Chapter 5

RESULTS

This chapter presents the findings from both internal measures and a questionnaire that focused on various aspects of the system, including the procedure, robot design, listening and speaking experiences, stress reduction, and the robot's ability to enhance work productivity, efficiency, and accuracy. The results obtained in each category of the questionnaire are presented and discussed using graphs generated from the Google form.

5.1 Sample Composition

It was a sample of 12 people over the age of 20 who works remotely. The feedback form that was provided at the conclusion of the experiment was used to obtain a sample composition. The response distribution for the feedback form's age field is shown in Figure 5.1.

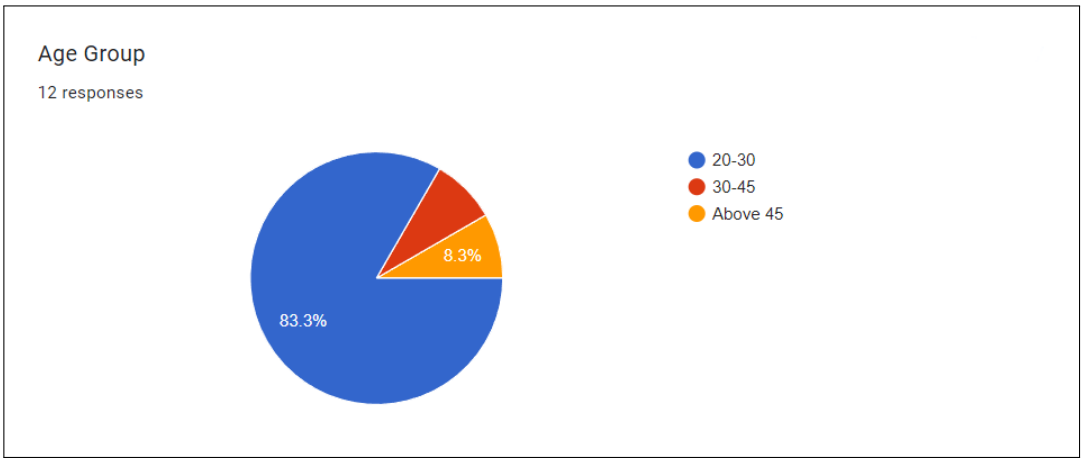

Figure 5.1: Responses for the age field of the feedback form

5.2 Results of the Speech Misdetections

Table 5.1 shows the results that were obtained for the number of speech mid-sections at the end of each experiment.

Table 5.1: Observed speech misdetections

Participant	**Number of Speech Misdetections**
Participant 01	0
Participant 02	2
Participant 03	1
Participant 04	0
Participant 05	2
Participant 06	0
Participant 07	0
Participant 08	1
Participant 09	0
Participant 10	0
Participant 11	0
Participant 12	0

Participants were asked to note how many speech misdetections occurred during the experiment. It was observed that the maximum number of speech misdetections in the system was 2, resulting in a total of 6 speech misdetections throughout 12 experiments. Therefore, the average speech misdetections per experiment were $0.5 < 1$. These results indicate that the implemented system is accurate in capturing user conversations.

It has been observed that the primary reason for speech misdetections is the time delay between the robot conversations and obtaining user input. Most participants who experienced speech misdetections provided their user input before the detection period, resulting in a misdetection.

5.3 Results of the Facial Expressions Analysis

Table 5.2 shows the results that were obtained for the facial expressions analysis of participants

It can be observed that neutral emotions were observed most of the time from almost all of the participants. Other than neutral emotions, participants showed happy emotions. The total amount of neutral emotions was 741 which resulted in an average of 62 neutral emotions per participant. The total amount

Table 5.2: Results of the facial expression analysis

Participant	Anger emotions	Disgust emotions	Scared emotions	Happy emotions	Neutral emotions	Sad emotions	Surprised emotions
Participant 01	Nil	Nil	Nil	30	66	Nil	Nil
Participant 02	Nil	Nil	Nil	12	55	Nil	Nil
Participant 03	Nil	Nil	Nil	27	60	Nil	Nil
Participant 04	Nil	Nil	Nil	10	57	Nil	Nil
Participant 05	Nil	Nil	Nil	3	63	Nil	Nil
Participant 06	Nil	Nil	Nil	16	70	Nil	Nil
Participant 07	Nil	Nil	Nil	12	75	Nil	Nil
Participant 08	Nil	Nil	Nil	10	69	Nil	Nil
Participant 09	Nil	Nil	Nil	34	56	Nil	Nil
Participant 10	Nil	Nil	Nil	25	49	Nil	Nil
Participant 11	Nil	Nil	Nil	33	53	Nil	Nil
Participant 12	Nil	Nil	Nil	21	68	Nil	Nil

of happy emotions was 233 which resulted in an average of 19 happy emotions per participant. Even though the emotion classification algorithm was capable of identifying all the other mentioned emotions, none of them were identified from all the participants.

It is observed that some participants took higher engagement time with the robot due to speech misdetections and they tend to have a higher amount of emotion count than other participants. Further, due to the lighting conditions and placement of the web camera of the laptop, some of the captured emotions were not able to classify properly.

5.4 Results of the User Feedback

This section illustrates the results of the user feedback obtained on areas related to the user experience of the implemented system.

5.4.1 Results of the Procedure

To obtain the feedback of the user on the procedure of the implemented system, it was given a scale between 1-5 that representing 1 forpoor and 5 for very excelent. 4 participants out of 12 have given their feedback with a scale value of 4 and the rest of the participants have given their feedback with a scale value of 5 which represents excelent. Fugure 5.2 shows the responses for ratings of overall procedure.

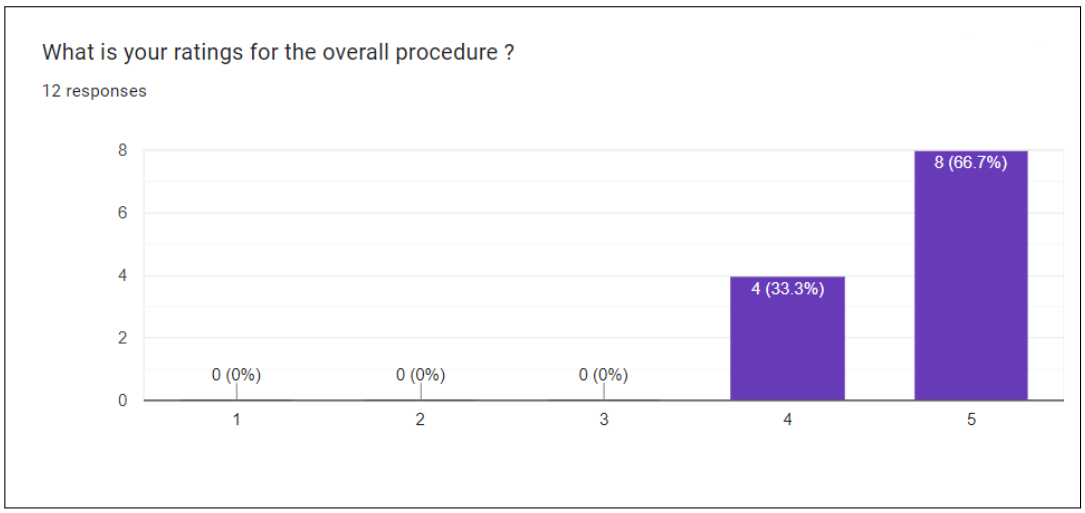

Figure 5.2: Responses for the overall procedure

5.4.2 Results of the Robot Design

To gather user feedback on the design of the implemented companion robot, participants were asked whether they found the robots to be friendly in appearance. Out of the 12 participants, 7 strongly agreed that they had a friendly impression of the robots, while the remaining participants simply agreed that they found the robots to be friendly. Figure 5.3 show the response for friendly impression based on the design.

5.4.3 Results of Robot Eyes Animation

user feedback was collected regarding the suitability of the robot's eye animations. Participants were asked to rate the appropriateness of the animations using a scale ranging from 1 (Highly unsuitable) to 5 (Highly suitable). Out of the 12 participants, 2 rated the animation as a 3 on the scale, 4 participants

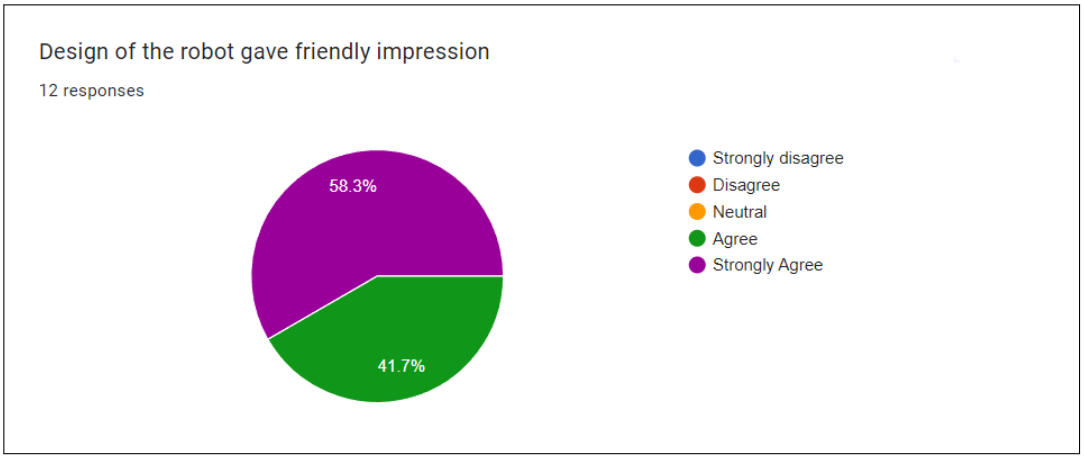

Figure 5.3: Responses for the Friendly impression based on the design

rated it as a 4, and the remaining participants, totaling 6, rated it as a 5, indicating that the animation was highly suitable. These results provide insights into the effectiveness of the robot's eye animations in the context of stress reduction. Figure 5.4 shows the responses for the suitability of robot eyes animation.

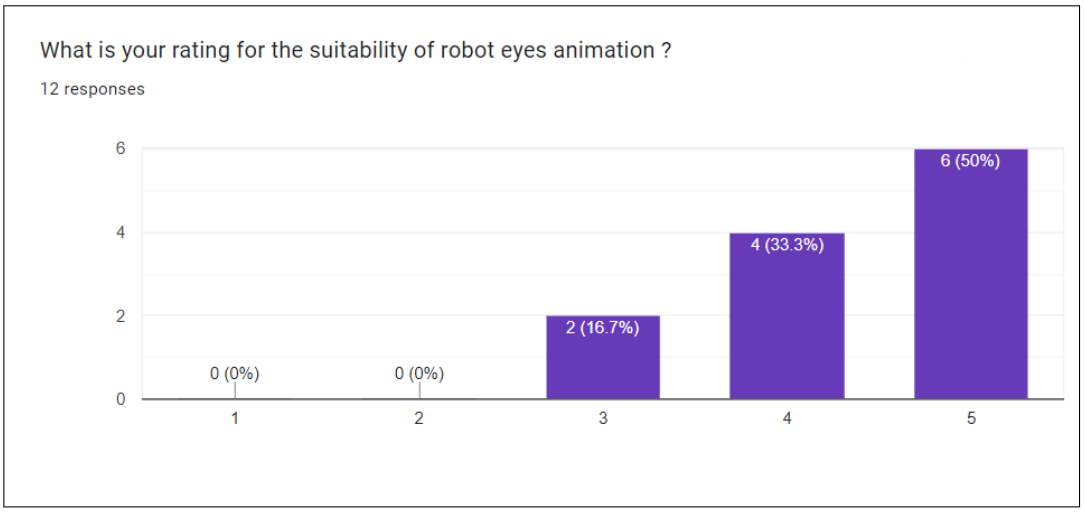

Figure 5.4: Responses for the Robot eyes animation

5.4.4 Results of Robot Conversation Process

Feedback was gathered from participants regarding the robot conversation process. They were asked to rate how effectively the robot communicated with them on a scale of 1 to 5, where 1 represented poor communication and 5 indicated excellent communication. Out of the 12 participants, 2 rated the robot's communication as a 4, while the remaining participants rated it as a 5, indicating

excellent communication. These results demonstrate that the robot is capable of accurately communicating with the users. Figure 5.5 show the results of rating robot conversation

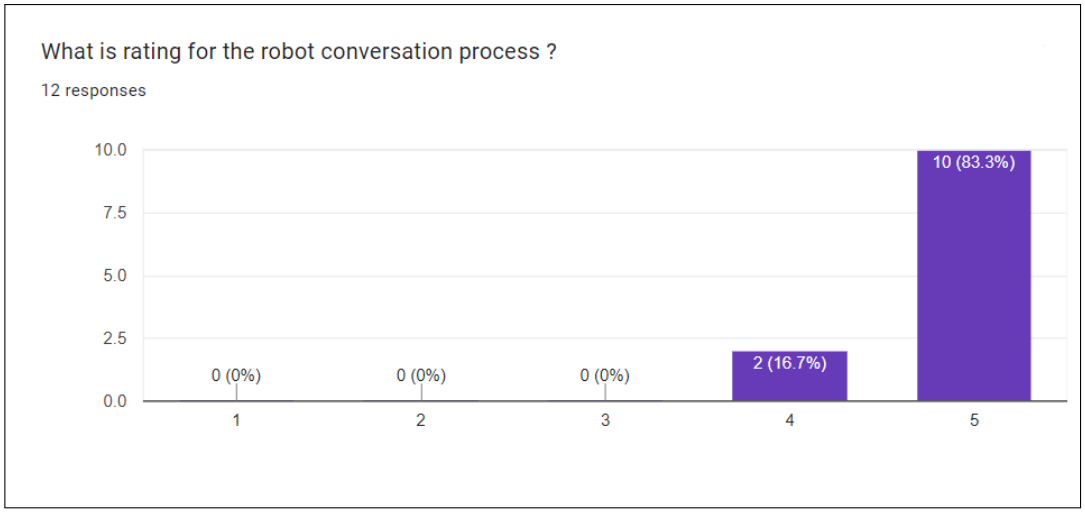

Figure 5.5: Results of rating robot conversation

5.4.5 Stress Reduces and Increases Workability

A scale ranging from 1 to 10 was provided, where 1 represented no stress and 10 indicated high stress. Out of the 12 participants, 1 individual rated their stress level as 1 (no stress), another participant rated it as 3, one participant rated it as 4, three participants rated it as 5, three participants rated it as 7, one participant rated it as 8, one participant rated it as 9, and only one participant rated their stress level as 10 (high stress). These results provide insights into the robot's ability to effectively detect stress levels. Figure 5.6 shows the results of stress prior to the procedure.

After the procedure, participants were asked to rate their stress levels on a scale of 1 to 10, where 1 indicated no stress and 10 indicated high stress. One participant rated their stress level as 1, indicating no stress. Three participants rated their stress level as 2, while three others rated it as 3. One participant rated their stress level as 4, and another participant rated it as 5. Two participants rated their stress level as 8, and one participant rated it as 9. No participants rated their stress level as 10, indicating high stress. These results suggest that participants experienced reduced stress and anxiety after engaging with the implemented companion robot. Figure 5.7 shows the results of stress levels after engaging with the implemented companion robot.

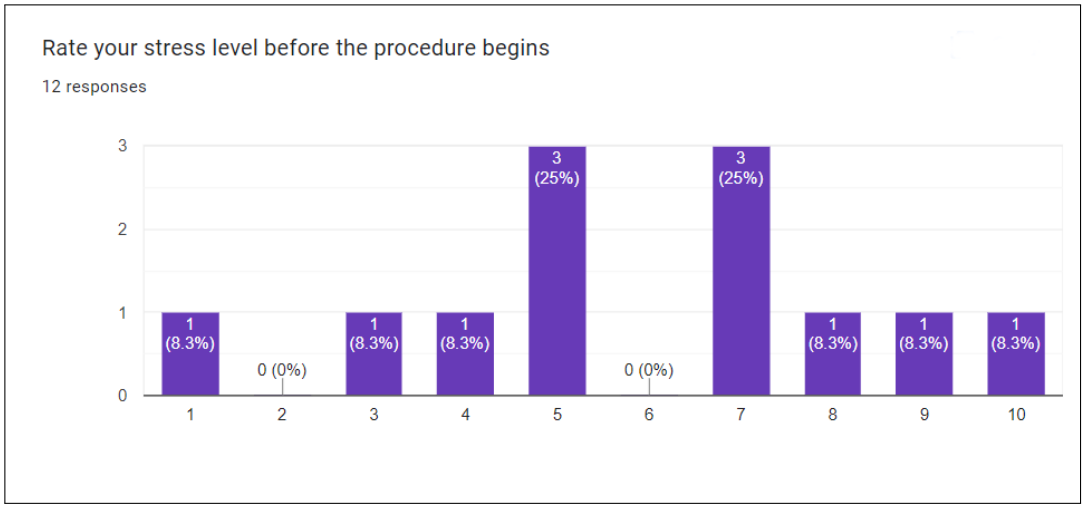

Figure 5.6: Result of stress prior to the procedure

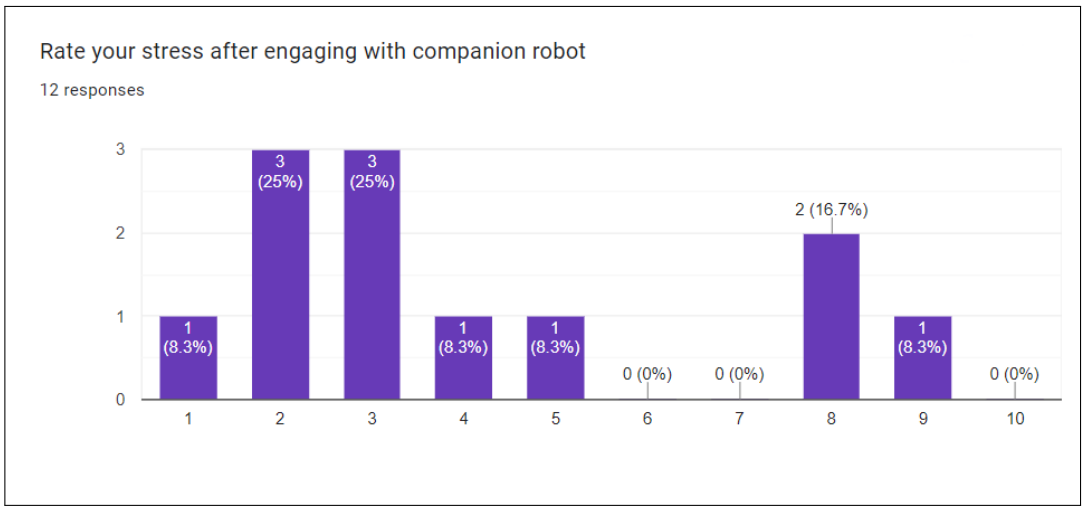

Figure 5.7: Result of stress after the procedure

5.4.6 Results on Efficiency and Accuracy

The efficiency and accuracy of the implemented system were evaluated from the users' perspective using three questions on the feedback form. The first question asked participants to rate how well the robot detected stress and provided suggestions to reduce it based on their preferences. The response options were: strongly disagree, disagree, neutral, agree, and strongly agree. The results showed that 75% of the participants agreed with the robot's stress detection and suggestions, while 25% strongly agreed. This indicates a generally positive perception of the system's efficiency and accuracy among the participants. Figure 5.8 shows the result of users' ratings of the efficiency and accuracy of implemented companion robot.

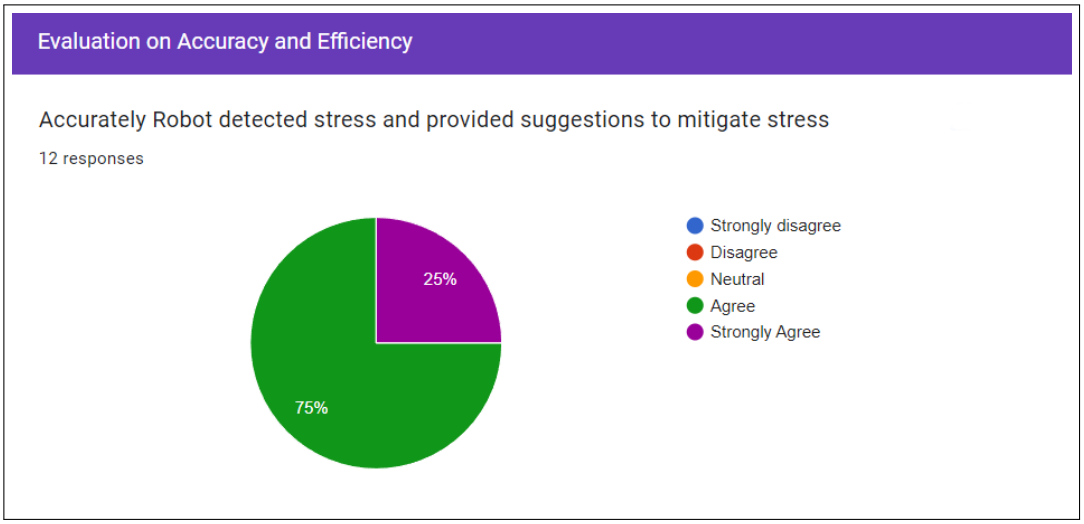

Figure 5.8: Result of Efficiency and Accuracy of implemented companion robot

The quality of information or suggestions provided by the robot was assessed by asking participants to rate it on a scale of 1 to 5, where 1 indicated poor quality and 5 indicated excellent quality. Out of the 12 participants, 1 participant rated it as 3, 7 participants rated it as 4, and 4 participants rated it as 5, indicating excellent quality. These results suggest that the majority of participants found the information and suggestions provided by the robot to be of high quality. Figure fig. 5.9 shows the result of The quality of information or suggestions provided by the robot

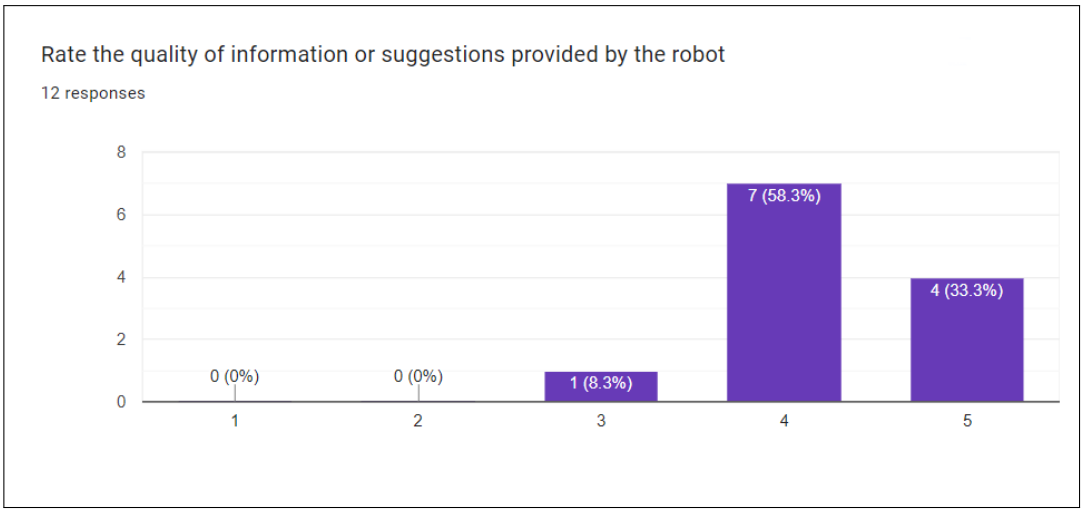

Figure 5.9: Result of The quality of information or suggestions provided by the robot

Finally, participants were asked to rate the effectiveness of the robot in reduc-

ing stress. They were asked to rate it on a scale of 1 to 5, where 1 indicated poor effectiveness and 5 indicated excellent effectiveness. Out of the 12 participants, 7 rated the robot's effectiveness as 4, while 5 participants rated it as 5, indicating excellent effectiveness in reducing stress. These results highlight the positive impact of the robot in effectively reducing participants' stress levels, further confirming its efficacy in providing stress reduction support. Figure fig. 5.10 shows the results of the effectiveness of the robot in reducing stress.

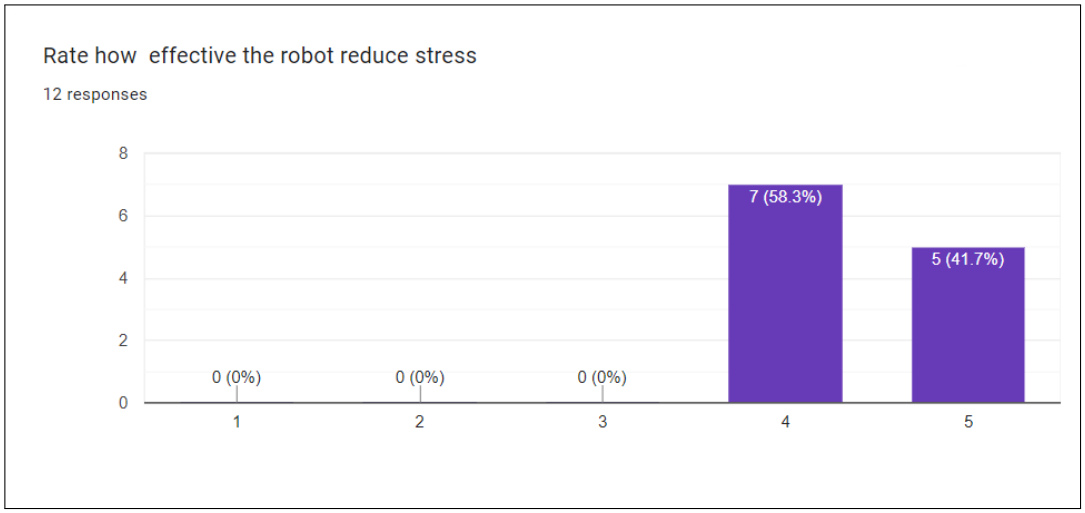

Figure 5.10: Result of The effectiveness of the robot in reducing stress

5.4.7 Analysis of Stress

To statistically determine whether stress and anxiety are reduced after engaging with the implemented companion robot for remote workers, a hypothesis test can be conducted. The relevant hypotheses are as follows:

H_0: participants do not experience reduced stress and anxiety after engaging with the implemented companion robot

H_1: participants experience reduced stress and anxiety after engaging with the implemented companion robot

To perform the above hypothesis, we can choose either a parametric test or a non-parametric test. The selection of the appropriate test depends on checking whether the population follows a normal distribution and considering the sample size that has been selected in here the sample size is 12. Here it is used 5% level of significant for all hypothesis tests. The Shapiro-Wilk normality test is commonly used to assess whether a given dataset follows a normal distribution or not. It is a statistical test that evaluates the null hypothesis that the data were drawn

from a normally distributed population. By applying the Shapiro-Wilk test, we can determine if the population data, in the context of the hypothesis, can be assumed to be normally distributed. This helps in making an informed decision about whether to use a parametric test (if the data are normally distributed) or a non-parametric test (if the data are not normally distributed) for the hypothesis testing. statistical analysis was done using the R programming language within the RStudio platform. R and RStudio provide a comprehensive environment for conducting statistical tests and analyzing data,

Shapiro-Wilk normality test

$$data : before\_stress\_level–after\_stress\_level \tag{5.1}$$

Table 5.3: after applying the Shapiro-Wilk normality test results were obtained as follows

W (Test Statistic)	**p-value**
0.86982	0.06501

Since the p-value 0.06501 is greater than 0.05, It demonstrates that the difference between the two populations is normally distributed. The sample size also less than 30 because we perform the entire analysis using 12 participants as the sample. Even though the two population difference is normally distributed but the sample size is considerably small (< 30) we should use a nonparametric test. Here data are collected, stress before engaging with the implemented companion robot and after engaging with the implemented companion robot from each participant. So here we have two dependent samples, each size 12. So to perform the above hypothesis using the non-parametric test which can take as the two sample dependent test is, Wilcoxon Matched- Pairs Signed Ranks Test.

The Wilcoxon Matched-Pairs Signed Ranks Test is a statistical test used to compare paired observations from the same sample or group. It is a non-parametric test and is typically used when the data do not meet the assumptions of a parametric test, such as when the data are not normally distributed or the sample size is small.

In the context of hypothesis testing, the Wilcoxon Matched-Pairs Signed Ranks Test is appropriate when we want to assess whether the one relate variable is significantly less than other related variable. It does not assume a specific distribution of the data and instead compares the ranks of the paired differences.

By applying the Wilcoxon Matched-Pairs Signed Ranks Test, we can determine if the stress and anxiety levels after significantly less than before engaging with the companion robot, supporting or rejecting the alternative hypothesis.

Wilcoxon Matched-Pairs Signed Ranks Test

$$data : before\_stress\_level–after\_stress\_level \tag{5.2}$$

Table 5.4: after applying the Wilcoxon Matched-Pairs Signed Ranks Test results were obtained as follows

V (Test Statistic)	**p-value**
55	0.002617

The alternative hypothesis: true location shift is not equal to 0.

Since p – value of the above test is 0.002617 less than 0.05 reject the null hypothesis (H_0) and conclude that the participants experience reduced stress and anxiety after engaging with the implemented companion robot at 5% level of significance.

Chapter 6

DISCUSSION

It is intended to discuss analyzed results and limitations under this chapter.

6.1 Analysed Results

Users have provided strongly positive feedback on the research done on the companion robot for stress reduction and workability enhancement. They have expressed satisfaction with the robot's success in reducing stress and increasing productivity at work. Users prefer a streamlined approach to information gathering where only the necessary details are requested, according to a notable finding from the research. This strategy has shown to be more successful than requesting a lot of unnecessary information from users. It is advised to keep using this process and give priority to workflow optimization in order to further improve the user experience. Despite the positive reviews, there is still room for improvement as a significant number of users gave it a rating of 4, suggesting possible improvements.

A significant number of users agreed that the robot's design should be revised to increase user attraction and create a more friendly first impression, so this is an important factor that should be taken into account. It has been determined that modifying the design of the robot's eyes is one potential way to enhance the user's perception of friendliness and approachability.

According to the research's findings, the majority of users strongly agreed that the implemented system's conversation flow was understandable and well-planned. This positive feedback offers insightful information for future improvement. It is advised to improve the conversation's flow in order to make the conversations more interesting and appealing to users. The distribution of user

feedback should be considered during this optimization process to ensure that the issues that come up most frequently are addressed.

Importantly, the research revealed that engaging with the companion robot resulted in a significant reduction in anxiety and stress levels for the majority of users. This outcome underscores the positive impact of the robot's presence in promoting emotional well-being. It further highlights the potential of the companion robot to effectively alleviate stress and contribute to users' overall mental health.

6.2 Results on Efficiency and Accuracy of the Implemented System

When results on speech misdetections in the implemented system are taken into account, it becomes clear that the system has improved speech detection accuracy. But it was found that the majority of speech misdetections happened because it took too long to get user input after the robot had a conversation. It was further noted that user delays were based on how well the system performed.

The analysis of facial expressions exhibited by participants yielded interesting results.The majority of users exhibited neutral facial expressions while interacting with the robot, suggesting that it did not cause any distractions. Furthermore, a few users showed happy emotions, indicating a positive user experience and a positive effect of the robot.

When considering user feedback, the majority of participants in the experiment expressed agreement with the system's ability to accurately detect user stress levels. However, the research also identified areas for improvement based on the majority of user suggestions.

In conclusion, research on incorrect speech recognition and facial expression analysis offers useful suggestions for improving the system that has been put in place. Although the increased speech detection accuracy is a good development, user input delays should be addressed. The robot does not interfere with user engagement, according to the positive results of facial expression analysis. The performance of the system can be further improved by taking into account user feedback and responding to suggestions. The ultimate goal is to develop a system that accurately measures user stress levels and reacts to them while offering a seamless and enjoyable user experience to reduce stress.

6.3 Limitations

The research has certain limitations, including a lack of focus on enhancing stress detection methods, design components, and conversational workflow. A thorough understanding of the system's performance in this area was hindered by the study's incomplete exploration of sophisticated techniques for precisely detecting and assessing user stress levels.

Additionally, the research did not offer in-depth explanations or suggestions for enhancing the companion robot's design, such as by making it more aesthetically pleasing or user-friendly. Additionally, there was little discussion of how to make conversational workflows more engaging and satisfying. A companion robot that is more user-centric and effective at reducing stress and improving workability would benefit from addressing these limitations.