Perceived Intelligence and Protégée Effect in a Techable Agent Software

Kristian Månsson[1], Magnus Haake[1], Agneta Gulz[1]

[1]*Lund University Cognitive Science*

tkman85@gmail.com

1 Introduction

Studies have shown that the "teachable agent" paradigm, i.e. "learning-by-teaching" using teachable virtual agents in educational software, benefits learning by increasing students' sense of responsibility and supporting metacognition (see for instance Schwartz et al. (2009) and Biswas et al. (2005)). The "protégée effect" is a theoretical concept that describes the beneficial factors of the teachable agent paradigm in that the student makes larger learning efforts when the goal is to teach an agent than when the goal is to learn for themselves (Chase, Chin, Oppezzo & Schwartz, 2009). The authors hypothesize that the following three mechanisms lead to the increased effort of learning: a feeling of responsibility towards the teachable agent, an adoption of an incrementalist view of knowledge, and a protection of their ego since it is the agent that is tested for its learning and potentially fails.

Elaborating on the teachable agent paradigm there is a difference between an agent that can learn and an agent that can be taught (Pareto, Schwartz & Svensson, 2009), leading to different approaches for the design of the AI in teachable agent software. This, in turn, spurs an underlying question of how much effort should be put into the development of the underlying artificial intelligence of a teachable agent. It needs to be teachable, but to what extent does it need to learn versus seem to learn. In order to pursue this question, this study aims to investigate the connection between the protégée effect, perceived intelligence of the teachable agent, and students' learning outcomes.

2 Experiment

The teachable agent (hence TA) educational software used in this study is called Guardians of History (GoH). GoH is aimed at middle school history education (year 4 to 6 in the Swedish educational system) and is developed by the Educational Technology Group at Lund University and Linköping University. In GoH, the student makes time travels to gather information by exploring scenes and interacting with historical characters from different eras in order to subsequently teach the TA in so-called classroom activities (see Figure 1). For this study, a subset of the possible time travels with associated teaching activities were selected for the experiment. After being taught the TA conducts a test, where it provides answers to questions depending on facts it has learned (see Figure 1). The TA was implemented with two different settings: TA_R (as in recency) and TA_A (as in associative). The recency setting (TA_R), corresponds to the original implementation of the TA, where the agent learns (and unlearns) the latest fact that it has been exposed to in each learning activity. That means that for every new learning activity, the agent overwrites all the previous facts learned from previous learning activities. The associative agent (TA_A), developed as a part of a university project course (Bäckström, Månsson, Persson, Sakurai & Karåker Sundström, 2017), is implemented with a basic associative model using weighted concept relations representing the agent's certainty of different facts. The weights of the concept associations increase or decrease depending on the results of the learning activities. Furthermore, the TA_A asks for confirmation of learned facts at random intervals.

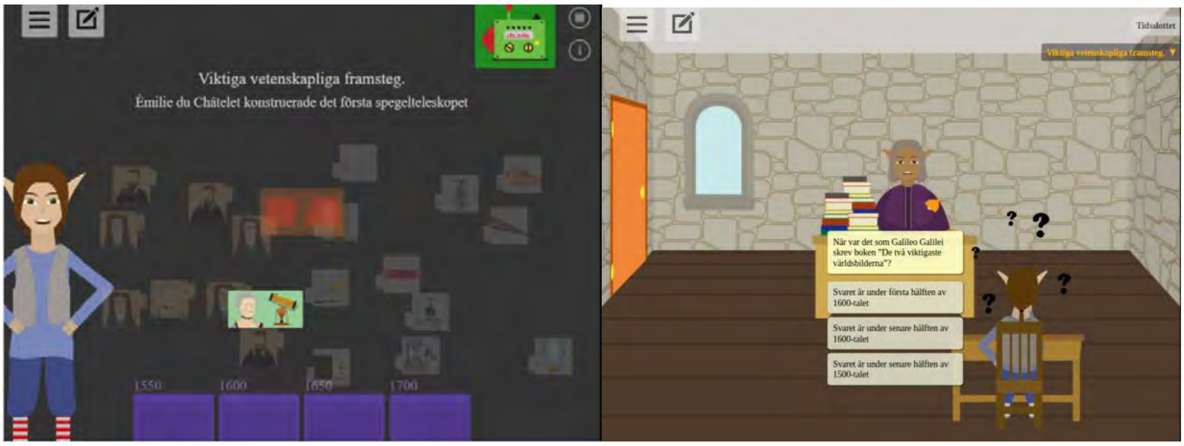

Figure 1. Learning activities with the TA to the left. To the right the test setting.

For the experiment, 94 Swedish grade 5 and 6 students from 5 classes from the same Swedish school were recruited. The students were randomly assigned to one of two groups of 47 students each. In order to play down intervention effects, no information on sex or age was explicitly collected for this study – relying on the random assignment of half the students to one group and half to the other in every class to control for sex or age differences. For the analyses, 3 students were excluded from the dataset due to language difficulties. Another 3 students were excluded as they (for unknown reasons) didn't finish the game. An additional 3 students didn't fill out or hand in their questionnaires, resulting in a data set of $N = 85$ participants with 41 in group TA_R (with 24 in 5th grade and 17 in 6th grade) and 44 in TA_A (with 25 in 5th grade and19 in 6th grade).

The Protégée-effect (PE) was measured using 5 Likert-scale items operationalized by Kirkegaard (2016) for studies with GoH. Perceived intelligence (PI) was measured by 6 items on a semantic difference scale, adapted and translated from Bartneck, Kuli & Croft (2009) who used it for measuring perceived intelligence of a robot. A knowledge test of 10 multiple choice questions based on the content of the game was used to assess the knowledge gained.

3 Results

An independent samples *t*-test was conducted to compare perceived intelligence (PI) in the TA_R and TA_A conditions. Results showed no significant difference in PI between the TA_R ($M = 16.4$, $SD = 4.5$) and the TA_A ($M = 17.8$, $SD = 4.3$) conditions ($t(83) = -1.32$, $p = .19$), i.e. there was no significant difference between the TA_R group and the TA_A group with regard to perceived intelligence as measured by the questionnaire items.

A calculation of the Pearson product-moment correlation coefficient to assess the relationship between PI and PE suggested a large effect (Cohen, 1988) positive correlation ($r = .64$, $p < .001$). The strong correlation indicates that the students' self-reported perceived intelligence of their TA and their self-assessed evaluation of a protégée effect goes hand in hand.

A Matt-Whitney's U test was used to compare the score on the knowledge test between the TA_R group ($M = 0.6$; *Range* = 1-10) and the TA_A group ($M = 0.65$; *Range*: 0-9). No significant difference between the groups ($W = 843$, $p = .60$) could be found, i.e. neither the TA_R group nor the TA_A group performed better as measured by the knowledge test.

Using a Matt-Whitney's U test, no significant correlation between PE and performance score in the knowledge test could be established. No additional interaction effects could be found. No other significant correlations except for the ones stated were found.

4 Discussion

The students did not perceive TA_A as more intelligent. The lack of any significant difference between the groups regarding the perceived intelligence and protégée-effect might point to other factors - such as narrative or the explicitly stated role for the student as the teacher - to elicit the sought-after protégée effect.

The strong positive correlation between perceived intelligence and protégée effect could be interpreted as either that one strongly influences the other or they strengthen each other reciprocally. Another possible explanation is that student's general positive attitude towards the game is what is measured. The strong correlation between how the student either actively ascribe or passively perceive the TA as a thinking and learning agent and the elicited protégée effect, points to one having a strong influence on the other, or both are a manifestation of an underlying phenomena. This correlation should be of interest for researchers as well as designers of teachable agent software.

A lack of correlation between the protégée effect and the score in the knowledge test is somewhat surprising. As the protégée effect is constructed as a theoretical attempt to explain the positive learning outcomes from teachable agents, it might be too coarse to be measured in this way i.e. the result might be sensitive to false negatives. Another possible explanation is that the measurement does not actually reflect the protégée effect; the measurement might rather reflect, or be strongly influenced, by a student's general positive - or negative - attitude towards the software.

A longer study where the students would have more time to interact with the TA might provide further insight towards how the students' perception of the TA varies over time. Validation of the measurements is also of great importance, as they are newly adapted for this study.

References

Bartneck, C., Kuli, D., & Croft, E. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International journal of social robotics, 1*(1), 71–81.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at Vanderbilt University. (2005). Learning By Teaching: a New Agent Paradigm for Educational Software. *Applied Artificial Intelligence, 19* (3–4), 363–392.

Bäckström, J., Månsson, K., Persson, M, Sakurai, E. & Sundström, K. (2017). Designing and Implementing an Associative Learning Model for a Teachable Agent. In Balkenius, C., Gulz, A., Haake, M. & Wallergård, M. (Ed.), *Intelligent, socially oriented technology III: Projects by teams of master level students in cognitive science and engineering.* Lund University

Chase, C., Chin, D. B., Oppezzo, M., Schwartz D.L. (2009) *Teachable Agents and the Protégé Effect: Increasing the Effort Towards Learning*, Stanford University School of Education.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Oxford, England: Routledge

Kirkegaard, C., (2016). *Adding Challenge to a Teachable Agent in a Virtual Learning Environment*, Diss., Linköping University

Pareto, L., Schwartz, D. L., & Svensson, L. (2009). Learning by Guiding a Teachable Agent to Play an Educational Game. *Frontiers in Artificial Intelligence and Applications, 200* (1), 662–664.

Schwartz, D.L., Chase, C., Chin, D., Oppezzo, M., Kwong, H., Okita, S., Roscoe, R., Hoyeong, J. Wagster, J. Biswas, G., (2009) Interactive Metacognition: Monitoring and Regulating a Teachable Agent. In D.J. Hacker, J. Dunlosky, \& A.C. Graesser (Eds.), *Handbook of Metacognition in Education*, 340-358