Towards Effective Robot Tutoring for Skills Acquisition

Giulia Belgiovine*[1,2], Francesco Rea[1], Jacopo Zenzeri[1], Alessandra Sciutti[1]

***Abstract*—Effective tutoring during skill learning requires to provide the appropriate physical assistance to the learners, but at the same time to assess and adapt to their affective state to avoid frustration. With the aim of endowing robot tutors with these abilities, we designed an experiment to evaluate how training with a humanoid social robot influenced the performance, experience, and communicative behavior of naive learners. Participants had to learn to balance an unstable inverted pendulum with the support of physical assistance. The presence of the humanoid robot increased the involvement in the task. Moreover monitoring participants' facial expressions proved helpful to recognize when the task was particularly challenging. These findings will enable the robot to adapt its tutoring strategy to the learner's needs in real-time.**

***Index Terms*—Robot Tutor, Skills Learning, Social Robot**

I. Introduction

The role of the expert tutor in the skill transfer process is of critical importance and often relies on a series of implicit signals involving several communication channels. Considering for example the rehabilitation context, beyond the selection of the appropriate physical assistance, the expert tutor should be mindful of the state of their patients, keeping them committed to the task while monitoring the stress, anger, or other reactions that might be triggered by the lengthy and often challenging rehabilitation process. Given the widespread adoption of robotics in the context of rehabilitation, it would be desirable that also robot tutors exhibit a similar ability of understanding and adapting to learners' needs. Several studies have demonstrated the potential of social robots to positively contribute to users' learning and experience in the field of skills acquisition [1]. It has been shown that the presence of physical robots may have advantages in sensing and using affective data, by inducing higher degrees of emotional expressiveness [2]. With the final aim to design an optimal tutoring architecture for social robots, we conducted a study to address the following aspects: (1) understanding the effect of the behavior of the robot tutor on the performance and experience of naive participants [3], and (2) investigating different informative cues that would allow the robot to build a comprehensive model of the learners' skills and state, and to act accordingly. In particular, we deepened the understanding of the implicit affective signals detectable by the robot and whether the behavior of the social humanoid tutor changed the communicative behavior of the participants [4].

This work was supported by a Starting Grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme. G.A. No 804388, wHiSPER. [1] Istituto Italiano di Tecnologia, Genova, Italy. [2] Università degli Studi di Genova, Genova, Italy. * giulia.belgiovine@iit.it

Fig. 1: The robot looking at the pendulum with a focused expression while the participant tries to balance it by using the Wristbot.

II. Methods

We recruited 32 participants (18 females, 14 males, age 26.1±3.5). They had to learn the right strategy to accomplish a complex motor task, i.e. keep in balance an inverted pendulum for as long as possible by using the Wristbot, a robotic manipulandum used as a joystick. To have a task that met an optimal challenge level, we implemented virtual dynamics that determined the angular orientation of the pendulum based on the angular position of the Wristbot. The dynamics included a non-linear spring, virtually connecting the Wristbot to the pendulum, and a viscous force-field in which the pendulum moved. The protocol comprised 3 different sessions: i. baseline (1 trial), ii. training (5 trials) and iii. test (3 trials). During the training phase, the task dynamics were facilitated by the assistance that wanted to simulate the help of a trainer that intervenes in the task by dampening the fall of the pendulum, making it easier to control. Half of the participants were tested in the Control group, and the remaining in the iCub group. Both groups performed the training with the same physical assistance implemented through the Wristbot. However, while for the Control group the assistance was attributed to the Wristbot, for the iCub group, the humanoid robot iCub pretended to control the pendulum and to provide assistance. It also presented social behaviors such as talking, looking at participants' faces with a happy expression, focusing on the pendulum. However, its behavior did not adapt to participants' performance, to emulate the condition in the Control group where no feedback was provided by the Wristbot.

III. Results

In the training session, the iCub and the Control groups showed significantly higher performance than their baseline ones (Fig. 2). The removal of the assistance caused a deterioration of the average performances that, however, remained significantly higher than the initial baseline values. This revealed that the training was effective and that it led to the ability to generalize the skill also in the absence of assistance. The average improvements in the performance of the 2 groups were not statistically different. Since the designed task required participants to monitor and act on the pendulum almost continuously, we can conclude that the presence of iCub embodying the assistance and showing social behaviors did not interfere with the concentration but led to improvements similar to the one of the Control group. We evaluated participants' experience through the Intrinsic Motivation Inventory (IMI), comprising the sub-scales *Competence*, *Effort/Importance*, and *Interest/Enjoyment*; the latter is considered the self-report measure of intrinsic motivation. The results indicated that the score of the sub-scale *Interest* of the iCub group was significantly higher than the one of the Control group ($Z = 3.29$, $p < .001$). Instead, there was no significant difference for the *Competence* and *Effort* scales, meaning that participants' perception of their ability and commitment to the task did not change among conditions. In summary, for both groups, the task was similarly challenging, but the presence of iCub influenced the way participants approached the task, making the training experience more enjoyable for them. To analyze participants' affective state, we used the FaceChannel neural network [5], which describes facial expressions using a continuous representation of Arousal and Valence within the range of -1 and 1, representing calm/negative and excited/positive, respectively. A Mixed Model ANOVA revealed that there was a significant effect of sessions for Valence, ($F(1.34, 28.1) = 8.54$, $p < .01$), while there was no significant effect of groups, nor of the interaction between the two. Indeed the Valence recorded in the baseline was significantly higher than the one recorded in the training and the test sessions. Probably, the expressions of the participants reflected a reaction of sarcasm and disbelief in facing such a difficult task, and they were associated with a feeling of frustration. Therefore, starting from facial expressions, we can infer the difficulties experienced and the challenge level. For Arousal, there was no significant effect of groups or sessions. Only the interaction approached significance, with a tendency for Arousal to decrease over trials towards negative values in the Control group.

IV. Discussion

We proposed a novel experimental design where it is possible to record and send both haptic signals by using the robotic device Wristbot, and social cues, through the humanoid robot iCub. We have demonstrated that learning in the context of interaction with a humanoid robot assistant led subjects to increased motivation and a more enjoyable training experience, without negative effects on attention and perceived effort.

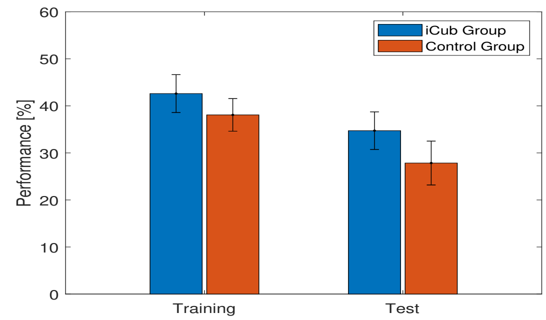

Fig. 2: Mean and SE of performance improvements with respect to the baseline, in training and test sessions.

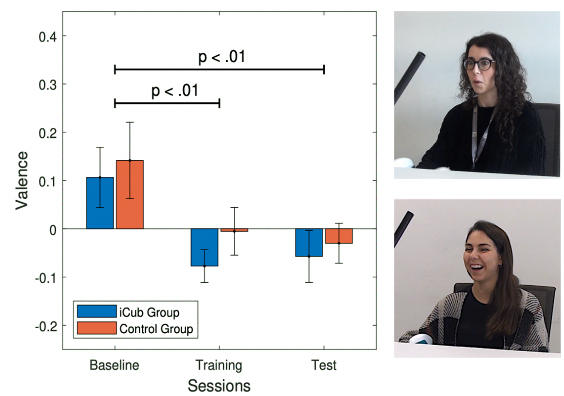

Fig. 3: Mean and SE of Valence in the different sessions (left). Examples of facial expressions in the baseline session (right).

Moreover, the Valence of the naive participants seemed to be mostly driven by the difficulty experienced in the task rather than by the presence of the humanoid tutor. In the future, we want to develop an adaptive assistive policy by integrating different communicative cues and let the robot learn how to improve the tutoring strategy while interacting with naive learners.

References

[1] J. Fasola and M. J. Matarić, "Robot motivator: Increasing user enjoyment and performance on a physical/cognitive task," in *2010 IEEE 9th International Conference on Development and Learning*, pp. 274–279, IEEE, 2010.

[2] S. Spaulding, G. Gordon, and C. Breazeal, "Affect-aware student models for robot tutors," in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 864–872, 2016.

[3] G. Belgiovine, F. Rea, J. Zenzeri, and A. Sciutti, "A humanoid social agent embodying physical assistance enhances motor training experience," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 553–560, IEEE, 2020.

[4] G. Belgiovine, F. Rea, P. Barros, J. Zenzeri, and A. Sciutti, "Sensing the partner: Toward effective robot tutoring in motor skill learning," in *International Conference on Social Robotics*, pp. 296–307, Springer, 2020.

[5] P. Barros, N. Churamani, and A. Sciutti, "The facechannel: A fast and furious deep neural network for facial expression recognition," *SN Computer Science*, vol. 1, no. 6, pp. 1–10, 2020.