition, the feedback was either **neutral** ("correct", "incorrect") or **affective** (e.g., "well done", "too bad"). The GPS location of all participants was tracked as they carried out the treasure hunt activity. At the end of the study, all students answered a set of subjective questions regarding the robot and the task.

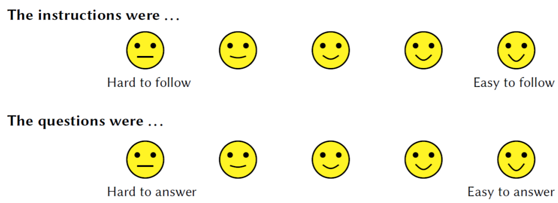


Figure 3: Excerpt from subjective questionnaire

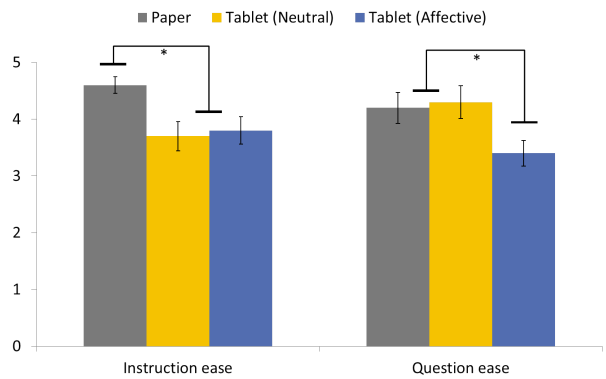


Figure 4: The influence of condition on questionnaire responses

2. INITIAL RESULTS AND DISCUSSION

We gathered two classes of dependent measures in this study: objective and subjective. The objective measures were derived from the GPS logs: we counted the number of intended waypoints that the group reached, as well as the total time taken to complete the route. We could not assess the accuracy of the question responses for technical reasons. The subjective measures were based on the post-study questionnaire, which consisted of three parts:

1. Three questions regarding the student's opinion of the robot agent before the treasure hunt;
2. Four questions regarding the treasure hunt itself; and
3. Seven questions addressing the students' opinion of the virtual agent during the treasure hunt, based on the Likeability section of the Godspeed questionnaire series [1].

Students in the paper-based condition responded only to the first two sections, while students in the tablet conditions also completed the third section. The questionnaire items were presented using a five-point Smileyometer [5], as in Figure 3; this instrument has been shown to be a good tool for evaluating child-computer interactions.

When we analysed the study results, we found that the main effect of the experimental manipulation was on two items on the subjective questionnaire where the students were asked to assess how easy it was to follow the instructions and to answer the questions (the exact items are shown in Figure 3). Figure 4 shows the responses from all participants to these three questions, grouped by the interaction condition. In summary, the students who did the paper-based treasure hunt rated the instructions as significantly easier than the students in either of the tablet-based conditions ($p < 0.05$, two-way Mann-Whitney test). On the other hand, the students who used the tablet with the affective robot found the questions significantly harder than the students in the other two conditions (also $p < 0.05$).

We hypothesise that two different factors contributed to these differences in perceived ease. For the instructions, a crucial aspect of the paper-based treasure hunt was that the students were able to read ahead as much as they wanted in the route; on the other hand, the tablet application presented only one navigation instruction at a time, and only gave the next instruction after the previous one had been completed. This ability to get a global overview of the entire route likely contributed to the differences in perceived instruction ease. On the other hand, we suspect that the difference in the perceived question difficulty is likely due to the particular affective feedback that was selected: saying "brilliant" every time a question was answered correctly could have made the students feel that the question was particularly difficult, while saying "too bad" for every incorrect response may have emphasised the wrong answers and again increased the perceived difficulty. Unfortunately, we are unable to assess the relationship between perceived difficulty and actual question-answering performance, because—as mentioned above—for technical reasons, the latter information is not available.

3. CONCLUSIONS AND FUTURE WORK

We have presented a study designed to assess the impact of affective feedback from a virtual robot in the context of a real-world treasure hunt. The subjective responses suggest that the addition of the virtual robot increased the perceived difficulty of the instruction-following task, while the use of affective robot feedback made the questions seem more difficult to answer. We have proposed potential explanations for both of these results, and are currently carrying out follow-up studies designed to obtain a more complete understanding of the factors involved. Once the targeted follow-up studies are completed, we will use the overall findings of this study to help design the empathic behaviour of the robot tutor in the context of the overall EMOTE system, particularly in the area of affective feedback generation. We will also develop a new version of the tablet app taking into account these results, and will evaluate the new app in the context of the same real-world treasure hunt activity.

4. ACKNOWLEDGEMENTS

This work was partially supported by the European Commission (EC) and was funded by the EU FP7 ICT-317923 project EMOTE. The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein. We thank the pupils and teachers at Edinburgh Academy for their help in carrying out this study.

5. REFERENCES

[1] C. Bartneck, E. Croft, and D. Kulic. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[2] J. Kędzierski, R. Muszyńki, C. Zoll, A. Oleksy, and M. Frontkiewicz. Emys – emotive head of a social robot. *International Journal of Social Robotics*, 5(2):237–249, 2013.

[3] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva. Modelling empathic behaviour in a robotic game companion for children: An ethnographic study in real-world settings. In *Proceedings of HRI 2012*, pages 367–374, 2012.

[4] M. Y. Lim, M. E. Foster, S. Janarthanam, A. Deshmukh, H. Hastie, and R. Aylett. Let's go for a treasure hunt. In *Proceedings of the Affective Agents Workshop (AAW) at IVA 2014*, pages 1–8, Boston, Aug. 2014.

[5] J. Read and S. Macfarlane. Endurability, engagement and expectations: Measuring children's fun. In *Interaction Design and Children*, pages 1–23. Shaker Publishing, 2002.