Affective Personalization of a Social Robot Tutor for Children's Second Language Skill

Goren Gordon[a,b,c], **Samuel Spaulding**[a,c], **Jacqueline Kory Westlund**[a,c], **Jin Joo Lee**[a,c],

Luke Plummer[a], **Marayna Martinez**[a], **Madhurima Das**[a], **Cynthia Breazeal**[a,c]

[a] Personal Robots Group, MIT Media Lab,
20 Ames Street E15-468, Cambridge, MA 02139
[b] Curiosity Lab, Industrial Engineering Department, Tel-Aviv Univerisity, Israel
[c] {ggordon,jakory,samuelsp,jinjoo,cynthiab}@media.mit.edu

Abstract

Though substantial research has been dedicated towards using technology to improve education, no current methods are as effective as one-on-one tutoring. A critical, though relatively understudied, aspect of effective tutoring is modulating the student's affective state throughout the tutoring session in order to maximize long-term learning gains. We developed an integrated experimental paradigm in which children play a second-language learning game on a tablet, in collaboration with a fully autonomous social robotic learning companion. As part of the system, we measured children's valence and engagement via an automatic facial expression analysis system. These signals were combined into a reward signal that fed into the robot's affective reinforcement learning algorithm. Over several sessions, the robot played the game and personalized its motivational strategies (using verbal and non-verbal actions) to each student. We evaluated this system with 34 children in preschool classrooms for a duration of two months. We saw that (1) children learned new words from the repeated tutoring sessions, (2) the affective policy personalized to students over the duration of the study, and (3) students who interacted with a robot that personalized its affective feedback strategy showed a significant increase in valence, as compared to students who interacted with a non-personalizing robot. This integrated system of tablet-based educational content, affective sensing, affective policy learning, and an autonomous social robot holds great promise for a more comprehensive approach to personalized tutoring.

Introduction

Socially assistive robotics (SAR) is an emerging field which strives to create socially interactive robots that aid people in different areas of their lives, such as education and care for the elderly Tapus, Maja, and Scassellatti (2007); Fasola and Mataric (2013). Educational assistive robots are designed to support children's learning and development, e.g., in the classroom Movellan et al. (2009); Chang et al. (2010) or in one-on-one tutoring sessions Saerbeck et al. (2010); Kory, Jeong, and Breazeal (2013); Fridin (2014). However, children may learn in different ways and at different paces. In order to teach children most effectively, one must personalize the educational interaction to each child VanLehn (2011). While many intelligent tutoring systems personalize the curriculum to each student (e.g., the order and type of questions asked or the content presented), fewer have addressed the equally critical aspect of personalizing the tutoring interaction to the affective state of the student. If a student is discouraged by the material and disengages from the tutoring system, personalization of the educational content may be ineffective - the student is not attending to it all. Recent innovations in affective sensing technology, such as McDuff, Kaliouby, and Picard (2012), have allowed researchers to begin using students' affective responses as input to intelligent tutoring systems, such that the system can respond to students' affective states Woolf et al. (2009); Arroyo et al. (2004); Vanlehn et al. (2011). However, so far, these systems generally incorporate simple rule-based extensions to deal with students' affective states, and are based on virtual tutors, not physically embodied robots Nye, Graesser, and Hu (2014). Additional prior work has shown that a physically embodied tutor may be more effective than a virtual tutor Leyzberg, Spaulding, and Scassellati (2014).

In this work, we present an integrated affective tutoring system that uses an integrated child-tablet-robot setup Gordon and Breazeal (2015); Jeong et al. (2014); Kory and Breazeal (2014). The supportive affective behavior of a robotic tutor is autonomously learned and personalized to each student over multiple interactive tutoring sessions. The system is composed of four primary components:

1. a novel, fully autonomous social robot platform (called Tega), which was specifically designed to be engaging for children, is robust enough to work continuously for several hours, and is portable in order to be deployed in the field;
2. a novel, educational Android tablet app that allows for general curriculum generation and seamless integration with the social robot;
3. an Android smartphone that uses the commercial Affdex SDK to automatically analyze facial expressions in real-time McDuff, Kaliouby, and Picard (2012); and
4. a cognitive architecture that integrates and feeds affective information from Affdex and educational information from the tablet into an affective reinforcement learning al-

gorithm, which determines the social robot's verbal and non-verbal behavior.

The integration of all these components was enabled by using the Robot Operating System (ROS) throughout the setup Quigley et al. (2009).

We evaluated this system in a real world experimental paradigm. Native English-speaking preschool children (ages 3-5) interacted with the system to learn second language vocabulary (Spanish) in their own classroom over a two-month period. We first show that the tutoring setup facilitates these children's learning of new Spanish words. We then analyze the performance of the affective reinforcement learning algorithm to show that it personalized to specific children - i.e., the algorithm adapted in different ways for each child. An analysis of the effects of the robot's behavior on children's detected valence and engagement shows that only children's positive valence is robustly and significantly changed immediately following a set of non-verbal actions by the robot. Consequently, we compared the valence of children who interacted with either a personalized or a non-personalized robot tutor. We found that the change in valence between the first and last sessions over the two-month period was significantly different between the two conditions. That is, in the non-personalized condition, positive valence decreased, while in the personalized condition, positive valence increased. These results, obtained using an integrated system that combines educational content, affective sensing, and an expressive social robot deployed in a real-world, long-term interaction study, shows that affective personalization of social robotic tutors can positively influence the students' affect in constructive and meaningful ways.

Related Work

Intelligent Tutoring Systems (ITSs) refer to a wide variety of computer-based educational tools. Common features of an ITS include the ability to change its behavior in response to student input, provide help in the form of a hint or additional instruction, and conduct some form of evaluation of the user. VanLehn VanLehn (2011) distinguishes between two broad classes of computer-based tutors. First, 'Computer-Aided Instruction' (CAI) systems are characterized by a 'digital workbook' style that provides hints or feedback on students' answers. Second, 'Intelligent Tutoring Systems' are characterized by interactivity, open-response answers, and feedback on students' *process* towards a solution, rather than just the solution itself.

ITSs are already in use outside of the lab, in schools or daycare classrooms. But, as is the case in many of the applied sciences, the deployed systems typically lag behind the cutting edge of research. Thus, while commercial tutoring systems rarely consider students' affective or emotional states, the research community has begun to address these problems. The subfield of "affect-aware tutors" Woolf et al. (2009) seeks to design more effective ITSs that explicitly sense, model, and reason about students' affective states. Inspired by psychological theories of emotion and learning, affect-aware tutors seek to foster engagement and learning from data-driven estimates of students' affective states. For example, the Wayang geometry tutor is a system that features a virtual agent which helps students solve geometry problems Arroyo et al. (2004). In order to foster engagement, the tutor uses an empathy-based affective behavior system: the emotional actions of the tutor are intended to mirror the (estimated) emotional state of the user. For example, if a child appears bored, the tutor might also display signs of boredom before suggesting a new topic or problem to keep the student engaged.

Recent efforts to develop affect-aware tutoring systems have culminated in a number of major systems that have been extensively studied, including the Wayang Tutor Arroyo et al. (2004) and Affective Meta-Tutor Vanlehn et al. (2011) projects) which have been extensively studied. Yet much of the work on affect and modeling in the ITS literature focuses on models to *infer* affect. Typically, once affective states are detected or identified, they trigger simple behavioral rules - a tutor might change its facial expression or offer a supportive comment. However, these rules are hard-coded by the developers and remain fixed throughout the deployment.

On the other hand, the research and development of social robot tutors has recently been flourishing Movellan et al. (2009); Leyzberg, Spaulding, and Scassellati (2014); Deshmukh et al. (2015); Kanda et al. (2004). For example, RUBI-4 is a humanoid robot with articulated arms, an expressive face, and a tablet embedded in its midsection that played simple vocabulary games with preschool children Movellan et al. (2009). Personalization of robot tutors, even via simple algorithms, has been shown to greatly increase tutoring effectiveness compared to non-personalized robot tutors as well compared to virtual robots Leyzberg, Spaulding, and Scassellati (2014). While more sophisticated learning approaches to personalized robot tutors have been studied Gordon and Breazeal (2015), understanding how to effectively personalize tutor behavior to support positive student affective is still an open problem.

Interaction Design

We created a complete interaction scenario for evaluating our affective social robotic tutor. The educational goal was learning new words in a second language, in this case, Spanish. We used a child-tablet-robot scenario, wherein the tablet provided a shared context for the child-robot interaction. The child and the robot worked together to help a Spanish-speaking in-game virtual character (a 2D animated Toucan) travel to Spain. The game had content revolving around a trip to Spain: packing for the trip, visiting a zoo, having a picnic with friends, and so forth. These different adventures provided the opportunity both to learn new words and to review previous content, consistent with best educational practices of spaced repetition. While a game in which the robot provides both the curricular content and affective support could be envisioned, here, the robot and Toucan were created as separate characters. This allowed the robot to be presented as a peer tutor and teammate, on the child's level, rather than as a teacher. In addition, since the curriculum

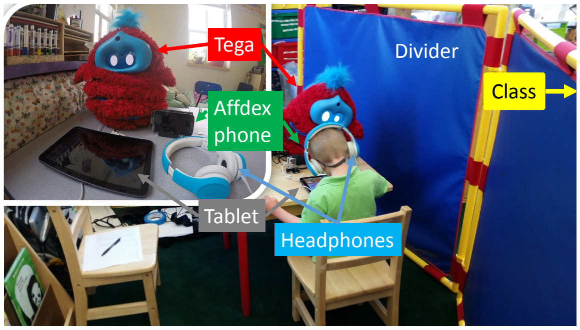

Figure 1: In our evaluation study, the child sat across the table from the robot. The tablet was placed on the table between them. The smartphone collecting data via Affdex sat beside the robot. The child wore headphones that were connected to the tablet and to the robot.

was reinforced, but not presented, by the robot, this better mimics some tutoring scenarios, in which children work with a tutor to learn new information that has been presented by a third party. The Toucan, on the other hand, did not communicate supportive/affective information, only curricular/instructional cues. Both the robot and the toucan communicated via pre-recorded speech and animations in order to convey a realistic and engaging personality, which cannot be easily achieved with an artificial voice. Game scripts determined the dynamics of the interaction between the tablet game, the virtual Toucan, and the game-related responses of the robot.

Platform Design

Tega is a new robot platform that is designed and developed specifically to enable long-term interactions with children. The complete integrated system was comprised of a Tega robot, a novel Android tablet app that enabled seamless communication between the game content and the robot, an Android smartphone app using the commercial Affdex SDK for automatic facial expression, and a portable Ubuntu workstation and router for the local communication. All components of the system were either designed or adapted to work with the Robot Operating System (ROS) that alleviated the synchronization problem often associated with such complex and integrated systems. All data was continuously logged in the form of ROS messages in synchronized ROS-bag files for later analysis.

Robotic Platform

The Tega robot platform is a part of a line of Android-based robots Setapen (2012) that leverage smart phones to graphically display the robot's animated face as well as drive computation, including behavioral control, sensor processing, and motor control for its five degrees of freedom (DoF).

Tega's expressive joints are combinatorial and consist of five basic DoFs: head up/down, waist-tilt left/right, waist-lean forward/back, full-body up/down, and full-body left/right. The robot is designed for robust actuator movements such that the robot can express consistent behaviors over long periods of time. For example, the foundation of its kinematic chain is based on a lead-screw design that enables the robot to rapidly and reliably exercise its primary squash-and-stretch joint.

For long-term continual use, the robot has an efficient battery-powered system that can run for up to six hours before needing to be recharged. The robot's electronic design extends the smartphone's abilities with on-board speakers and an additional high-definition camera that has a wide field of view of the scene and users. Aesthetically, the robot is the size of a teddy bear (about 11 inches tall), brightly colored, and with a furry exterior to appeal to young children Kory, Jeong, and Breazeal (2013).

ROS-Integrated Educational Tablet App

We designed an educational Android-based tablet game to provide the second-language learning curriculum. This component of the integrated system was written with the Unity 3D game engine (version 4.6.2), designed to be an extensible game platform that could plug in to any child-robot interaction requiring a tablet. For each interaction session, experimenters could load custom images or sounds required for the specific scenario, on demand and in real-time. The game included a virtual character - a 2D animated toucan - who was shown in the lower left hand corner of the screen on the tablet. The toucan presented the Spanish words and gave cues regarding the curriculum and instructions on how to play the game, but did not communicate any supportive or affective information.

Because the Unity app on the tablet could not natively run a ROS node, the app was designed to send and receive JSON messages over a WebSocket connection to the Rosbridge server, which provided a communication link to the other ROS nodes in the system. These capabilities allowed the game logic in the interaction between the robot, child, and Toucan to be abstracted from the app running on the tablet. Furthermore, the game logic and the game content could be modified without having to re-compile or re-deploy the app on the tablet. This allows the game to be easily reused in future studies.

Automatic Facial Expression Detection

In order to analyze children's emotional expressions, we used the Affdex mobile SDK. This is a commercial tool marketed by Affectiva, Inc. to enable developers to develop affect-aware mobile applications. Affdex uses real-time face detection and analysis algorithms to extract estimates of four physical facial expression features (Smile, BrowFurrow, BrowRaise, and LipDepress) and two hidden affective features (Valence and Engagement), either from offline video or from real-time video streams of people's faces. For each of these six metrics, the SDK produces a number in the range of $[0, 100]$, with the exception of Valence, which ranges from $[-100, 100]$.

Due to the nature of the current SDK, which runs only on Android or iOS devices, a separate Android phone was integrated into the setup for running Affdex. However, as the technology develops, the Affdex SDK could potentially be

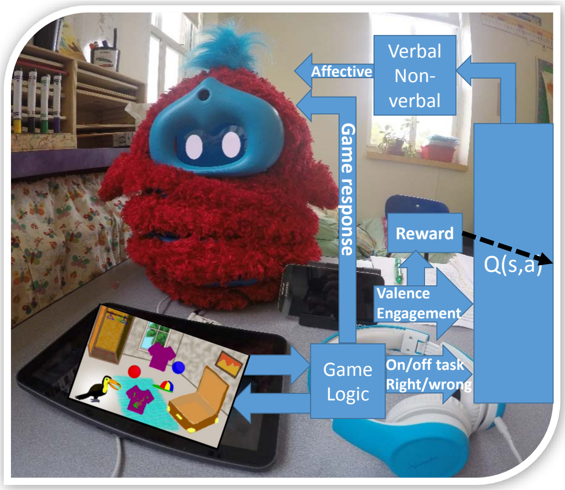

Figure 2: This diagram shows the robot's behavior and affective policy components, with a superimposed screenshot from the tablet app. The Toucan on the left-bottom corner on the tablet was a constant virtual character that provided the curriculum, i.e., the Spanish words. The rest of the screen was dynamically loaded for each game interaction that the child and robot played.

integrated directly into the robot. The source code was modified to allow transmission of the Affdex outputs via ROS over the local network. The raw data we used were the valence and engagement values, which ranged from $\hat{s}_{val} \in [-100, 100]$ and $\hat{s}_{eng} \in [0, 100]$, respectively. The affective data was received at 20 frames per second, but only when a face was detected. Due to the sporadic nature of the raw data, we applied a median smoothing window of 20 frames.

Robot's Behavior and Affective Policy

The social robot companion's behavior was dictated by two distinct components, namely, the game logic and the affective policy, Fig. 2. During the games, the robot supplied instructions (English game) and hints (Review and Spanish games). The robot also supplied general encouragement when the child was correct, e.g. "Good job!" or "You're working hard!", as well as hints when the child was wrong, e.g. "I think it was that one." Finally, the robot's gaze followed the required response. For example, the robot would gaze at the tablet when the child was expected to interact with it, or would gaze at the child while talking to her.

The affective policy augmented the robot's behavior in order to increase its affective role as a supporting companion in this tutoring setup. The policy was active whenever the child completed a specific required task or a task's maximal duration expired without the child's interaction. Thus, for example, when the child was correct, the robot responded both with the game-related response ("You're working hard!") and with the appropriate affective response (described below).

The policy was formulated as a $Q(s, a)$ matrix, where s represent the affective- and game-states, whereas a represent the affective responses. The state-space consisted of four dimensions: (1) valence, supplied by the Affdex software, smoothed and discretized to three values $s_{val} = \{Neg, Med, Pos\}$ for $\hat{s}_{val} = \{[-100, -33], (-33, 33], (33, 100]\}$, respectively; (2) engagement, supplied by the Affdex software, smoothed and binarized to two values $s_{eng} = \{Lo, Hi\}$ for $\hat{s}_{eng} = \{[0, 50), [50, 100]\}$, respectively; (3) on/off-task, a binary value representing interaction with the tablet in the previous 5 seconds; e.g., if the child interacted with the tablet in the previous 5 seconds prior to the requirement of the an affective response, this state will be $s_{task} = 1$; and (4) right/wrong, a binary value representing the correctness of the last requested response for the child during the game, e.g., $s_{cor} = 1$ when the child was correct. In total, the state-space consisted of $3 \times 2 \times 2 \times 2 = 24$ states.

The action space of the affective policy represented the robot's affective state, also divided into engagement and valence representations. Thus, the robots action-space consisted of $3 \times 2 + 1 = 7$ actions, namely, $a = \{\{Neg, Med, Pos\} \times \{Lo, Hi\}, NoAction\}$, where the latter is an important addition, since sometimes not responding to a student's mental state may be the best supporting action. These affective states were then translated to generic non-verbal behaviors and game-specific verbal utterances. Below is a non-exhaustive list of these affective-states paired with example non-verbal (italics) and verbal responses for a scene in which the child helps the Robot and Toucan pack for the trip to Spain: $a = \{Pos, Hi\} \rightarrow$ *Excited* and "Woo hoo, you're trying so hard!"; $a = \{Med, Hi\} \rightarrow$ *Thinking* and "The suitcase looks heavy."; $a = \{Neg, Hi\} \rightarrow$ *Frustrated* and "I can't decide what to pack next!"; $a = \{Pos, Lo\} \rightarrow$ *Laugh* and "This is so much fun!"; $a = \{Med, Lo\} \rightarrow$ "I wish I could fly."; $a = \{Neg, Lo\} \rightarrow$ *Yawn* and "I'm tired.".

The initial policy represented a mirroring social robot companion, i.e., reflecting the child's own affective state, with a bias towards negative valence when the child is either wrong or off-task. However, due to the complex nature of the affective dynamics between a child and a robot companion, we implemented an affective reinforcement learning algorithm so as to personalize this affective policy to each child. In order to achieve this, we used a standard SARSA algorithm, where the reward was the weighted sum of the valence and engagement $r = 0.4(\hat{s}_{val} + 100)/2 + 0.6\hat{s}_{eng}$, with the goal of maximizing engagement and valence. We implemented an ϵ-greedy algorithm, where ϵ was decreased with each successive session $\epsilon_i = \{0.75, 0.5, 0.25, 0.25, 0.25, 0.25, 0.25\}$, and the learning rate also decreased $\alpha_i = \{0.5, 0.4, 0.3, 0.2, 0.1, 0.1, 0.1\}$.

Study Design

We performed a formal evaluation study using the setup described above. We tested two conditions: personalized affective response from the robot, in which the affective reinforcement learning algorithm was implemented, versus non-personalized affective responses, in which the initial Q-matrix was used throughout the sessions.

The study was performed inside three preschool classrooms, during normal activity hours of the classroom. Thus, children participants were called to play with the social robot companion and then returned to their class. This unique situation required specific adjustments to a standard robot-child interaction scenario, suc as using a divider in order to create an "interaction space" in which the children were physically separated from the rest of the classroom, and using headphones to minimize the chances of interference from non-participants. Despite these restrictions, children were eager to come and play with the robot repeatedly for a duration of two months.

Participants

Thirty-four children ages 3-5 participated in the study (19 boys, 15 girls, age$= 4(\pm 0.7SD)$. They were recruited from three preschool classrooms in a single school located in the Greater Boston Area. The parents of all children signed informed consent forms.

Out of these initial thirty-four subjects, only the 27 children who completed at least 3 sessions with the robot were included in the analysis. Out of these, 9 were English Language Learners (ELL), i.e. their native language was not English, and while they interacted with the tablet and robot, they did not follow the verbal directions of the robot or experimenter, and were thus excluded from the analysis. Of the remaining 18 children, 9 (5 boys, 4 girls, age$= 4(\pm 0.7SD)$ were in the personalized condition and 9 (5 boys, 4 girls, age$= 4(\pm 0.5SD)$ were in the non-personalized condition.

Results

We first present the results of the words learned by the students. We follow with the analysis of the affective personalization of the robot's affective policy. We then show the "instantaneous" effects of the robot's non-verbal behaviors on the child's affective state and consequently show the results of long-term personalization on the change in affective states of the children.

Preschool children learned new Spanish words

While many tutoring systems were developed for teaching children new languages, our setup was unique in several ways. First, the participants age-range was relatively low (3-5). Second, it involved an interaction during the regular activities of the class, as opposed to being set up in a specialized location and time. Third, it involved a fully autonomous social robot companion, i.e. no intervention of the teacher was required. Despite these challenging conditions and requirements, Fig. 3 show that children learned at least some of the words presented during the interaction. In order to test learning, we conducted a digital pre- and post-assessment test of the Spanish words introduced in each session. This test was formatted like the Peabody Picture Vocabulary Test (Dunn and Dunn, 2007): for each word tested, four pictures were shown on the screen. The child was asked to touch the picture corresponding to the word. The words most frequently learned were those repeated most often during the interaction, such as "blue," "monkey," and "clean."

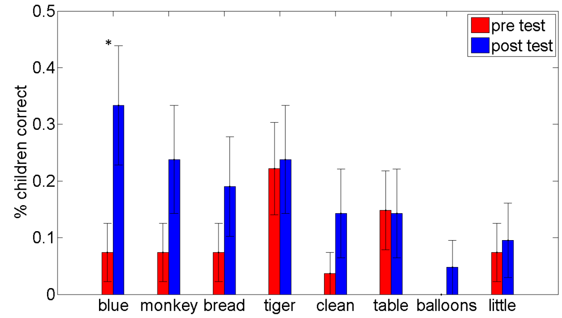

Figure 3: Children were more likely to learn the Spanish words that were used most frequently during the seven sessions.

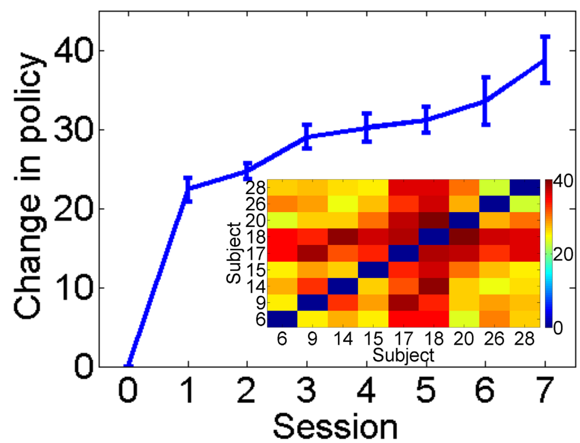

Figure 4: The affective policy d_t changed greatly over the seven sessions. Inset: distance matrix for all 9 subjects in the personalization condition.

Affective policy personalized to students

While we implemented an affective reinforcement learning algorithm, there was no guarantee that the policies would change dramatically, or personalize to each child (i.e., change in different ways for each child). In order to analyze the personalization of the affective policy, we first computed the RMS distance from the original policy: $d_t = \sqrt{(1/(12 \times 7)) \sum_{s,a} (Q_t(s,a) - Q_0(s,a))^2}$. Figure 4 shows this distance at the end of each session, averaged over all subjects in the personalized condition. As can be seen, the distance increases with time, with little variance, suggesting that all subjects experienced new and adapting affective policies. However, it is evident that the policy did not converge after seven sessions. This is not surprising due to the fact that each session lasted only several minutes, while affective interactions are extremely complex and dynamic, and a policy governing them may take a long time to learn.

Did the policies change to a single underlying affective policy, or did the affective social robot companion learn a different policy for each child? The resultant policies for

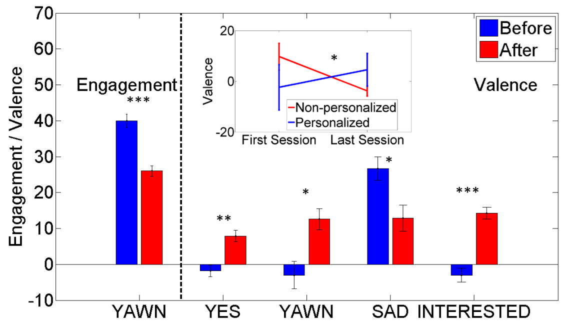

Figure 5: There was significant change in children's engagement and valence as measured via Affdex following non-verbal robot behaviors. $* - p < 0.05, ** - p < 0.01, *** - p < 0.001$. Inset: From the first session to the last session, positive valence increased in the personalized condition but decreased for the non-personalized condition.

each child at the end of the last session were extremely different, as can be seen in a representative sample of the $Q(s, a)$ matrices. One important aspect of note is the lack of change of the negative valence state. This surprising result stemmed from the lack of consistent negative valence states of the child, i.e., the robot never sensed children in a continuous negative valence state, and thus could not adjust to it. Figure 4(inset) shows that indeed the robot personalized to each child, ending up with extremely variable $Q(s, a)$ matrices, as evident by the distance matrix between all subjects.

Children's valence, but not engagement, changes in response to non-verbal behaviors

The whole premise of the affective social robot companion is that the robot's behavior can affect children's own affective states. While the verbal component of the affective policy was complex and variable, depending on the specific game and session, the unique robot platform we developed has the capacity for many non-verbal behaviors. In contrast with some commercially-available robot platforms that are frequently used in child-robot interaction research (e.g., the Nao from Aldebaran Robotics), our social robot platform has a highly expressive smart-phone based animated face, as well as uniquely designed movements based on animated expressions. Thus, the integrated setup implemented in this study enabled us to directly quantify the effects of non-verbal behaviors on children's affective states. We time-aligned the Affdex valence and engagement time-series values to the execution of specific non-verbal behaviors of the robot. We then compared a 10 seconds window before and after the execution of the behavior.

We found that engagement changed significantly only for the *Yawn* behavior, which also instigated a change in valence, Fig. 5. Engagement decreased whereas valence increased for this specific long and expressive behavior. To our surprise, for all other behaviors, engagement did not significantly change. However, valence was significantly affected by several expressive behaviors: (a) it increased after *Yes* nodding behavior, an *Interested* lean-over, and a positive non-verbal utterance; (b) it dramatically decreased after the very expressive *Sad* behavior. These results confirm the premise that children's affective states respond to the non-verbal behavior of the social robot.

Affective personalization increases long-term valence

Although we used a weighted sum of engagement and valence as the reward, the results just discussed suggest that only valence is affected by the robot's non-verbal behaviors. Thus, we hypothesized that personalization would affect valence but not engagement. To test this hypothesis, we compared the change in valence from the first to last session in the personalized and non-personalized condition, Fig. 5:Inset. The results supported our hypothesis: while valence was not significantly different in both conditions in the first and last sessions, the change was ($\Delta s_{val}^{per} = 7(\pm 6SEM), \Delta s_{val}^{non} = -18(\pm 6SEM)$, $p = 0.0251$ Kruskal-Wallis).

Conclusions

We have presented an integrated system that addresses the challenge of affective child-robot tutoring. The system detects a child's affective state and learns how to properly respond over a long-term interaction between child and a fully autonomous affective social robotic companion. In this system, we combined: (i) a novel fully autonomous social robot platform with an engaging and affective social interaction; (ii) a novel interactive and generic tablet app that allows an engaging co-play of robot and child; (iii) commercial automatic facial expression software and SDK that seamlessly integrates with robot affect response; and (iv) an affective reinforcement learning algorithm for affective long-term personalization.

We deployed our setup in the natural environment of three preschool classrooms during regular activity hours. These unique settings combined with a fully autonomous social robot companion necessitated constraining modifications to our setup, e.g., headphones and physical dividers. Nevertheless, we showed that the tutoring setup was effective in helping the children learn new Spanish words. Furthermore, we found that our affective reinforcement learning algorithm personalized to each child and resulted in a significantly increased long-term positive valence, compared to non-personalized interactions. These results can only be obtained in an integrated system, with an affective fully autonomous social robot, an affective automatic sensing system, and a long-term interaction in classrooms.

We believe that this paper presents critical steps towards an effective and affective robot tutor for young children. Future work will work towards a more compact setup. It will include changes such as incorporation of the Affdex software into the robot using the built-in camera in the Tega platform, a longer-term interaction (e.g., six months to a year) that will show either convergence of the affective policy or dynamical adaptation to changes in children's affective responses, and integration of speech recognition in order to faciliate a more natural interaction between child and robot.

Acknowledgments

This research was supported by the National Science Foundation (NSF) under Grant CCF-1138986 and Graduate Research Fellowship Grant No 1122374. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not represent the views of the NSF.

References

Arroyo, I.; Beal, C.; Murray, T.; Walles, R.; and Woolf, B. P. 2004. Web-based intelligent multimedia tutoring for high stakes achievement tests.

Chang, C.; Lee, J.; Chao, P.; Wang, C.; and Chen, G. 2010. Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Educational Technology & Society* 13(2):13–24.

Deshmukh, A.; Jones, A.; Janarthanam, S.; Hastie, H.; Ribeiro, T.; Aylett, R.; Paiva, A.; Castellano, G.; Ellen Foster, M.; Corrigan, L. J.; et al. 2015. An empathic robotic tutor in a map application. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, 1923–1924. International Foundation for Autonomous Agents and Multiagent Systems.

Dunn, L. M., and Dunn, L. M. 2007. *Peabody Picture Vocabulary Test. 4th ed.* Pearson Assessments.

Fasola, J., and Mataric, M. 2013. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction* 2(2):3–32.

Fridin, M. 2014. Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers & Education* 70(0):53–64.

Gordon, G., and Breazeal, C. 2015. Bayesian active learning-based robot tutor for childrens word-reading skills.

Jeong, G.-M.; Park, C.-W.; You, S.; and Ji, S.-H. 2014. A study on the education assistant system using smartphones and service robots for children. *Int J Adv Robot Syst* 11:71.

Kanda, T.; Hirano, T.; Eaton, D.; and Ishiguro, H. 2004. Interactive robots as social partners and peer tutors for children: A field trial. *Hum.-Comput. Interact.* 19(1):61–84.

Kory, J., and Breazeal, C. 2014. Storytelling with robots: Learning companions for preschool children's language development. In *2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 643–648.

Kory, J. M.; Jeong, S.; and Breazeal, C. L. 2013. Robotic learning companions for early language development.

Leyzberg, D.; Spaulding, S.; and Scassellati, B. 2014. Personalizing robot tutors to individuals' learning differences.

McDuff, D.; Kaliouby, R.; and Picard, R. W. 2012. Crowdsourcing facial responses to online videos. *Affective Computing, IEEE Transactions on* 3(4):456–468.

Movellan, J.; Eckhardt, M.; Virnes, M.; and Rodriguez, A. 2009. Sociable robot improves toddler vocabulary skills.

Nye, B. D.; Graesser, A. C.; and Hu, X. 2014. Autotutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education* 24(4):427–469.

Quigley, M.; Conley, K.; Gerkey, B.; Faust, J.; Foote, T.; Leibs, J.; Wheeler, R.; and Ng, A. Y. 2009. Ros: an open-source robot operating system.

Saerbeck, M.; Schut, T.; Bartneck, C.; and Janse, M. D. 2010. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor.

Setapen, A. 2012. *Creating Robotic Characters for Long-term Interaction.* Thesis.

Tapus, A.; Maja, M.; and Scassellatti, B. 2007. The grand challenges in socially assistive robotics. *IEEE Robotics and Automation Magazine* 14(1).

Vanlehn, K.; Burleson, W.; Echeagaray, M.-E. C.; Christopherson, R.; Sanchez, J. G.; Hastings, J.; Pontet, Y. H.; and Zhang, L. 2011. The affective meta-tutoring project: How to motivate students to use effective meta-cognitive strategies.

VanLehn, K. 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist* 46(4):197–221.

Woolf, B.; Burleson, W.; Arroyo, I.; Dragon, T.; Cooper, D.; and Picard, R. 2009. Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology* 4(3-4):129–164.