Does embodiment affect tutoring behavior?

Katrin Lohan, Sebastian Gieselmann, Anna-Lisa Vollmer, Katharina Rohlfing, Britta Wrede

2010

Preprint:

This is an accepted article published in Proc. IEEE 9th Int Development and Learning (ICDL) Conf. The final authenticated version is available online at: https://doi.org/[DOI not available]

Does embodiment affect tutoring behavior?

Katrin Solveig Lohan[x], Sebastian Gieselmann[||], Anna-Lisa Vollmer[||],
Katharina J. Rohlfing[||], and Britta Wrede[||]
[x ||]Bielefeld University, CoR-Lab, Applied Informatics Group, Bielefeld, Germany, http://www.cor-lab.de

I. Introduction

In recent years, research has moved towards the learning by interaction paradigm [1] suggesting that interaction with an artificial agent is facilitated when characteristics of a social interaction are considered. It is envisioned that agents will learn from humans by simply interacting with each other. In such a scenario, learning by interaction " goes beyond common supervised or unsupervised strategies by taking into account wider feedback and assessments for the learning processes " ([1] p.140). So far, little is known about interactional processes and feedback strategies involved. Yet, in order to learn, a learner will typically need to be provided with information given by a teacher who not only gives certain structure to the interaction but also instructs for and demonstrates the learning contents. The given information can only be effective if the learner is receptive. To assure this, the tutor makes use of interactive regularities checking the learner's behavior. The term contingency has been suggested to encompass such regularities in interaction. More specifically, it refers to a temporal sequence of behavior and reaction ([2], [3]). It has been shown that contingency is an important factor in interactions with infants and contributes to the cognitive development of infants [4]. In the interaction with an artificial agent, contingency has been operationalized by eye-gaze bouts [5]. So far, it has been shown that while in a situation with a child, eye-gaze bouts in total, average and frequency is much higher than in interaction with an adult as learner. In a situation with an artificial agent a decrease of eye-gaze bouts could be observed [5]. It could thus be reasoned that tutor's monitor behavior is impaired when interacting with a robot. However, so far only an interaction with a virtual robot has been investigated. In contrast, an embodied robot could evoke a more natural tutor behavior. In this study, we therefore investigated a tutoring situation with an embodied robot and focused on tutor's monitoring behavior. We followed the minimal definition of embodiment by K. Dautenhahn et al. to quantify the difference in embodiment between these two systems [6]. Accordingly the *Degrees of Embodiment* (DOM) are calculated as: $DOM_{S,E} = f(x, y, t)$, where system *S* in respect to an environment *E* is calculated by a function f of the vectors *x* and *y* and the time *t*. *x* describes the number of sensors, the detected modalities of the sensors and the channels of information provided by the sensors, *y* describes the *degree of freedom* (DoF) of the robot. In our experiments E remained unchanged and thus the DOM is not affected by this factor. We argue that in our case, the function f is only dependent on the degrees of freedom of the robot. Possibly, we can give a value which represents the *difference in the degree of embodiment* (DDOM). Since our goal was to study some characteristics of contingency ([7], [5]), our dependent variable was the tutor's monitoring behavior operationalized by the eye-gaze.

II. Technical Setup

Following the idea of the difference in the degree of embodiment (DDOM), the degrees of freedom (DoF) in our two setups are as follows: *The simulated robot Akachan* mainly consists of two components: the robot simulation and a model of a saliency-based visual attention system. This simulation has four degrees of freedom (DoF): the eyes have two DoF, the eyelids and the mouth, but only the eyes are controlled by the outcome of the saliency-based visual attention system. The model of the visual attention system, inspired by the behavior and the neuronal mechanism of primates, can detect salient locations in a scene, based on changes the intensity, the orientation, and the motion i.e. optical flow, see [8]. *The physical embodied iCub robot* has fiftythree DoF for the whole body. We applied two different embodiment conditions: One where the iCub is with moving the whole *Head* and another one where it was only shifting the eyes (*NoHead*). We designed the gazing behavior of the two robots similar as it has been done by T. Farroni et al. [9] who demonstrated a face, in different conditions, to 4 to 5 month old infants. In their study, they showed that even young infants show a faster saccadic reacting time, when there was a shift in the demonstrated face than if there was a shift in the eye-gazing behavior. To be able to compare this robot with the Akachan simulation, we were only using six DoF for the *Head* movement condition and three for the *NoHead* movement condition.

III. Experiment

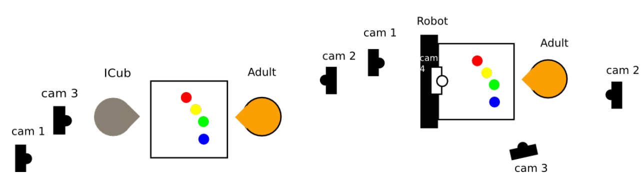

Fig. 1. This graphic shows both setups we used for our experiments.

The data of a group of 14 participants were conducted in both of our two experiments. The first one was with the simulated robot called Akachan [5]. The second one was an infant-like (2-4 year old) looking embodied robot, the iCub [10]. Both robots were equipped with the same visual attention system.

TABLE: Design and Subjects

Robot/Variable	Head	NoHead	Sum
Akachan	6	8	14 (9 female and 5 male)
iCub	6	8	14 (9 female and 5 male)

TABLE I

The group of participants in the Akachan condition was split into two further groups. One of the sub-group was paired with the iCub condition.

Setting: The participants were instructed to explain six objects to the robot. For the analysis in this paper, only the stacking cups were chosen. The robot's behavior was controlled by the same salience system which was used for the Akachan experiments, but there were two different controlled behaviors. One behavior was that only the robot eyes were following the most salient point of the scene (*NoHead*). In the second behavior the whole head and the eyes of the robot were following the most salient point of the scene (*Head*).

Fig. 2. These three pictures show the difference between looking to the object (left), looking to the interaction partner (middle) and looking somewhere else (right).

Annotation: For analyzing the data, (1) the action of the stacking-cups task, and (2) the sub-actions (a1-a3) of grasping one cup until releasing it into the end position were marked in the video. We defined action as the whole process of transporting all objects to their goal positions, and subaction as the process of transporting one object to its goal position. The dependent variable *eye-gaze* was annotated in three categories (Fig. 2): looking at the interaction partner, looking at the object and looking somewhere else. We then calculated our measures ([7], [5]): *Frequency of eye-gaze bouts to interaction partner*, i.e. eye-gaze bouts per minute, was computed from the *Interact* annotations. Also, the *average length of eye-gaze bout to interaction partner* and the *total length of eye-gaze bouts to interaction partner* as the percentage of time of the action spent gazing at the interaction partner were computed.

IV. Results

TABLE: Akachan vs iCub

variable	*M*	*SD*	t(5)	p
frequency of eye-gaze bouts to object	5.61	2.23	2.51	0.054
average length of eye-gaze bout to object	-1.69	1.03	-3.998	0.010
total length of eye-gaze bout to interaction partner	2.27	8.62	2.63	0.046
total length of eye-gaze bouts to object	-2.41	8.14	-2.961	0.031

TABLE II

For all measures we calculated a ***student T-test****. Equating the Akachan vs. Head we used a paired* ***student T-test****. Because of the design it was not possible to calculate a Annova. All variables where normally distributed.*

V. Discussion and Conclusion

In this study, we investigated a tutoring situation with an embodied robot and focused on tutor's monitoring behavior. We reasoned that if there is a difference in the tutoring behavior between a physical embodied robot and a simulated robot, we will find significant differences in the eye gazing behavior of the tutor. We further studied embodiment in the *Head* and *Nohead*-condition. With respect to the question whether people teach an embodied robot in a different manner than a robot simulation, we conclude that yes, an embodied robot is taught differently. However, this might be due to the DDOM rather than simply embodiment. Concerning our findings in the *Head* condition, iCub was gazed at longer and the shifts of the gazing were higher towards the object in this condition. In the Akachan condition, we found that during the task, the tutor was looking longer towards the object than in the *Head* condition. The data indicate that there is an increase of tutor's monitoring behavior in the *Head* condition, because there was more "checking" towards the iCub. This gives us the opportunity to think about the differences in the perception of an embodied robot like the iCub and additionally about possible cues that guide the tutors perception of the robot.

References

[1] B. Wrede, K. Rohlfing, M. Hanheide, and G. Sagerer, "Towards learning by interacting," *Creating Brain-Like Intelligence*, pp. 139–150, 2009.

[2] G. Gergely and J. Watson, "Early socio-emotional development: Contingency perception and the social-biofeedback model," *Early social cognition: Understanding others in the first months of life*, pp. 101–136, 1999.

[3] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. Nehaniv, K. Fischer, J. Tani, B. Belpaeme, G. Sandini, L. Fadiga, B. Wrede, K. Rohlfing, E. Tuci, K. Dautenhahn, J. Saunders, and A. Zeschel, "Integration of action and language knowledge: A roadmap for developmental robotics."

[4] T. Striano and D. Stahl, "Sensitivity to triadic attention in early infancy," *Developmental Science*, vol. 8, no. 4, pp. 333–343, 2005.

[5] A. Vollmer, K. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K. Rohlfing, and B. Wrede, "People modify their tutoring behavior in robot-directed interaction for action learning," in *Proceedings of the International Conference on Development and Learning*, 2009.

[6] K. Dautenhahn, B. Ogden, and T. Quick, "From embodied to socially embedded agents–implications for interaction-aware robots," *Cognitive Systems Research*, vol. 3, no. 3, pp. 397–428, 2002.

[7] R. Brand, W. Shallcross, M. Sabatos, and K. Massie, "Fine-grained analysis of motionese: eye gaze, object exchanges, and action units in infant-versus adult-directed action," *Infancy*, vol. 11, no. 2, pp. 203–214, 2007.

[8] Y. Nagai, C. Muhl, and K. Rohlfing, "Toward designing a robot that learns actions from parental demonstrations," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, 2008, pp. 3545–3550.

[9] T. Farroni, M. Johnson, M. Brockbank, and F. Simion, "Infants' use of gaze direction to cue attention: The importance of perceived motion." *Visual Cognition*, vol. 7, no. 6, pp. 705–718, 2000.

[10] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub humanoid robot: an open platform for embodied cognition," in *PerMIS: Performance Metrics for Intelligent Systems Workshop. Aug*, 2008, pp. 19–21.