When to Provide Feedback? Exploring Human-Co-Robot Interactions in Engineering Enviroments

Christian Enmanuel Lopez, The Pennsylvania State University

Christian Lopez Bencosme, is currently a Ph.D. student at Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at the Pennsylvania State University. He has worked as an Industrial Engineer in both the Service and Manufacturing sectors before pursuing his Ph.D. His current research focused on the design and optimization of systems and intelligent assistive technologies through the acquisition, integration, and mining of large scale, disparate data. He is currently working on a project that ambition to design a system capable of providing students customized motivational stimuli and performance feedback based on their affective states.

Dr. Conrad Tucker, Pennsylvania State University, University Park

Dr. Tucker holds a joint appointment as Assistant Professor in Engineering Design and Industrial Engineering at The Pennsylvania State University. He is also affiliate faculty in Computer Science and Engineering. He teaches Introduction to Engineering Design (EDSGN 100) at the undergraduate level and developed and taught a graduate-level course titled Data Mining–Driven Design (EDSGN 561). As part of the Engineering Design Program's "Summers by Design" (SBD) program, Dr. Tucker supervises students from Penn State during the summer semester in a two-week engineering design program at the École Centrale de Nantes in Nantes, France.

Dr. Tucker is the director of the Design Analysis Technology Advancement (D.A.T.A) Laboratory. His research interests are in formalizing system design processes under the paradigm of knowledge discovery, optimization, data mining, and informatics. His research interests include applications in complex systems design and operation, product portfolio/family design, and sustainable system design optimization in the areas of engineering education, energy generation systems, consumer electronics, environment, and national security.

When to Provide Feedback? Exploring Human-Co-Robot Interactions in Engineering Environments

Abstract

Co-robots are robots that work alongside their human counterparts towards the successful completion of a task or set of tasks. In the context of engineering education, co-robots have the potential to aid students towards the successful completion of an engineering assignment by providing students with real-time feedback regarding their performance, technique, or safety practices. However, determining when and how to provide feedback that advances learning remains an open research question for human-co-robot interactions. Towards addressing this knowledge gap, this work describes the data types available to both humans and co-robots in the context of engineering education. Furthermore, this works demonstrates how these data types can be potentially utilized to enable co-robot systems to provide feedback that advances students' learning or task performance.

The authors introduce a case study pertaining the use of a co-robot system capable of capturing students' facial keypoint and skeletal data, and providing real-time feedback. The co-robot is created using commercially available, off-the-shelf components (e.g., Microsoft Kinect) in order to expand the reach and potential availability of these systems in engineering education. This work analyzes the facial expressions exhibited by students as they received instructions about how to complete a task, and feedback about their subsequent performance on that task. This allows the authors to explore the influence that co-robot visual feedback systems have in changing students' behavior while performing a task. The results suggest that students' facial keypoint data is statistically significantly different, depending on the feedback provided (p-value<0.005). Moreover, the results suggest there is a statistically significant relationship between students' facial keypoint data while receiving instructions on how to complete a task, and their subsequent performance on that task (p-value<0.005). These findings suggest that students' facial keypoint data can be utilized by a co-robot system to learn about the state changes in students, as they complete a task, and provide interventions when certain patterns are discovered that have the potential to reduce students' learning or task performance.

The outcomes of this paper contribute to advancing the National Academy of Engineering's Grand Challenge of personalized learning [a] by demonstrating how students' facial expression data can be utilized in an effective manner to advance human-co-robot interactions and improve the capability of co-robot systems to provide feedback that advances students' performance.

[a] http://www.engineeringchallenges.org

1. Introduction

The term *Collaborative robots* (i.e., co-robots) is defined as a class of robots that work alongside their human counterparts towards the completion of a common task [1]. In the context of engineering education, co-robots have the potential to aid students during tasks that may require real-time observation and feedback [2]. Figure 1 presents a scenario that involves a student (left) and a co-robot (right) working together in an engineering design workshop. Here, the task to be completed is the design of an engineering concept/idea/prototype, created using tools such as a hammer, wood, and other engineering laboratory equipment and materials. The primary objective of the student is to complete this task in a manner that optimizes performance (e.g., time to completion, usefulness of the prototype), while the primary objective of the co-robot is to ensure that the student completes the task in a safe and effective manner that results in i) tangible learning outcomes (e.g., improved performance) and ii) the desire to perform such tasks in the future. Measuring whether a student has the desire to perform similar tasks in the future is related to their affective state (i.e., emotions) during task performance [3]. Ekman has proposed six basic emotions that include happiness, sadness, disgust, surprise, fear, and anger [4]. However, given the heterogeneity of individuals, determining how and when to provide feedback that advances students' performance remains an open research question. In an effort to personalize the feedback provided to the students, having a "*one-size-fits-all*" definition of affect or performance, may not be well suited [5]. Therefore, by observing students while they perform a task, co-robot systems could potentially provide interventions when certain patterns are discovered that have the potential to reduce students' task performance. Facial keypoint data has been shown to correlate with an individual's affective state [6], thereby potentially serving as a means for real-time assessment of a student's affective state during task performance. Recent studies have shown that the acquisition of facial keypoint data, requires the least amount of interruption when compared to other data collection options such as surveys and body position mapping [7].

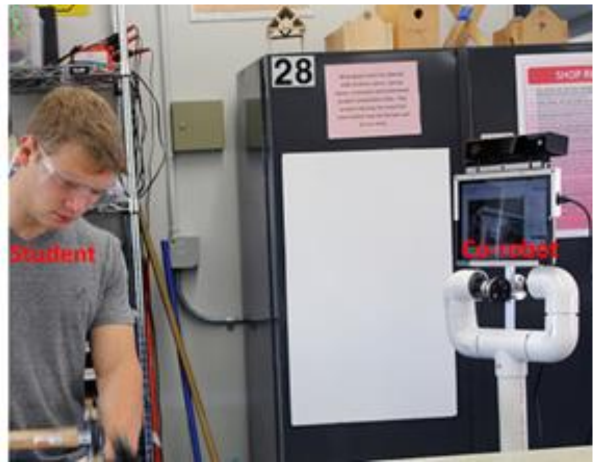


Figure 1: Student-Co-robot Collaboration towards the Successful Completion of an Engineering Prototype

This paper presents data types captured by both the co-robot (i.e., facial keypoint data, skeletal data) and student (i.e., visual performance feedback), and a method that demonstrates how each of these data types can be utilized in an effective manner to advance human-co-robot interactions.

2. Literature Review

2.1 Sensing Systems and Information Exchange

Interactions between humans and co-robots depend on the acquisition and processing of data communicated (intentionally or unintentionally) by the human or co-robot. In order for a co-robot to know when and how to provide feedback to students during engineering-related activities, it must first be able to sense its environment. While this may be a trivial problem for humans, co-robot sensing is a non-trivial problem and is an active area of research [8–17]. Off-the-shelf sensing systems, such as the Microsoft Kinect, have been successfully employed in the field of robotics and machine learning for sensing and navigation tasks. From an engineering design perspective, the Microsoft Kinect sensor has been utilized as a 3D scanner for capturing digital representations of engineered artifacts [18,19]. Specifically relating to co-robot sensing, Choi and Christensen propose an RGB-D object sensing and tracking approach that utilizes GPUs to efficiently identify objects [20]. For co-robot tasks that extend beyond sensing objects, Burns and Samanta propose an approach for co-robots to sense humans, based on a Carmine depth sensing system coupled with open-source middleware (i.e., NITE from OpenNI) [21]. For tasks involving human-co-robot interactions, Morato et al. propose the use of multiple Kinect sensors to achieve safe human-robot collaboration during assembly tasks [22]. The Kinect sensing system has also been employed in the MobilAR platform to augment a user's perception of their environment [23]. Research into sensors and sensing systems is extensive and cuts across a wide range of domains [24–30].

While advancements in sensing systems have enabled co-robot systems to sense objects and humans in an environment, a knowledge gap exists in terms of how co-robot systems synthesize the data captured by its sensors in an effort to help provide their human counterparts with meaningful feedback during a task or set of tasks. In this work, the authors employ the Microsoft Kinect sensor to capture students' skeletal and facial keypoint data, and demonstrate how this data can be utilized to help co-robots understand students' behavior.

2.2 Co-Robot Feedback and Decision Making Methods

Research into how and when co-robots decide to make decisions or provide feedback is an active area of scientific inquiry [31–39]. A hierarchical decision-making approach to co-robot feedback requires the co-robot system to elicit feedback from its human counterpart. Kaupp et al. use probabilistic models to determine when co-robots should query humans for information [40]. Cao et al. propose a hierarchical decision-making model, wherein the role of each co-robot in a group, is assigned by its human supervisor, based on each co-robot's report of resources needed [41]. In each of these approaches, the request for information by the co-robot from its human counterpart, risks the interruption of the task being performed.

In this work, the acquisition of information by the co-robot from its human counterpart is achieved automatically by constantly acquiring facial keypoint and skeletal data that can then be used to make inferences about the complexity of the task and the performance characteristics needed in order to successfully complete the task. There have been several works that aim to make co-robot data acquisition and intervention more seamless. For example, Rani et al. propose to enhance human decision-making outcomes by designing co-robots that can detect human anxiety during human-robot collaborations [42]. Hinds et al. investigated whether humans have preferences towards the visual likeness of co-robots by varying the extent to which the robot had human-like appearances (e.g., human, human-like, and machine-like) [43]. Stanton and Stevens discovered that the gaze of a robot resulted in a faster response by their human counterpart, suggesting a correlation between robot gaze and social facilitation. However, the same work also found that the bodily movements of the robot did not impact participants' behavior [44].

For educators seeking to design practical co-robot systems that they can deploy in their classrooms, the above studies are insightful, as they suggest that less time should be spent on the physical design and movement of co-robot systems. Instead, more time should be spent on creating a visually-engaging interaction between humans and co-robots. The method section presented next, therefore, focuses on the acquisition of facial and body cues by a co-robot system from its human counterpart (i.e., in the form of facial keypoint and skeletal data), and its human counterpart's capture of the visual representation displayed by the co-robot system (i.e., in the form of an avatar representation of the student that includes reward and penalty visualizations). Moreover, this work presents how this data can potentially be utilized in an effective manner to advance human-co-robot interactions**.**

3. Method

The method presented in this work outlines how a student's facial keypoint and skeletal data can be captured by a co-robot system, and subsequently utilized to provide feedback to students that advance their learning or task performance. In this work, the authors assume no predefined classification of affective states (e.g., as in the case of works such as Ekman [4]). Instead, the affective state of a student is assumed to be on a continuous scale that is captured using facial keypoint data. Instead of mapping affective states into predefined categories, such as happiness, anger, etc., the authors explore the variation between students' facial keypoint data, while holding the task being performed constant. For example, Figure 2

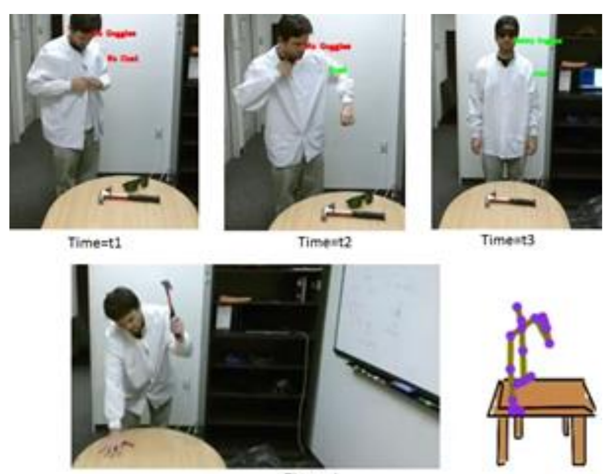


Figure 2: Visual Data Captured and Presented by a Co-robot During Student Task Completion

presents a time series representation of a student hammering on a table. The co-robot takes as input, this video data, primarily focusing on the facial keypoints and body movement of the student. In Figure 2, a skeletal representation of the student (bottom right) can be seen, which represents the visual feedback that the co-robot system displays back to the student, with the color of the skeletal image varying, depending on how well the student is performing a task. In this work, the co-robot is created using a commercially available, off-the-shelf Microsoft Kinect sensor in order to expand the reach and potential availability of these systems in engineering education. The capability of the co-robot to provide real-time visual feedback (e.g., via a mounted display, see Figure 1) enables researchers to study the differences in facial expressions exhibited by students as they received instructions on how to perform a task and feedback of their subsequent performance on that task. This knowledge can enable co-robots systems to improve the timing and context of their feedback. Additionally, it enables co-robots systems to provide interventions when certain patterns are discovered that have the potential to reduce a student's task performance.

3.1. Data Acquisition

3.1.1. Facial Keypoint Data

The multimodal Microsoft Kinect sensor mounted on the co-robot (see Figure 1) coupled with the Kinect Software Development Kit (SDK)[45] allows it to collect a student's facial keypoint data. Table 1 shows the 10 facial keypoints from a student's face that is captured by the sensor. The facial keypoint values are measured on a range from 0-1, indicating a relative weight from an Action Unit. This approach resembles the Facial Action Coding System [46], in which expert observers manually code the facial displays of an individual. These facial displays are referred to as Action Units [47]. The Microsoft Kinect's algorithms allow it to automatically capture these Action Units with the use of video and depth data [48].

Table 1. Facial keypoint data collected by the Microsoft Kinect

1	*Upper Lip Raised*	6	*Jaw Lowered*
2	*Left Lip Stretched*	7	*Right Brow Lowered*
3	*Right Lip Stretched*	8	*Left Eyelid Closed*
4	*Left Brow Lowered*	9	*Right Eyelid Closed*
5	*Left Lip Corner Depressor*	10	*Right Lip Corner Depressor*

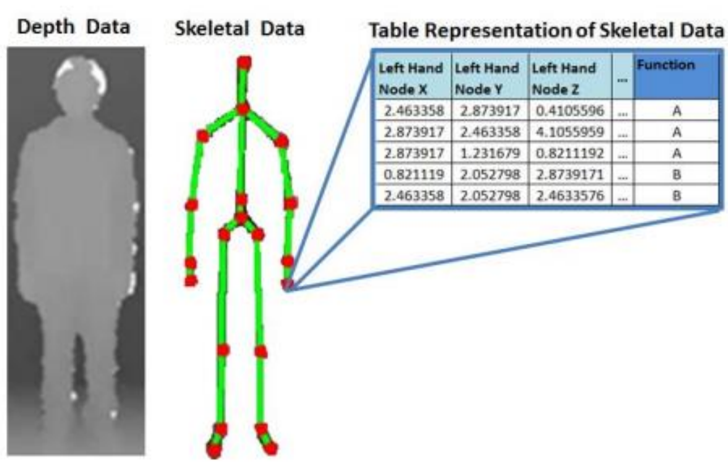
Figure 3 shows a set of actors displaying particular *Action Units* (AU) that relate to the facial keypoints that can be captured by the Microsoft Kinect sensor. Therefore, if a student has his/her eyes closed (e.g., similar to the second actor shown from right to left in Figure 3), the *Right Eyelid Closed* and *Left Eyelid Closed* facial keypoint values will shows as 1; while 0 if the eyes are completely open. An advantage of these facial keypoint data, captured by the Microsoft

Kinect sensor, is that they are robust against movement of students' faces and variations of students' facial structure [49]. This is an essential attribute that allows collecting students' facial data from a wider population even in environments where students move constantly (e.g., engineering laboratory environments).

Figure 3. Representation of Facial Keypoint data, from Ekman 1978 [46]

3.1.2. Visual Feedback Data

Besides capturing facial data, the Microsoft Kinect sensor allows the co-robot to capture students' skeletal data while performing a task. This data enables the robot to assess a student's performance on a task and subsequently provide visual feedback to the student about his/her performance. The Microsoft Kinect sensors capture data from a student's body joints, which consist of *X*, *Y*, and *Z* positions, relative to a fixed reference point (i.e., the location of the Kinect). Figure 4 shows a representation of a student's skeletal data in which the first instance of data captured for the Left Hand Node is *X* =2.4634, *Y*=2.8739, and *Z*=0.4105. This data allows the co-robot to construct a three-dimensional representation of a student's skeletal system while performing a task. If the co-robot is trained with ground truth data acquired from multiple students correctly and incorrectly performing a series of physical tasks (e.g., hammering a nail, cutting wood), the co-robot would be capable of determining (i.e., through the use data mining algorithms), what type of task the students is performing. Furthermore, the co-robot will also be able to determine whether a student's actions deviate from the movements historically associated with correctly performing a given task.



Left Hand Node X	Left Hand Node Y	Left Hand Node Z	...	Function
2.463358	2.873917	0.4105596	...	A
2.873917	2.463358	4.1055959	...	A
2.873917	1.231679	0.8211192	...	A
0.821119	2.052798	2.8739171	...	B
2.463358	2.052798	2.4633576	...	B

Figure 4. Representation of a human skeletal data, from Tucker & Kumara 2015 [2] .

This machine learning approach of "*teaching*" the co-robot what type of movements are required to correctly perform a task, involves the use of supervised machine learning algorithms and the use of ground truth data [2]. Another approach is for the co-robot to be trained with predefines performance models. These performance models can consist of a series of *if-else* rules that will enable the co-robot to automatically assess if a student is correctly performing a task or not. For example, Figure 5 illustrates a student's skeletal system while performing a physical task in a virtual environment, with the body joints highlighted. The joints highlighted in red indicate the ones that do not comply with the predefined performance model for that task, while the ones in green indicate the joints that comply. The student's skeletal data captured while he/she is performing a task enables the co-robot to identify if the student is correctly performing the task or not.

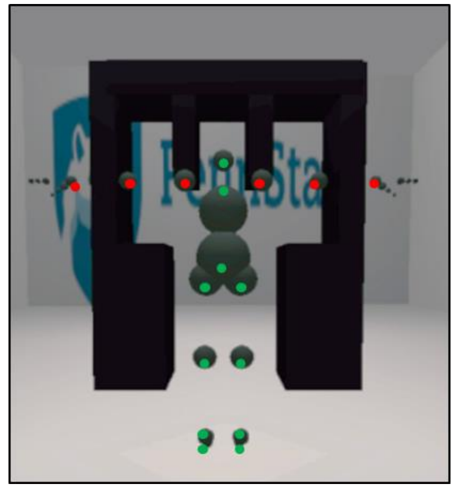


Figure 5. An example representation of a student's skeletal system while performing a task in the virtual environment

Figure 6, shows an example of a predefined performance model consisting of a series of *if-else* rules. The algorithms check if the coordinates position of a student's body joints X_j,Y_j,Z_j, for j in the set of *Joints* **{**1 trough K}, comply with the predefined performance model (X_p,Y_p,Z_p). Therefore, for a student to perform a task correctly, all of his/her joints have to be within the area predefined by the performance model of that specific task. If at least one joint is not within the predefined area (e.g., $[X_1,Y_1,Z_1] \neq [X_p,Y_p,Z_p]$), this suggest that the student has not performed the tasks correctly. In the example shown in Figure 5, it can be seen that 6 of the joints did not comply with the predefined performance model (i.e., joints highlighted in red). Therefore, in this example, the student did not perform the task correctly. Furthermore, the total number of joints K tracked can vary between sensors. Table 2 shows the student's body joints data captured by the Microsoft Kinect sensor used in this work. This skeletal data enables the co-robot to assess a student's performance on a task and subsequently, provide visual feedback to the student about his/her performance via the display mounted on the co-robot (see Figure 1).

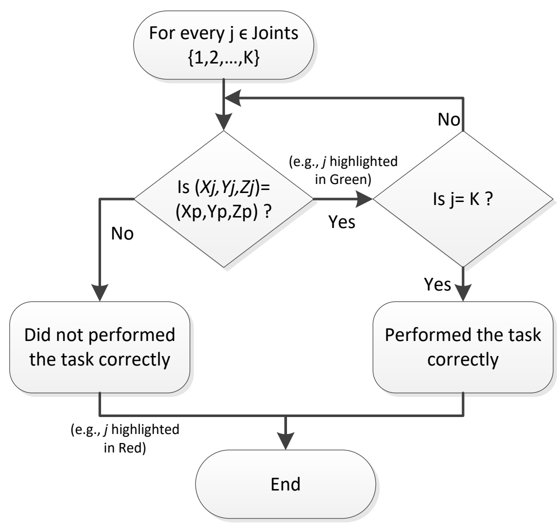


Figure 6. Example of predefine performance model

Table 2. Student' Joints Tracked by the Microsoft Kinect, from Lopez and Tucker [50]

1	*Head*	**10**	*Left toe*
2	*Neck*	**11**	*Left ankle*
3	*Pelvis*	**12**	*Left knee*
4	*Left wrist*	**13**	*Right knee*
5	*Left elbow*	**14**	*Right ankle*
6	*Left shoulder*	**15**	*Right toe*
7	*Right shoulder*	**16**	*Right hip*
8	*Right elbow*	**17**	*Left hip*
9	*Right wrist*		

4. Application

A case study is presented that demonstrates how facial keypoint and skeletal data can be used to potentially improve human-co-robot interactions and provide feedback to students. The case study involves the use of a co-robot system to assist students while performing tasks on a physically-interactive application in a virtual environment. The co-robot system is capable of capturing students' facial keypoint and skeletal data, and providing real-time visual feedback to students on how to perform a task. This physically-interactive application allows researchers to "*teach*" the co-robot predefined performance models for each task. Therefore, enabling the co-robot to assess a student's performance and provide visual feedback to students without the use of supervised machine learning algorithms. In physically-interactive applications, individuals are required to use full body motions (e.g., *jump*, *move side to side*, *bend*) to perform a physical task [50]. In this work, the students had to perform certain motions to pass through obstacles without touching them (see Figure 5). The application consisted of 12 different tasks (i.e., *obstacle avoidance*).

This application allows the authors to demonstrate how facial keypoint and skeletal data can be used by a co-robot to provide real-time feedback to students in a controlled environment, while still maintaining some of the characteristics present in a real world application of a co-robot system in an engendering laboratory environment. For example, this application allows the authors to control for the start and completion time of a task, while still providing students with real-time visual feedback about the characteristic of the tasks and his/her performance after completing the task. This enables researchers to control for the time in which a student is introduced to a task, the time when he/she start performing a task, and the time when the co-robot provides the feedback to the student. This helped reduce any nuisance factor related to the timing of the tasks or feedback while quantifying the effects of the co-robot visual feedback on a student's facial keypoint data. Data from 31 undergraduate students (males: 28, females: 3) from the Pennsylvania State University was collected. Students' ages ranged from 18 to 22 years old (mean: 19.9 years, standard deviation: 1.2 years). The participants were from different academic programs (e.g., engineering, humanities) and from different levels of studies (e.g., freshmen to seniors). Each participant was introduced to the application and the experimental setup, after the completion of the informed consent documents.

4.1. Data Acquisition

4.1.1. Facial Keypoint Data.

The facial keypoint data capture by the co-robot allows the authors to gain a better understanding of the differences in facial expressions exhibited by students as they receive feedback pertaining how to perform a task and their subsequent performance on that task. The co-robot can potentially use this facial keypoint data to find patterns that will enable it to improve the timing and content of its feedback. The results from the case study suggest that students that correctly performed the task, exhibited unique patterns on their facial keypoint data while receiving instructions on how to perform the task. This pattern was found to be significantly different between the students that did not perform the task correctly (e.g., Fail) and those who did (e.g., Pass). Figure 7 shows a *Spider chart* of the student's average facial keypoint data values while the co-robot provides them with instructions on how to perform the task. From Figure 7, it is evident that there are significant differences in some facial keypoint data between the students that correctly performed a task and those who did not. To test this hypothesis, a series of two-sample *t*–test were performed. The hypotheses tests can be express as:

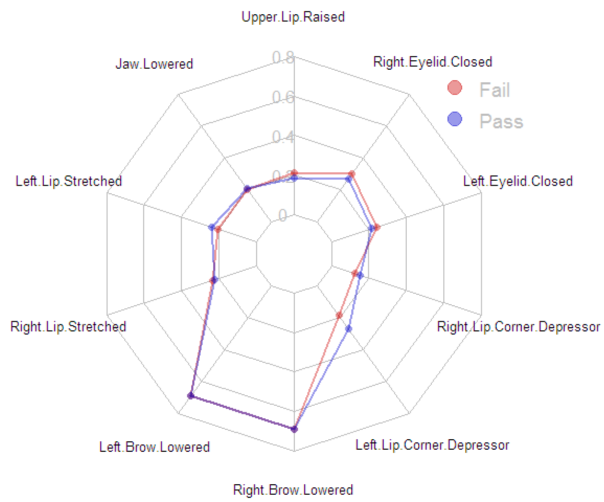


Figure 7. Spider Chart of students' average facial keypoint data while receiving instructions on how to perform a task, grouped by their subsequent performance on the task,

$$\text{(i)} \qquad H_o: \boldsymbol{\mu^b}_{\boldsymbol{y,P}} = \boldsymbol{\mu^b}_{\boldsymbol{y,F}} \;\; vs \;\; H_a: \boldsymbol{\mu^b}_{\boldsymbol{y,P}} \neq \boldsymbol{\mu^b}_{\boldsymbol{y,F}}$$

Where :

- $\boldsymbol{\mu^b}_{\boldsymbol{y,P}}$ is the mean values for the facial keypoint data *y* while receiving instructions on how to perform a task. For *y* ϵ the set of facial keypoints, and *P* ϵ the set of students that subsequently performed the task correctly (e.g., Pass).

- $\boldsymbol{\mu^b}_{\boldsymbol{y,F}}$ is the mean values for the facial keypoint data *y* while receiving instructions on how to perform a task. For *y* ϵ the set of facial keypoints, and *F*ϵ the set of students that subsequently did not perform the task correctly (e.g., Fail).

A Bonferroni correction for family-wise error rate is used to test for the statistical significance of the results (Bonferroni correction= α/m, where m is the number of tests). Therefore, a p-value of 0.005 (Bonferroni correction=0.05/10) on the two-sample *t*–test will provide statistically significant evidence against the null hypotheses. Table 3 present the *t-statistics* and *p-values* of the hypotheses tests conducted. The results suggest that the facial keypoints values of *Upper Lip Raised, Left Lip Stretched, Right Lip Stretched, Left Lip Corner Depressor, Right Lip Corner Depressor, Left Eyelid Closed, and Right Eyelid Closed* on average are significantly different between the group of students that correctly performed a task and those who did not. The co-robot could use this new knowledge to improve the timing and content of its feedback, and automatically predict when a student might need a different type of feedback, based his/her facial keypoint data projected while receiving instructions on how to perform a task.

Table 3. Summary statistics of the two-sample t-test results of the facial keypoint data before performing a task

Facial Keypoints	***t-statistic***	***p-value***
Upper Lip Raised	7.29	<0.005
Jaw Lowered	-1.38	0.1686
Left Lip Stretched	-8.25	<0.005
Right Lip Stretched	3.36	<0.005
Left Brow Lowered	0.64	0.5242
Right Brow Lowered	0.64	0.5242
Left Lip Corner Depressor	-15.43	<0.005
Right Lip Corner Depressor	-7.05	<0.005
Left Eyelid Closed	9.03	<0.005
Right Eyelid Closed	7.64	<0.005

4. 1.2. Visual Feedback Data

Thanks to the capabilities of the Microsoft Kinect sensor, the co-robot can capture students' skeletal data while performing a task. This data enables the co-robot to assess a student's performance on a task and subsequently provide visual feedback about his/her performance. By analyzing students' facial keypoint data while the co-robot provides performance feedback, the authors can gain a better understanding of the differences in facial expressions exhibited by students as they receive this feedback about their performance on a task.

The results from the case study suggest that students who received a positive performance feedback (e.g., Pass), exhibited unique patterns on their facial keypoint data, which is significantly different from the students who did not correctly perform the tasks and hence, received a negative performance feedback (e.g., Fail). Figure 8 shows a *Spider chart* of the students' average facial keypoint data values while the co-robot displays a visual feedback about their performance on a task, grouped by the type of feedback provided (e.g., Fail or Pass). From this chart, it is clear that students' facial keypoint data significant varies depending on the feedback provided. To test these hypotheses, a series of two-sample *t*–test were performed. The hypotheses tests can be express as:

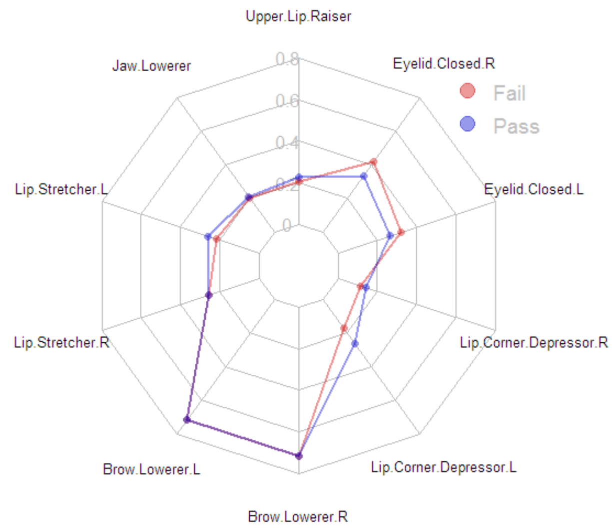


Figure 8. Spider Chart of students' average facial keypoint data while receiving performance feedback, grouped by type of feedback.

$$\text{(ii)} \qquad H_o: \boldsymbol{\mu^a}_{\boldsymbol{y,P}} = \boldsymbol{\mu^a}_{\boldsymbol{y,F}} \;\; \boldsymbol{vs}\; H_a: \boldsymbol{\mu^a}_{\boldsymbol{y,P}} \neq \boldsymbol{\mu^a}_{\boldsymbol{y,F}}$$

Where :

- $\boldsymbol{\mu^a}_{\boldsymbol{y,P}}$ is the mean values for the facial keypoint data y while receiving performance feedback. For $y \in$ the set of facial keypoints, and $P \in$ the set of students that performed the task correctly and received positive performance feedback (e.g., Pass).

- $\boldsymbol{\mu^a}_{\boldsymbol{y,F}}$ is the mean values for the facial keypoint data y while receiving performance feedback. For $y \in$ the set of facial keypoints, and $F \in$ the set of students that did not perform the task correctly and received negative performance feedback (e.g., Fail).

Table 4 present the *t-statistics* and *p-values* of the hypotheses tests conducted. The results suggest that the facial keypoints values of *Upper Lip Raised, Left Lip Stretched, Left Lip Corner Depressor, Right Lip Corner Depressor, Left Eyelid Closed, and Right Eyelid Closed* on average are statistically significantly different between the group of participants' that correctly performed a task and received positive performance feedback, and those who did not and received negative performance feedback.

Table 4. Summary statistics of the two-sample t-test results of the facial keypoint data before performing a task

Facial Keypoints	***t-statistic***	***p-value***
Upper Lip Raised	-3.27	<0.005
Jaw Lowered	-0.71	0.4757
Left Lip Stretched	-6.11	<0.005
Right Lip Stretched	-0.15	0.8804
Left Brow Lowered	0.32	0.7491
Right Brow Lowered	0.32	0.7491
Left Lip Corner Depressor	-11.17	<0.005
Right Lip Corner Depressor	-4.68	<0.005
Left Eyelid Closed	9.69	<0.005
Right Eyelid Closed	12.75	<0.005

The previous results suggest that students project different facial cues as they received instructions on how to perform a task and feedback regarding their performance after completing the task. Additionally, it suggests that these differences in facial cues related to how well the student performed on the task. As Figure 7 and 8 illustrate, there are some facial keypoints of a student's face that change more significantly than others. For example, Figure 9 shows a time series plot of how students' facial keypoints values of *Jaw Lowered* and *Upper Lip Raised* change over time. The plots show how the average values of these facial keypoints change as the co-robot provides instructions to the student on how to perform the task (i.e., time 0sec- 1.5sec, before Task Start Time), and feedback regarding their performance after completing the task (i.e., time 2.5sec- 3sec, after Feedback Time). The plot illustrates how the facial keypoint of *Jaw Lowered* does not change significantly, while the facial keypoint of *Upper Lip Raised* does change significantly. This *Upper Lip Raised* facial keypoint on average changes significantly after the co-robot provides a student with performance feedback. Moreover, these changes are correlated to the type of performance feedback the co-robots provides. Figure 9 shows that if the student received a positive performance feedback (e.g., Pass) the *Upper Lip Raised* facial keypoint values increases, while it decreases if negative performance feedback (e.g., Fail) is provided. More importantly, it shows that while the co-robot provides instructions on how to perform a task, the *Upper Lip Raised* facial keypoint values are on average, significantly different for students who subsequently performed the task correctly and those who did not. The co-robot could use this facial keypoint data to find patterns that will enable it to improve the timing and content of its feedback, with the objective to maximize student's task performance.

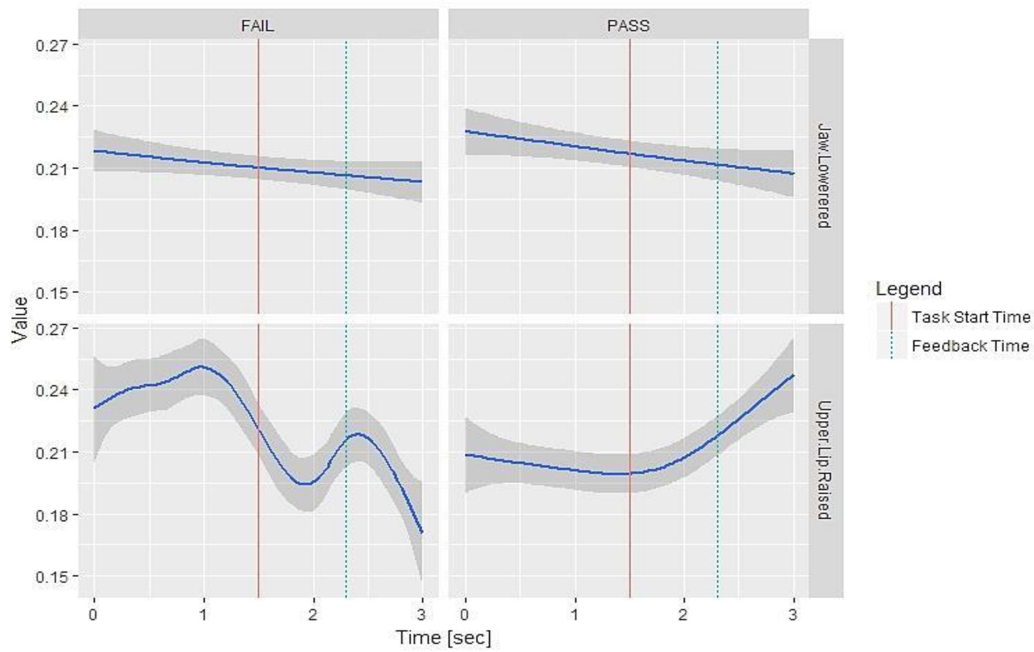


Figure 9. Time series plot of student facial keypoint data

5. Conclusions and Path forward

Determining when and how to provide feedback that advances learning remains an open research question for human-co-robot interactions. The authors propose a method that seeks to address this knowledge gap through the design of a co-robot system that is capable of capturing student's facial keypoint and skeletal data, and providing real-time feedback and instructions to students. The co-robot is created using a commercially available, off-the-shelf Microsoft Kinect sensor, in order to expand the reach and potential availability of these systems in engineering education. The ability of the co-robot to provide real-time visual feedback enables researchers to study the differences in facial expressions exhibited by students as they received instructions about how to perform a task and feedback regarding their subsequent performance on that task. The results of this work reveal that students' facial keypoint data acquired while they perform a task, can be utilized by co-robot systems to learn about the state changes in humans, as they complete a task. This work also explores the influence that co-robot visual feedback has in changing students' behavior while performing a task. The results suggest that co-robots systems could use students' facial keypoint data to find patterns that will enable them to improve the timing and content of their feedback. Future studies should keep exploring the effects of co-robots' feedback and the relationship between students' facial keypoint data and their task performance. Additionally, future works should focus on implementing this new knowledge to improve the human-co-robot interactions and the capability of co-robots to assist and improve the performance of students in engineering laboratory environments.

Acknowledgement

This research is funded by the National Science Foundation NSF NRI #1527148. Any opinions, findings, or conclusions found in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

References

1. National Robotics Initiative 2.0: Ubiquitous Collaborative Robots (NRI-2.0) (nsf17518) | NSF - National Science Foundation.
2. Tucker C, Kumara S. An Automated Object-Task Mining Model for Providing Students with Real Time Performance Feedback. In: ; 2015:26.178.1-26.178.13.
3. Hu Q, Bezawada S, Gray A, Tucker C, Brick T. Exploring the Link Between Task Complexity and Students' Affective States During Engineering Laboratory Activities. In: *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers; 2016:V003T04A019-V003T04A019.
4. Ekman P, Rosenberg EL. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS).*; 2012. doi:10.1093/acprof:oso/9780195179644.001.0001.
5. Calvo RA, D'Mello S. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Trans Affect Comput*. 2010;1(1):18-37. doi:10.1109/T-AFFC.2010.1.
6. Metaxas D, Zhang S. Review article A review of motion analysis methods for human Nonverbal Communication Computing. *Imavis*. 2013;31(6-7):421-433. doi:10.1016/j.imavis.2013.03.005.
7. Bezawada S, Hu Q, Gray A, Brick T, Tucker C. Automatic Facial Feature Extraction for Predicting Designers' Comfort With Engineering Equipment During Prototype Creation. *J Mech Des*. 2017;139(2):21102. doi:10.1115/1.4035428.
8. Palacin J, Salse JA, Valgañón I, Clua X. Building a mobile robot for a floor-cleaning operation in domestic environments. *Instrum Meas IEEE Trans*. 2004;53(5):1418-1424.
9. Singh A, Batalin MA, Stealey M, et al. Mobile robot sensing for environmental applications. In: *Field and Service Robotics*. Springer; 2008:125-135.
10. Schlegl T, Kroger T, Gaschler A, Khatib O, Zangl H. Virtual whiskers—Highly responsive robot collision avoidance. In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE; 2013:5373-5379.
11. Do Y, Kim J. Infrared range sensor array for 3D sensing in robotic applications. *Int J Adv Robot Syst*. 2013;10.
12. Leonard JJ, Durrant-Whyte HF. *Directed Sonar Sensing for Mobile Robot Navigation*. Vol 175. Springer Science & Business Media; 2012.
13. Fredslund J, Mataric MJ. A general algorithm for robot formations using local sensing and minimal communication. *Robot Autom IEEE Trans*. 2002;18(5):837-846.
14. Üneri A, Balicki MA, Handa J, Gehlbach P, Taylor RH, Iordachita I. New steady-hand eye robot with micro-force sensing for vitreoretinal surgery. In: *Biomedical Robotics and Biomechatronics (BioRob), 2010 3rd IEEE RAS and EMBS International Conference on*. IEEE; 2010:814-819.
15. Edsinger-Gonzales A, Weber J. Domo: a force sensing humanoid robot for manipulation research. In: *Humanoid Robots, 2004 4th IEEE/RAS International Conference on*. Vol 1. IEEE; 2004:273-291.
16. Lilienthal A, Zell A, Wandel M, Weimar U. Sensing odour sources in indoor environments without a constant airflow by a mobile robot. In: *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*. Vol 4. IEEE; 2001:4005-4010.
17. Wijesoma WS, Kodagoda KS, Balasuriya AP. Road-boundary detection and tracking using ladar sensing. *Robot Autom IEEE Trans*. 2004;20(3):456-464.
18. Vasudevan N, Tucker CS. Digital Representation of Physical Artifacts: The Effect of Low Cost, High Accuracy 3D Scanning Technologies on Engineering Education, Student Learning and Design Evaluation. In: *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers; 2013:V001T04A013-V001T04A013.
19. Tucker CS, Saint John DB, Behoora I, Marcireau A. Open Source 3D Scanning and Printing for Design Capture and Realization. In: *ASME 2014 International Design Engineering Technical Conferences and*

Computers and Information in Engineering Conference. American Society of Mechanical Engineers; 2014:V01BT02A013-V01BT02A013.
20. Choi C, Christensen HI. RGB-D object tracking: A particle filter approach on GPU. In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE; 2013:1084-1091.
21. Burns B, Samanta B. Human Identification for Human-Robot Interactions. November 2014:V04BT04A044. doi:10.1115/IMECE2014-38496.
22. Morato C, Kaipa KN, Zhao B, Gupta SK. Toward Safe Human Robot Collaboration by Using Multiple Kinects Based Real-Time Human Tracking. *J Comput Inf Sci Eng*. 2014;14(1):011006-011006. doi:10.1115/1.4025810.
23. Bernier E, Chellali R, Thouvenin IM. The MobilAR Robot, Ubiquitous, Unobtrusive, Augmented Reality Device. July 2012:375-381. doi:10.1115/ESDA2012-82794.
24. Behoora I, Tucker C. Quantifying Emotional States Based On Body Language Data Using Non Invasive Sensors. In: *Proceedings of the 2014 ASME IDETC/CIE, DETC2014*. Buffalo, NY; 2014.
25. Castro-Giráldez M, Fito PJ, Toldrá F, Fito P. Physical sensors for quality control during processing. *Handb Meat Process Wiley-Blackwell, Oxford*. 2010:443-456.
26. DiCicco DS, Gong X, Jeffers JL, Kennedy Jr HJ, Rosser RJ, Tan Y. *Image Insertion in Video Streams Using a Combination of Physical Sensors and Pattern Recognition*. Google Patents; 2000.
27. Dyrbye K, Brown TR, Eriksen GF. Packaging of physical sensors for aggressive media applications. *J Micromechanics Microengineering*. 1996;6(1):187.
28. Henderson T, Shilcrat E. Logical sensor systems. *J Robot Syst*. 1984;1(2):169-193.
29. Hill J, Szewczyk R, Woo A, Hollar S, Culler D, Pister K. System architecture directions for networked sensors. In: *ACM SIGOPS Operating Systems Review*. Vol 34. ACM; 2000:93-104.
30. Jakoby B, Vellekoop MJ. Physical sensors for water-in-oil emulsions. *Sensors Actuators A Phys*. 2004;110(1):28-32.
31. Riek LD. The social co-robotics problem space: Six key challenges. *Robot Challenges Vis*. 2014.
32. Casper J, Yanco H. AAAI/Robocup-2001 robot rescue. *AI Mag*. 2002;23(1):43.
33. Galbraith B V., Guenther FH, Versace M. A neural network-based exploratory learning and motor planning system for co-robots. *Front Neurorobot*. 2015;9.
34. Kok JR, Spaan MT, Vlassis N. Multi-robot decision making using coordination graphs. In: *Proceedings of the 11th International Conference on Advanced Robotics, ICAR*. Vol 3. ; 2003:1124-1129.
35. Stone P, Sridharan M, Stronger D, et al. From pixels to multi-robot decision-making: A study in uncertainty. *Rob Auton Syst*. 2006;54(11):933-943.
36. Huang H-P, Liang C-C. Strategy-based decision making of a soccer robot system using a real-time self-organizing fuzzy decision tree. *Fuzzy Sets Syst*. 2002;127(1):49-64.
37. Lin ZM, Yang J-S, Yang CY. Grey decision-making for a billiard robot. In: *Systems, Man and Cybernetics, 2004 IEEE International Conference on*. Vol 6. IEEE; 2004:5350-5355.
38. Mihaylova L, Lefebvre T, Bruyninckx H, Gadeyne K, De Schutter J. A comparison of decision making criteria and optimization methods for active robotic sensing. In: *Numerical Methods and Applications*. Springer; 2002:316-324.
39. Rao RV. *Decision Making in the Manufacturing Environment: Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods*. Springer Science & Business Media; 2007.
40. Kaupp T, Makarenko A, Durrant-Whyte H. Human–robot communication for collaborative decision making—A probabilistic approach. *Rob Auton Syst*. 2010;58(5):444-456.
41. Cao M, Stewart A, Leonard NE. Integrating human and robot decision-making dynamics with feedback: Models and convergence analysis. In: *47th IEEE Conference on Decision and Control, 2008. CDC 2008*. ; 2008:1127-1132. doi:10.1109/CDC.2008.4739103.
42. Rani P, Sarkar N, Smith CA, Kirby LD. Anxiety detecting robotic system–towards implicit human-robot collaboration. *Robotica*. 2004;22(1):85-95.
43. Hinds PJ, Roberts TL, Jones H. Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interact*. 2004;19(1):151-181.
44. Stanton C, Stevens CJ. Robot pressure: the impact of robot eye gaze and lifelike bodily movements upon decision-making and trust. In: *Social Robotics*. Springer; 2014:330-339.
45. Microsoft Research. Kinect for Windows SDK beta. *Program Guid*. 2011:1-33.
46. Ekman P, Friesen W V. Manual for the facial action coding system. *Consult Psychol Press*. 1978.
47. Tian YL, Kanade T, Conn JF. Recognizing action units for facial expression analysis. *IEEE Trans Pattern Anal Mach Intell*. 2001;23(2):97-115. doi:10.1109/34.908962.

48. Han J, Shao L, Xu D, Shotton J. Enhanced computer vision with Microsoft Kinect sensor: A review. *IEEE Trans Cybern*. 2013;43(5):1318-1334. doi:10.1109/TCYB.2013.2265378.
49. Sato, K., Nose, T., & Ito A. Synthesis of Photo-Realistic Facial Animation from Text Based on HMM and DNN with Animation Unit. *Adv Intell Inf Hiding Multimed Signal Process Proceeding Twelfth Int Conf Intell Inf Hiding Multimed Signal Process*. 2016;2:29-3.
50. Lopez, C. E., & Tucker CS. A quantitative method for evaluating the complexity of implementing and performing game features in physically-interactive gamified applications. *Behav Comput Hum*. 2017;71:42-58.