Face to Robot Interaction with Anki Vector to Support Wellness and Mental Health

1st Junior Ben Tamo
School of Electrical and Computer Eng.
Georgia Institute of Technology
Atlanta, GA

2nd Koushik Karan Geetha Nagaraj
School of Electrical and Computer Eng.
Georgia Institute of Technology
Atlanta, GA

3rd Adithya Mohanavel
School of Electrical and Computer Eng.
Georgia Institute of Technology
Atlanta, GA

***Abstract*—This study investigates the potential benefits of using Anki Vector, a small robot, for face-to-robot interaction to improve mental health and wellness. The robot's friendly and engaging manner can promote mental well-being by uplifting mood and facilitating relaxation. The research develops an intelligent vision system to detect human facial emotions and responds to these human emotions, and examines how these interactions facilitate relaxation compared to watching relaxation videos. The results suggest that face-to-robot interaction with Anki Vector has the potential to be a useful tool in facilitating relaxation. Overall, this paper highlights the role that robots can play in supporting mental health and wellness and suggests areas for further research and development in this field.**

***Index Terms*—Anki Vector, Face-to-robot**

I. Introduction

Mental health is an important part of total well-being since it affects our emotional, psychological, and social functioning. It has an effect on our capacity to deal with stress, relate to people, and make decisions, among other things. Connecting with people, keeping cheerful, and engaging in physical exercise are just a few of the tactics for sustaining excellent mental health. There has been growing interest in the potential of social robots to improve mental health and well-being in recent years. Several studies have found that social robots can help with stress reduction, emotional control, and social support, as well as generating happy feelings in people of all ages. This paper aims to review the existing literature on the potential benefits of social robots in mental health care and develop a model for anki vector to address this problem statement and collect user study using a robot model trained by the team.

II. Related Works

There has been a surge of interest in the employment of technology, notably robots, in the field of mental healthcare in recent years. Several research have been conducted to evaluate the potential advantages of utilizing socially assistive robots (SAR) and robot-based psychotherapy in mental health therapies.

A. Intelligent Visual System

Within Intelligent Visual system for SAR several research have been conducted to evaluate the use of socially assistive robots (SAR) in mental healthcare therapy. Particularly Lahariya et al [1], looked at the ability of SAR to discern emotions and gender on single-face and multiple-face photos. The suggested model was made up of two CNN models: Mini-Xpection CNN and 4-Layer CNN. According to the findings of this study, SAR has the potential to improve the quality and efficacy of mental health therapies by providing patients with tailored, 24-hour support, removing the stigma associated with mental health treatment, and improving patient engagement and motivation. The study did, however, discover a flaw in the model's low accuracy in recognizing emotions.

In another research David et al [2] looked at predicting mood and gender from single-face and multiple-face photographs. The suggested model combined two CNN models: Mini-Xpection CNN and 4-Layer CNN. The model performed well in terms of gender categorization, with an accuracy of 95%, but performed poorly in terms of emotion classification, with 68% accuracy across seven classes (angry, fear, sad, pleased, surprise, neutral, disgust).

B. Wellness and Mental health

Geva et al [3] research performed a review of the literature on the use of SAR in mental health therapy and made recommendations for its growth. The authors reviewed the literature on the use of SAR in mental health treatments and discussed the benefits and drawbacks of using SAR in mental health therapy. The findings indicated that SAR has the potential to improve the quality and efficacy of mental healthcare interventions; nevertheless, the authors also identified challenges to the adoption of robot-assisted psychotherapy, such as efficacy issues and ethical considerations.

The sMFQ was validated as a reliable measure of depressive symptoms in young adults in a research conducted by Rawana and colleagues [4], in addition to its acknowledged use in childhood and adolescence. The Avon Longitudinal Study of Parents and Children (ALSPAC) data were utilized in the study to find sMFQ cut-points that best represent DSM-5 depression diagnosis at age 25. The sMFQ is an effective instrument for screening and monitoring depression in young adults, according to the study, with the teenage cut-point of 12 performing well in terms of sensitivity and specificity.

III. Hypothesis and Evaluation

Based on the prior work done in the context of human-robot interaction and Emotion detection a hypothesis is formulated to check if the robot improves one's mood better.

- People who visually interact with Anki Vector report increased relaxation feelings compared to watching relaxation videos.

To check the validity of this hypothesis three metrics will be used: heart rate, Data from a Comparative study, and Likert scale analysis on the Short Mood and Feelings Questionnaire (SMFQ). Heart Rate: The first evaluation metric that provides results excluding the bias of participant's view in the User study is the readings of the Heart Rate. For this, the heart rate is collected at every interval of the study in order to check if one's mood is improved or not. As elucidated by Wu Yan et al [10] there is a relation that amusement tends to have a lower heart rate when compared to other emotions. Also, from data portrayed by Shu L et al [11] it is clear that happiness has a lower heart rate reading when compared to sadness which has a higher heart rate reading. We aim to utilize this fact and correlate that a higher heart rate reading compared to a baseline heart rate reading would mean increased stress and thus deteriorated mood and a lower heart rate means reduced stress that corresponds to improved mood. This evaluation metric is one of the concrete metrics used in validating the hypotheses and is used as one of the factors to derive which procedure is better in terms of elevating/improving one's mood. Data from Comparative study: In order to check whether the robot can improve one's mood and feelings there should be a comparison method implemented. Based on findings from Peters et al [5] a comparative study identifies patterns and variations. The two scenarios, the experiment video, and the interaction with Anki-vector are compared in order to arrive at conclusions. The participant fills in a comparative questionnaire with a numerical scale that comprises 8 questions and 2 questions of preferences. A paired sample t-test is performed on the 1-10 scale of the comparative questionnaire. Based on the results of the paired sample t-test conclusions can be made. Likert Scale Analysis on SMFQ: The Mood and Feelings Questionnaire is utilized from the original questionnaire from Angold, A et al [9]. The SMFQ is used which comprises 13 questions to estimate one's mood. In the Likert scale NOT TRUE = 0, SOMETIMES = 1, and TRUE = 2 values are given to obtain a value for estimation. A total of 0 on this scale indicates a good mood and a 26 indicates a worst mood. There are two papers that use this method of concluding Likert scaling to derive outcomes. One is by Olga Eyre [12] et al that compared the difference from the base mood. The other one is by Thabrew et al [13] which uses the total method to define points of consideration of mood. In the end, using all the 3 metrics we can conclude if the stated hypothesis can be validated or not.

IV. Methods

A. Intelligent Visual System

Figure 1 is our proposed workflow for the facial emotion detection system. We proposed to train our model on the FER2013 dataset, as the pixel resolution of these images is low. Therefore training our model on low-quality images will generalize well on Anki Vector images which are slightly better than FER2013 data.

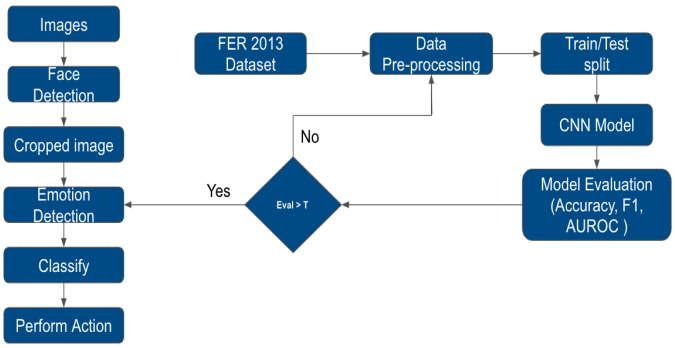

Fig. 1. Intelligent Visual System Workflow

The dataset we are using to train and test our model is the FER2013 dataset. The data consists of 48x48 pixel grayscale images of faces with facial expressions into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples. Figure 2 shows sample images of the dataset.

Fig. 2. FER2013 Dataset

Our proposed model is a simple convolutional neural network (CNN) as depicted in figure 3. In the conv2d layers in yellow, a set of 3x3 filters are applied to the image to extract features. The output from each conv2d is a set of feature maps that highlight different aspects of the image, such as edges, shapes, or textures. The red layer is the maxpooling2d layer. A pooling layer is a layer in a CNN that reduces the spatial size (width and height) of the input volume while retaining the depth. The pooling layer is generally inserted after one or more convolutional layers. The green layer is the dropout layer. A dropout layer is used as a regularizer to prevent overfitting. It works by randomly dropping out (i.e., setting to zero) a certain percentage of the nodes in the layer during each training iteration. The light blue layer is the flatten layer used to convert the multidimensional feature maps into a single-dimensional output. Lastly, the dark blue layer is the dense layer, also known as the fully connected layers to make predictions.

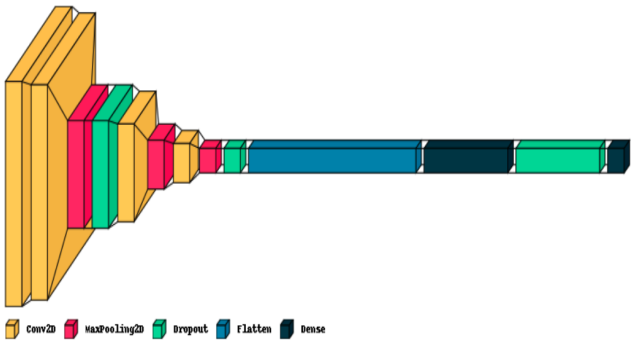

Fig. 3. Model Architecture

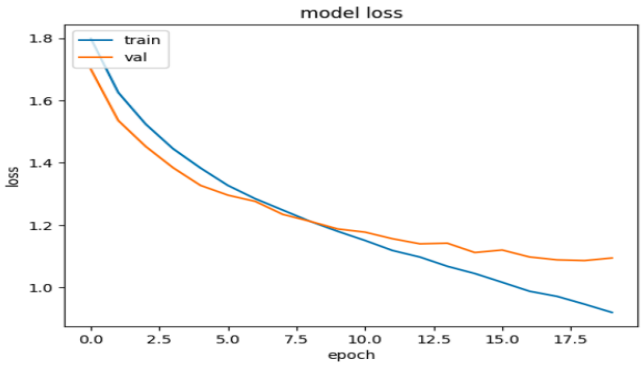

Fig. 6. Model Loss

Our model was trained through 20 epochs for over 3 hours. Figures 5, 6 show the performance of the model. The accuracy obtained is 70.1%. This low accuracy is due to the class imbalance problem shown in figure 4. However, we are only focusing on 5 facial emotions, sad, happy, surprised, neutral, and fear. From the classification report, all these 5 classes achieve a recall above 90%.

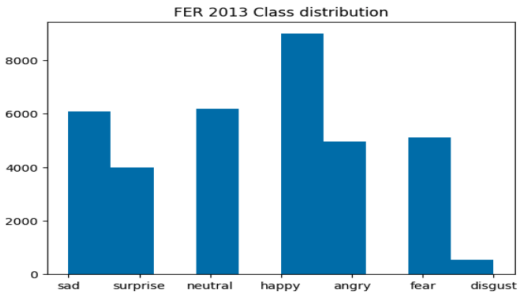

Fig. 4. FER2013 Class Distribution

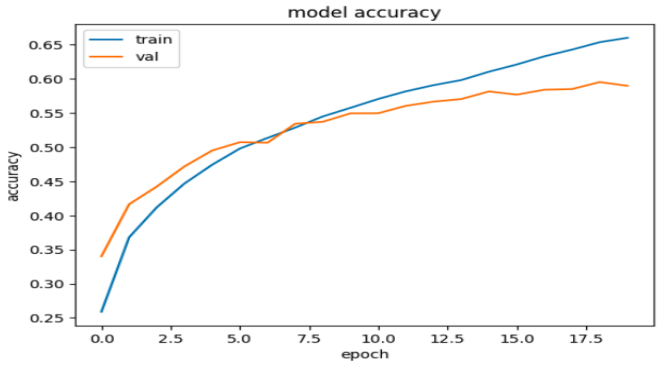

Fig. 5. Model Accuracy

B. User Study Design

Overview:

In order to arrive at conclusions for the hypothesis a User study for Human-Robot Interaction is performed. The goal of this study is to obtain conclusive evidence to prove whether the hypothesis is validated or not. The structure of the user study will involve conducting a series of operations involving filling out questionnaires and interacting with the robot by the participant. The study design steps are portrayed in the flow chart below.

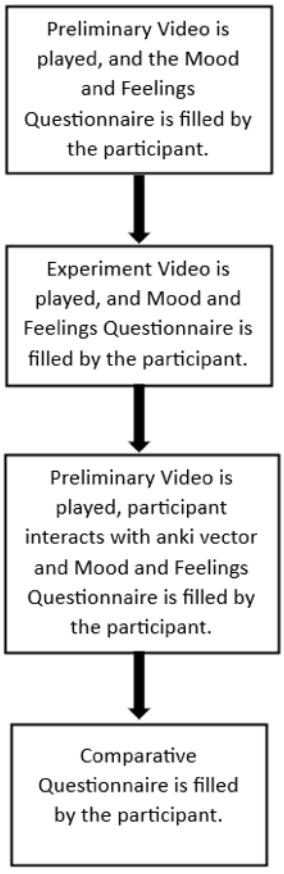

Fig. 7. User Study Design.

Prior to diving into the process of a user study in detail, it is important to put forth the details regarding the participants. The user study group aims to be between 15-25 and the

participants are college students between the ages of 20 and 27. The optimal time for a video to take impact and maximize output is less than 6 minutes derived from Brame CJ [8]. It was concluded that the video length be fixed to be between 2-3 minutes. The instructions provided to the participants would be:

- Please watch the 3-minute Preliminary video 1 and fill in the Mood and Feelings Questionnaire based on your feelings.
- Please watch the 3-minute experiment video and fill in the Mood and Feelings Questionnaire based on your feelings.
- Rewatch the 3-minute Preliminary video 2, interact with anki vector for 3 minutes, and fill in the Mood and Feelings Questionnaire based on your feelings.
- Please fill out the Comparative Questionnaire and optional feedback.

The approximate time to complete the user study for one participant is 16 minutes. The estimated time for the entire user study to be completed is 400 minutes. The robot used for this user study is anki Vector.

User Study Process:

The preliminary video 1 chosen is a clip of "Conjuring 1" and the preliminary video 2 chosen is a clip of "Conjuring 2". A 3-minute experiment video is taken from the internet that has one of the most views in order to accommodate all kinds of people for more standardization. The experiment video chosen is a clip of "Nature's best relaxing video". The steps involving the user study is below:

Step 1: As portrayed in Figure 7, the first step of the User study involves showing a Preliminary video 1 to the participant. Preliminary Video 1 does not contain an overview of the project rather it contains content that aims to increase the stress level of the participant. The point of a preliminary video is to standardize the participant's mood. The preliminary videos are different as playing the same video may not cause any change to one's mood. This is inspired by the work of Droit-Volet Sylvie et al [7] who concluded that movies impact viewers' moods. Once the video is played the participant is made to fill out a Mood and Feelings Questionnaire. The Mood and Feelings Questionnaire plays a vital role in the study's design. The existing Mood and Feelings questionnaires were developed by Angold, A. et al [9]. This followed a 3-scale Likert scale system. As elucidated by Olga Eyre [12] et all the total points are 26 for the questionnaire and a cutoff of greater than 12 indicated a possibility of depression in the paper. This concept is adapted and taken such that a lower value indicates a good mood and a higher value means a bad mood. This concludes the first step in the user study where the mood of a participant is standardized, and Short Mood and Feelings Questionnaire (SMFQ) is filled.

Step 2: The main aim of this project is to check if the implemented module of emotion detection in the robot (anki-vector) elevates one's mood. In order to perform a comparative analysis an experiment video is used. The next step here is to play a relaxing video to the participant to elevate the participant's mood. The previous statement is supported by the findings of Droit-Volet Sylvie et al [7] and by findings from Peters et al [14]. The participant then watches the experiment video and fills out the Mood and Feelings Questionnaire.

Step 3: The first segment of the user study ends with the experiment video. Preliminary video 2 is then played again in order to standardize the mood again because the experiment video may have brought about changes to one's mood and feelings. The participant is then made to interact with the robot (Anki vector) with implemented emotion recognition. The interaction time with a robot (Anki-Vector) is 3 minutes, similar to the experiment video's duration to maintain standardization. After the interaction, the participant is made to fill out the Mood and Feelings Questionnaire.

Step 4: The final step involves the participant filling out a Comparative Questionnaire that involves questions checking the difference in impact between the experiment video and the interaction with the robot (Anki Vector). The formulated Comparative Questionnaire follows a 10-point Numerical Scale for 8 questions and 2 questions of word answers for preferences. The points in the future the points will be used for evaluation.

Additional Step: The participant's heart rate will be recorded every time a questionnaire is filled in. As discussed earlier this will be used to assess one's mood.

V. Results and Discussion

A. Results from Heart Rate Readings

The primary results are obtained from the Heart rate readings taken from all participants for each step.

Table I depicts the Heart rate readings after watching the preliminary video indicated by start, Heart rate readings after watching the Experiment Video, and Heart Rate readings after interacting with the robot (Anki-Vector). The Difference - Experiment Video indicated is the calculation obtained by subtracting the initial heart rate from the Heart rate obtained after watching the Experiment Video. The Difference Anki-Vector is obtained by subtracting the initial heart rate from the Heart rate obtained after interacting with the robot with emotion detection. The values in the last two columns are compared in order to arrive at conclusions. A higher positive value indicates a good increase of mood in the positive direction and a negative value indicates deteriorated mood as elucidated by Wu Yan et al [10] and by Shu L et al [11].

The study involved 15 participants and the metric for comparing is the baseline heart rate which is indicated by start. In Figure 8, the trend is that the Heart rate readings taken after the interaction with the Robot (Anki-Vector) are the least among the three classifications. To be more specific Vector better improves mood for $\frac{10}{15}$ (66.66%) occasions and the experiment video proves to be better $\frac{5}{15}$ (33.33%) times in terms of Heart Rate metric. **So, it is evident here that the robot (Anki Vector) in terms of mood improvement with respect to heart rate is better when compared to the experiment video thus proving the hypothesis to be true**.

TABLE I
HEART RATE READINGS ANALYSIS

S. No	GENDER	START	EXPERIMENT VIDEO	ANKI VECTOR	Difference -Exp-Video	Difference -Anki Vector
P1	MALE	70	65	59	+5	+11
P2	MALE	70	65	74	+5	-4
P3	MALE	97	74	65	+23	+32
P4	MALE	90	89	82	+1	+8
P5	MALE	89	70	69	+19	+20
P6	MALE	88	75	60	+13	+28
P7	MALE	84	79	76	+5	+8
P8	MALE	92	85	86	+7	+6
P9	FEMALE	80	69	68	+11	+12
P10	MALE	70	68	73	+2	-3
P11	MALE	82	80	77	+2	+5
P12	FEMALE	104	94	91	+10	+13
P13	MALE	88	77	86	+11	+2
P14	MALE	71	76	75	-5	-4
P15	MALE	89	88	90	+1	-2

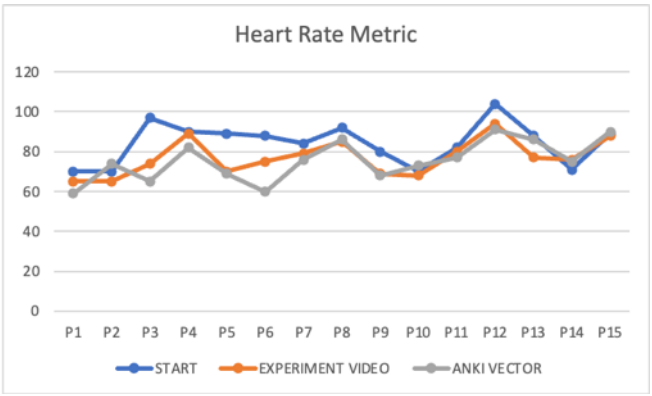

Fig. 8. Heart Rate Recordings Comparison

B. Results from Questionnaire

In order to obtain a conclusion from the comparative study questionnaire Paired-Sample t-test (Kim et al [15]) was used with the data analysis done in Table II. The first conclusion obtained from the 2 Inference questions in the Questionnaire is that in terms of increased enjoyability, all the participants chose Anki Vector. In terms of ease of use, 86.66% of participants chose the experiment video. There are 8 more questions and 4 each for the Anki vector and 4 for the Experiment video. These are used to conclude using the paired sample t-test. The first step is to Define the null and alternative

TABLE II
HEART RATE READINGS ANALYSIS

S.no	ANKI VECTOR (MEAN1)	EXPERIMENT VIDEO (MEAN 2)	DIFFERENCE
P1	39	36	3
P2	23	25	-2
P3	34	35	-1
P4	30	30	0
P5	38	36	2
P6	40	36	4
P7	24	26	-2
P8	30	29	1
P9	34	32	2
P10	27	23	4
P11	25	25	0
P12	36	34	2
P13	38	34	4
P14	26	29	-3
P15	38	36	2
MEAN	32.133	31.067	1.067
STANDARD DEVIATION	6.022	4.667	2.314

hypotheses: Null hypothesis $H_0 : \mu 1 - \mu 2 = 0 (\mu d = 0)$.The difference between the paired population is equal to 0. Here μ_1 is the Anki-Vector mean and μ_2 is Experiment Video mean. Alternative hypothesis $H_1 : \mu 1 - \mu 2 \geq 0 (\mu d \geq 0)$ The difference between the paired population means is not equal to 0. In Table II, Mean $1(\mu_1)$ (column 2 of Table II) is the mean of responses with respect to the robot (anki vector) and Mean 2 (μ_2) is the mean of responses with respect to the experiment video (column 3 of Table II). The next step is to total both sets of the values of the data and find the difference between each pair. Then find the mean and standard deviation.

From Table II the standard deviation is **2.314**, the sample mean is **1.0667**, and the sample size is 15 (group size). The estimated standard error of the mean is 0.5964. The value of t is $\frac{1.0667}{0.5964}$ which is 1.789. So, $t = 1.786$. As

commonly set the significance value α is 0.05 (Kim et al [15], which corresponds to a 5% probability of rejecting the null hypothesis when it is true. The critical value is found using the degree of freedom and significance value. The critical value here is 2.145. The critical value for a left-tailed test is $t_c = -1.761$. The rejection region $R = t : t < -1.761$. Since $t = 1.786 \geq t_c = -1.761$, it is concluded that the null hypothesis is not rejected. Using the P-value approach: p-value $p = 0.9521$ and since $p = 0.9521 > 0.05$, it is concluded that the null hypothesis (H_o) is not rejected. This proves that there is no statistical difference between the two samples. **This metric hence does not provide a solid conclusion in terms of comparative study that the participants felt the Anki vector is greater in improving one's mood when compared to the Experiment video thus giving an inconclusive result to the hypothesis in terms of this metric taken.**

C. Results from Likert Scale Analysis of SMFQ

As obtained from [13] Thabrew et al and the Child Outcome research consortium the total of the Likert scales is taken and compared to see the mood and feelings. First, a table is created with a tally of values.

TABLE III
LIKERT SCALE ANALYSIS ON SMFQ

S.no	BASE	EXPERIMENT VIDEO	ANKI VECTOR	BASE-EXPERIMENT VIDEO	BASE-ANKI VECTOR
P1	10	4	0	6	10
P2	5	3	3	2	2
P3	3	2	1	1	1
P4	2	1	1	1	1
P5	2	0	0	2	2
P6	10	0	7	10	3
P7	13	2	0	11	13
P8	6	3	10	3	-4
P9	10	7	4	3	6
P10	1	1	0	0	1
P11	13	10	10	3	3
P12	10	6	8	4	4
P13	3	1	1	2	2
P14	1	4	0	-3	-1
P15	10	8	7	2	3

From the above Table III, the mood with respect to the analysis on SMFQ when measured by watching the experiment video is better $\frac{2}{15}$ (6.66%) of the time. The Anki Vector with respect to the analysis on SMFQ seems to improve mood by $\frac{6}{15}$ (40%) of the time. Meanwhile, both of these methods improve the mood at the same rate $\frac{7}{15}$ (46.6%) of the time. **Hence the maximum percentage belongs to change of mood at the same rate, a conclusive result is not obtained by using this method to analyze as this is a neutral result. Hence, the Hypothesis can neither be accepted nor rejected with the Likert Scale Analysis on SMFQ.**

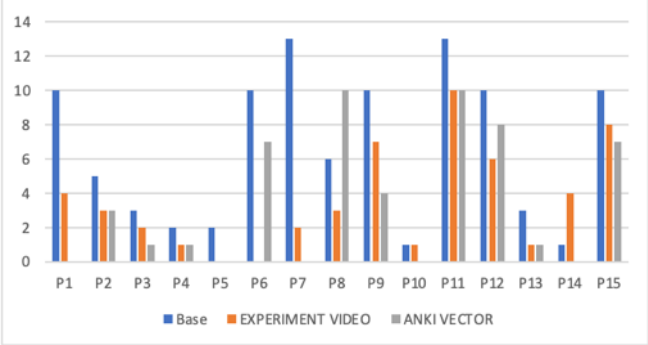

Fig. 9. Likert Scale Totals at 3 situations

Overall, one of the three evaluation metrics (heart Rate) validates the hypothesis but the other (Likert Scale Analysis on SMFQ and Paired Sample T-test on Comparative Questionnaire) does not provide conclusive results on the hypothesis. Overall, there is no conclusive evidence that people who visually interact with Anki Vector report increased relaxation feelings compared to watching relaxation videos.

VI. CONCLUSION

Mental health includes our emotional, psychological, and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make choices. Someone experiencing mental health problems could have their thinking, mood, and behavior affected. There are several ways to maintain positive mental health that include connecting with others, staying positive, and getting physically active, . . . In this project, we explored the wellness and mental health effects of face-to-robot interaction with Anki vector. Face to Robot interaction with Anki Vector refers to making facial expressions (happy, sad, . . .) to the robot. Our Findings suggest that there is no conclusive evidence that face-to-robot with an Anki vector robot can uplift the user's mood or relaxation feeling but one of the metrics validates it to a basic level. There are many effective treatments available, including therapy, medication, and lifestyle changes. Our long-term vision is to use Anki Vector to supplement the current means of treatment. A future extension of this project will be performing a study on people dealing with mental health issues. Another extension of the approach of the project would be to explore various evaluation metrics for a solid conclusion.

REFERENCES

[1] Lahariya, A., Singh, V., Tiwary, U. S. (2021). Real-time Emotion and Gender Classification using Ensemble CNN. arXiv preprint arXiv:2111.07746.

[2] David, D., Matu, S. A., David, O. A. (2014). Robot-based psychotherapy: Concepts development, state of the art, and new directions. International Journal of Cognitive Therapy, 7(2), 192-210.

[3] Geva, N., Hermoni, N., Levy-Tzedek, S. (2022). Interaction matters: The effect of touching the social robot PARO on pain and stress is stronger when turned ON vs. OFF. Frontiers in Robotics and AI, 9, 926185.

[4] Olga Eyre, Rhys Bevan Jones, Sharifah Shameem Agha, Robyn E Wootton, Ajay K Thapar, Evie Stergiakouli, Kate Langley, Stephan Collishaw, Anita Thapar, Lucy Riglin,Validation of the short Mood and Feelings Questionnaire in young adulthood,Journal of Affective Disorders,Volume 294,2021.

[5] Chernova, S., Thomaz, A.L. (2014). Designing and Evaluating an LfD Study. In: Robot Learning from Human Teachers. Synthesis Lectures on Artificial Intelligence and Machine. Springer, Cham.

[6] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. 2006. Common metrics for human-robot interaction. In Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (HRI '06). Association for Computing Machinery, New York, NY, USA, 33–40.

[7] Droit-Volet Sylvie, Fayolle Sophie, Gil Sandrine (2011). Emotion and Time Perception: Effects of Film-Induced Mood. Frontiers in Integrative Neuroscience, Vol.5.

[8] Brame CJ. Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. CBE Life Sci Educ. 2016 Winter;15(4): es6.doi :10.1187/cbe.16-03-0125. PMID: 27789532; PMCID: PMC5132380.

[9] Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., Silver, D. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents (Vol. 5, pp. 237–249).

[10] Wu Yan, Gu Ruolei, Yang Qiwei, Luo Yue-jia. How Do Amusement, Anger, and Fear Influence Heart Rate and Heart Rate Variability. (Vol. 13, 2019).

[11] Shu L, Yu Y, Chen W, Hua H, Li Q, Jin J, Xu X. Wearable Emotion Recognition Using Heart Rate Data from a Smart Bracelet. Sensors (Basel). 2020 Jan 28;20(3):718. doi 10.3390/s20030718. PMID: 32012920; PMCID: PMC7038485.

[12] Olga Eyre, Rhys Bevan Jones, Sharifah Shameem Agha, Robyn E Wootton, Ajay K Thapar, Evie Stergiakouli, Kate Langley, Stephan Collishaw, Anita Thapar, Lucy Riglin, Validation of the short Mood and Feelings Questionnaire in young adulthood, Journal of Affective Disorders, Volume 294,2021,

[13] Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., Merry, S. (2018). Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. International Journal of Methods in Psychiatric Research, 27(3), 1–9.

[14] Peters, B. G., Pierre, J. (2007). Comparative analysis as a research tool: A case study of the use of comparative analysis in public administration research. Journal of Comparative Policy Analysis: Research and Practice, 9(3), 227-244.

[15] Kim, Tae. (2015). T test as a parametric statistic. Korean Journal of Anesthesiology. 68. 540. 10.4097/kjae.2015.68.6.540.