The Eighth AAAI Symposium on Educational Advances in Artificial Intelligence 2018 (EAAI-18)

Gesturing and Embodiment in Teaching: Investigating the Nonverbal Behavior of Teachers in a Virtual Rehearsal Environment

Roghayeh Barmaki
Laboratory for Computational Sensing and Robotics
Johns Hopkins University, Baltimore, MD, USA
rl@jhu.edu

Charles Hughes
Department of Computer Science
University of Central Florida, Orlando, FL, USA
ceh@cs.ucf.edu

Abstract

Interactive training environments typically include feedback mechanisms designed to help trainees improve their performance through either guided or self-reflection. In this context, trainees are candidate teachers who need to hone their social skills as well as other pedagogical skills for their future classroom. We chose an avatar-mediated interactive virtual training system–TeachLivE–as the basic research environment to investigate the motions and embodiment of the trainees. Using tracking sensors, and customized improvements for existing gesture recognition utilities, we created a gesture database and employed it for the implementation of our real-time gesture recognition and feedback application. We also investigated multiple methods of feedback provision, including visual and haptics. The results from the conducted user studies and user evaluation surveys indicate the positive impact of the proposed feedback applications and informed body language. In this paper, we describe the context in which the utilities have been developed, the importance of recognizing nonverbal communication in the teaching context, the means of providing automated feedback associated with nonverbal messaging, and the preliminary studies developed to inform the research.

Introduction

Interpersonal communication involves a variety of modes and components in communication. We might think that actual words are the primary part of communication; however, there are other nonverbal components that are also integral (Albert 1971). These nonverbal elements include both non-vocal (e.g. body language) and vocal components (e.g. voice pitch and intonation). Body language by itself includes several aspects: facial expressions, eye contact, posture or stance, gestures, touch and appearance (Richmond, McCroskey, and Payne 1991).

The interactive Teaching and Learning Environment (TLE TeachLivE) is a virtual classroom setting used to prepare teachers for the challenges of working in K-12 classrooms (Dieker et al. 2014). Its primary use is to provide teachers the opportunity to improve their classroom management, pedagogical and content delivery skills in an environment that neither harms real children nor causes the teacher to be seen as weak or insecure by an actual classroom full of students. We chose this virtual classroom as our basic research environment, and developed assessment and feedback methods on top of it because of the pressing need for training novice and pre-service teachers with the required skills, prior to their entering a real classroom. The TeachLivE virtual classroom is based on a multi-client-server structure provided by its underlying digital puppetry architecture (Nagendran et al. 2014). The system uses a Kinect sensor facing the participant and a TV screen showing the virtual students. A human-in-the-loop (called an interactor) orchestrates the behavior of the virtual students in real-time based on each character's personality and backstory, a teaching plan, various genres of behaviors and the participant's input. The system also supports reflective learning following an experience in the virtual classroom. In this research, aligned with the reflective learning and assessment horizon of the project, we explore the gestures and embodiment of the teacher candidates with the analysis of their recorded teaching sessions.

In the teaching effectiveness content, one might observe different categories to investigate for teaching competency: purely objective ones that require no actual understanding of intent of the behavior, for example, proximity and talk time; those with shallow semantics associated with intent, for example, open versus closed posture or use of open versus closed words, (such as why versus what) that can be part of a deeper semantic analysis, for instance, showing respect/disrespect or asking open versus closed questions; and yet others that require deep analysis to acquire any semantic meaning, for example, providing encouragement versus closing off a conversation. Even in the latter, deep semantic situations, cues can be observed through embodiment that may indicate a desire to continue a dialogue versus a desire to close off all further communication. This research addresses postures and gestures as two of the key assessment components of the nonverbal communication in the teaching context.

We have developed a real-time gesture recognition and feedback application on top of the Microsoft Kinect Software Development Kit based on existing recordings from teachers who have had an interactive experience in the virtual classroom. To evaluate this feedback application, we conducted two preliminary studies. The hypothesis is that our developed feedback application has positive impacts on the participants' body language, leading to more open and fewer closed gestures.

This paper is organized as follows. In the next section, we introduce related work. Following that, we provide the details of our method, studies, results, and conclusions.

Background and Related Work

The literature on nonverbal communication in learning environments indicates the importance of nonverbal behaviors in successful student-teacher interactions and student learning. Teachers use gesture and embodiment to be effective in several fundamental aspects of their profession, including communication, assessment of student knowledge, and the ability to instill a profound understanding of abstract concepts in traditionally difficult domains such as language learning and mathematics (Alibali and Nathan 2007; Alibali et al. 2014; Roth 2001).

For instance, Kelly et al. showed that employing iconic gestures or illustrations during teaching had a significant impact on English-speaking adults who were learning new Japanese words (Kelly, McDevitt, and Esch 2009). Scholars who study classroom communication emphasize the teacher's nonverbal behavior as information to the students (Smith 1979; Woolfolk and Brooks 1985). This means that the way that a teacher sends messages with her body is important and all the students are able to perceive the nonverbal signals of their surrounding people according to Smith et al. (1979). To support this, Caswell and Neill (2003) mention that competent teachers have open rather than closed body gestures while teaching or interacting with students. An open stance has arms and legs that are not crossed.

A key factor to development and improvement of one's skills is feedback, which is one of the most influential interventions in learning (Hattie and Timperley 2007). The means to present feedback vary greatly and several dimensions of feedback have been identified. One of these dimensions refers to the timing of feedback, which can be delayed or immediate (Hattie and Timperley 2007; Mory 2004). Most of the studies conducted comparing both types of feedback concluded that, for most learning situations, the impact of immediate feedback is more positive, since delayed feedback tends to defer the acquisition of needed information (Mory 2004; Schneider et al. 2015).

There are some research projects that are working on feedback provision in interactive environments. A tool to improve nonverbal communication skills for public presentation has been introduced by Schneider et al. (2015). They use the Microsoft Kinect V2 to track the trainee's body movements during the elevator pitch rehearsal and provide visualized or haptic immediate feedback about each subject's nonverbal communication skills. Another project called MACH is designed for interview practice and provides both immediate and delayed feedback to the trainee (Hoque et al. 2013). Tanveer et al. (2016) introduce an automated interface to make public speakers aware of their mannerisms; mannerism means exhibiting unconscious body movements while speaking. Chollet et al. (2014) present an interactive virtual audience platform for public speaking training. Each user's public speaking behavior is automatically analyzed using audiovisual sensors. The virtual characters display indirect feedback depending on each user's behavior descriptors correlated with public speaking performance.

Research in public speaking training has informed our efforts, as existing research in education has shown that, like public speaking, performance feedback is an essential component of effective professional development and staff training programs that target workplace behavior change (Leach and Conto 1999); however, the context of public speaking is slightly different than teacher preparation. Using the TeachLivE, we are dealing with teachers who stand, pace and approach students, unlike most systems designed for interview training and other social skills training, e.g. TARDIS (Anderson et al. 2013) and MACH (Hoque et al. 2013).

Designing effective mechanisms for training, and assessment, in general, has been challenging from several points of views. Presently, most of the trainees receive feedback from subject-matter experts (SME). SMEs must observe the whole recorded interactive sessions carefully (or do onsite observation) in order to provide an assessment to the trainees. In addition, most of the proficiency measures are very subjective, and there is always a potential problem of biased evaluation from the human expert. The other problem of using subject-matter experts is the cost associated with human resources. In an effort to better assist SMEs as one of our aims, we have conducted a case study with a follow up. We recorded and analyzed different types of information, including event logs, video, audio and depth sensor data from multiple devices to recognize high-level features required for a semi-automated reflection tool, i.e., one in which a subject-matter expert is provided suggestions to aid in the event annotation process. Such a hybrid approach can lead to an optimized use of the human resources in the assessment process.

Research Design

This research evolved based on the existing literature expressing the importance of open body gesturing in successful interactive teaching. We started by reviewing existing recordings of teaching sessions to have a baseline about the way teachers use their body in the virtual classroom. We observed that most of the teachers were not thoughtful of their body movements at all. Many of the teachers exhibited closed stances most of the time in their recorded teaching sessions. The recognized frequent closed postures (hereafter gestures) were arms folded in front (unreceptive), hands clasped in back (seductive), hands placed on hips, (skeptical), hands clasped in lower front (protective), and hands clasped in upper front (submissive). These gestures are noted as closed or "not-recommended" gestures (Barmaki 2015, Caswell and Neill 2003, Schneider et al. 2015). Moreover, according to the bounding volume (BV) measure –as the volume taken by the person in the space– introduced by Niewiadomski et al. (2013), these are contracted postures, thus the BV value is low, denoting less openness. In an effort to develop immediate feedback mechanisms, we assume that a percentage of of closed gesture employment from a user depend on the feedback condition.

We are interested in creating an automated system to detect these closed gestures, reminding the trainees about their closed body language stance. With the assistance of the Kinect as an affordable non-intrusive motion-capture device and its Software Development Kit (2016) for recording and analysis of human motions (including Visual Gesture Builder (VGB) and Kinect Studio), we created a corpus of closed gestures. We recorded some clips from five (3M, 2F) students using the Microsoft Kinect V2.

These students were only involved in corpus collection and they did not participate in the follow-up case studies. The summary of system design, from details of recording, annotation, analysis and the generation of gesture detector engines, evaluating the accuracy of those detectors to our implemented immediate feedback user interfaces is presented in Figure 1.

To create the corpus, the recorded clips and the targeted postures were hand-annotated and used as training data for the gesture recognition method. This method uses the human full-body information (25 joints positions and orientations from the tracked body in each frame) along with the annotated tags indicated as labels. The VGB utility builds several binary classifiers for each gesture and the top-1000 decision-tree classifiers are ensembled together (using Adaboost algorithm) to create the ultimate closed gesture recognition database. The accuracy of our gesture detection engine was 96.5% ± 2.1 % for all five closed gestures. The detection algorithm was predicting, and storing the labels of every frame (30 fps) of the RGBD data from the Kinect in real-time. In the following sections, we describe the details of the conducted case study and a follow-up for gesture assessment.

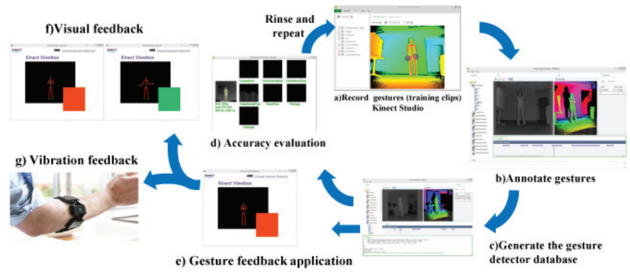


Figure 1. The data-driven and incremental process of creating the gesture detector system: a) training data capture, b) annotation, c) generation of gesture detector engines, d) evaluate the accuracy of gesture detector, e) the gesture feedback user interface generation f) visual feedback system in case study, and g) vibration feedback system using an armband in the follow-up study.

Case Study: Visual Feedback Evaluation

The developed system is designed to provide immediate feedback either in the form of visual or haptic signals (vibrations) any time that the participant exhibits a closed stance. The case study reported in this section is based on visual feedback within the virtual rehearsal classroom. The vibration feedback method is also tested in a follow-up study and more details are described at the end of this section.

Participants

In this case study for visual feedback evaluation, 30 (24F) participants were recruited. Only those UCF College of Education and Human Performance students over the 18 years of age who had experienced the teaching rehearsal environment before the experiment were eligible to participate. All the participants were from the students registered for a classroom management and strategies undergraduate course in the TESOL (Teaching English to Speakers of Other Languages) program.

Participation in our user study is defined as a classroom activity because of its suitable overlap with the course objectives. Students did not get compensation or credit for their participation.

Apparatus

For this study, we collected full-body tracking and teaching session videos. Full body of the participant teacher was recorded for body language and pose recognition using the Microsoft Kinect.

The room for client-site was equipped with required laptops, displays, auditory equipment, and wired network connection for the experiment. As mentioned earlier, the participant teacher stands in front of a large TV display and talks and interacts with the virtual classroom shown in the TV.

In the visual feedback setting, in addition to the virtual classroom, participants could see their motion–tracked by the Microsoft Kinect–and visual cues on the monitor facing them (Barmaki and Hughes 2015). Figure 1.f shows an example of a participant's posture and her feedback depicted on the screen.

In the server-side, the interactor, who controls and animates virtual students, was also asked to give consistent performances for all instances of the case study sessions, to minimize the inconsistencies that may occur because of interactor changes.

Study Procedure and Design

The case study was a 2×1×2 counter-balanced within-subjects study (two settings; TLE with and without visual feedback; one trial, and two groups). It means that all the participants attended both study settings, but the order of the sessions was flipped for some subjects. Individuals were expected to spend approximately 30 minutes for the recruitment. There were two 7-minute sessions for the teaching plus three 5-minute intervals for questionnaires (one pre- and two post-questionnaire). In the pre-questionnaire, participants were asked about their demographics and prior teaching experience both in real and virtual settings. After the teaching sessions with the theme of biology introductory course with given teaching plan, the participants completed a post-questionnaire soliciting their evaluation of the user experience and learning perception in both of the sessions.

The study design and participant assignment are shown in Figure 2 for both groups.

The feedback was in the form of a visual prompt each time the participant exhibited a closed, defensive stance. Visual feedback uses traffic light model, which means the color of the visual signal immediately changes from green to red when the participant exhibits a closed gesture as illustrated in Figure 1 (step f). The collected body tracking data from the participants was processed to extract a higher-level feature. The calculated feature value was the percentage of time that a subject exhibited closed gesture in the recorded clips of the study sessions. We call this variable CGP, standing for closed gesture percentage exhibition. The lower CGP is better as supported by previous work (Caswell and Neill 2003).

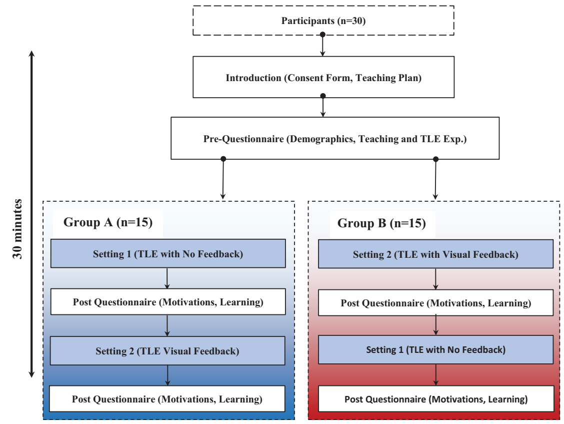


Figure 2. Overview and participant assignment of the case study. TLE indicates for the TeachLive Teaching and Learning Environment.

Follow-up Study: Visual vs. Haptic Feedback

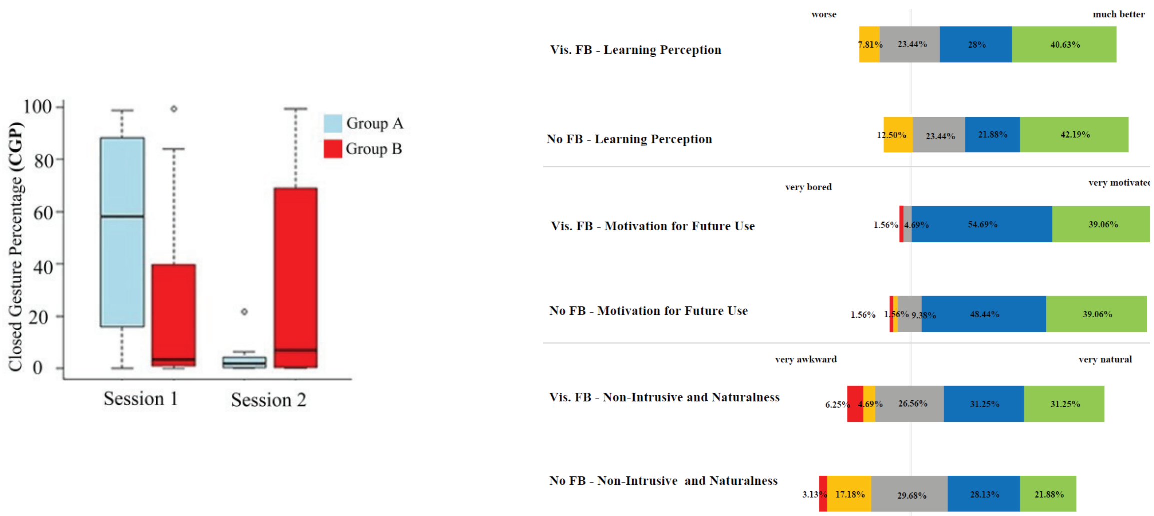
In the follow-up study, we worked on the form of feedback application. We modified the user interface for the feedback application, by exchanging visual prompts with haptic signals. We used a Myo armband for our study, though there are several alternative wrist band devices available with haptic commands. The hypothesis of this follow-up study is that the vibration feedback application is more comfortable, natural, and less distractive for immediate feedback provision to users in the rehearsal sessions. To investigate our hypothesis, we compared the vibration versus visual feedback provision in a follow-up study as follows.

Participants

We tested the vibration feedback application with three participants (N=3, 2F) who took part in the main case study and volunteered to participate in the follow-up study as well.

Study Procedure and Design

The study was a single session observation with a 2×1×1 design with no counterbalancing. All three participants experienced both settings. They first taught the same teaching plan for seven minutes in visual feedback setting. Afterwards, they completed the postquestionnaire, and participated in the second teaching session. The first setting was the TeachLivE session with visual feedback, and the second was teaching session with vibration or haptic feedback. We used timed vibration commands instead of visual prompts for closed gesture feedback in this setting. Timed feedback was used to leverage the naturalness of the experience. We asked participants to evaluate their experience after each session. They completed the same post-questionnaire that was used in the main case study.



(a) *(b)*

Figure 3. Results for the case study. a) Medians and interquartile ranges of CGP in two sessions for groups A and B. Circles represent outliers. b) Post-questionnaire analysis in three main categories of learning perception, motivation, and naturalness of the experience.

Results

We report the results of a case study and a follow-up study from two perspectives, based on the closed gesture employment or embodiment changes, and also self-reported usability analysis in post-questionnaires as follows.

Closed Gesture Evaluation

To evaluate the impact of our proposed visual feedback application on body language mindfulness, we calculated CGP for 60 recorded clips from 30 participants in the case study. The hypothesis is that, regardless of the group assignment (A or B), the effect of receiving posture feedback, either in the current session or from a prior session, is considerable on CGP. The boxplot in Figure 3.a presents the distribution of closed gesture percentage (CGP) between the two groups during the study.

Figure 3.a shows some key findings from this study. It presents the wide range of closed gesture employment for groups A and B in both sessions (excluding session 2 for group A). It also indicates that the average of CGP medians for group B participants is lower than for group A. Considering 3.5% and 6.9% as medians for two sessions for group B and 58%, 1.8% for group A, the average values were 5.24% for group B and 29.95% for group A. As reflected in Figure 3.a, the value for CGP among 30 participants of the study is not normally distributed. The CGP values include some extreme data points that seem to affect the CGP distribution.

Furthermore, according to the analysis of variance test, there was a significant effect of setting (visual feedback vs. no feedback) on CGP ($F_{(1,28)} = 13.66$, $p < .001$, $\eta^2 = 0.33$) in this study. As indicated by the results, the hypothesized statement is supported for the participants in the case study with large effect size. The average time that all the participants in group A exhibited closed gestures reduced significantly from their first session to their second teaching session. CGP was slightly increased in the second unaided session for group B (median= 6.9%).

Lastly, the results indicated a non-significant correlation between group assignment and CGP, which indicates the successful randomization and minimizing the grouping effect.

The analysis of the recorded full-body tracking data from each participant in the follow-up study was also completed. We calculated the closed gesture employment rate (CGP) for visual and vibration feedback sessions for six sessions from three participants. The CGP results for visual feedback (M=8.58%, SD = 1.6%) and vibration feedback (M=6.77 %, SD=3.81%) in follow-up sessions indicated that participants' embodiment was not considerably different in the visual versus vibration settings (t(2) =0.75, ns). As noted, the participants were exposed to the feedback application before this study. Hence, the participants' closed gesture employment remained similar with a slight difference for both settings at a very low level, and it did not reveal a significant difference.

Post-Questionnaire Evaluation

There were two main types of questions in the post-questionnaire. User experience was evaluated by four questions on the naturalness of the interaction, motivation and the likelihood of future participation. Learning perception was evaluated by two questions on improvements in the understanding of teaching strategies and comparison of the proposed application with traditional teacher training approaches.

To report the results from the post-questionnaires, we divided six 5-point Likert scale questions into three subtopics. These topics are naturalness and non-intrusiveness, the motivation for future use and learning perception. We analyzed the 360 responses ($30 \times 2 \times 6 = 360$) from 30 participants in case study and reported these in Figure 3.b. Figure 3.b presents valuable information about the case study. In general, we received very positive input from the participants (regardless of their assigned groups) as the right-skewness of the stacked bar charts supports this conclusion. Interestingly, participants' responses indicated that their experience in setting with visual feedback was natural and non-intrusive (62.5% versus 50% for no feedback setting). This is very encouraging to us from a usability perspective. The participants were also very motivated to participate in future studies. For the setting with no feedback, we had more responses as "neutral" or the "same" versus the setting with visual feedback that indicates more positive responses such as "motivated", and "better learning perception". Overall, participants enjoyed their interactive experience and they reflected their mindfulness about embodiment in future interactions, especially in their teaching sessions.

For the follow-up, in the usability evaluation questionnaire and informal reflection, all three subjects were motivated and supportive to use the vibration method for gesture feedback, even in real classroom scenarios. Two of them indicated that the armband and its feedback are not visible (if you wear a long-sleeve shirt), and they can confidently use it as a personal assistant in a class full of students. The other participant reflected his evaluation as follows:

> I felt as the armband was more helpful in providing feedback. It is a lot easier to tell when I am doing something wrong when the armband vibrates as opposed to trying to look at the Kinect screen while talking.

Discussion

Analysis of recorded embodiment data and postquestionnaire responses from the participants in the case study indicated the positive impact of informed body language and gesture in communication proficiency. That said, most of the participants were very motivated to use the gesture feedback application for their rehearsal sessions. The participants also mentioned their learning experience was improved; some participants reported that their experience with the system was better than the traditional learning resources.

For the follow-up study, we also observed interesting findings, especially from the qualitative analysis. Participants enjoyed their interactive experience with the vibration feedback method. Even with this small pool of participants, the results illustrate the potential of successful future large-scale studies with the aid of the proposed feedback system.

Conclusion and Future Research

In this paper, we described the importance of recognizing nonverbal communication behaviors in the teaching context, along with different applications and machine learning methodologies to design an automated posture feedback application. We reported results from the preliminary studies to evaluate the proposed feedback application. The goal is to use this application for teacher assessment as an automated tagging (behavioral annotation) system based on posture recognition and correlations to expert human tagging. We specifically focused on gesture recognition and different types of non-intrusive feedback provision techniques (including visual and haptics) for closed gesture stances exhibited by trainees.

The major limitation of the research reported here was related to the recruitment process. Recruiting the representative number of participants was a difficult process that adversely affected our research agenda. For example, we were not able to do counterbalancing to reduce the learning effect because we didn't have a sufficient number of participants. As a result, we noted mere-exposure effect in the follow-up study.

Several directions of future research are promising. While this work touched very briefly on the different methods of feedback provision, it did not provide strong quantitative evidence about their relative performance since the size of the participant pool for the conducted follow-up study was not large. However, the potential of study was successfully described with qualitative analysis. Thus, conducting an experiment with a representative number of participants to evaluate the haptic feedback method versus visual cues is one of the objectives of this research project over time.

In summary, although further work is required to gain a complete understanding of the embodiment in teaching, the findings indicate that using immediate feedback in teaching training sessions assist practicing trainees to master their communication skills not only for teacher-student interaction but also for different interpersonal human-centered interactions, such as police de-escalation training.

Acknowledgements

The authors wish to thank Drs. Lisa Dieker, Michael Hynes, Cynthia Hutchinson, Kathleen Ingraham and the entire TeachLivE team for their support during the data collection and analysis processes. Their insights and their assistance with subject recruitment were invaluable. We also wish to express our appreciation for the support of the interactors and the study participants.

Finally, we wish to acknowledge support for the evolution of the TeachLivE project from the Bill and Melinda Gates Foundation (OPP1053202), the National Science Foundation (CNS1051067, IIS1116615) and the Johns Hopkins Science of Learning Institute. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

Albert, M. 1971. Silent messages. Wadsworth: Belmont, CA, USA.

Alibali, M. W., and Nathan, M. J. 2007. Teachers' gestures as a means of scaffolding students' understanding: Evidence from an early algebra lesson. *Video Research in the Learning Sciences,* 349–365.

Alibali, M. W.; Nathan, M. J.; Wolfgram, M. S.; Church, R. B.; Jacobs, S. A.; Johnson Martinez, C.; and Knuth, E. J. 2014. How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction,* 32(1): 65–100.

Anderson, K.; André E.; Baur, T.; Bernardini, S.; Chollet, M.; Chryssafidou, E.; Damian, I.; Ennis, C.; Egges, A.; Gebhard, P.; Jones, H.; Ochs, M.; Pelachaud, C.; Porayska-Pomsta, K.; Rizzo, P.; and Sabouret, N. 2013. The TARDIS Framework: Intelligent Virtual Agents for Social Coaching in Job Interviews. In *Proceedings of the Tenth International Conference on Advances in Computer Entertainment*, 476–491.

Barmaki, R., and Hughes, C. E.; 2015. Providing real-time feedback for student teachers in a virtual rehearsal environment. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. 531–537.

Caswell, C., and Neill, S. 2003. *Body language for competent teachers:* Routledge.

Chollet, M.; Sratou, G.; Shapiro, A.; Morency, L.; and Scherer, S. 2014. An interactive virtual audience platform for public speaking training. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems,* 1657–1658.

Dieker, L. A.; Straub, C.; Hughes, C. E.; Hynes M. C.; and Hardin, S. E. 2014. Virtual environments can take us virtually anywhere. *Educational Leadership*, 71(8): 54–58.

Hattie, J., and Timperley, H. 2007. The power of feedback. *Review of Educational Research,* 77(1): 81–112.

Hoque, M. E.; Courgeon, M.; Martin, J.; Mutlu, B.; and Picard, R. W. 2013. Mach: My automated conversation coach. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing,* 697–706.

Kelly, S. D.; McDevitt, T.; and Esch, M. 2009. Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes,* 24(2): 313–334.

Leach, D. J., and Conto, H. 1999. The additional effects of process and outcome feedback following brief in-service teacher training. *Educational Psychology,* 19(4): 441–462.

Leong, C. W.; Chen, L.; Feng, G.; Lee, C. M.; and Mulholland, M. 2015. Utilizing depth sensors for analyzing multimodal presentations: Hardware, software, and toolkits. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction,* 547–556.

Mory, E. H. 2006. Feedback research revisited. Handbook of Research on Educational Communications and Technology.

Nagendran, A.; Pillat, R.; Kavanaugh, A.; Welch, G.; and Hughes, C. E. 2014. A Unified Framework for Individualized Avatar-Based Interactions. *Presence: Teleoperators and Virtual Environments*, 23(2): 109–132.

Niewiadomski, R.; Mancini, M.; and Piana, S. 2013. Human and virtual agent expressive gesture quality analysis and synthesis. *Coverbal Synchrony in Human-Machine Interaction,* 269–292.

Porter, A.; McMaken, J.; Hwang, J.; and Yang, R. 2011. Common core standards the new US intended curriculum. *Educational Researcher,* 40(3): 103–116.

Richmond, V. P.; McCroskey, J. C.; and Payne, S. K. 1991. Nonverbal behavior in interpersonal relations. Prentice Hall: Englewood Cliffs, NJ, USA.

Roth, W. 2001. Gestures: Their role in teaching and learning. *Review of Educational Research,* 71(3): 365–392.

Schneider, J.; Börner, D.; Van Rosmalen, P.; and Specht, M. 2015. Presentation trainer, your public speaking multimodal coach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction,* 539–546.

Smith, H. A. 1979. Nonverbal communication in teaching. *Review of Educational Research,* 49(4): 631–672.

Tanveer, M. I.; Zhao, R.; Chen, K.; Tiet, Z.; and Hoque, M. E. 2016. Automanner: An automated interface for making public speakers aware of their mannerisms. In *Proceedings of the Twenty-First International Conference on Intelligent User Interfaces,* 385–396.

Visual gesture builder: A data-driven solution to gesture detection. "http://aka.ms/k4wv2vgb", retrieved 4/3/2017.

Woolfolk, A. E.; and Brooks, D. M. 1985. The influence of teachers' nonverbal behaviors on students' perceptions and performance. *The Elementary School Journal,* 85(4): 513–528.