Microsurgery Study

Ashna Shah, Gustavo Arismendi , Shivani Arora
COSC 6323 - Statistical Methods in Research

May 4, 2018

A. INTRODUCTION

Microsurgery Study is a study about relationship of sympathetic arousal and skill in learning micro-surgical tasks. In this study 21 medical students participated after their consent. The subjects had to take part in 5 sessions of 1 hour each. In these sessions subjects performed 2 tasks on inanimate simulator. Theses tasks were cutting and suturing. In their first visit the subjects completed a biographic questionnaire, and a trait anxiety inventory. After each session they completed a post-study questionnaire i.e NASA-TLX questionnaire. The NASA-TLX instrument measures perceptions of subjects regarding difficulty of tasks. This is done using a sub scale which ranges from 0-20.

In each session, the subjects underwent the following treatments:

(i) Baseline: After listening to spa music for 5 minutes, the subjects were facially recorded using thermal and visual camera.
(ii) Cutting: In the inanimate simulator the subjects had to perform precision cutting . During this process they were facially recorded using thermal and visual camera.
(iii) Suturing: In the inanimate simulator the subjects had to perform suturing . During this process hey were facially recorded using thermal and visual camera.

In this project we have done statistical analysis and inference on the data of all the subjects who performed in this study. We have done this by doing Quality Analysis, Inference of Plots, Cumulative Plots, Linear Modeling and Random Effects on the data. We have based our analysis using different plots based on subjects, tasks and sessions.

B. SUMMARY OF PLOTS

1. Biographic Data

There are total 21 subjects who participated in this study. But out of 21 subjects, 15 subjects participated in all the 5 sessions of the experiment. For our analysis of biographic data we have considered 15 subjects. These subjects are medical students.

In order to find out the distribution of males and females among all the subjects, we have made a barplot to analyze this distribution, which is shown in Figure-1.

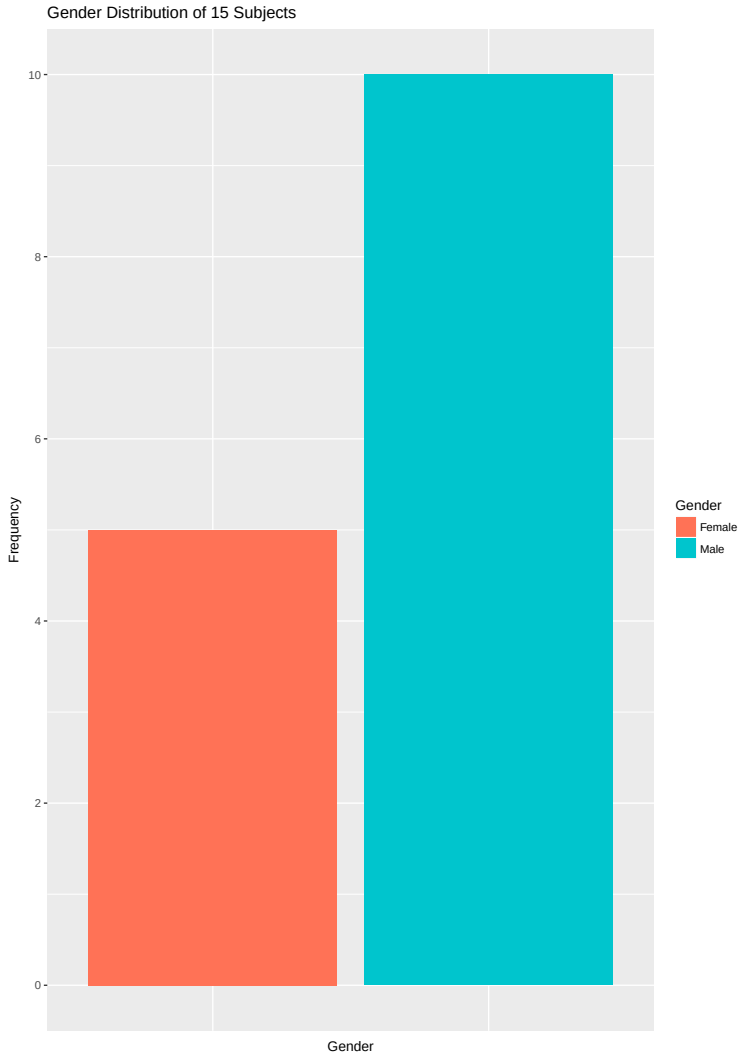


Figure 1: Barplot for Gender Distribution

From Figure-1 we can state that there are total 10 Males and 5 Female subjects. Clearly, there are more male subjects than female subjects. Thus, we can say that in com-

parison with females, male students are more willing to participate in medical experimental research as in our case is Microsurgery Study.

In order to find out the distribution of age among all the subjects, we have made a histogram to analyze this distribution, which is shown in Figure-2.

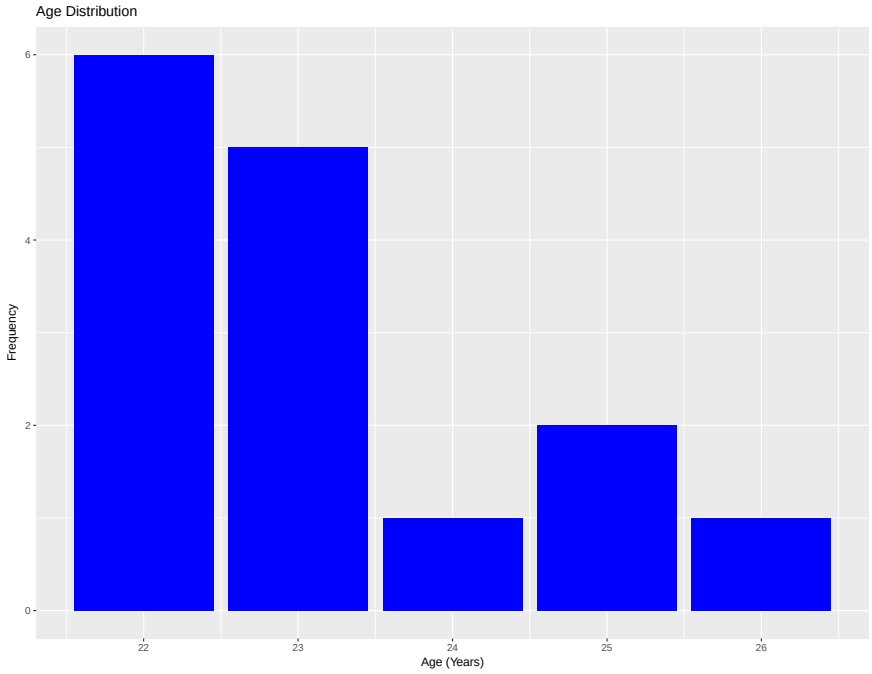


Figure 2: Histogram for Age Distribution

From this plot we can say that there are 6 subjects who are 22 years old, 5 subjects who are 23 years old, 2 subjects who are 25 years old and 1 subject each of the age 24 and 26. Thus, from this distribution we can say that as the subjects are medical students, we have majority of students in the age group of 22-24. As usually, students pursuing medical program are in this age range. For subjects who are 25, 26 and 27 years old we can say that they might have taken a gap before enrolling in medical program or might be part-time job students. The age gap between the eldest and the youngest subject is 5 years (27-22).

2. Trait Psychometric Data

TAI score takes values in the range 20-80, with scores up to low 40s considered normal, while higher scores considered indicative of overanxious individuals.
In order to analyze the TAI distribution received by all the subjects we have made a barplot which is shown in Figure 3.

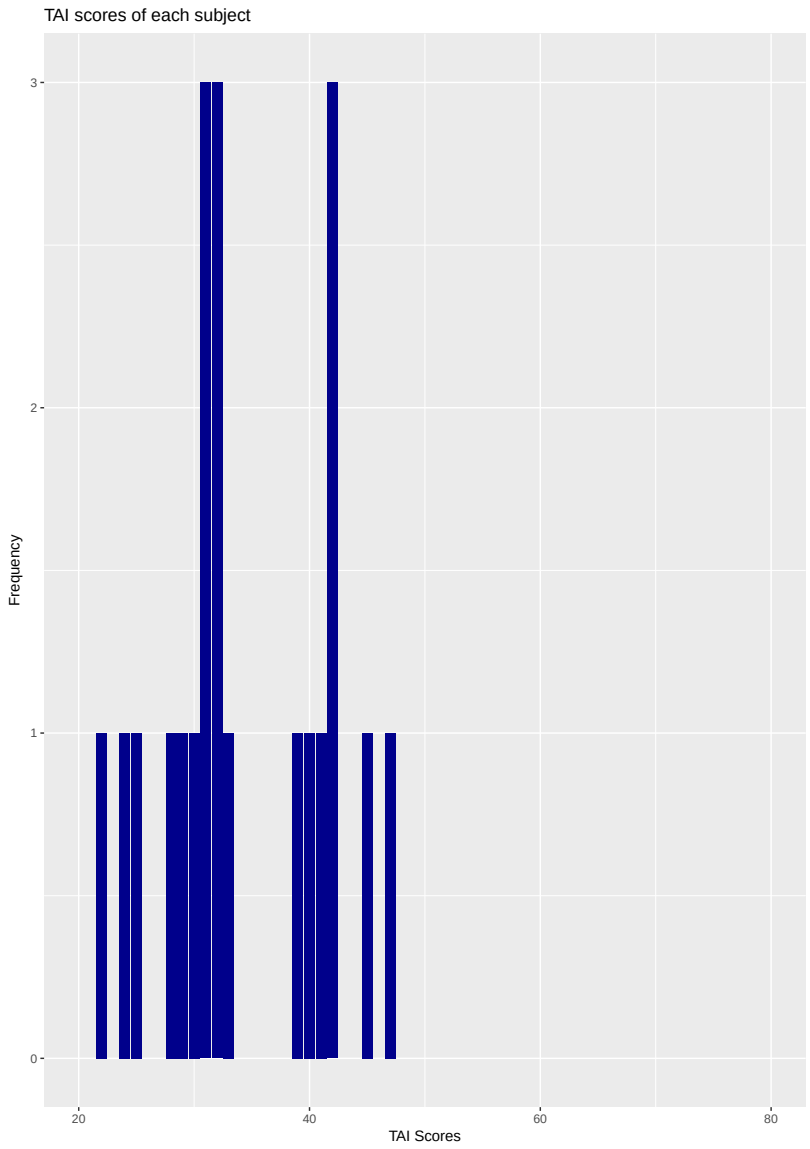


Figure 3: TAI scores of each subject

From this plot we can say that out of 21 subjects, 15 subjects have score count lower than 40. This signifies that 15 subjects are in the normal state of mind. They seem to be not stresses out or nervous for the experiment in which they participated.

While remaining 6 subjects are anxious and stressed out. They do not seem to be confident about their performance and seem to be little worried.

3. State Psychometric Data

In the experiment, after cutting and suturing tasks were completed by the subjects, the subjects filled out NASA-TLX questionnaire. The NASA-TLX instrument features 6 parameters of perceptions of subjects regarding task difficulty. These sub scales are scored in the range 1-20. The 6 parameters are-:

In order to analyze the state psychometric data of all the subjects for 6 factors we have made bar plots. Theses bar plots are shown in Figure 4,5 and 6-:

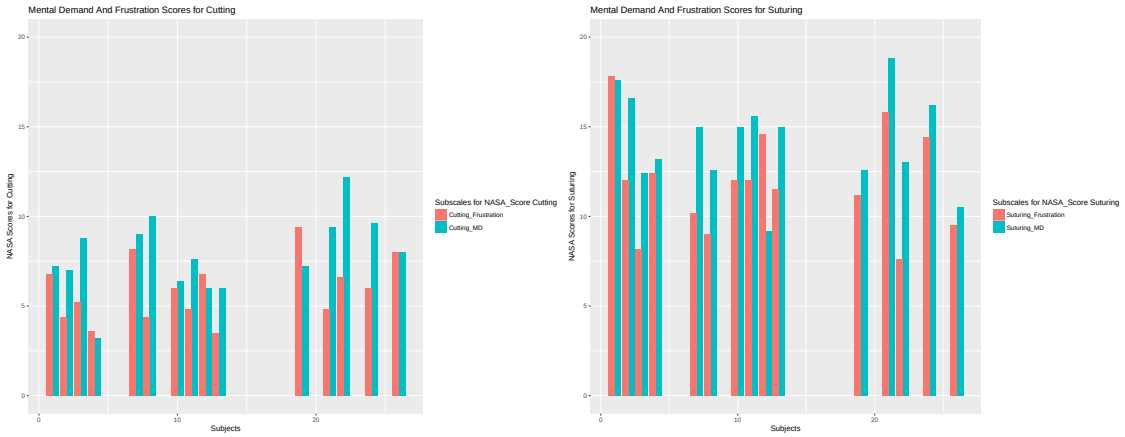


Figure 4: Mental Demand and Frustration NASA - TLX for Cutting and Suturing

In Figure 4 we can clearly say that Mental Demand and Frustration were much more during the task of Suturing than Cutting.
So we can say that cutting task was easy and simple, as mental activity and stressed level was very low.
While suturing task was complex and demanding,as stress activity and perceptual activity was very high.

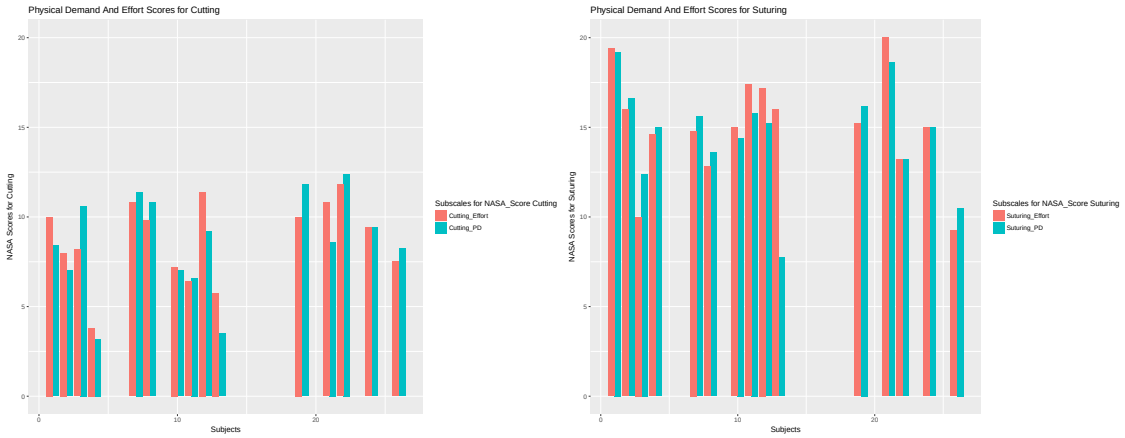


Figure 5: Physical Demand and Effort NASA - TLX for Cutting and Suturing

In Figure 5 we can clearly say that Physical Demand and Effort was much more during the task of Suturing than Cutting. So we can say that cutting task was less demanding, as it required less physical activity and effort.
While suturing task was strenuous, as it required much more physical activity and effort than cutting.

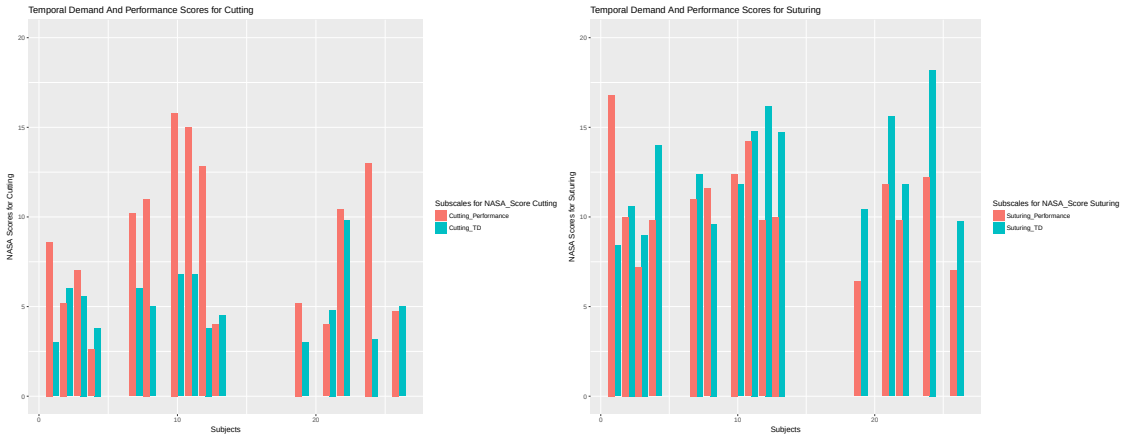


Figure 6: Temporal Demand and Performance NASA - TLX for Cutting and Suturing

In Figure 6 we can say that Temporal Demand and Performance both are high for Suturing.
In Cutting task, performance is low which is not good sign of learning. Also Temporal Demand is low, which means the subjects felt less pressure during the task.
But in Suturing, we find that all though it was overall demanding the performance was high than Cutting.
Also, Temporal Demand is high which meand they felt much high pressure and also that the pace was slow.

(4) Perinasal Perspiration (Stress) Signal

We have performed the mean Perinasal Perspiration (Stress) Signal of all the subjects.

(i) Average Perspiration (stress level in degrees Celsius squared) for 15 subjects in each session for Cutting-:

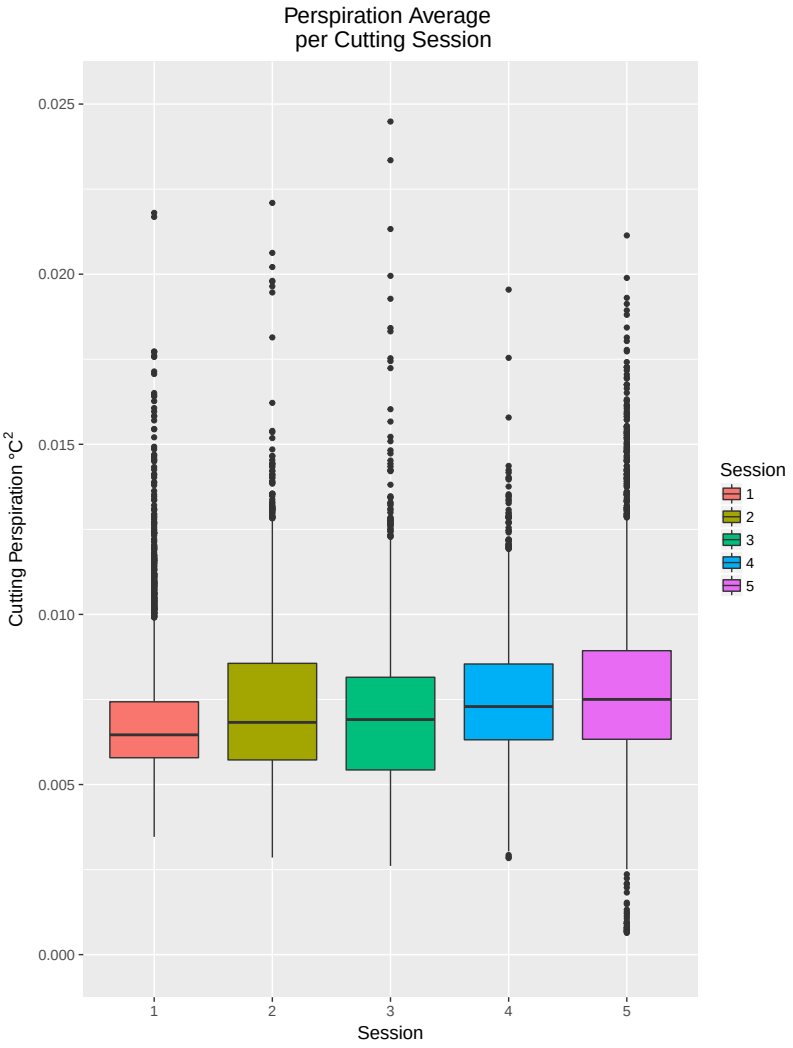


Figure 7: Average Perspiration Per Session - Cutting

From Figure 7 we can say that beginning with the perspiration average per cutting session graph we can observe a small increase in the perspiration median. The subjects had a session one value close to 0.0062 degrees Celsius, which steadily increased to a median of 0.0072 by session five. This tells us that the subjects had on average a higher perspiration levels for the cutting task in the last session than they did in the initial session. We can interpret this as subjects feeling higher levels of stress during the final session compared to other sessions.

(ii) Zoomed-In Plot for Average Perspiration (stress level in degrees Celsius squared) for 15 subjects in each session for Cutting-:

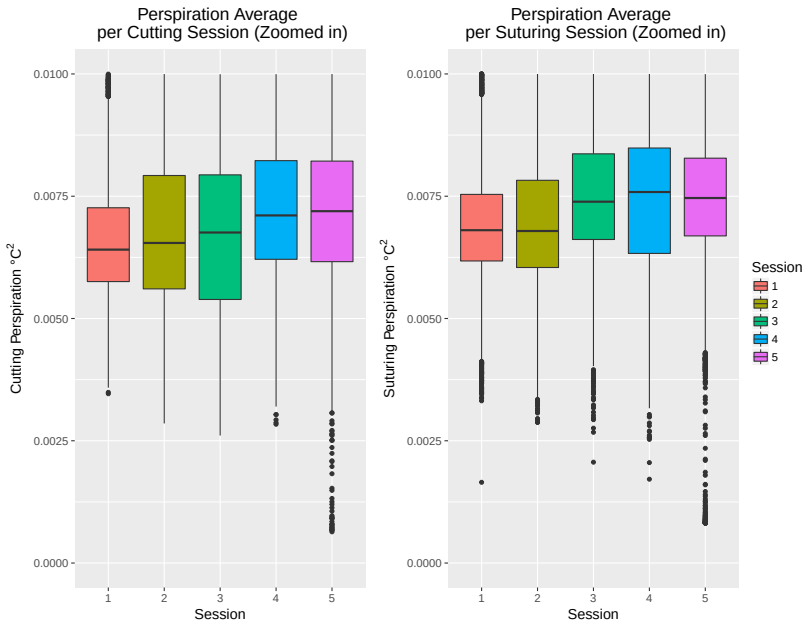


Figure 8: Zoomed-In Plot for Average Perspiration Per Session - Cutting

From Figure 8 we can say that the interquartile range shows a compact concentration of perspiration levels during the first session (0.0059 to 0.0072), an almost doubling in size during the second session (0.0057 to 0.0077) with the third quartile seeing the largest increase, the third session shows a box size (0.0055 to 0.0077) slightly similar to the second session but this time most of the values are in the second quartile, the fourth session box displays a smaller (0.0062 to 0.0078)and well balanced box size, and finally the last session reveals the third largest box size (0.0061 to 0.0078) for all the sessions with a median located in almost exactly halfway in the box (0.0071). The session box sizes overall divulge fluctuations in the second and third quartile of values halfway into the study and conclude the last sessions with symmetrical quartiles. The box sizes indicate that the subjects had increasing levels of perinasal perspiration throughout the study.

The box plot whiskers show an increasing pattern in the fourth quartile values (0.009 to 0.0013) and a decreasing pattern in the first quartile values of perspiration (0.0035 to 0.0025) during the five sessions. The outliers go from a significant number of high value outliers in the first session (with zero low end outliers) to a different distribution of outliers in the last session (a good number of outliers in the low end side). The interpretations we can make from these observations is that although the nominal levels of perinasal perspiration had an increase for most subjects, there was a number of small subjects that had a significant decrease in perinasal perspiration at the end of the five session study.

Overall the interpretation of the perspiration average per cutting session indicate an increase in median levels of stress by most subjects, and a decrease in stress levels for a small number of subjects. The box plots also show an overall increased level of disagreement between subjects (as show by the differences in second and third quartile sizes) from the first session to the last session.

(iii) Average Perspiration (stress level in degrees Celsius squared) for 15 subjects in each session for Suturing-:

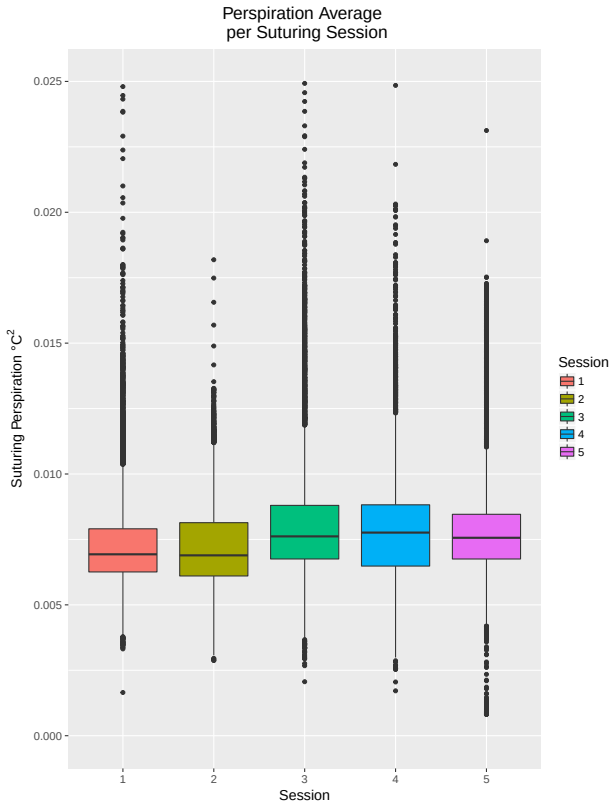


Figure 9: Average Perspiration Per Session - Suturing

From Figure 9 we can say that the median values show an increase of perinasal perspiration from session one (0.0065) to session five (.0075). This tells us that the subjects had higher levels of stress at the end of the study that they did at the beginning.

The interquartile range as shown by the box size displays an increasing concentration of values towards higher levels of perspiration with fluctuating ranges in the middle sessions (2, 3, and 4). In the first session the box interquartile range shows most of the subject values concentration between 0.0062 and 0.0075 while the fifth session reveals the subject values concentrating between 0.0064 and 0.0080. This increase tells us that overall subjects were more stressed in the fifth session than they were in the initial session of the study.

The boxplot whiskers reveal an increase in both the first and fourth quartile values for the suturing task. Overall the first quartile values increased from 0.0033 in session one to 0.004 in session five. The fourth quartile values increased from 0.0101 in the first session to 0.0105 in the fifth session. Another important fact that we can infer from the graph is the

change in the distribution of the number of outliers. In the first session there was a significant majority in the high end of the perspiration values. In the final session of the study there were close to one-third of all the outlier values in the lower end of the perspiration levels. These observations help us see that although there was an overall increase in the levels of perinasal perspiration, a significant number of subjects had lower levels of perspiration in the fifth session when compared to the first session.

Overall the interpretation of the perspiration average per suturing session indicate an increase in median levels of stress by most subjects, and a decrease in stress levels for a small number of subjects after the conclusion of the five sessions. The box plots also display an overall slight level of disagreement between subjects (as show by the differences in second and third quartile sizes) from the first session to the last session.

(iv) Comparison of Cutting and Suturing Task Perinasal Perspiration Results-:

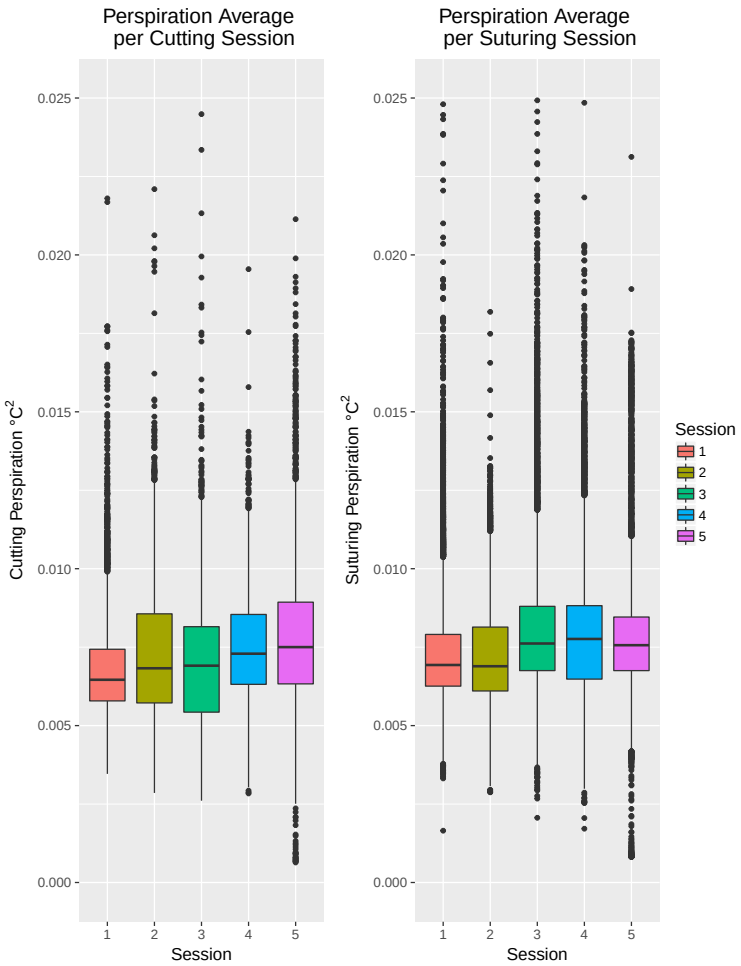


Figure 10: Average Perinasal Perspiration Per Session : Cutting Vs Suturing Task

From Figure 10 we can say that when comparing the levels of perinasal perspiration between the cutting and suturing tasks we can conclude that the majority of subjects found the suturing task more stressful than the cutting task as reflected by the median (0.0075 for suturing versus a 0.0068 for cutting). An interesting observation is that although the subjects found the suturing task more stressful overall, there seems to have been a significant shift from number of suturing high level outliers at the beginning of the study (first session) to lower valued outliers at the end of the study. A similar conclusion could be reach in the cutting tasks. These results help us conclude that overall stress levels increased throughout the study; that subjects found the suturing task more stressful than the cutting task; and that in both tasks there were a small group of subjects that was less stressed at the end of the study that they were at the beginning.

5. Performance Data

In order to analyze the time taken by the subjects to perform cutting and suturing tasks, we have drawn the time barplots for each subject. We have used Green color for Cutting and Red for Suturing. These plots are shown in the Figure 39 and Figure 40.

(a) Overall Cumulative Plots

(i) Time

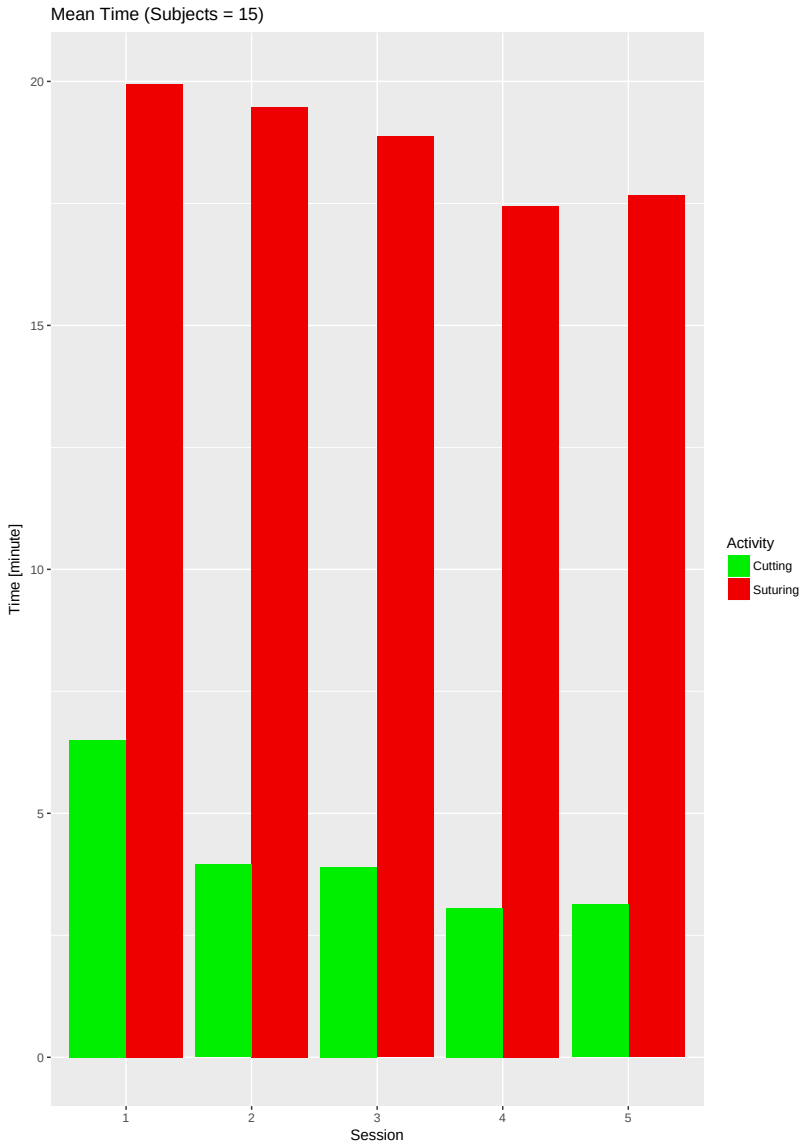


Figure 11: Average Time Taken by 15 subjects in each session for Cutting and Suturing

This figure shows the the total time taken by all 15 subjects for cutting and suturing in each session.

(i.1) Cutting

From this plot we can say that the performance of all the subjects improved after session 1. The cutting time decreased as they move from session 1 to session 5. The cutting time remains almost the same for session 2,3 and 4,5.
From this plot overall we can say that the subjects became more comfortable with the experiment and learned with experience. May be they were stressed or anxious or not confident in the beginning in session 1 that is why they took more time. But with each visit they improved on the skill and were able to finish the task faster.

(i.2) Suturing

The performance of the subjects improved after session 1. The suturing time decreased as they move from session 1 to session 5. The suturing time remains in 5th increased very slightly from 4th session. May be the subjects got slightly nervous as they were finishing the study. May be they were worried about the results or stressed out during the 1st session.
From this plot overall we can say that the subjects became more comfortable with the experiment and learned with experience. May be they were stressed or anxious or not confident in the beginning in session 1 that is why they took more time. But with each visit they improved on the skill and were able to finish the task faster.

(i.3) Comparison of Suturing and Cutting Time

From this plot we can clearly say that time taken for suturing is way more than time taken for cutting. For both the tasks we see a downward trend in time from session 1 to 5.
Overall, we can say that subjects found suturing task more difficult that cutting that is why they took my time in doing suturing than cutting. May be the subjects had done suturing before that is why they found it difficult. While they might have done cutting before that is why they were able to finish the task quickly.
The subjects are performing better as they move from session 1 to session 5.

(ii) Number of Sutures

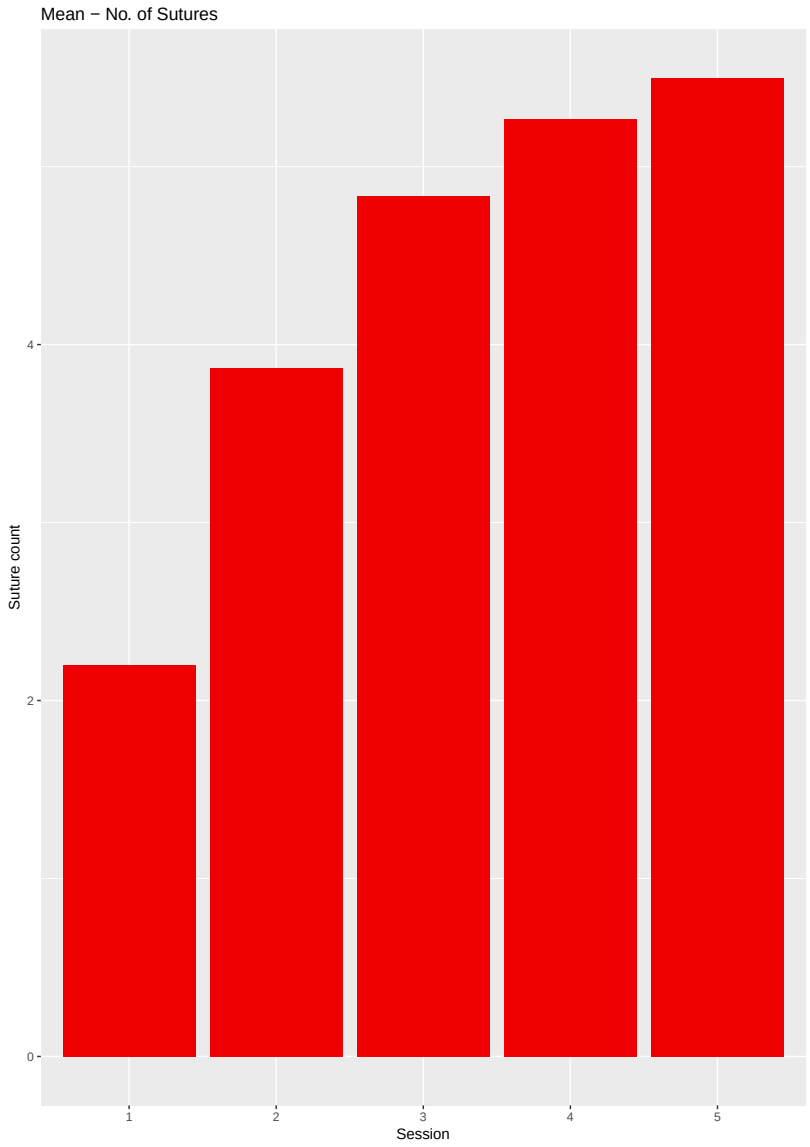


Figure 12: Average Number of Sutures in each session for 15 subjects

This figure shows the the total number of sutures by all 15 subjects in each session. From this plot we ca say that the performance of all the subjects improved as they moved from session 1 to 5. We can say after seeing the increase in number of sutures from session 1 to 5.
Also, we can say that the subjects might be anxious in the first session that is why they had very low number of sutures. But as they became more comfortable with the experiment and learned the skill with experience, they were able to perform better. With each session the subjects became more efficient in theor suturing task.

(iii) Score

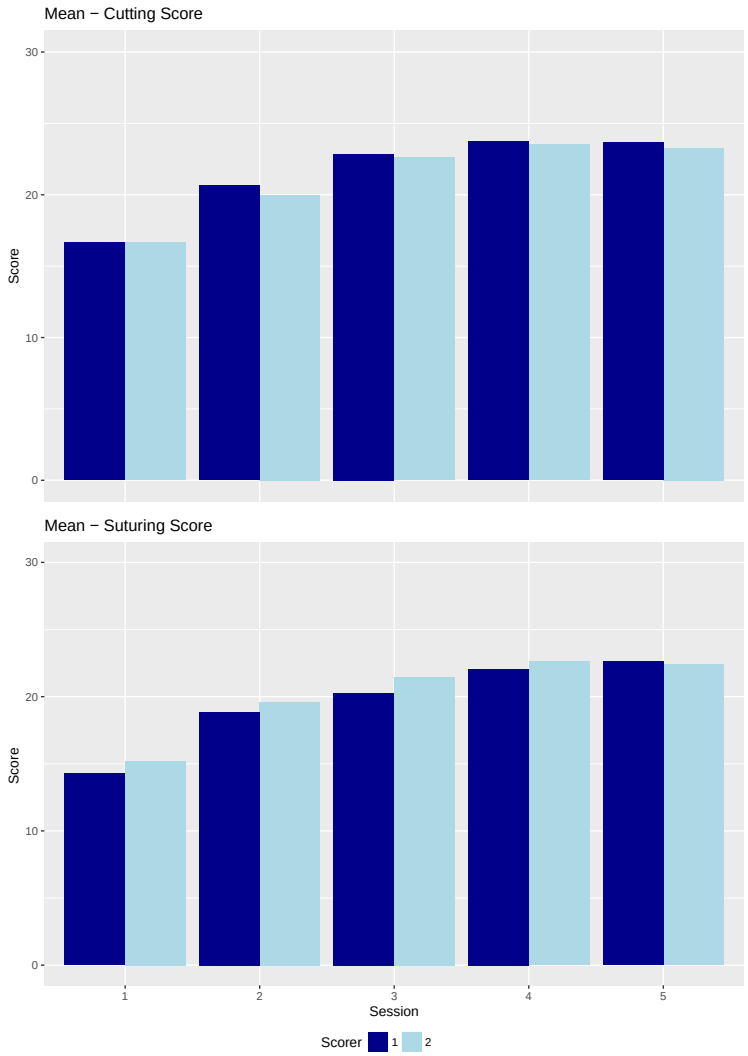


Figure 13: Average Score for 15 subjects in each session for Cutting and Suturing

From this plot we can see Score for all 15 subjects for 5 sessions.

(iii.1) Cutting

From this plot we an say that the scores of all the subjects increased as they moved from one session to fifth session. There is upward trend in scores. for session 4 and 5 the scores are almost constant.
Also, that for all 5 sessions, scorer-1 seems to give slightly more marks than scorer-2. Though there is not much difference between the scores given by both the scorers.

Overall, we can say that the subjects might be anxious in the first session that is why they had very low score. But as they became more comfortable with the experiment and learned the skill with experience, they were able to perform better. With each session the subjects became more efficient in task and thus got better scores.

(iii.2) Suturing

From this plot we an say that the scores of all the subjects increased as they moved from one session to fifth session. There is upward trend in scores.
Also, that for almost all 5 sessions, scorer-2 seems to give slightly more marks than scorer-1. Though there is not much difference between the scores given by both the scorers.
Overall, we can say that the subjects might be anxious in the first session that is why they had very low score. But as they became more comfortable with the experiment and learned the skill with experience, they were able to perform better. With each session the subjects became more efficient in task and thus got better scores.

(iii.3) Comparison of Cutting and Suturing Score

From this plot we can say that the overall scores of cutting are more than scores of suturing as give by both the scorers.
Thus, we can say that the subjects found cutting easier than suturing. That is why they were able to perform better in it and thus get better scores in each session for cutting. May be the subjects were new to suturing or were nervous to do it that is why they got lower suturing scores.

(b) Task Based Plots

(i) Time

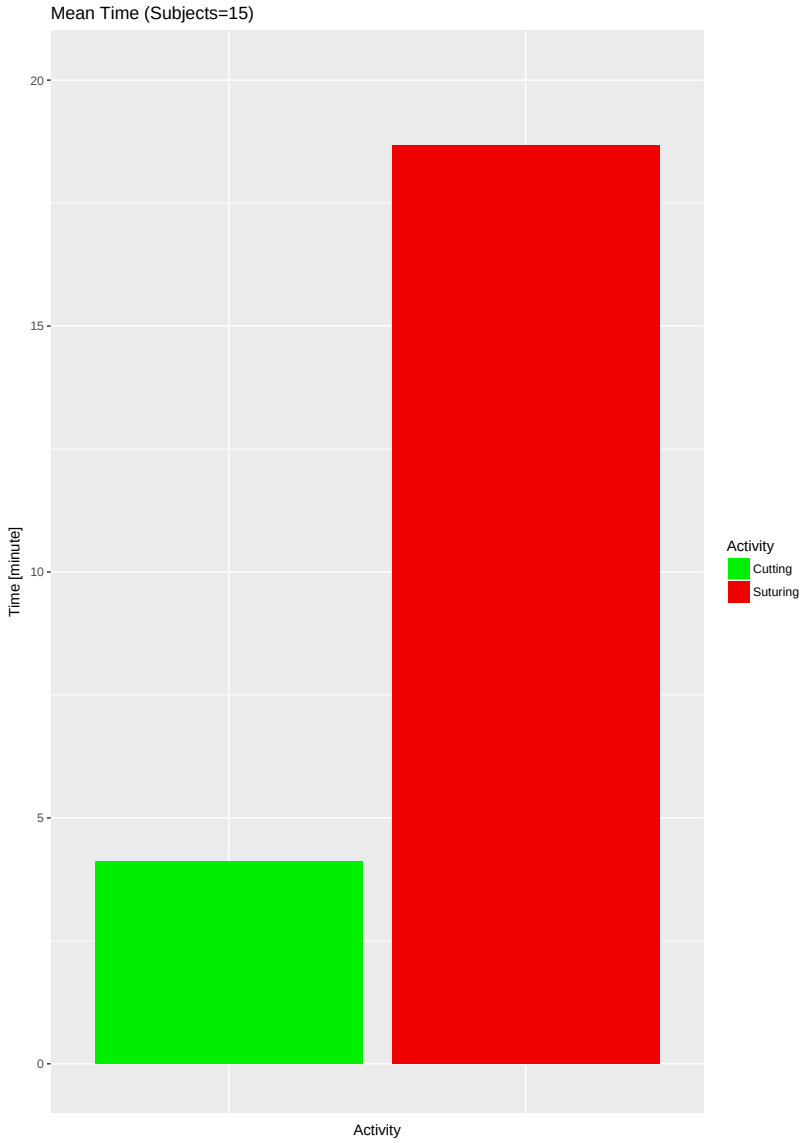


Figure 14: Average Time Taken by 15 subjects for Cutting and Suturing

From this plot we can clearly say that time taken for suturing is way more than time taken for cutting. Thus, we can say that subjects found suturing task more difficult than cutting.
May be the subjects had not done suturing before that is why they found it difficult, while they might have done cutting before that is why they were able to finish the task quickly. Or may be there are small points to be careful about in suturing and thus suturing task does consume more task.

(ii) Score

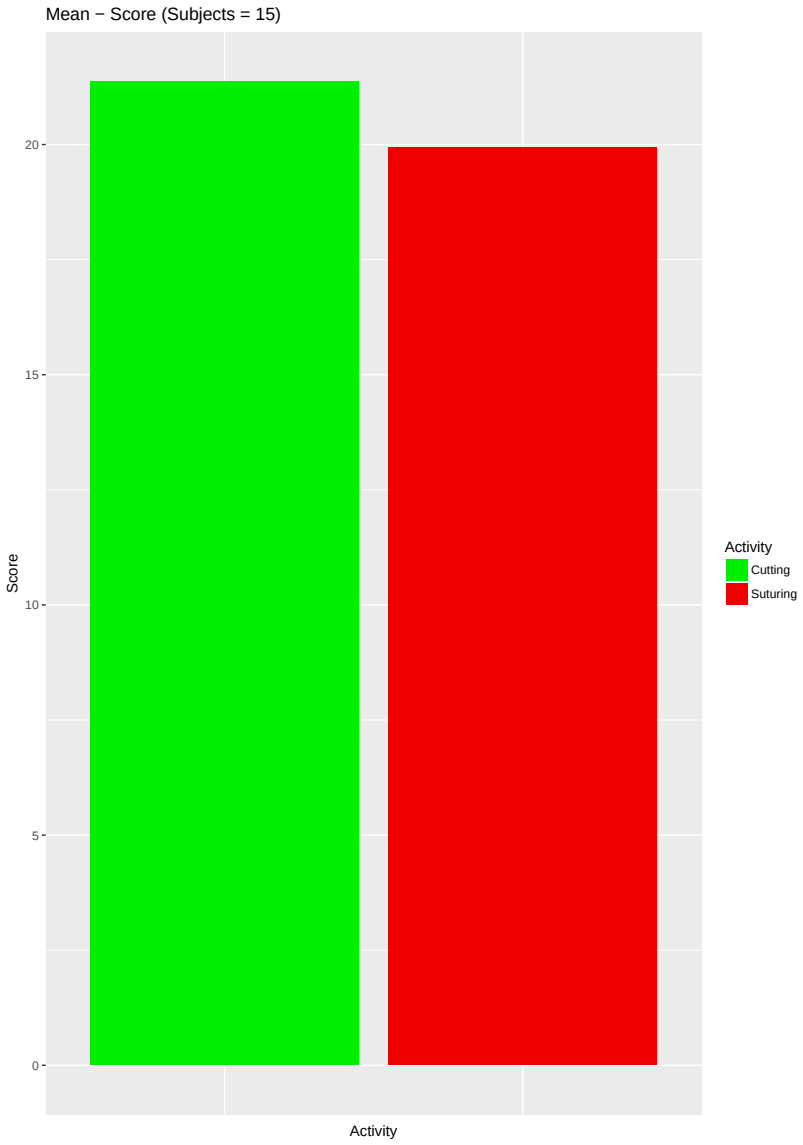


Figure 15: Average Score for 15 subjects for Cutting and Suturing

From this plot we can clearly say that score received by 15 subjects for cutting is more than the score received for suturing.
Thus, we can say that the subjects found cutting easier than suturing. That is why they were able to perform better in it and thus get better scores in it. May be the subjects were new to suturing or were nervous while performing the task and thus were not able to perform well. Thus, we can say that the subjects found cutting easier than suturing.

(c) Gender Based Plots

(i) Time

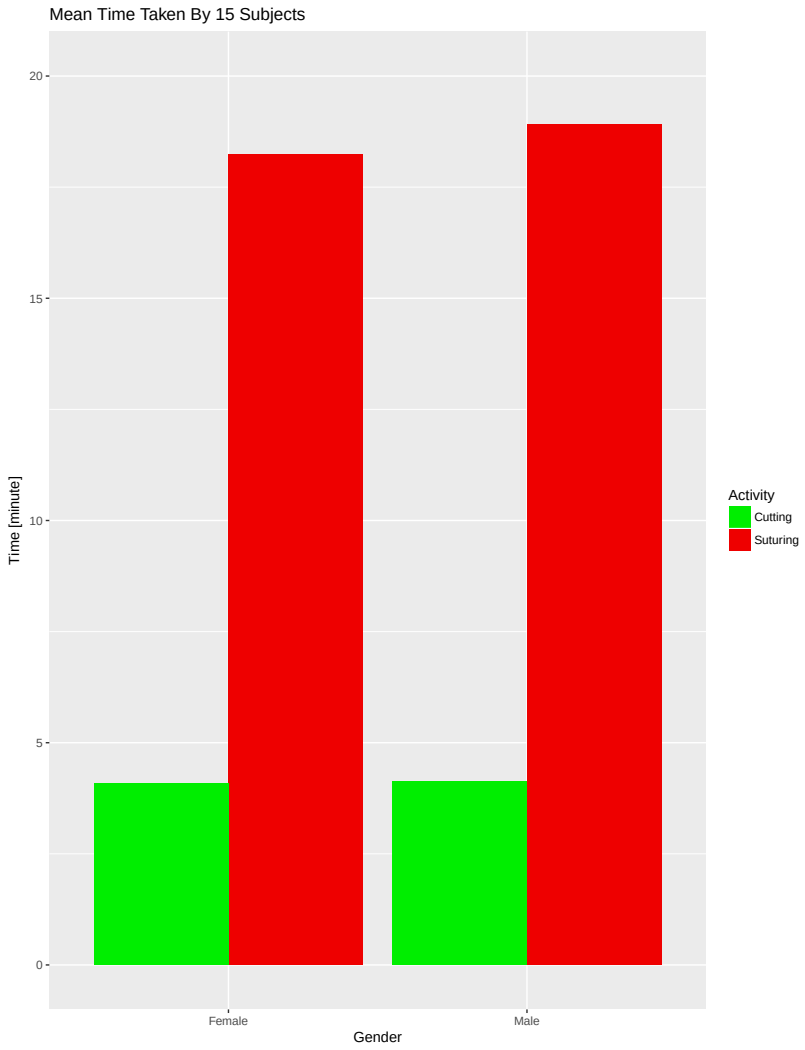


Figure 16: Average Cutting and Suturing Time : Male Vs Female

(i.1) Cutting

From this we can say that for both males and females, the time taken for cutting remains the same. They might have done cutting before that is why they were able to finish the task quickly. Thus, we can say that both male and female found cutting easier that is why they were able to complete the task earlier.

(i.2) Suturing

From this we can say that performance of Females is better than of Males as the suturing time for females is less than the suturing time for males. May be those females were quick learners and thus were able to finish the task earlier. May be the males were more careful

about small things and were taking it slow throughout the task that is why they took more time to finish the task.

(1.3) Comparison of Suturing Vs Cutting
Overall, we can say that females performed better than males as suturing time of females is lesser than of males. May be those females were quick learners and thus were able to finish the task earlier. May be in comparison with males the females were not worried and found suturing easier and thus were able to perform well in it.

(ii) Number of Sutures

Figure 17: Average Number of Sutures : Male Vs Female

From this plot we can clearly say that females performed better than males as the number of sutures is more than males. As seen in the time taken for suturing, females took lesser time and thus we can say they were relaxed and were good at this job. That is why they were able to perform well by taking out more number of sutures.

(iii)

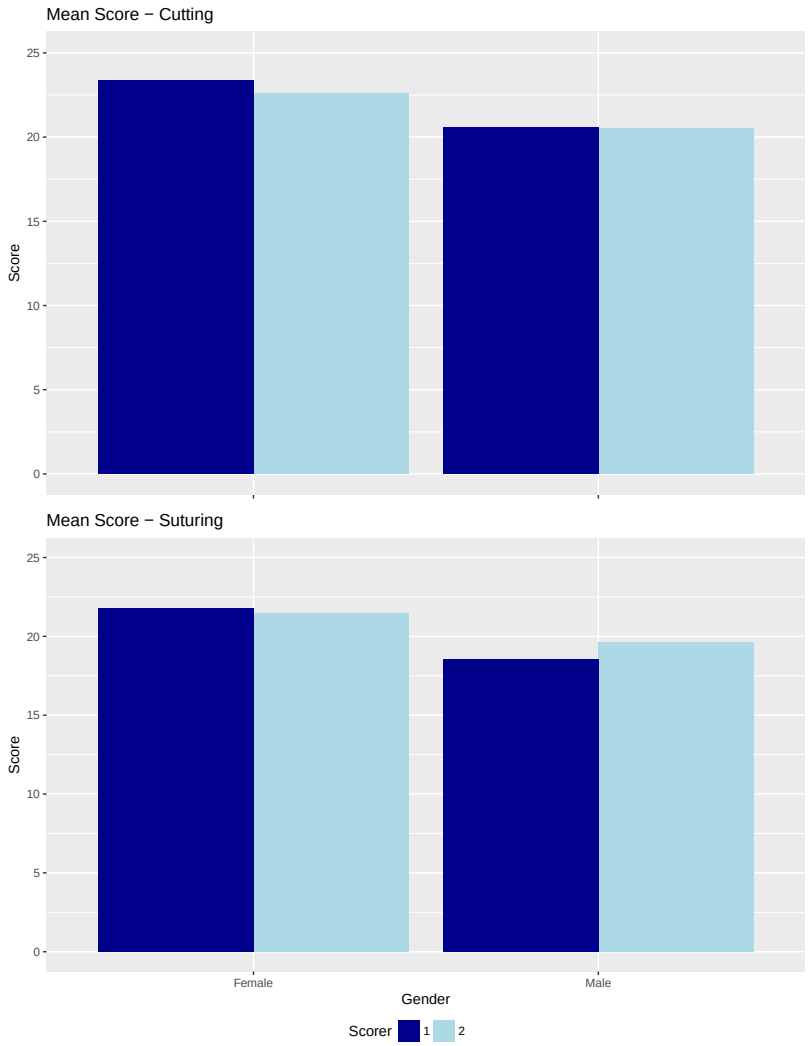


Figure 18: Average Score for Cutting and Suturing: Male Vs Female

(iii.1) Cutting

From this plot we can say that performance of females is better than males as the scores received by them from both the scorers are more than males.

Scores received from Scorer-1 are slightly more than Scorer-2 for females whereas for males they both receive the same scores.

(iii.2) Suturing

From tis plot we can say that performance of females is better than males as the scores received by them from both the scorers is more than males. The females took lesser time than males and took out more number of sutures thus we can say that females got higher scores than males.

Also, we can see that males received slightly more scores from Scorer-2 than Scorer-1 whereas females received slightly more scores from scorer1 than scorer2.

(d) Session Based Plots

(i) Time

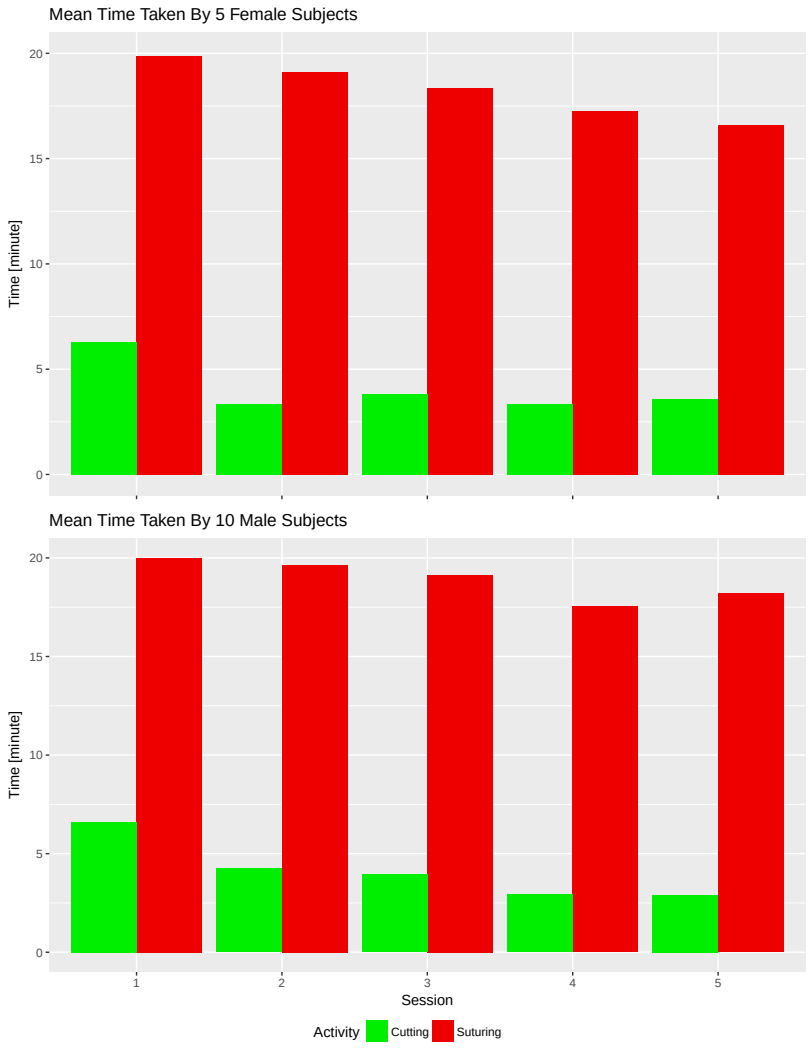


Figure 19: Average Cutting and Suturing Time for each session : Male Vs Female

(i.1) Cutting

Overall, almost for all sessions, cutting Time for Females is slightly more than the time taken by males. May be they were careful about small things and taking it easy that is why they took more time to perform the task in comparison with males.

(i.2) Suturing

Suturing Time for Females is slightly less than the time taken by males in each session. Thus we can say that performance of Females is better than of Males. May be those females were quick learners and thus were able to finish the task earlier. May be the males were more careful about small things and were taking it slow throughout the task that is why they took more time to finish the task.

(ii) Number of Sutures

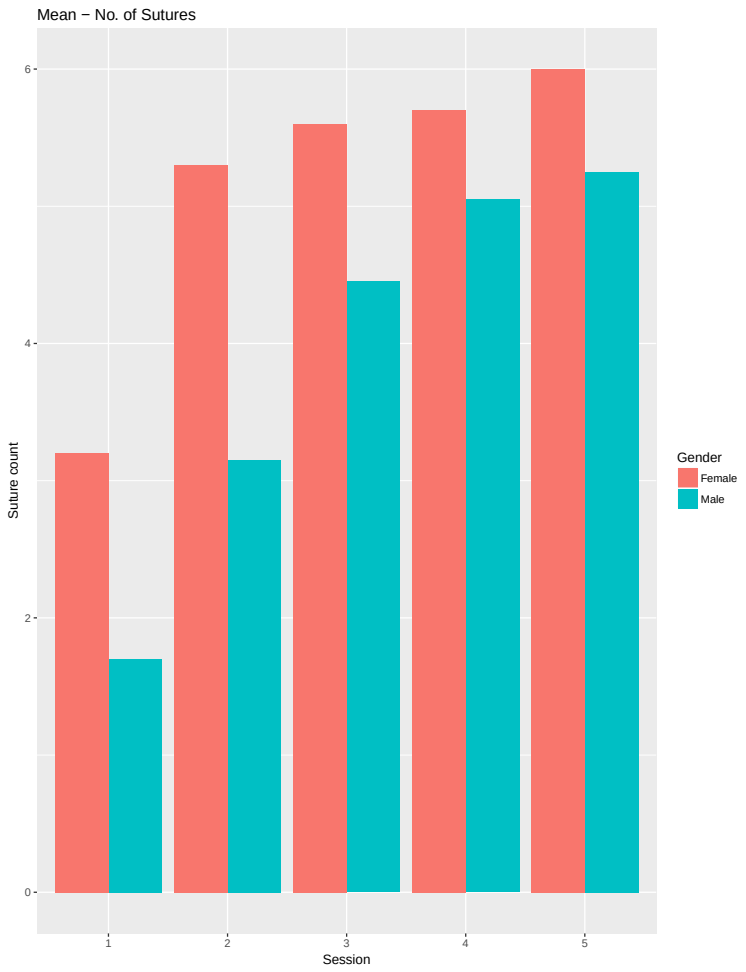


Figure 20: Average Number of Sutures for each session : Male Vs Female

From this plot we can clearly say that females performed better than males in each session. May be the females found the task easier and were more relaxed minded than males and were thus able to perform very well.
There is upward trend in performance of both males and females with progress in session.

(iii) Score

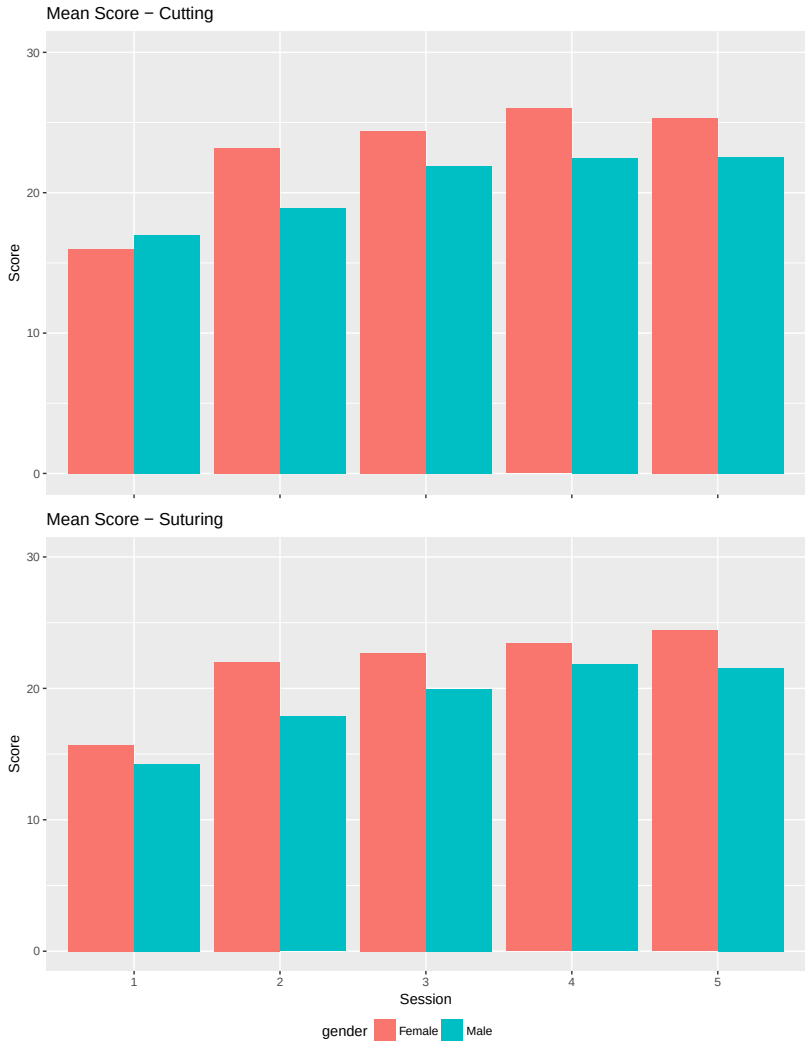


Figure 21: Average Score for each session for Cutting and Suturing : Male Vs Female

(i.1) Cutting

From this plot we can clearly say that females received more scores than males in each session. Thus, we can say that performance of females is better than males in each session. May be the females found the task easier and were more relaxed minded than males and were thus able to perform very well.
For session 3, 4 and 5 we do not see much difference in scores of males. While scores of

females slightly go down in session 5. May be

(i.2) Suturing

From this plot we can clearly say that females received more scores than males in each session. Thus, we can say that performance of females is better than males in each session. Overall there is upward trend in performance of both males and females with progress in session.

(e) Alloted Time for Suturing Task

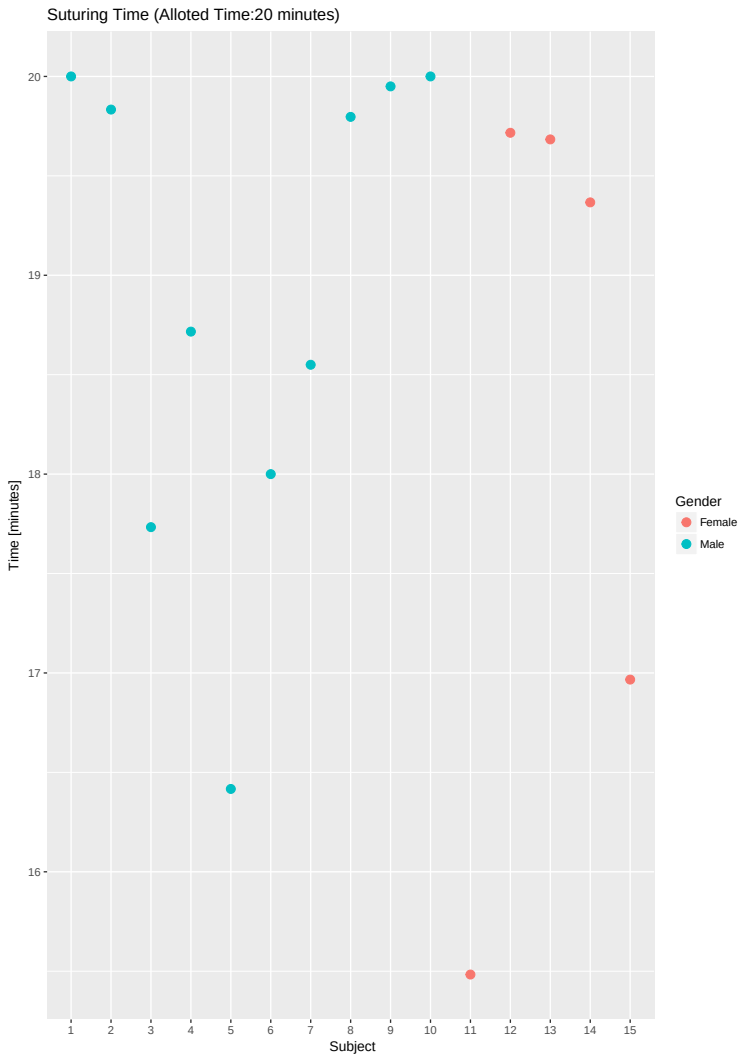


Figure 22: Average Suturing Time for each subject

The alloted timef or suturing was 20 mins. From this plot e can say that out of 10 males

and 5 female subjects 2 female and 1 male subjects were able to complete the task quite earlier than the alloted time. Almost 3 male subjevts completed the task in full 20 minutes. Thus, we can also say from this that overall males took longer time to finish the task of suturing.

C. INFERENCE

(i) Linear Model: We performed linear model on Model 1. The results of this model are as follows-:

Model 1:

Ho: There is no relationship between perspiration, sessions, task, scorer, age, gender and the score received by the subject.

```
Call:
lm(formula = time_taken_by_subjects ~ log(mean_perspiration_vector) +
    sessions + task + scorer + age + gender, data = table_data)

Residuals:
    Min      1Q  Median      3Q     Max
-532.30 -153.04   -3.42  141.34  451.52

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -2.122e+02  1.451e+03  -0.146  0.88385
log(mean_perspiration_vector) -1.880e+02  9.116e+02  -0.206  0.83674
sessionssession2             -1.046e+02  4.535e+01  -2.307  0.02175 *
sessionssession3              3.802e+02  4.582e+01   8.298 4.23e-15 ***
sessionssession4              8.345e+02  4.649e+01  17.950  < 2e-16 ***
sessionssession5              7.395e+02  4.538e+01  16.295  < 2e-16 ***
tasksuturing                  1.689e+02  5.767e+01   2.930  0.00367 **
scorer2                       1.592e-13  2.887e+01   0.000  1.00000
age                           6.194e+00  1.194e+01   0.519  0.60431
genderMale                    2.339e+01  3.176e+01   0.736  0.46207
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 248.4 on 286 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.7131,    Adjusted R-squared:  0.7041
F-statistic: 78.98 on 9 and 286 DF,  p-value: < 2.2e-16
```

Figure 23: Results for Linear Model of Model 1

The residuals show that our model predicted values are close to the actual observed values. The p-values of perinasal perspiration and scorer2 are all above 0.05, which tells us that there is no significant relationship between the scores and the these factors.
On the other hand, we can observe that sessions, tasksuturing, genderMale and age have p-values lower than 0.05. This makes these factors highly significant and they could thus

help us conclude that there is a significant relationship between sessions, tasksuturing, genderMale,age and the score received.

The residual standard error shows us that our model on average will deviate 0.0087% (3.607/286) from the true regression line. This means that our predicted score will be off from the actual score by 0.0087%.
The multiple R-squared of 0.4395 and the adjusted R-squared of .4219 tell us that our model explains close to 42.2% of the variance in the score received. This means that this model can only explain 42.2% of the variance of a subjects score.

The p-value which we are getting is 2.2e-16 and it is less than 0.05.
Given our data size of almost 300 samples, the resulting F-statistic of 24.92 on 9 and 286 DF, and a p-value less than 2.2e-16, we can reject the Ho.

We can thus conclude that there is a relationship between the response variable and the predictor variables. In other words, there is a relationship between the score a subject received , its perinasal pespiration, sessions, suturing task, second score, age, and gender.

(ii) Random Effects: We performed random effects on our Model 1. The results of this analysis are shown in Figure- -:

```
> table.lme<-lme(scores~log(mean_perspiration_vector)+sessions+task+scorer+age+gender,random = ~1|subjects,data =
 table_data,na.action = na.pass)
> summary(table.lme)
Linear mixed-effects model fit by REML
 Data: table_data
       AIC      BIC    logLik
  1674.075 1718.114 -825.0376

Random effects:
 Formula: ~1 | subjects
        (Intercept) Residual
StdDev:    2.684303 3.678291

Fixed effects: scores ~ log(mean_perspiration_vector) + sessions + task + scorer +      age + gender
                                Value Std.Error  DF   t-value p-value
(Intercept)                 -49.67611 19.400885 272 -2.560507  0.0110
log(mean_perspiration_vector) -46.01527 12.759122 272 -3.606461  0.0004
sessionssession2             -6.14541 17.659508 272 -0.347994  0.7281
sessionssession3            -61.09578  8.663060 272 -7.052447  0.0000
sessionssession4             71.25001 19.284391 272  3.694698  0.0003
sessionssession5             -1.16726  1.573936 272 -0.741618  0.4590
tasksuturing                  1.30310  0.760265 272  1.714012  0.0877
scorer2                       0.02330  0.097677 272  0.238558  0.8116
age                           0.02271  0.109386 272  0.207584  0.8357
genderMale                   -0.91418  0.149372 272 -6.120168  0.0000
 Correlation:
                            (Intr) lg(__) sssns2 sssns3 sssns4 sssns5 tskstr scorr2 age
log(mean_perspiration_vector)  0.990
sessionssession2               0.956  0.909
sessionssession3               0.305  0.421  0.026
sessionssession4              -0.982 -0.994 -0.895 -0.462
sessionssession5              -0.354 -0.331 -0.385 -0.104  0.319
tasksuturing                  -0.829 -0.829 -0.777 -0.300  0.823  0.284
scorer2                        0.010  0.071 -0.094  0.375 -0.067 -0.224 -0.069
age                           -0.054  0.079 -0.304  0.789 -0.087  0.136 -0.052  0.476
genderMale                     0.049  0.118 -0.117  0.521 -0.146  0.248 -0.067 -0.296  0.424
```

Figure 24: Results for Random Effect of Model 1

From this result we can say that AIC and BIC values of 1674.075 and 1718.114 respectively. Theses values are are high which indicate that our model doesnt fit the data well. In

other words, our response variable cant be predicted accurately by our predictor variables and random effects.

(iii) Linear Model: We performed linear model on Model 2. The results of this model are as follows-:

Ho: There is no relationship between perspiration, sessions, task, scorers, age, gender and the time it took the subjects to perform the task.

```
Call:
lm(formula = scores ~ log(mean_perspiration_vector) + sessions +
    task + scorer + age + gender, data = table_data)

Residuals:
    Min      1Q  Median      3Q     Max
-9.7219 -2.2865  0.4348  2.5484  8.6050

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                  35.0158    21.0768   1.661  0.09774 .
log(mean_perspiration_vector) 17.0810    13.2396   1.290  0.19804
sessionssession2              4.0999     0.6587   6.225 1.72e-09 ***
sessionssession3              5.9763     0.6655   8.980  < 2e-16 ***
sessionssession4              7.3082     0.6752  10.824  < 2e-16 ***
sessionssession5              7.3174     0.6591  11.103  < 2e-16 ***
tasksuturing                 -2.3204     0.8375  -2.771  0.00596 **
scorer2                       0.1554     0.4193   0.371  0.71119
age                           0.3811     0.1734   2.197  0.02879 *
genderMale                   -2.1035     0.4612  -4.561 7.56e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.607 on 286 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.4395,    Adjusted R-squared:  0.4219
F-statistic: 24.92 on 9 and 286 DF,  p-value: < 2.2e-16
```

Figure 25: Results for Linear Model of Model 2

The residuals show that we have a fairly symmetrical model. This means that our predicted values are close to our actual observed values.
The p-value of perinasal perspiration, scorer, age, and gender have p-values higher than 0.05 thus we can say that there is no relationship between these factors and time taken.
On the other hand p-values of session2, session3, session4, and session5 are tasksuturing are less than 0.05 which show that they are highly significant factors. In other words, there is a relationship between the these factors and time taken to complete the chosen task.

The residual standard error shows that our model on average will make predictions that are 268% off of the actual value. This means that on average the time our model predicts will be different than the actual time taken by 268%.

The multiple R-squared of .7131 and the adjusted R-squared of .7041 tell us that our model explains close to 70% of the variance in the response variable. This means that this

model can explain 70% of the variance of the time taken to complete the task.

Given our data size of almost 300 samples, the resulting F-statistic of 78.98 on 9 and 286 DF, and a p-value less than 2.2e-16 indicate that we have a significant relationship between the response and predictor variables.

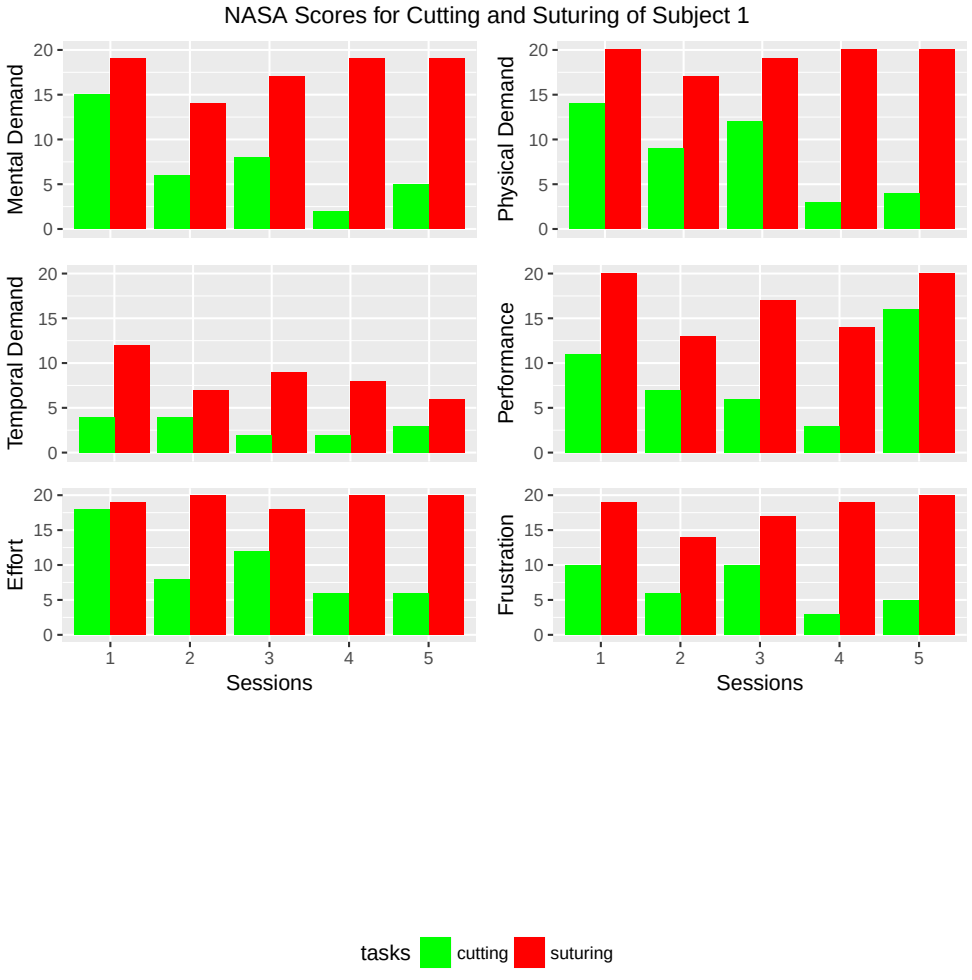
This means that we can reject Ho, and conclude there is a relationship between the time taken by a subject, its perinasal perpiration, sessions, suturing task, second score, age, and gender.

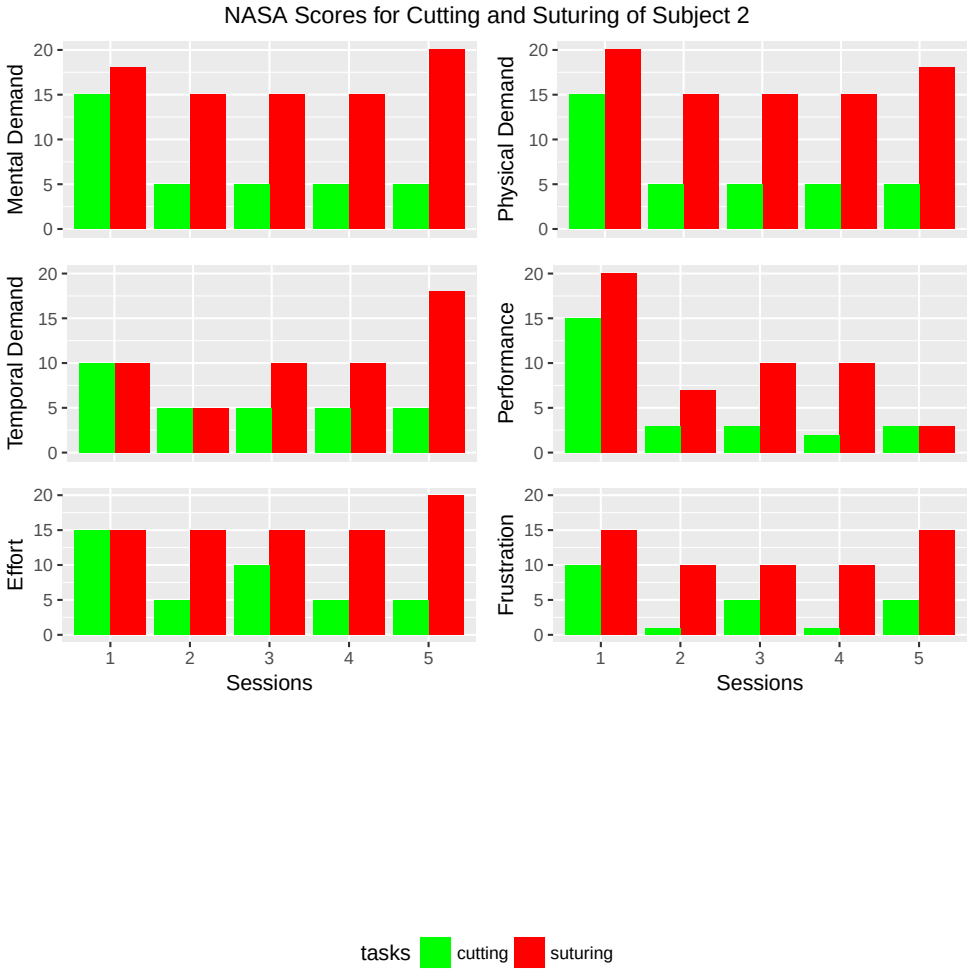
D. APPENDIX

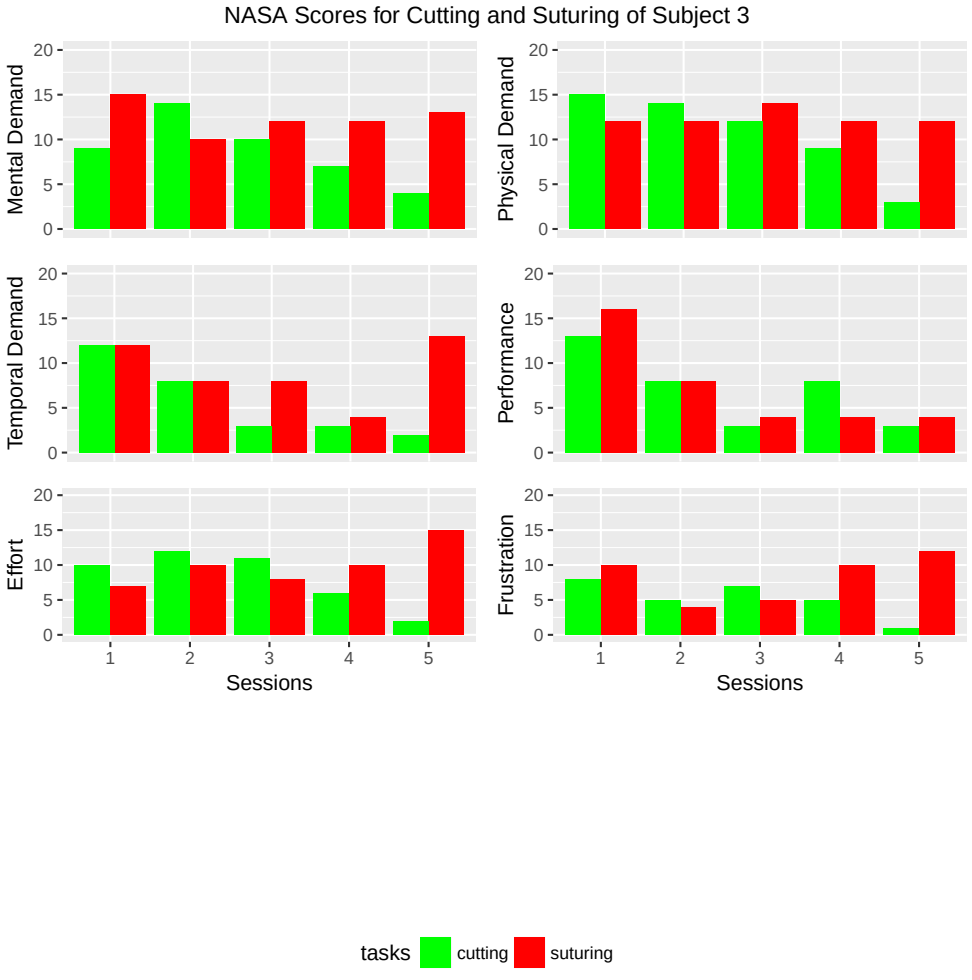
3. State Psychometric Data

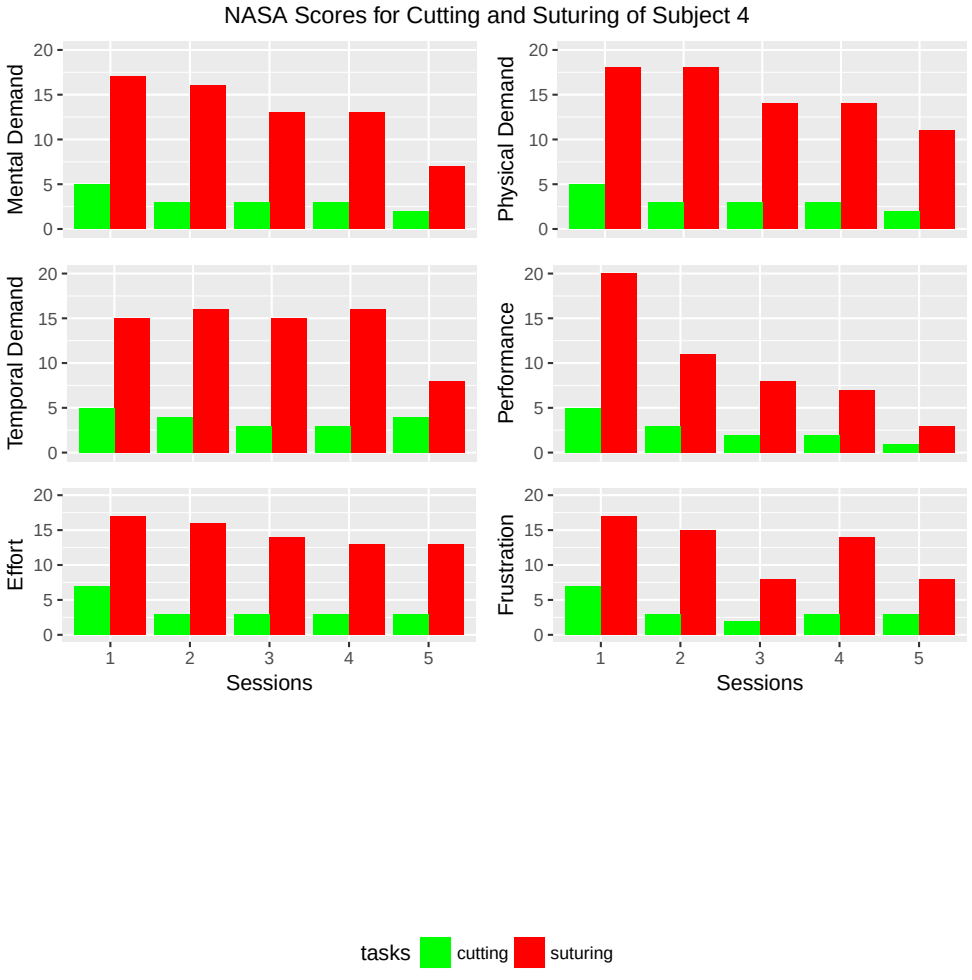
In the experiment, after cutting and suturing tasks were completed by the subjects, the subjects filled out NASA-TLX questionnaire. The NASA-TLX instrument features 6 parameters of perceptions of subjects regarding task difficulty. These sub scales are scored in the range 1-20. The 6 parameters are-:

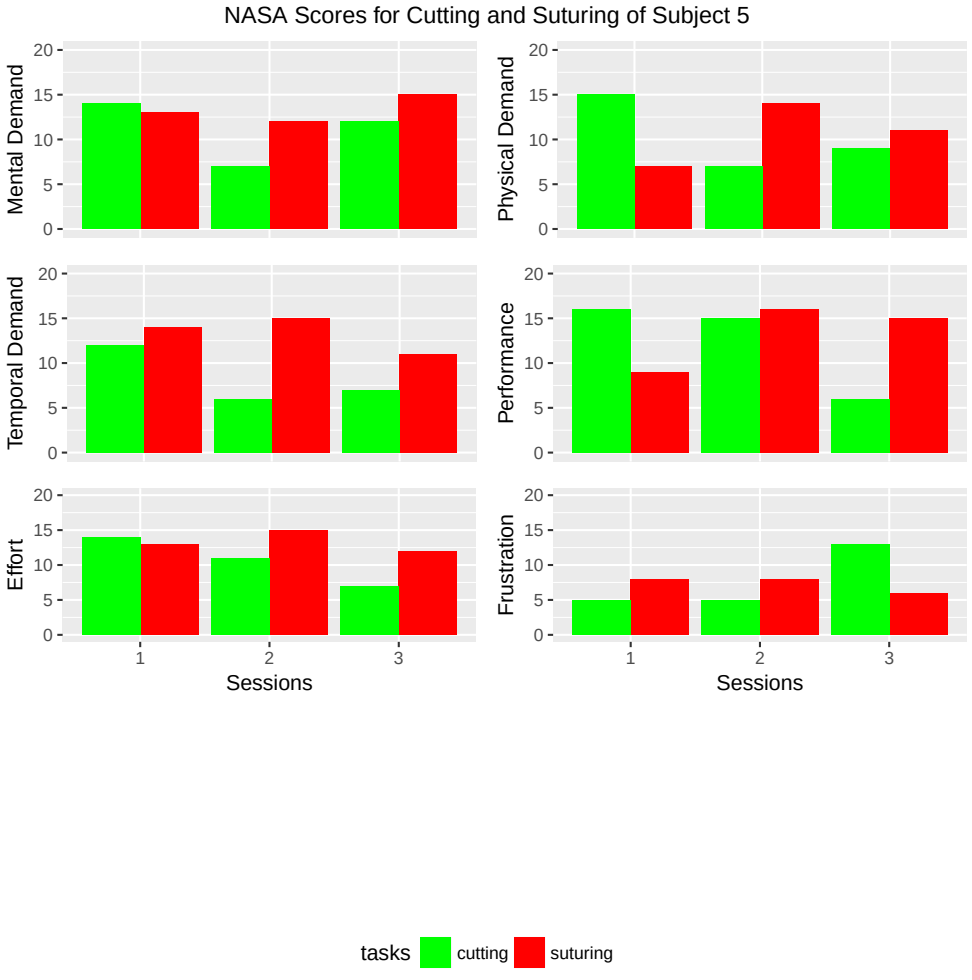
In order to analyze the state psychometric data of all the subjects for 6 factors we have made bar plots. Theses bar plots are shown in Figure 4,5 and 6-:

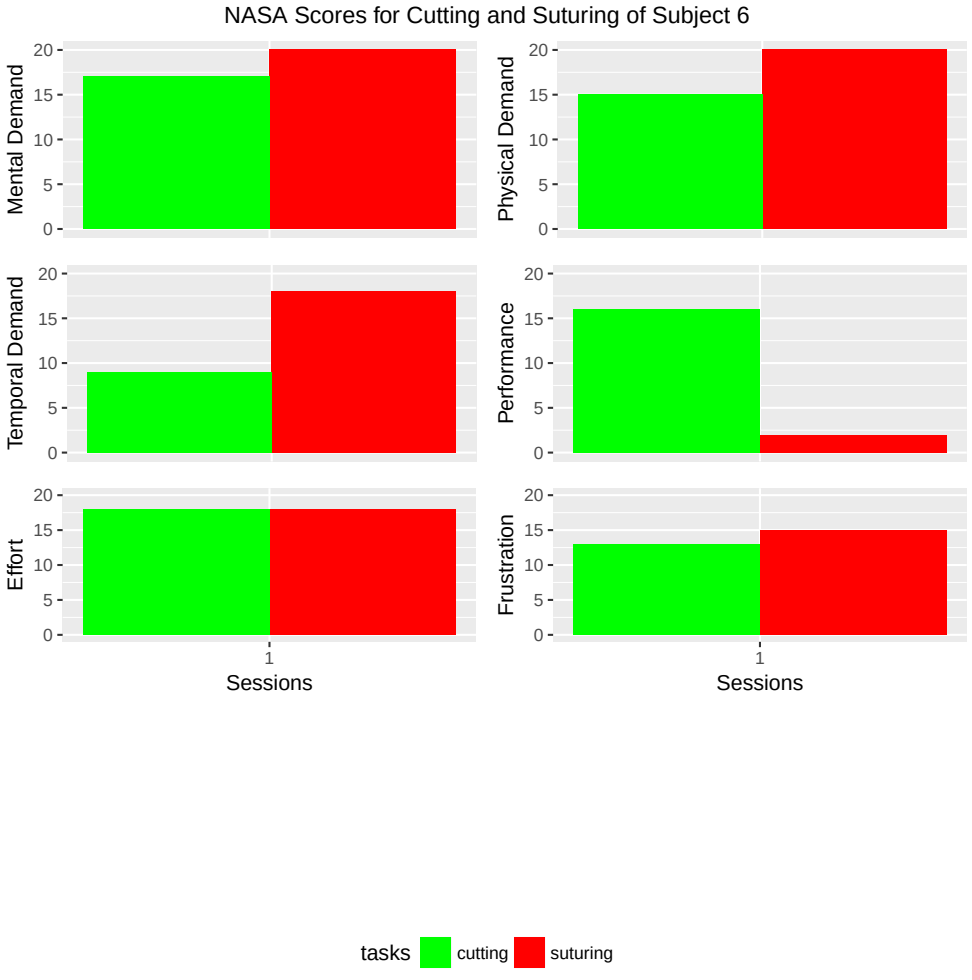

NASA Scores for Cutting and Suturing of Subject 1
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

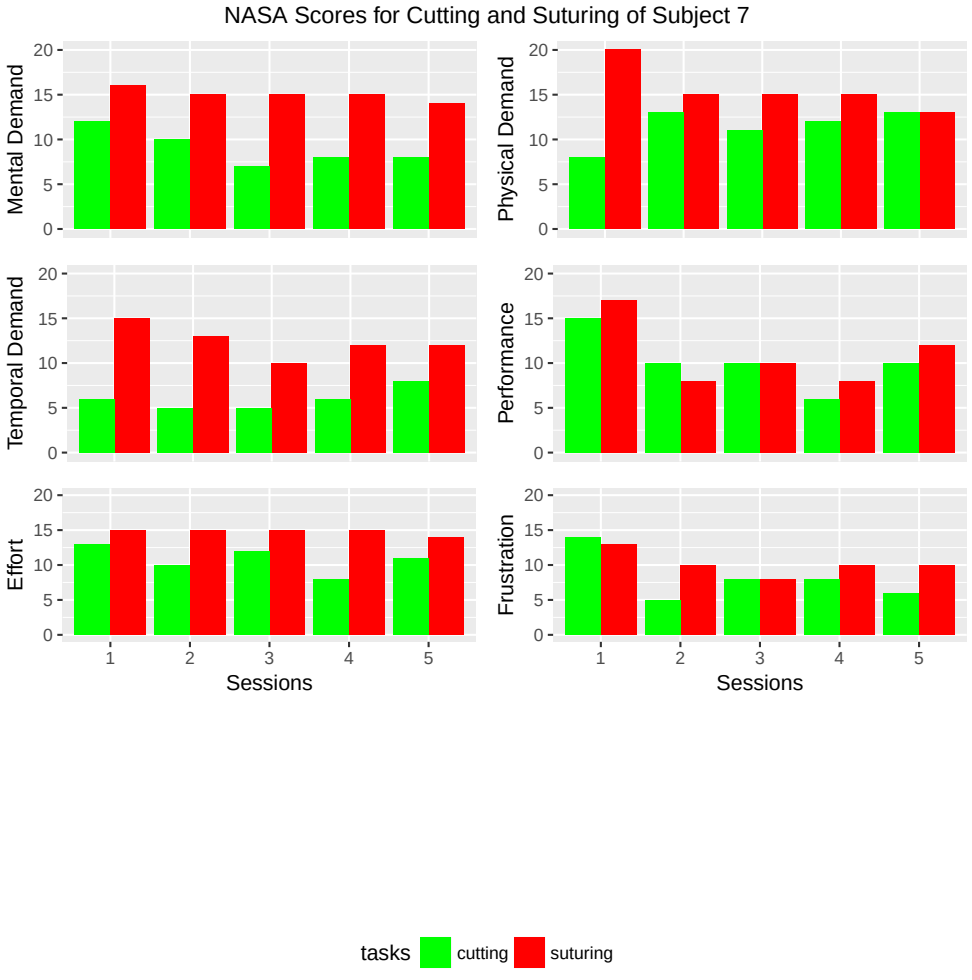

NASA Scores for Cutting and Suturing of Subject 2
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

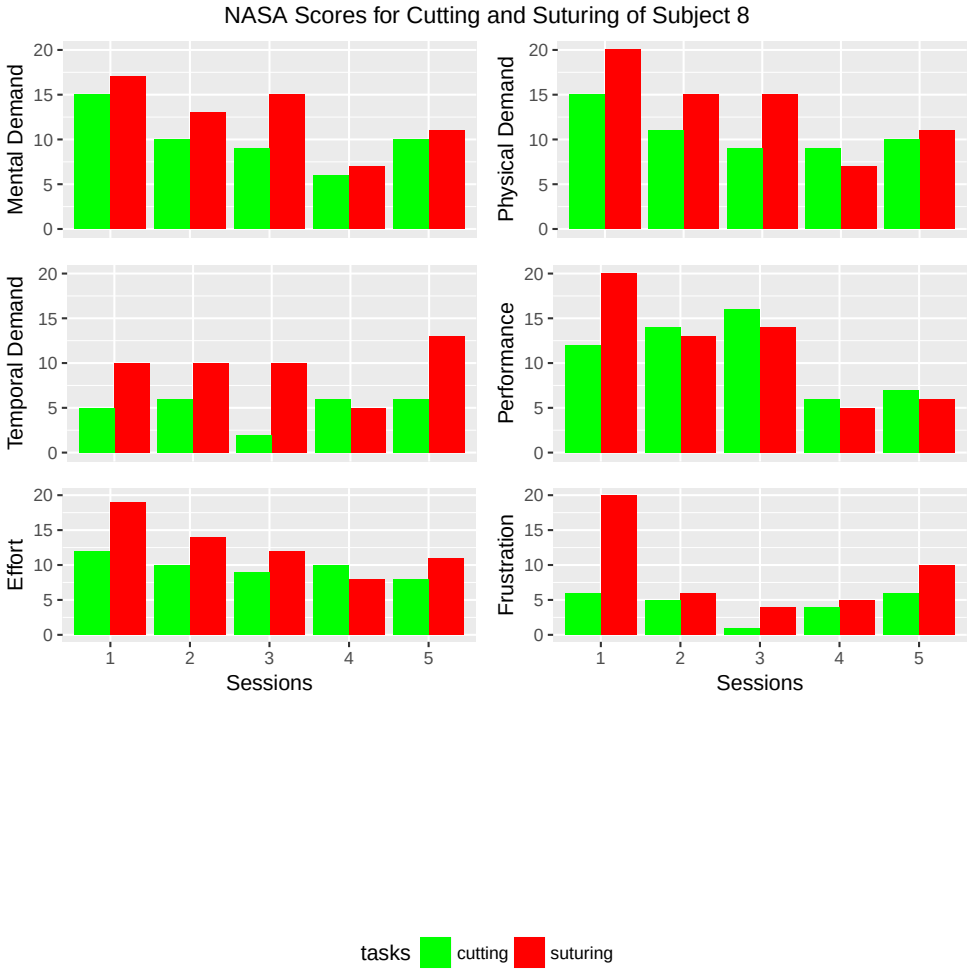

NASA Scores for Cutting and Suturing of Subject 3
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

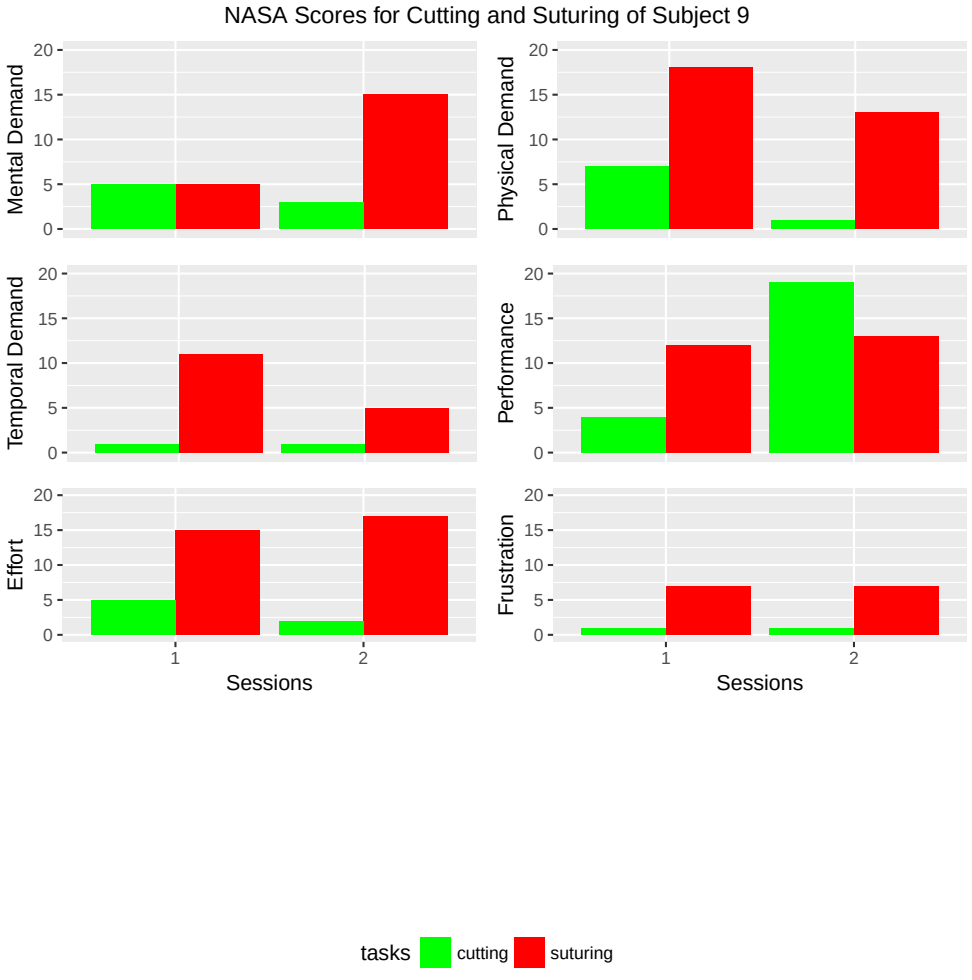

NASA Scores for Cutting and Suturing of Subject 4
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

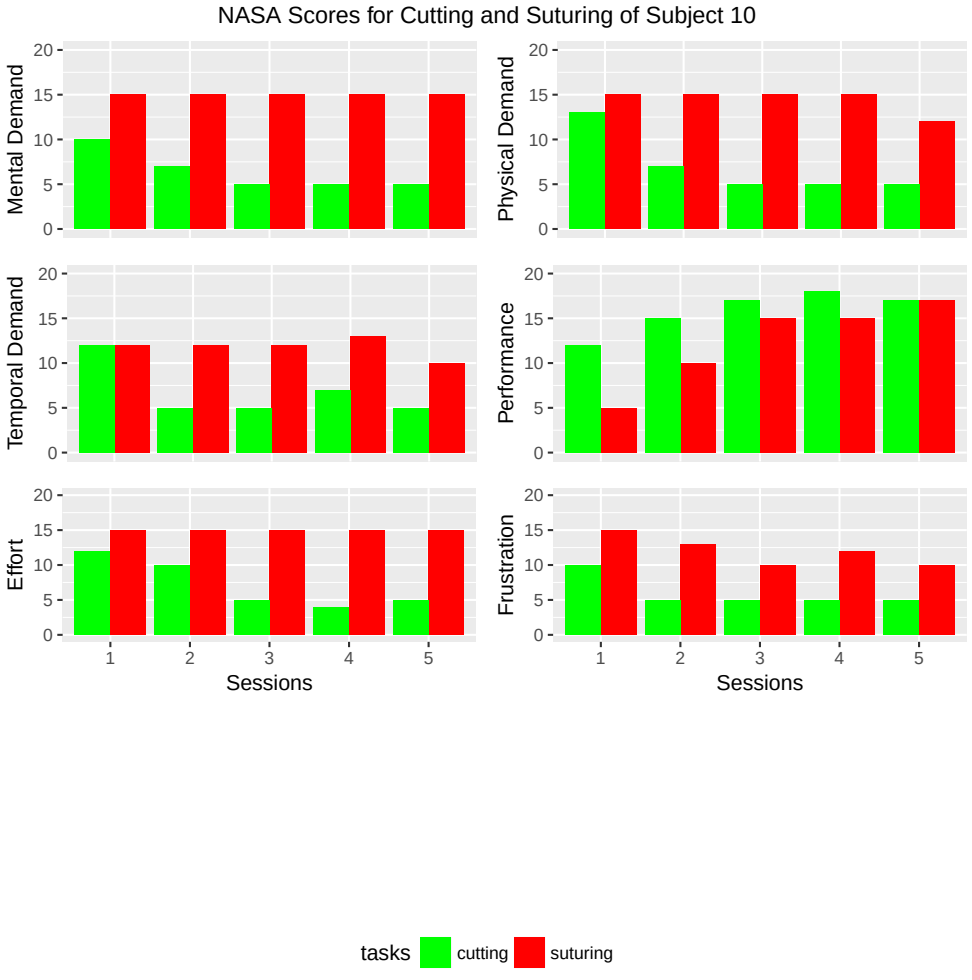

NASA Scores for Cutting and Suturing of Subject 5
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
Sessions
Sessions
tasks
cutting
suturing

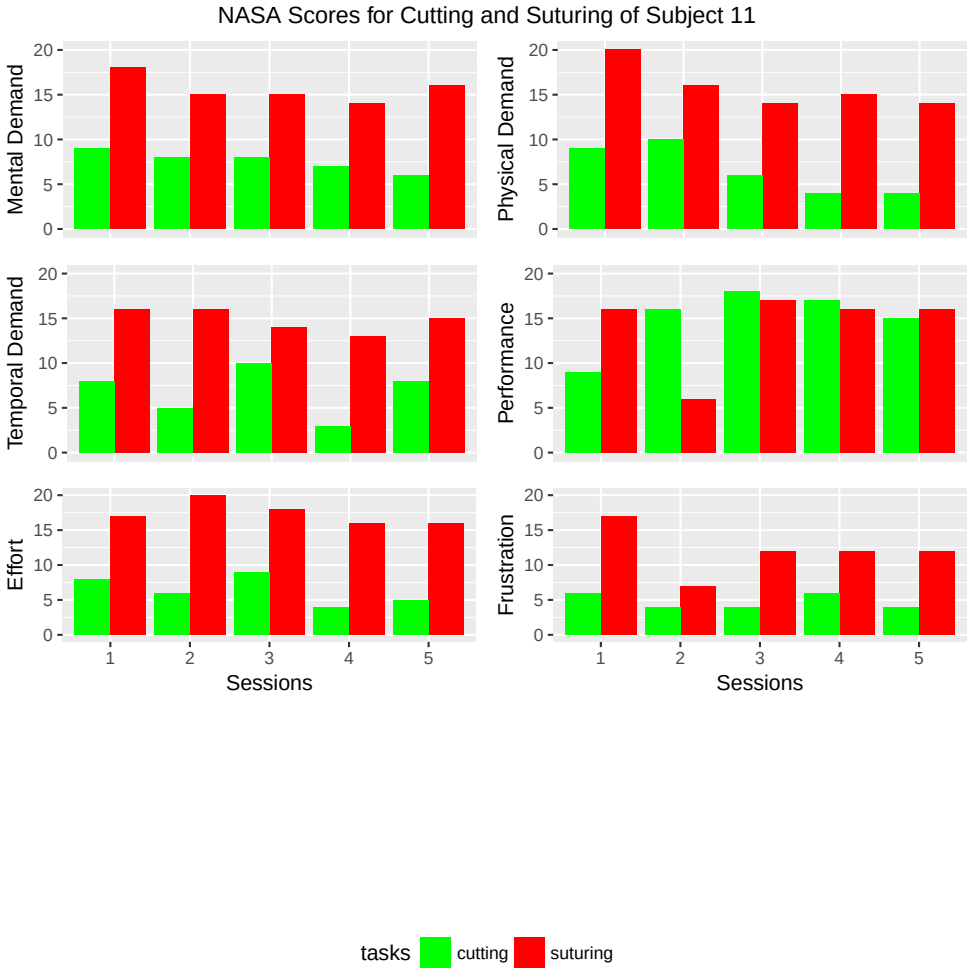

NASA Scores for Cutting and Suturing of Subject 6
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
20
15
10
5
0
1
Sessions
tasks
cutting
suturing

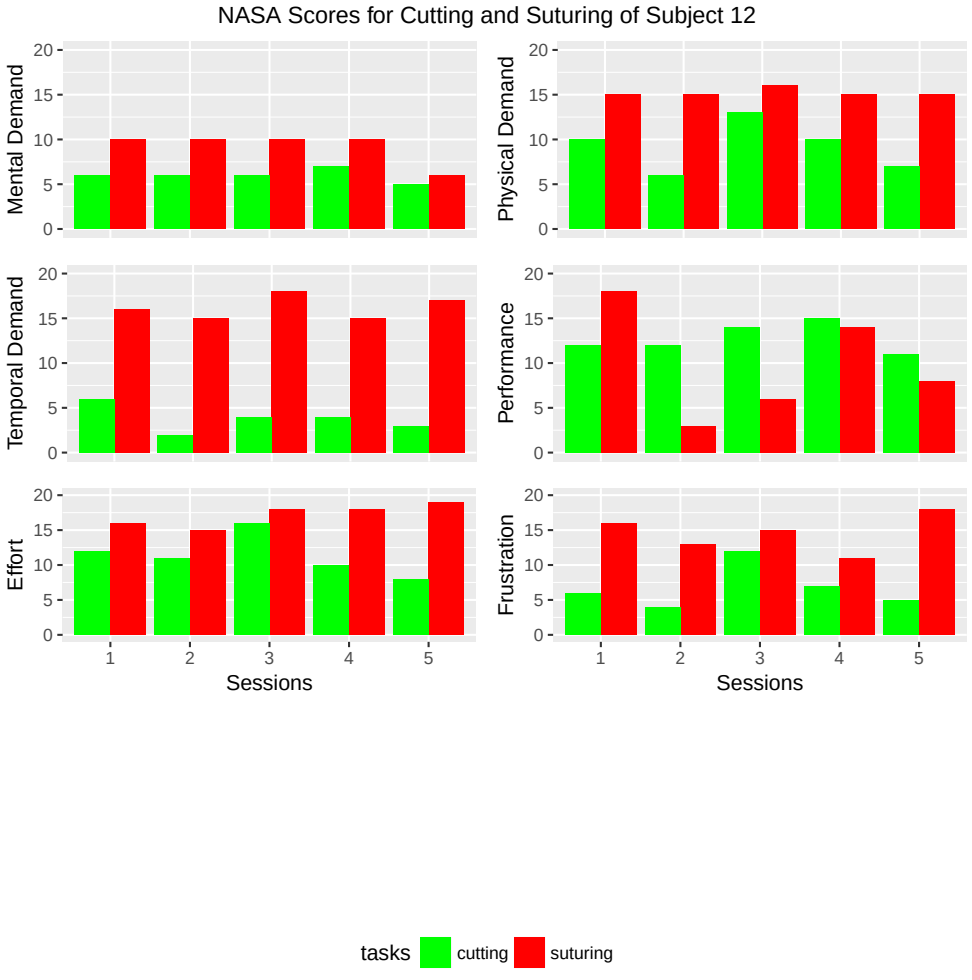

NASA Scores for Cutting and Suturing of Subject 7
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

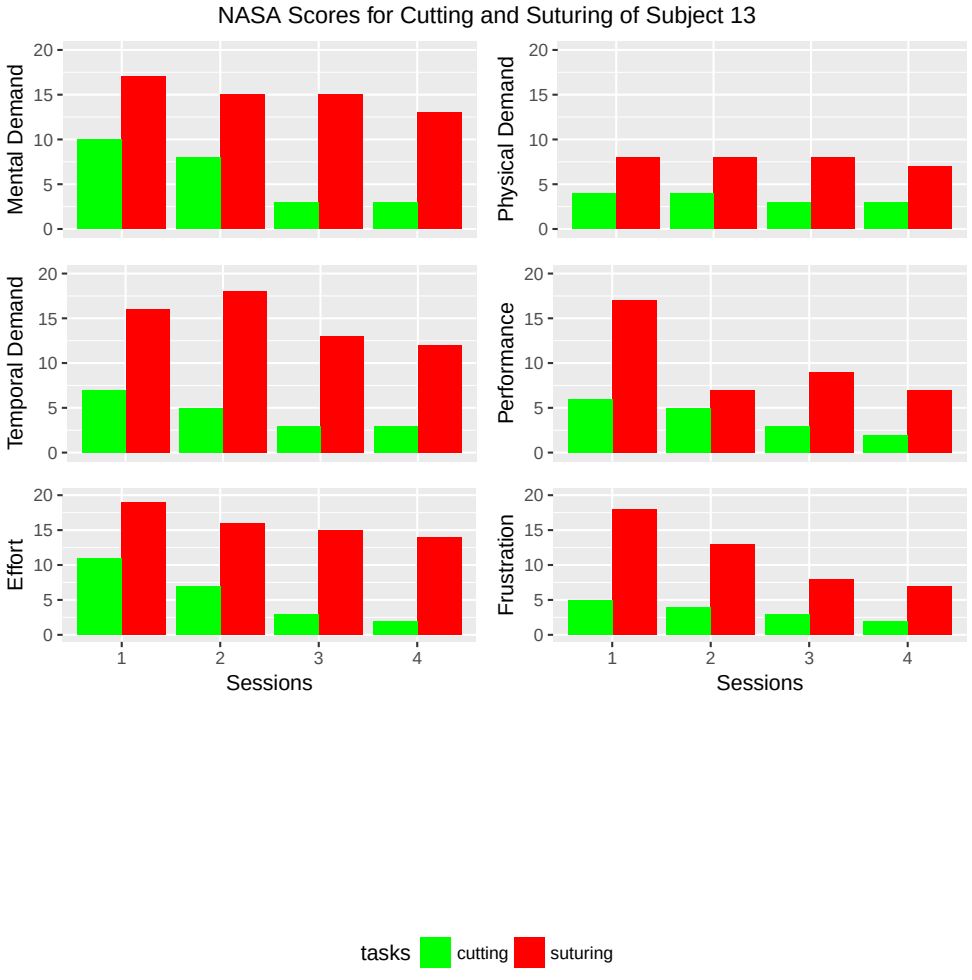

NASA Scores for Cutting and Suturing of Subject 8
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

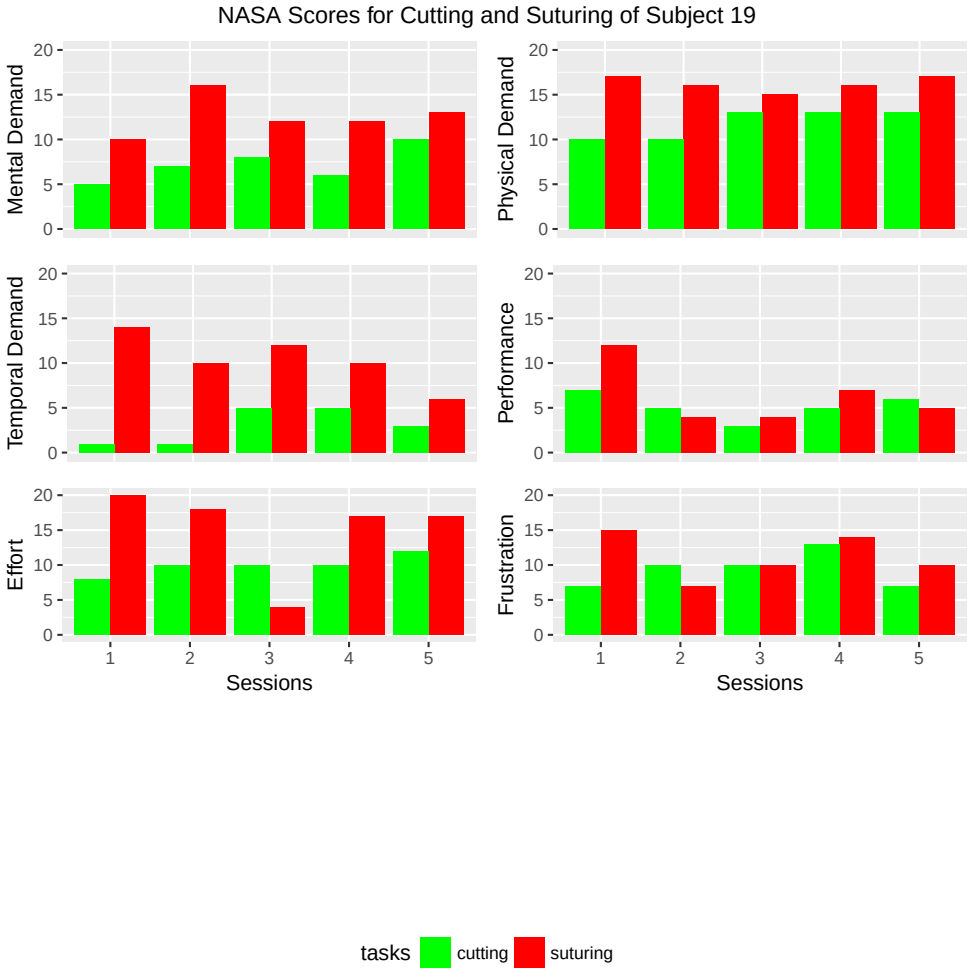

NASA Scores for Cutting and Suturing of Subject 9
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
Sessions
tasks
cutting
suturing

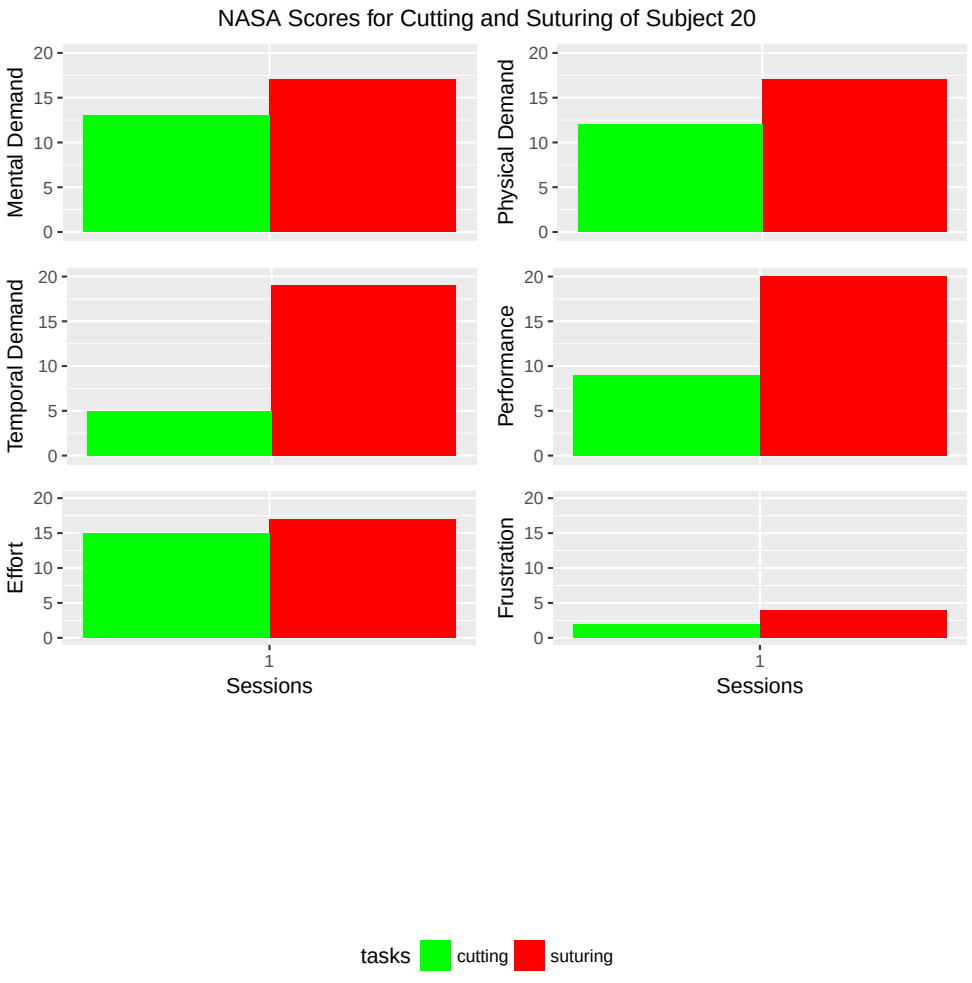

NASA Scores for Cutting and Suturing of Subject 10
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

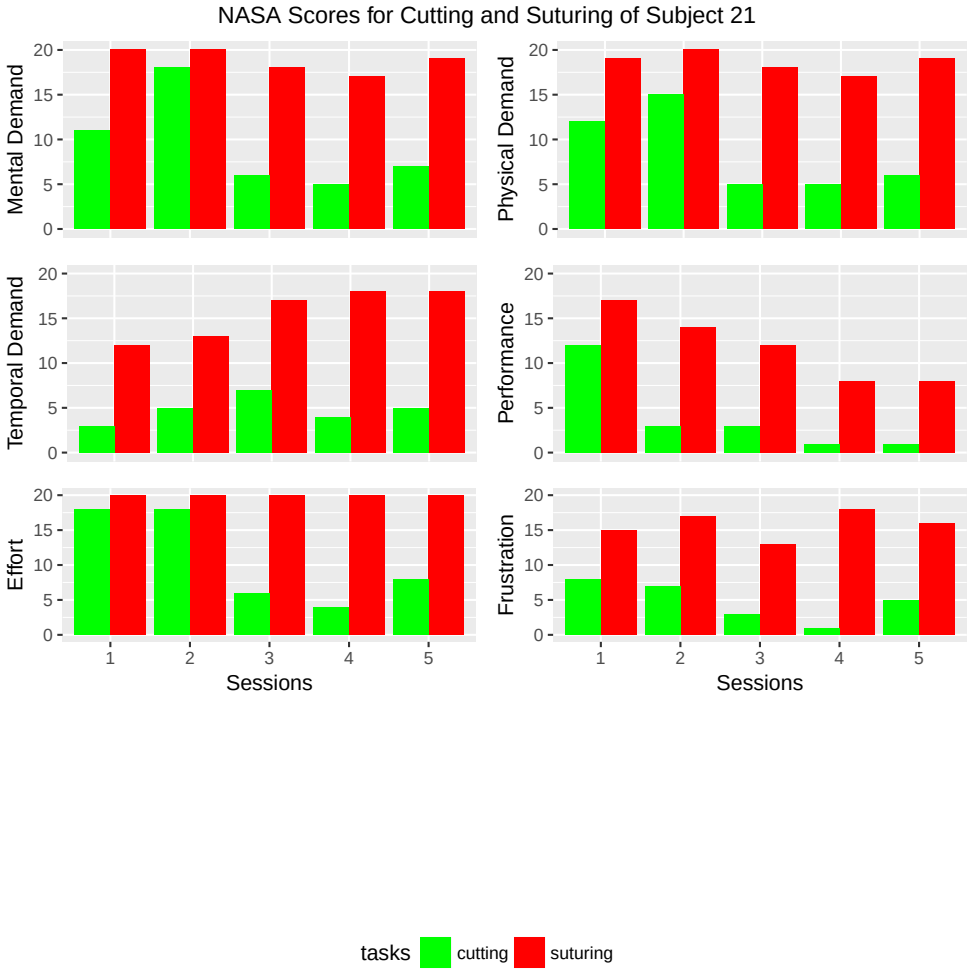

NASA Scores for Cutting and Suturing of Subject 11
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

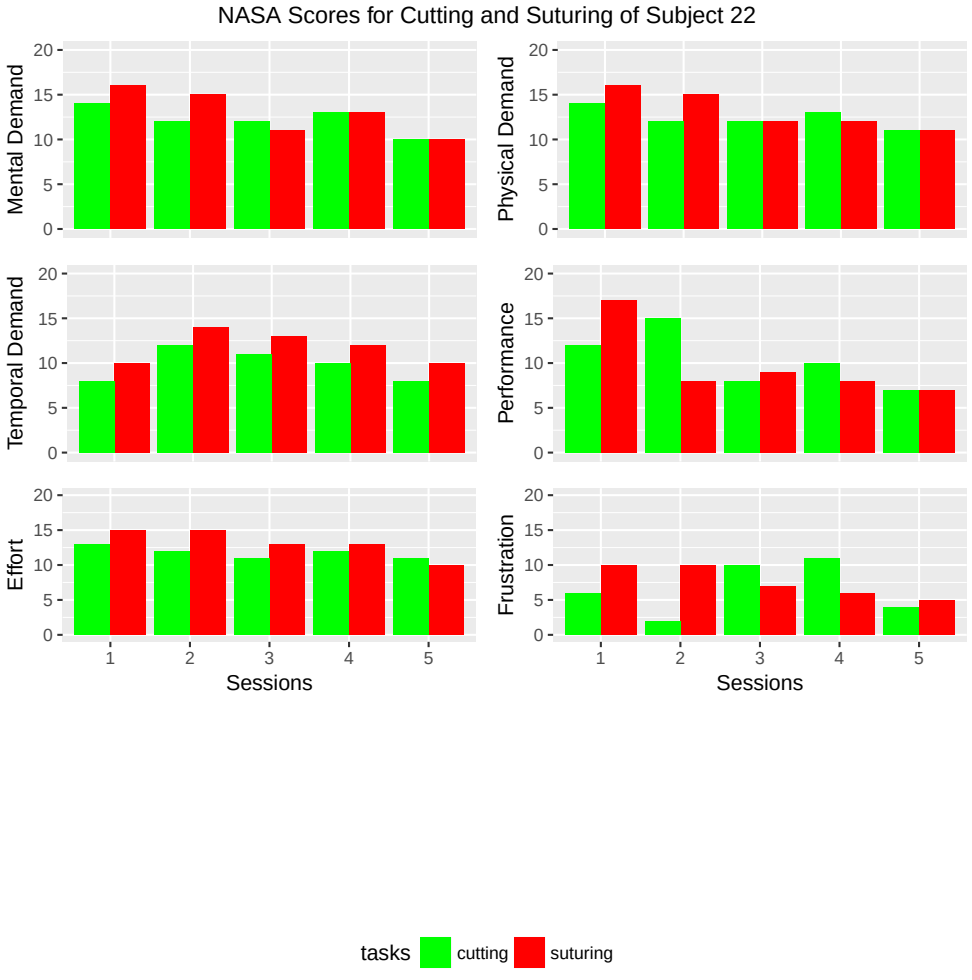

NASA Scores for Cutting and Suturing of Subject 12
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

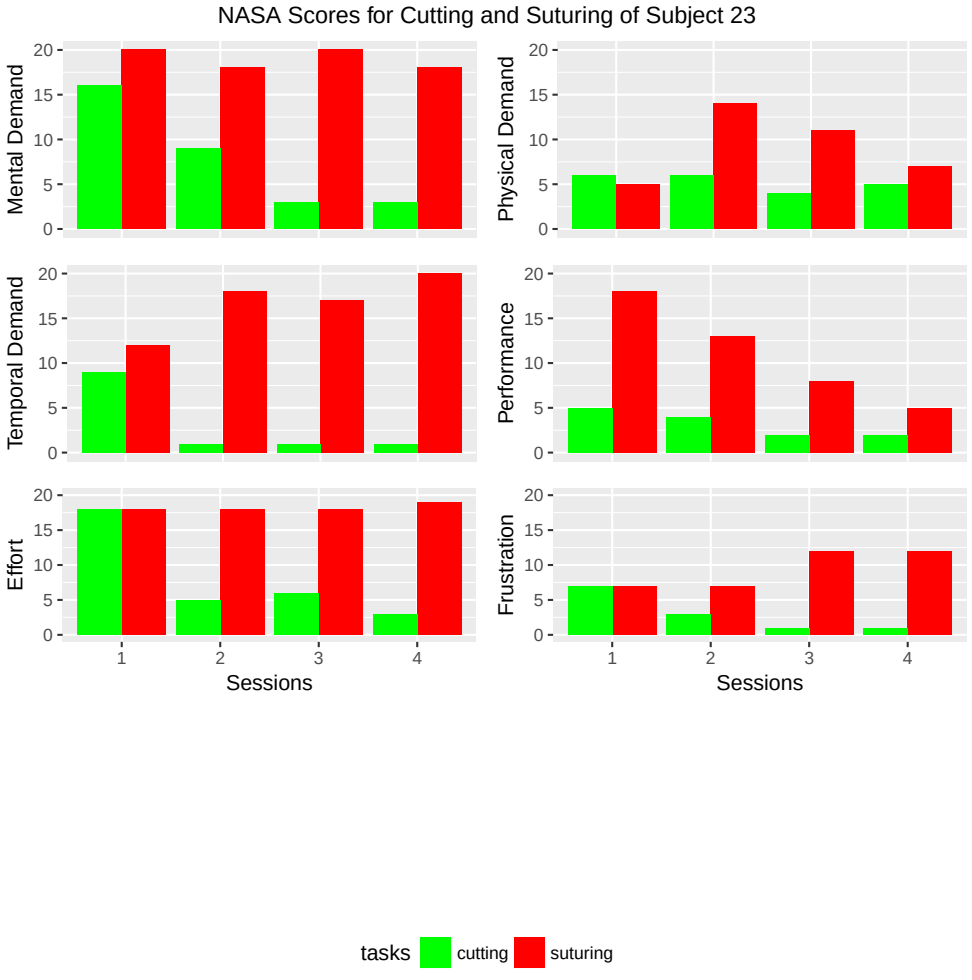

NASA Scores for Cutting and Suturing of Subject 13
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
Sessions
Sessions
tasks
cutting
suturing

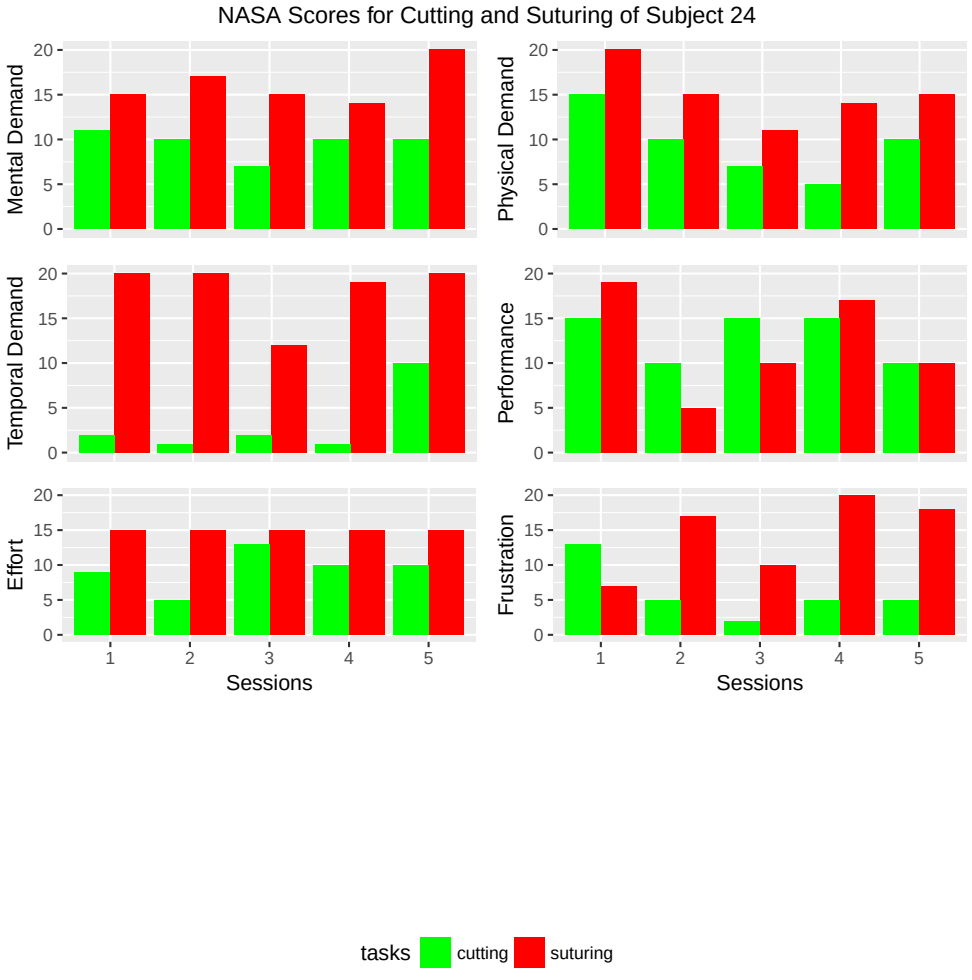

NASA Scores for Cutting and Suturing of Subject 19
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing

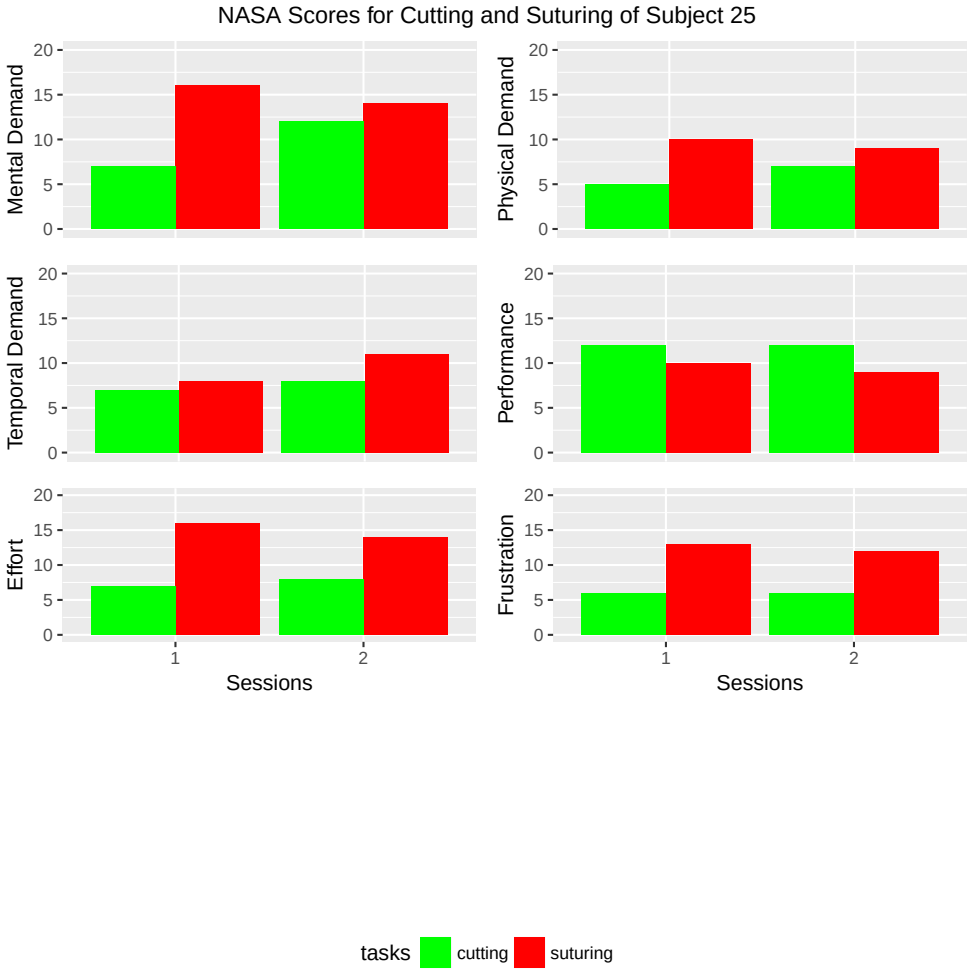

NASA Scores for Cutting and Suturing of Subject 20
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
20
15
10
5
0
1
Sessions
tasks
cutting
suturing

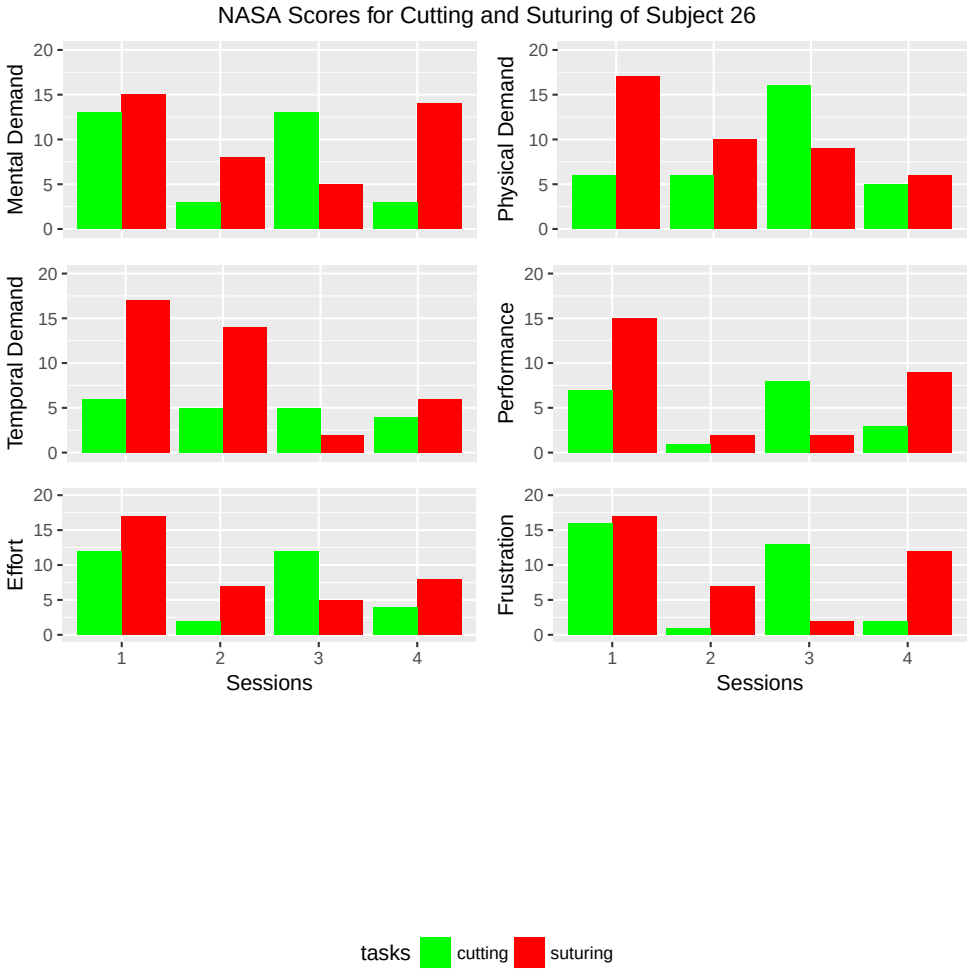

NASA Scores for Cutting and Suturing of Subject 21
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing


NASA Scores for Cutting and Suturing of Subject 22
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing


NASA Scores for Cutting and Suturing of Subject 23
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
Sessions
Sessions
tasks
cutting
suturing


NASA Scores for Cutting and Suturing of Subject 24
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
5
Sessions
Sessions
tasks
cutting
suturing


NASA Scores for Cutting and Suturing of Subject 25
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
Sessions
tasks
cutting
suturing


NASA Scores for Cutting and Suturing of Subject 26
Mental Demand
Physical Demand
Temporal Demand
Performance
Effort
Frustration
0
5
10
15
20
1
2
3
4
Sessions
Sessions
tasks
cutting
suturing

From the overall view of all the plots, we can see that in Performance parameter, in comparison with suturing, cutting was easier for almost all the students. This is because almost all the subjects got lower scores in the range 1-20, except for few subjects: subject 3, subject 5, subject 11 and subject 26.

We can also say that the overall score for suturing is very much high which means it was very difficult for everyone.

There were exceptions in some of the parameters such as Performance for Subject26, where she found suturing task easy. As we go session wise, we should see downward trend which means that they improved from first to last session.The most improved ones which we can see are subject 3,23 for both cutting and suturing.The subjects who did not improve through the sessions are Subject 1, Subject 21, Subject 24

If we see parameter wise, we can say the following analysis results-:

i) Frustration was very less in suturing compared to other parameters like Mental Demand, Physical Demand, Effort which are very high in suturing task.

ii) Physical demand and Effort are very high at the time of suturing, meaning they had more physical work effort during suturing.

iii) While during cutting mental demand and physical demand are very high, which means that they had mental stress and physical stress more during cutting.

iv) Performance is the only parameter which should go high as we pass through sessions, but in general it has decreasing trend except for Subject 10, for which we can say he performed well.

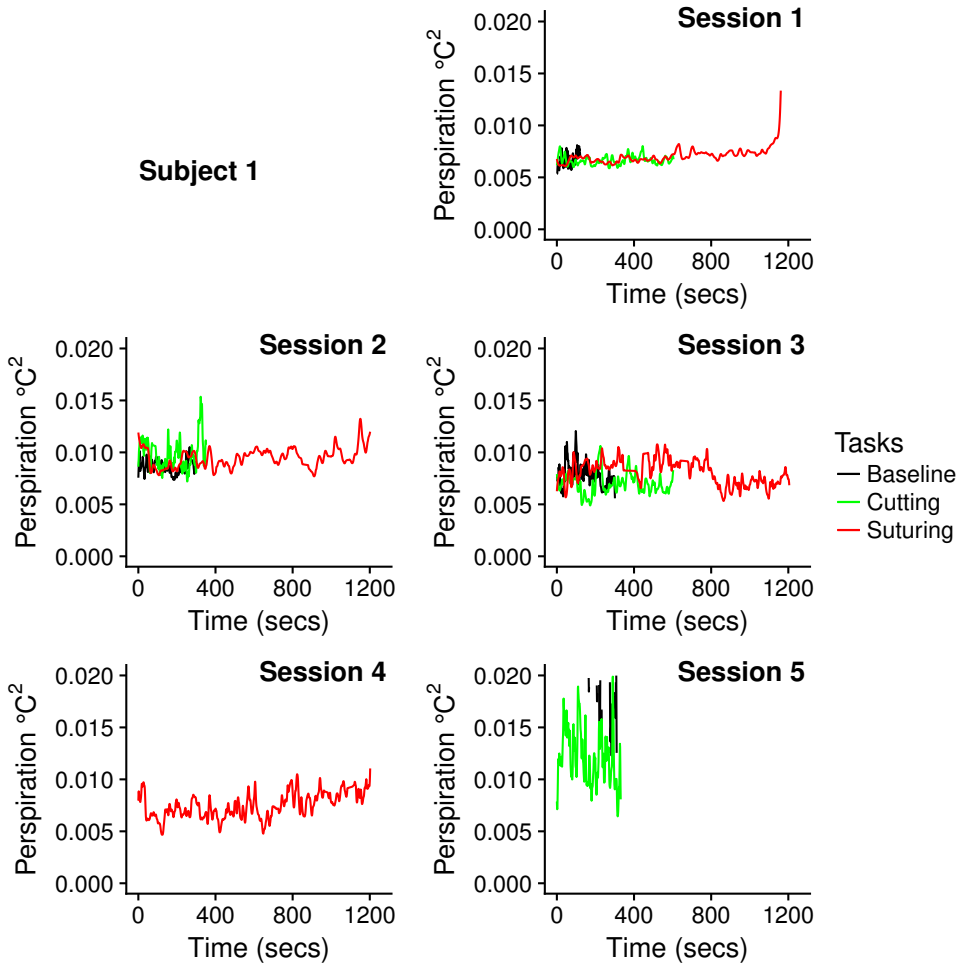
4. Perinasal Perspiration (Stress) Signal Data

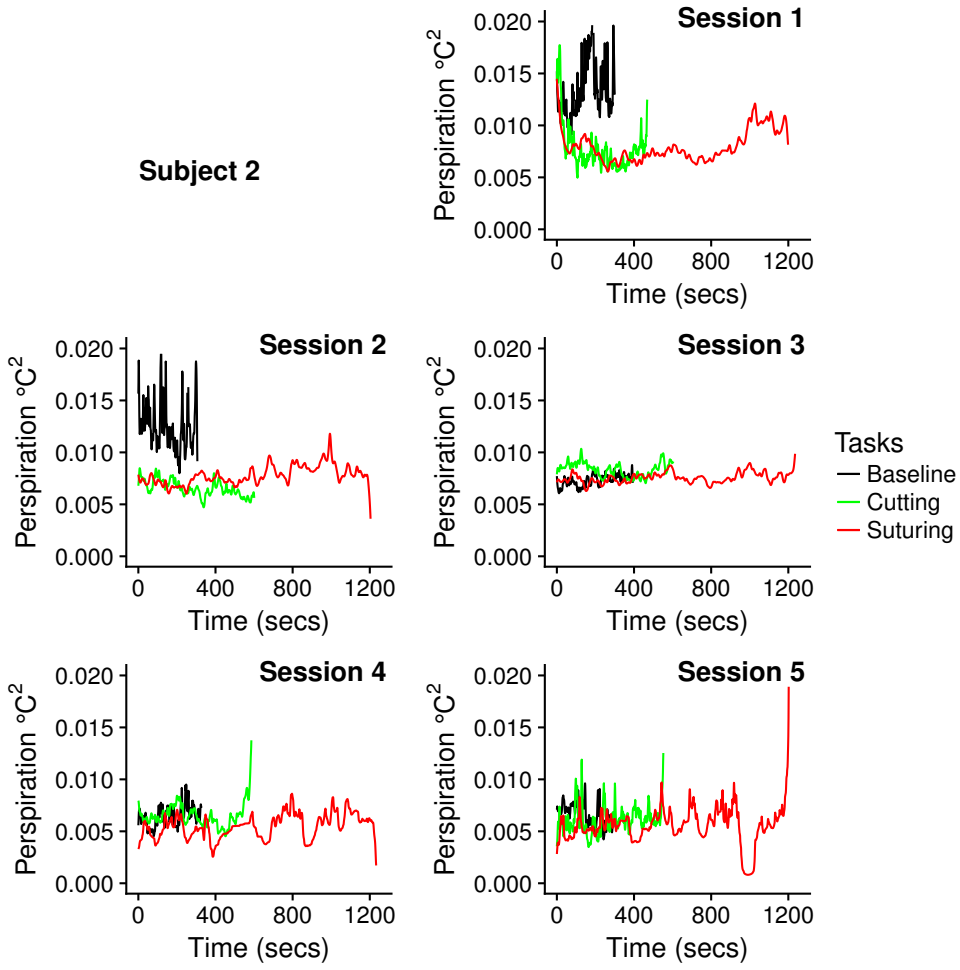
In order to find the Perinasal Perspiration (Stress) Signal, we have drawn plots for stress signals of each session for all 21 subjects.

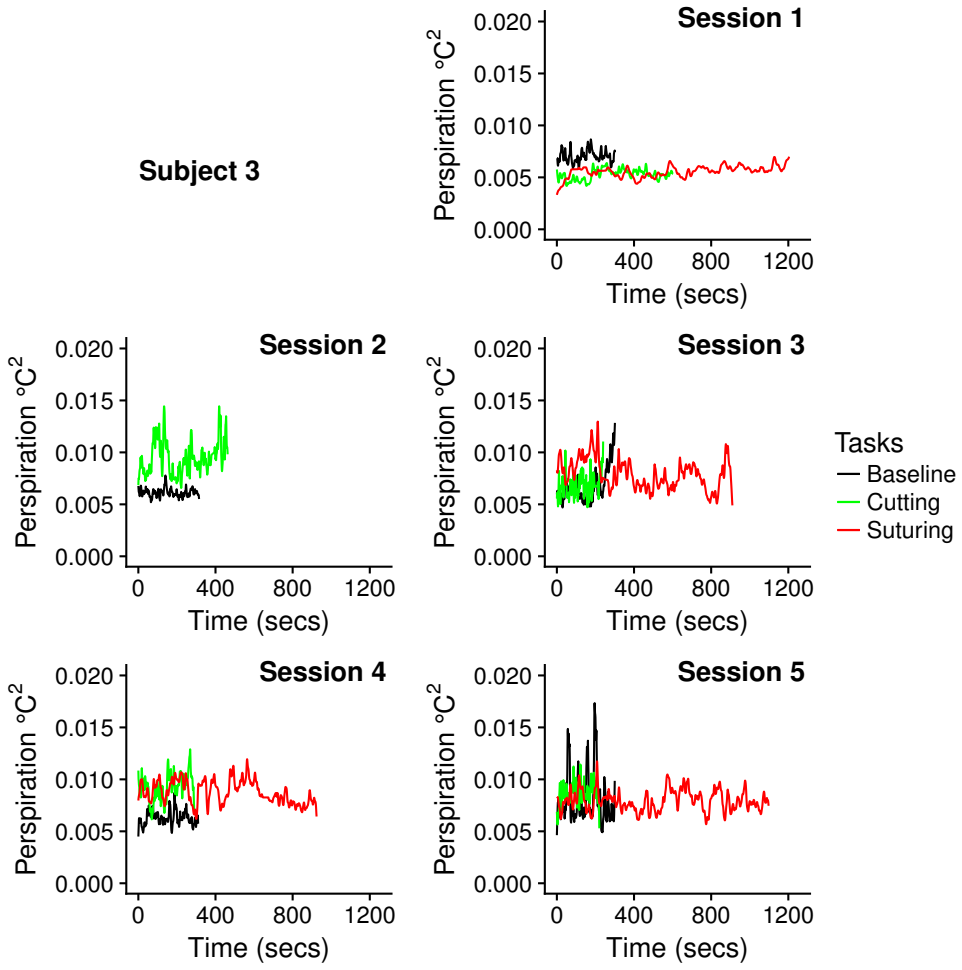
For all the subjects, the sampling rate fluctuates around 7 frames per second. We have reduced those frames to 1 frame per second and used these reduced signals to perform our analysis. Due to this average we get a smoother signal without affecting the validity of the data.

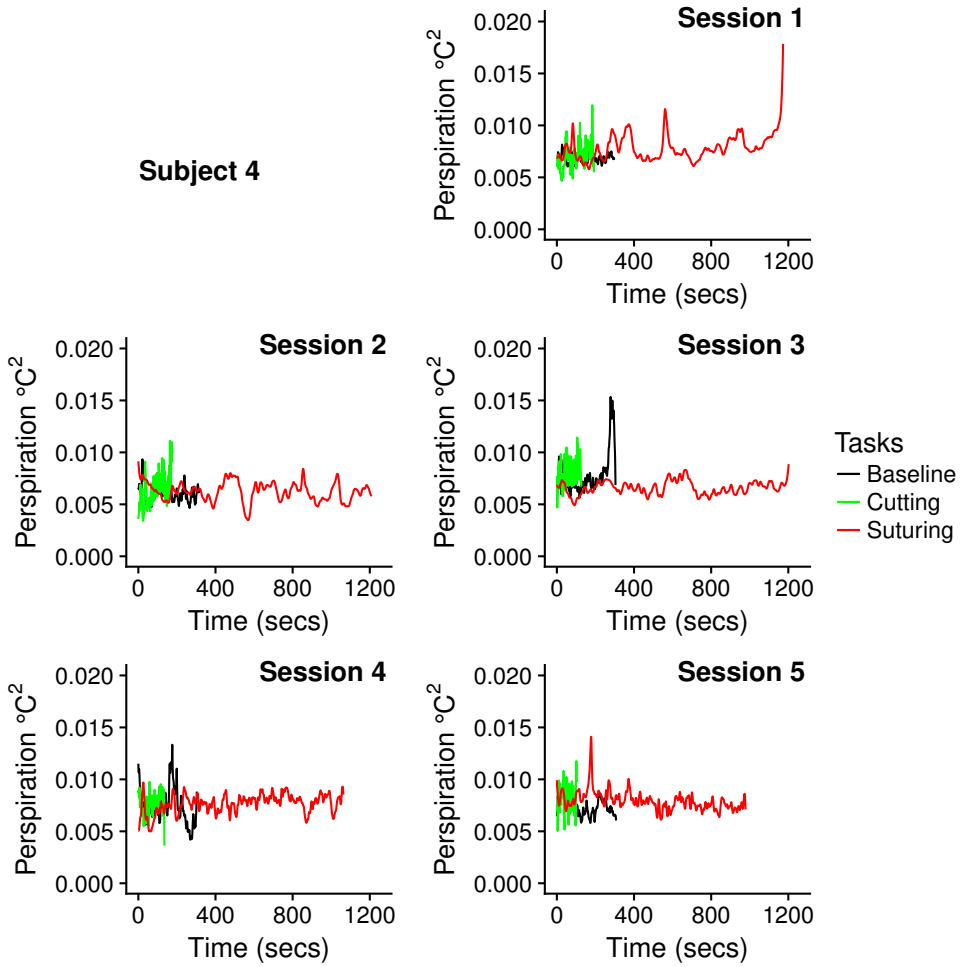
We have used following colors for the tasks-: Black for Baseline, Green for Cutting and Red for Suturing.

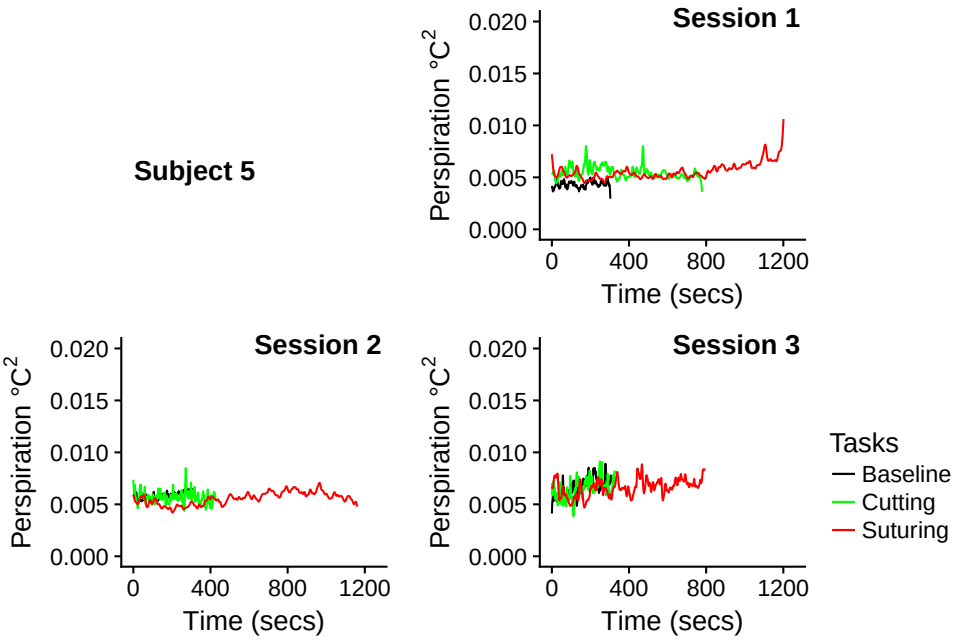
Our plots for Perinasal Perspiration (Stress) Signal are shown in Figure


Subject 1
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
Time (secs)
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Tasks
Baseline
Cutting
Suturing

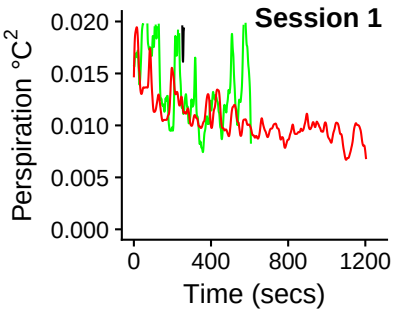

Subject 2
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 3
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
Time (secs)
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Tasks
Baseline
Cutting
Suturing

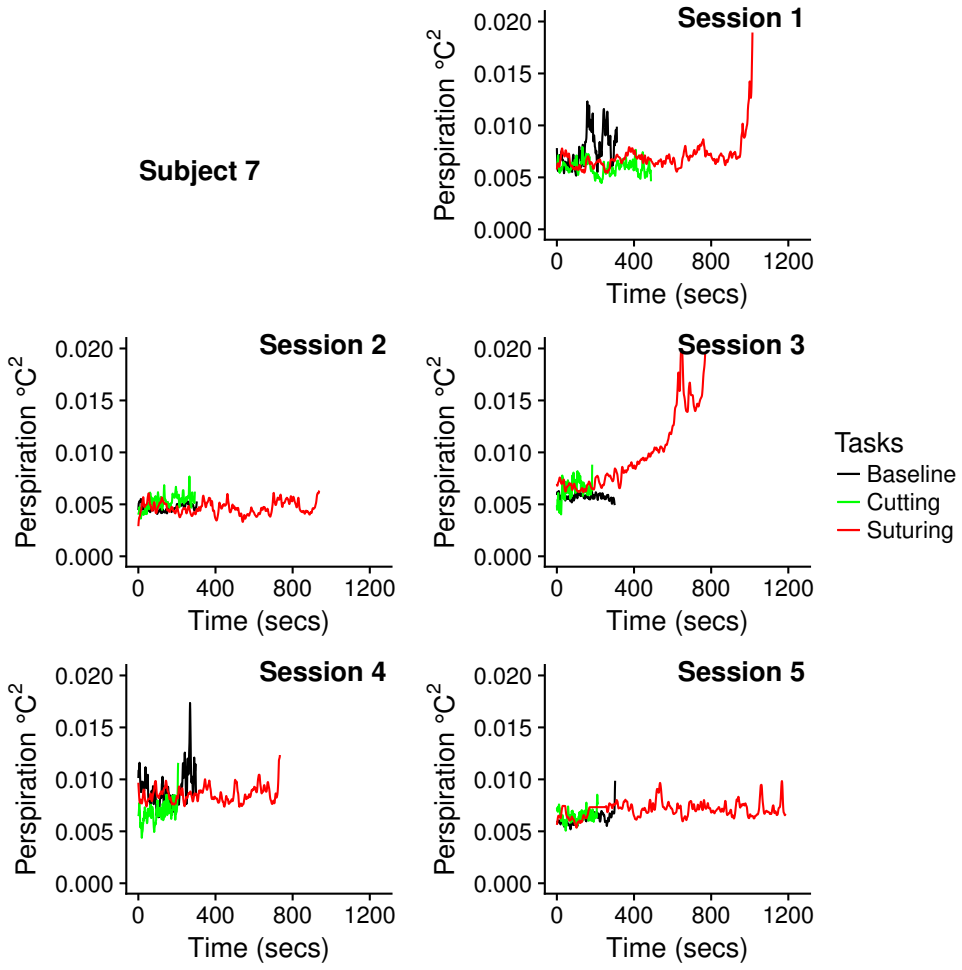

Subject 4
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
Time (secs)
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Tasks
Baseline
Cutting
Suturing

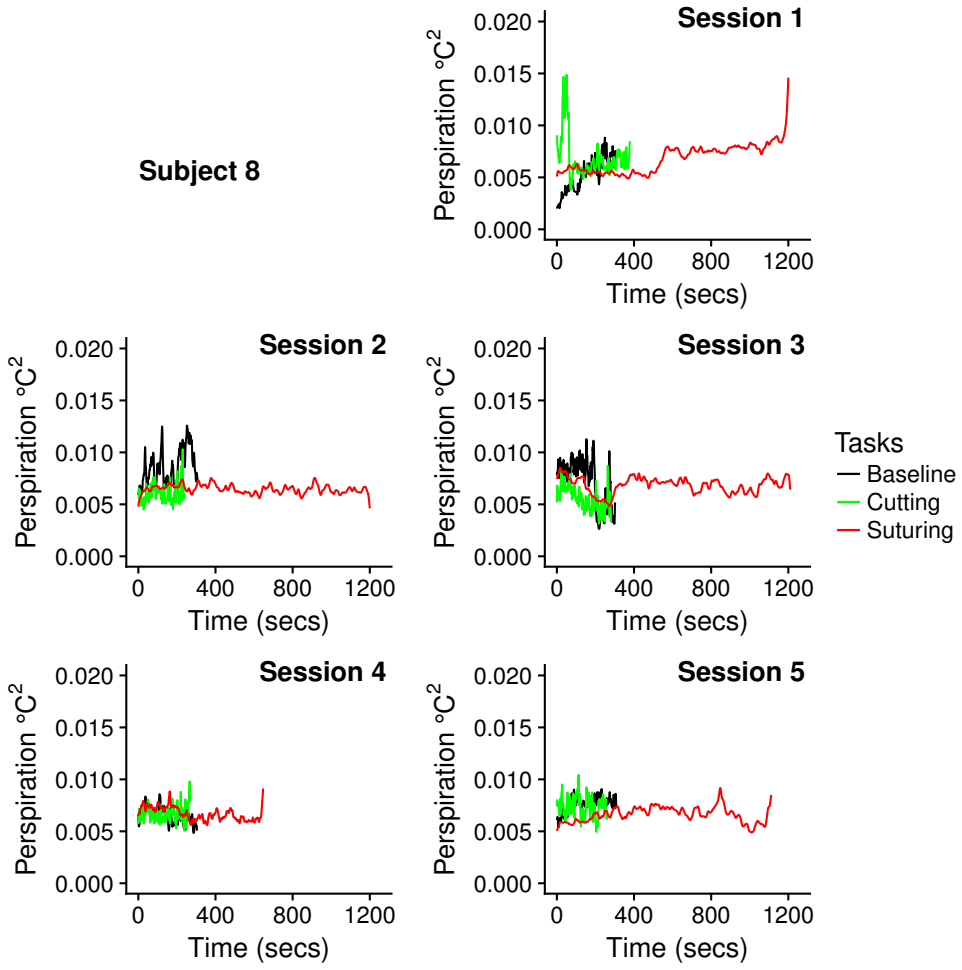

Subject 5
Session 1
Session 2
Session 3
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

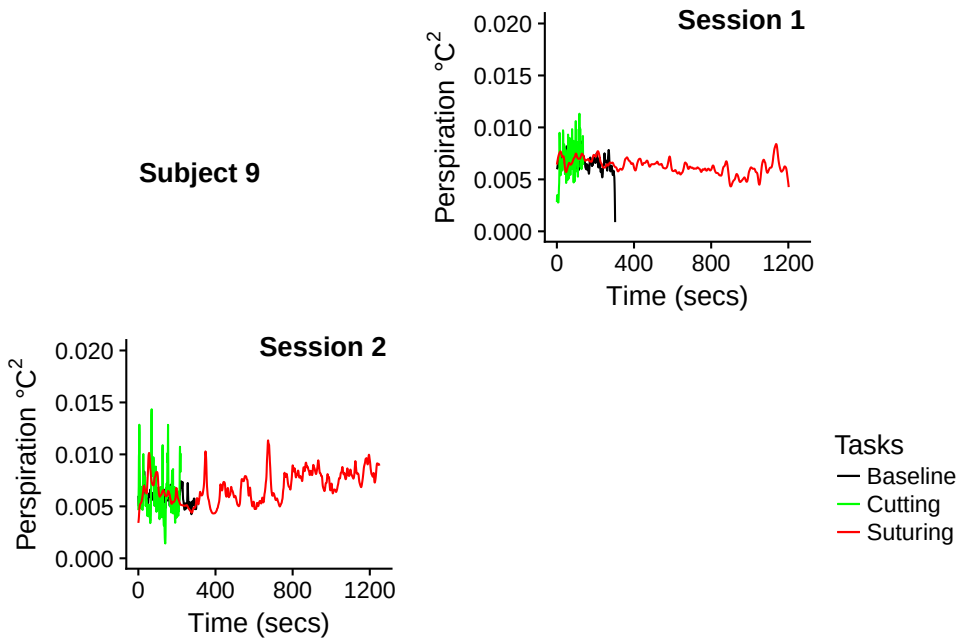
Subject 6

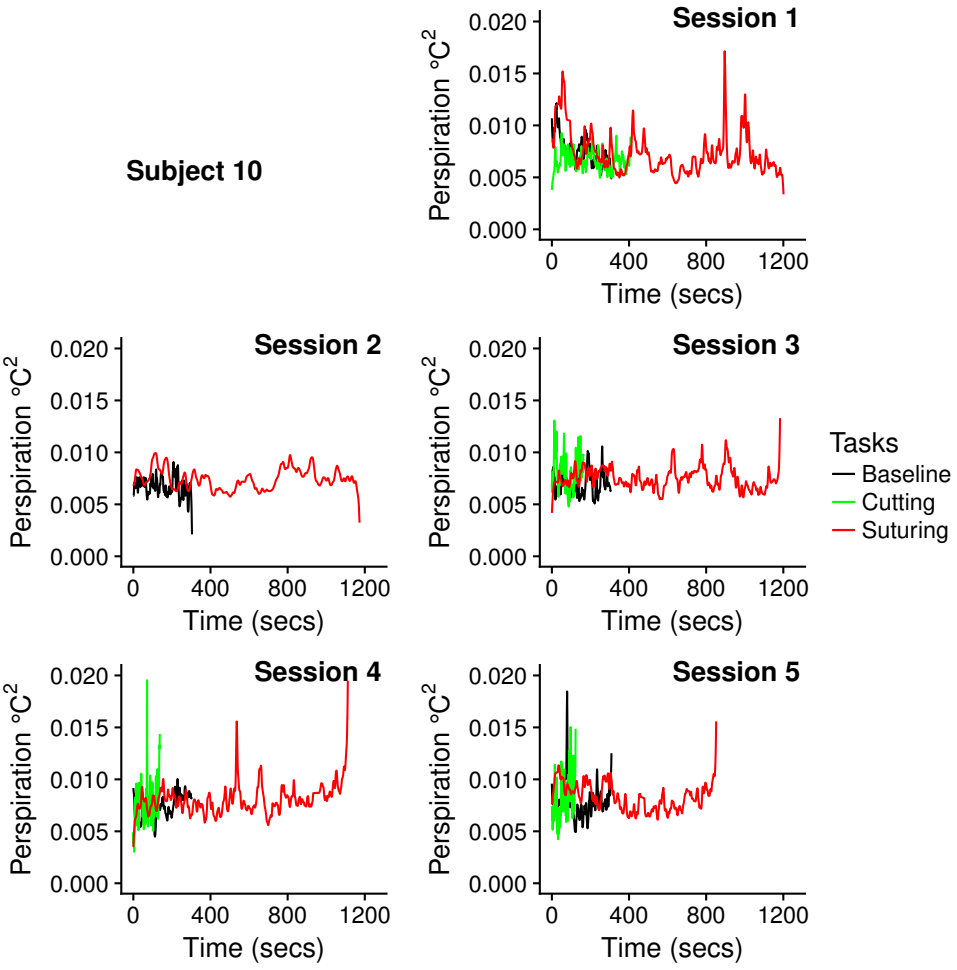


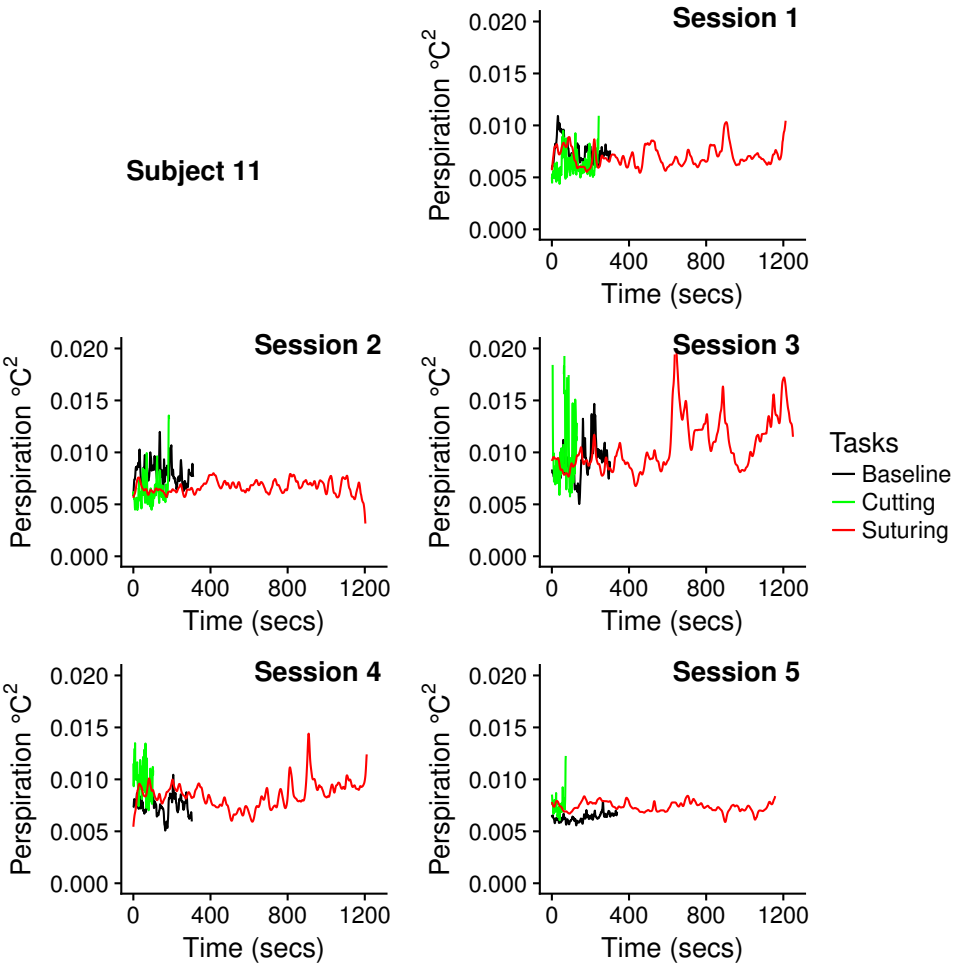
Tasks
— Baseline
— Cutting
— Suturing

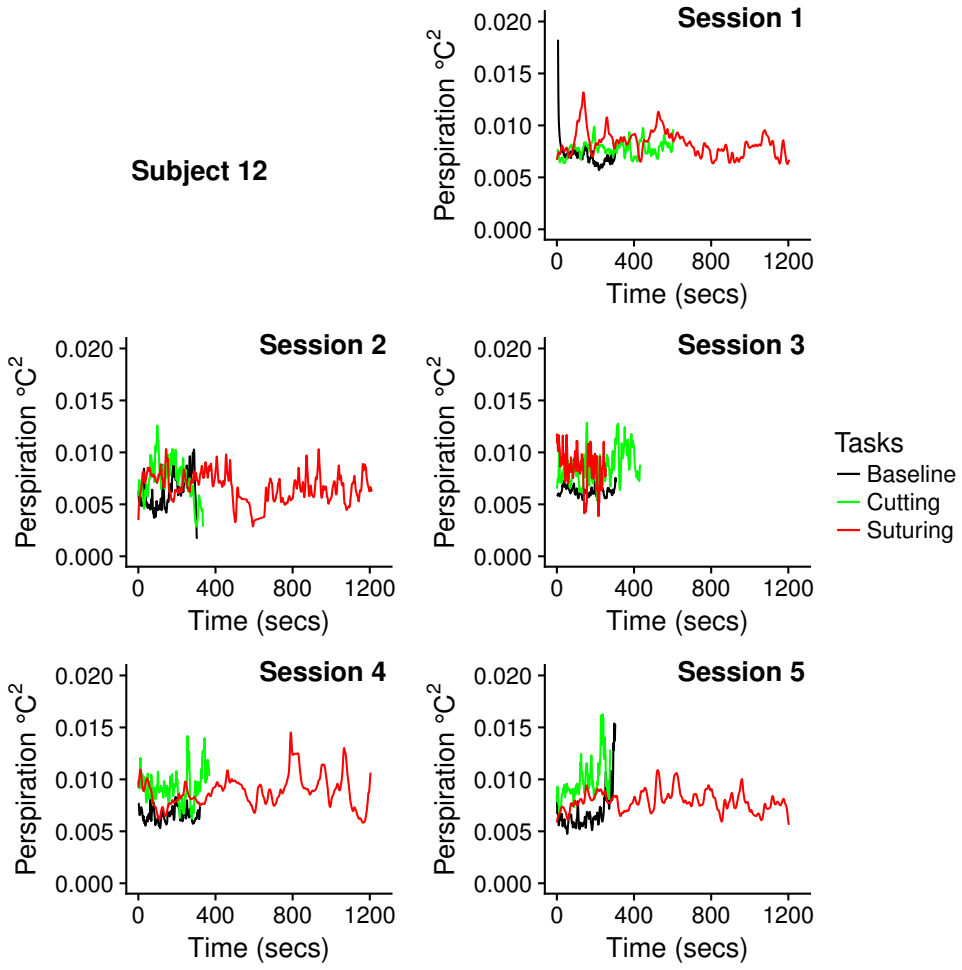

Subject 7
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

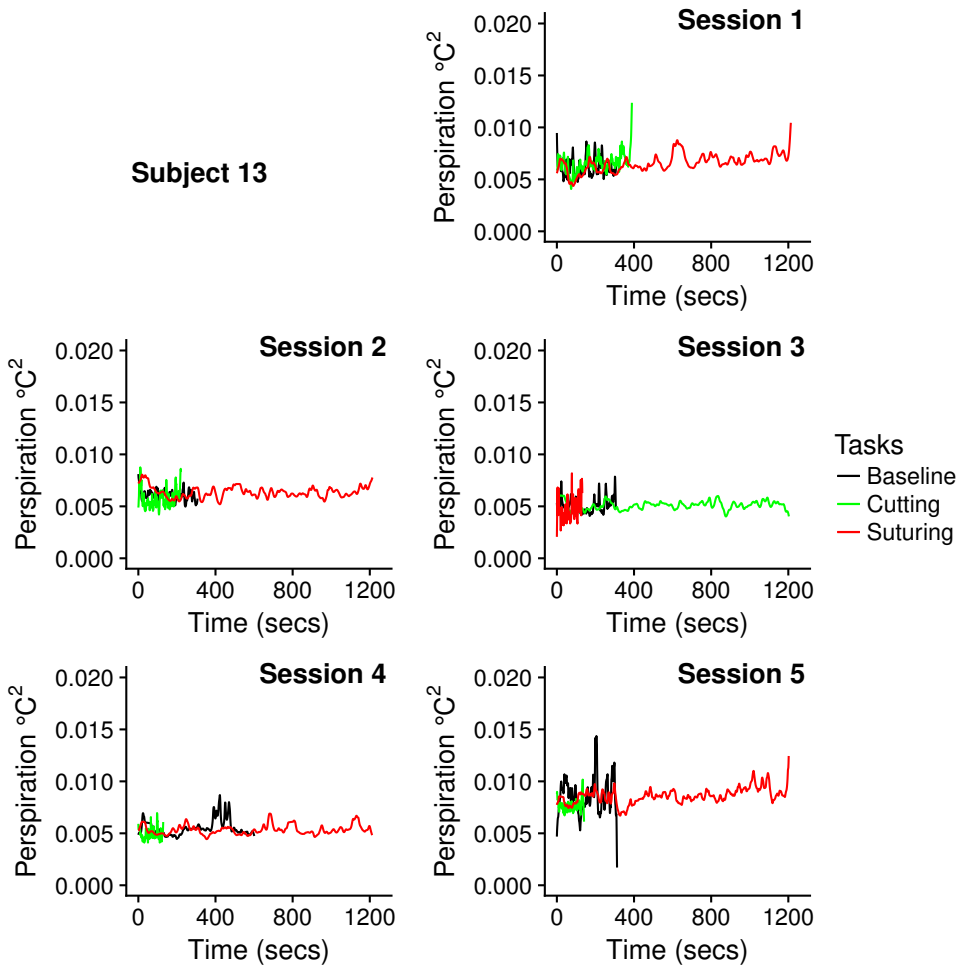

Subject 8
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

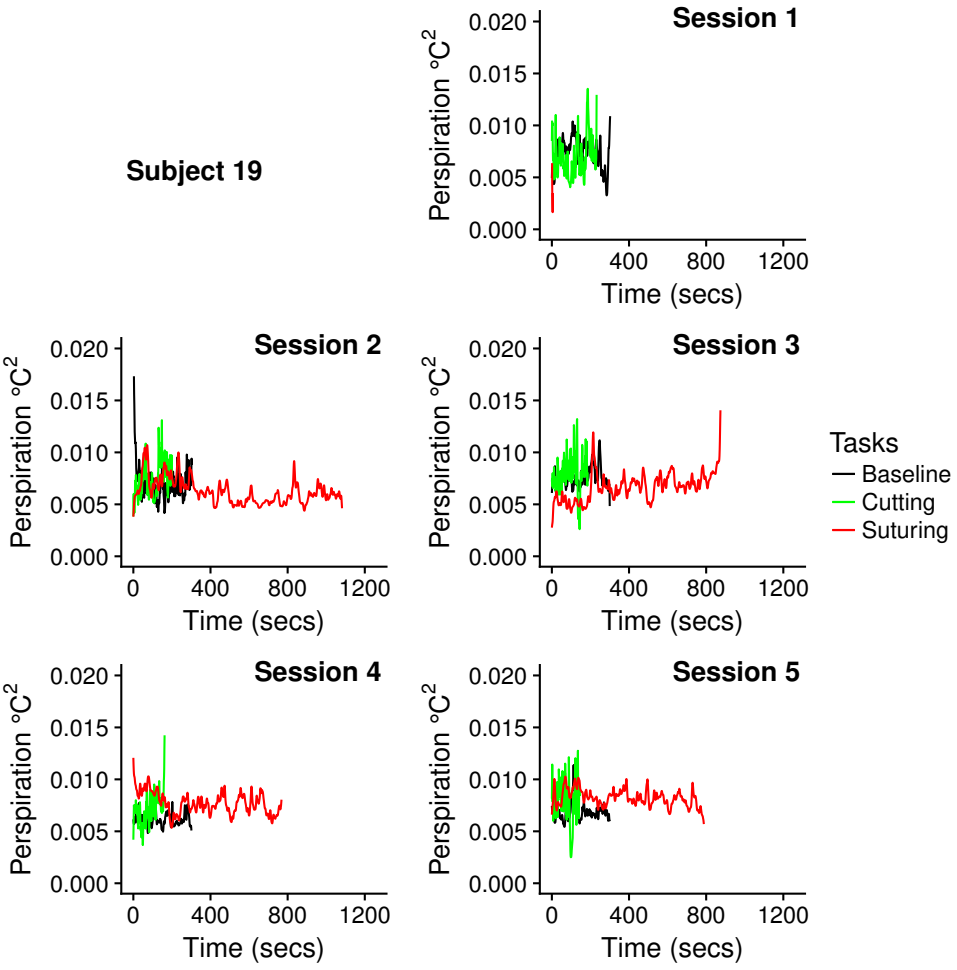

Subject 9
Session 1
Session 2
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 10
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

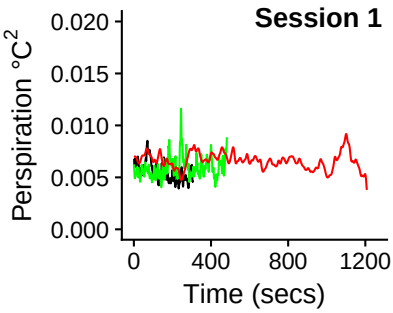

Subject 11
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 12
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

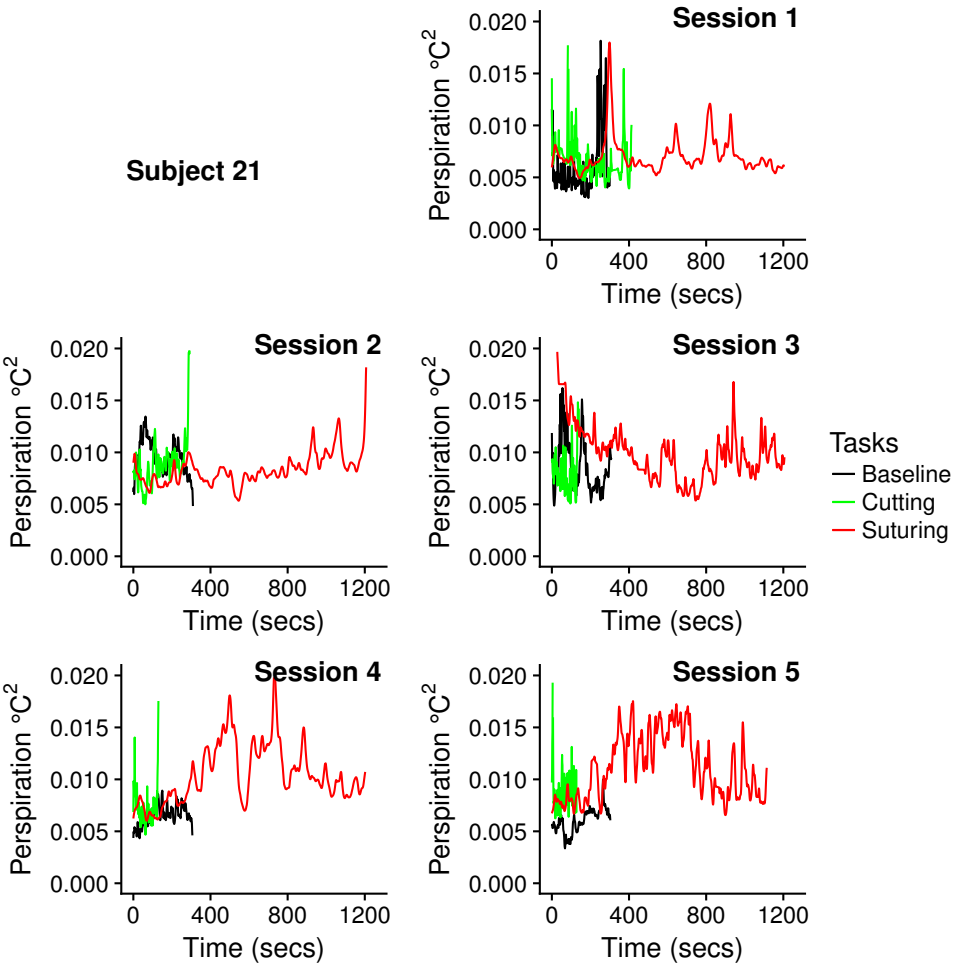

Subject 13
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

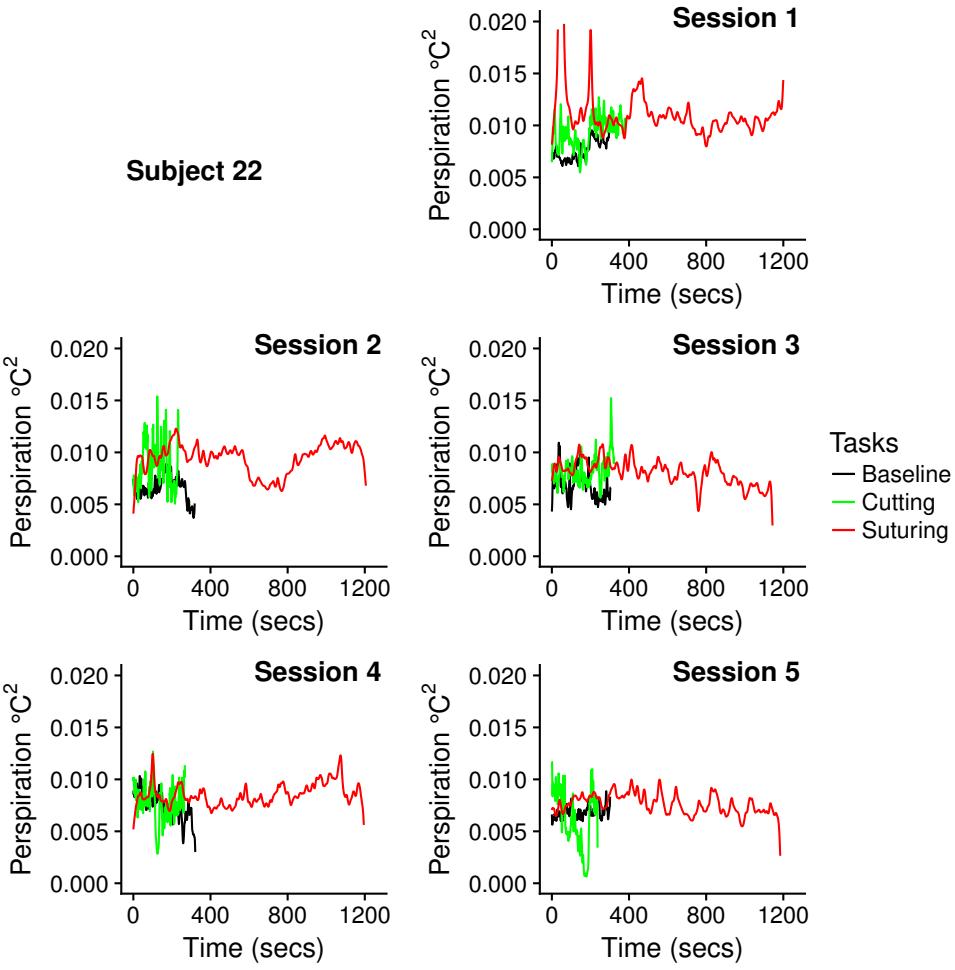

Subject 19
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

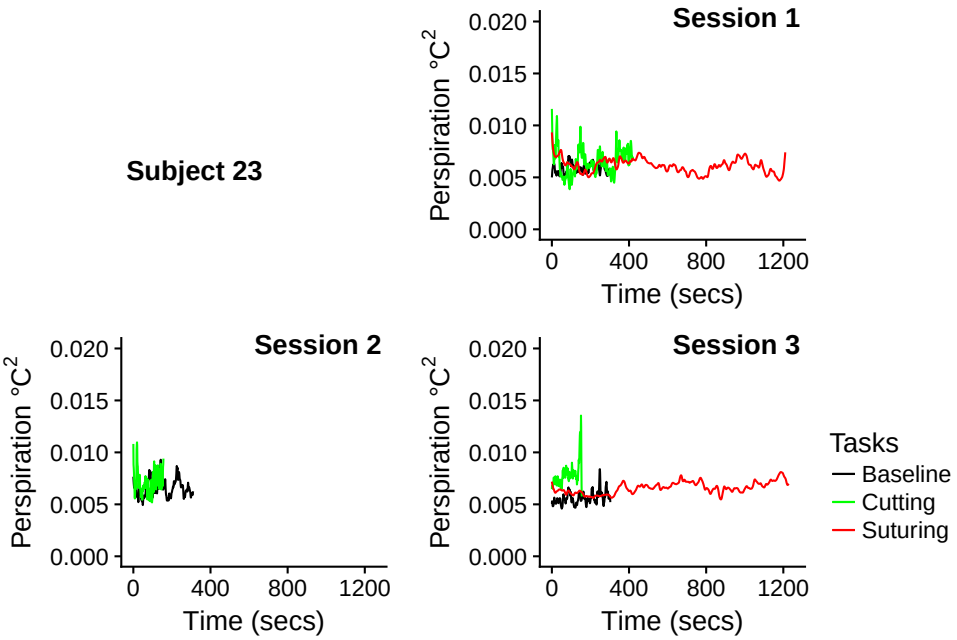
Subject 20

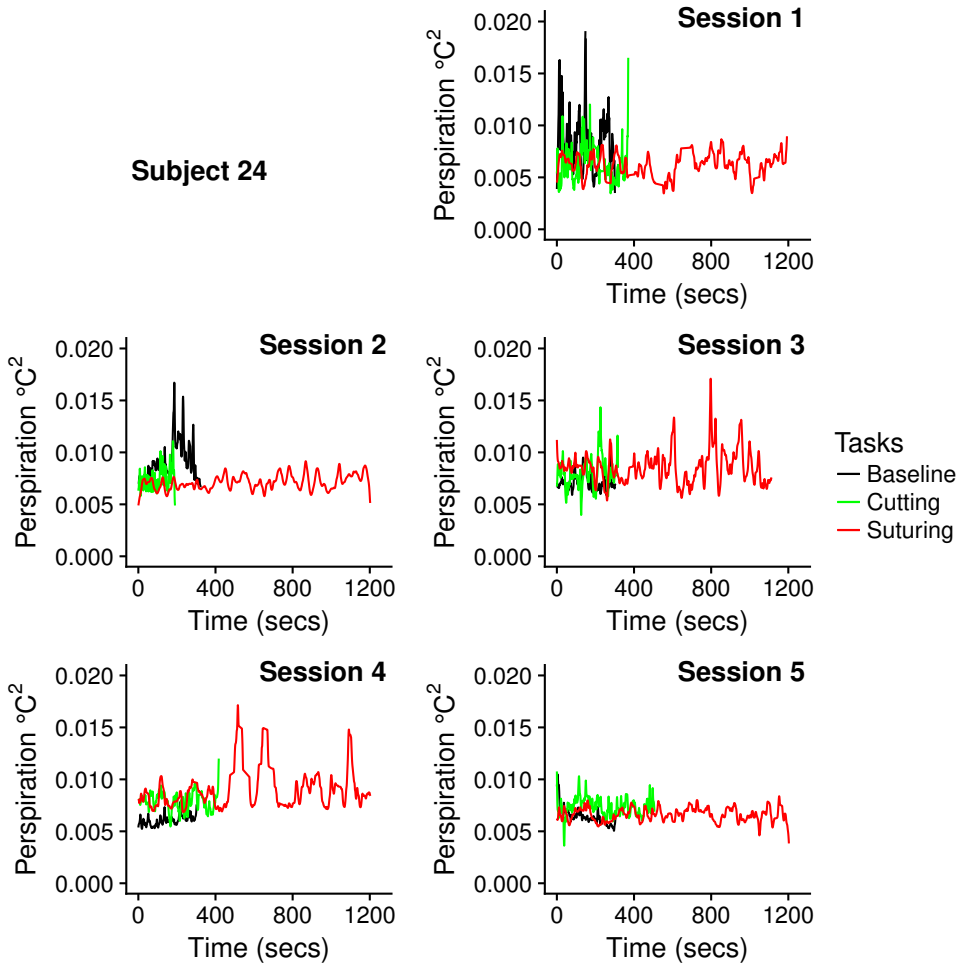


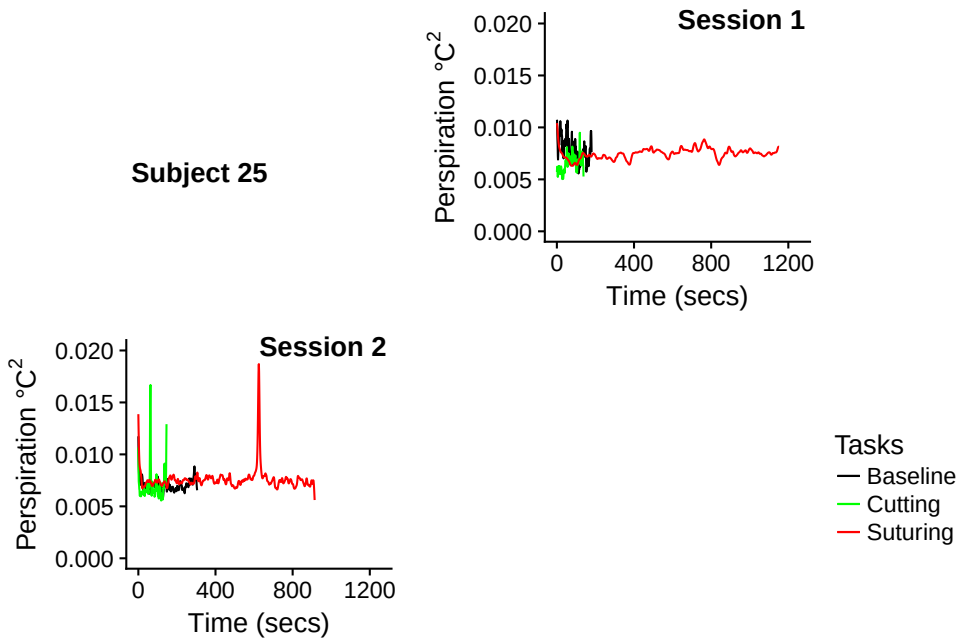
Tasks
— Baseline
— Cutting
— Suturing

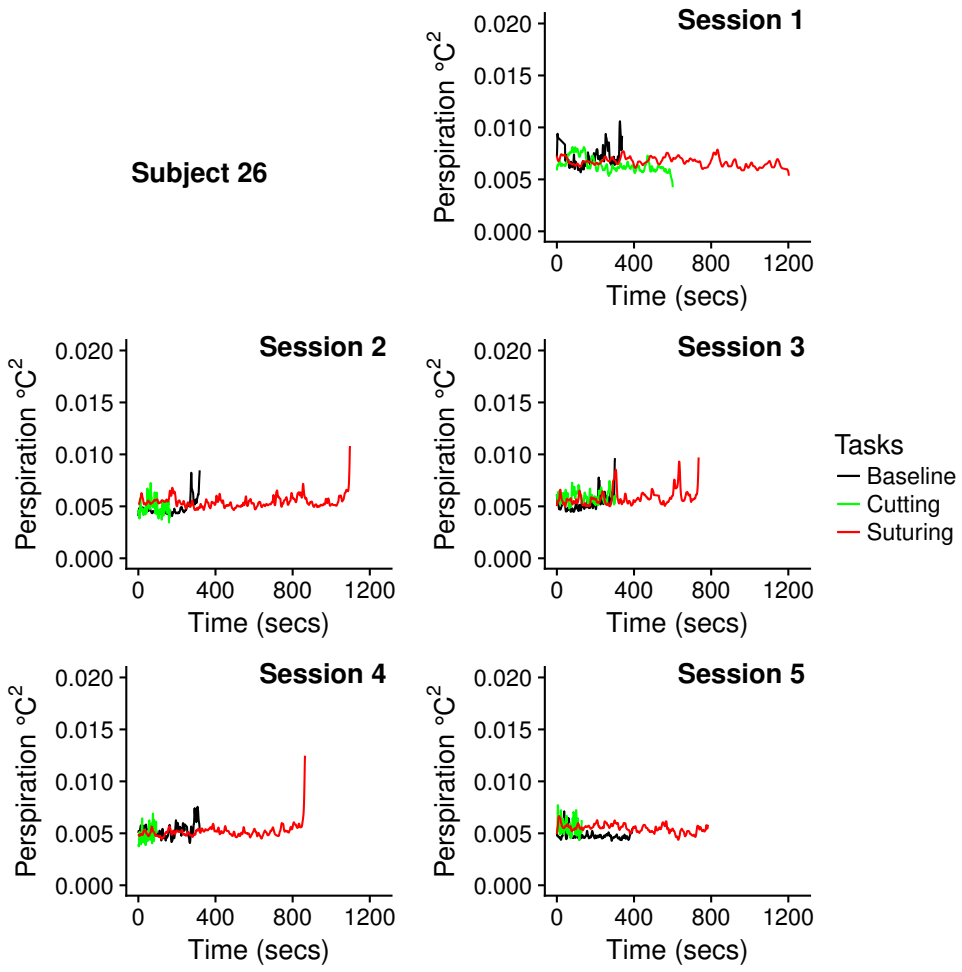

Subject 21
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 22
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 23
Session 1
Session 2
Session 3
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 24
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 25
Session 1
Session 2
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing


Subject 26
Session 1
Session 2
Session 3
Session 4
Session 5
Perspiration °C2
0.020
0.015
0.010
0.005
0.000
0
400
800
1200
Time (secs)
Tasks
Baseline
Cutting
Suturing

From these above plots we can have the following analysis-:
(i) Cutting:

After adjusting the session's average cutting by the session's average baseline we can observe some general trends in the graphs. Thirteen of the subjects had a higher average perspiration in the last session (session five) compared to their average perspiration in their first training session (session one). This tells us that a majority of the subjects were on average more stressed performing the cutting task in their last session than they were in their initial session. A good visual example of this generalization is the graph labeled "Cutting Session Average for Subject 21". Subject 21 had an average stress level of 0.00086 in session one, and after some improvements in the following sessions, he eventually saw an overall increase in stress level when comparing his initial average to the 0.00295 average in session five. There are many other subjects that follow a similar pattern as subject 21.

Another interesting observation revealed by the graphs is that some subjects had a consistent level of stress while performing the task while others had what seem as inconsistent jumps in their levels of stress as measured by perspiration. It seems like 8-10 subjects had a very consistent average (whether upward, downward trending or a mix of both) through all their training sessions. A example of this observation can be seen in subjects 4, 13, and 19.

We have performed the mean Perinasal Perspiration (Stress) Signal of all the subjects for Suturing task. The plots are from Figure 35 to Figure 38. From these plots we can analyze as following for the for Suturing -:

(ii) Suturing:

Overall, it seems that the suturing task was more difficult than the cutting task since a small number of subject (four of them) showed a decrease in their levels of stress during their last training session when compared against their first training session. The remaining subjects remained close to their initial level or showed an increase in their levels of stress as measured by the average perspiration. For an example please see "Suturing Session Average for Subject 8". Similarly to the cutting task there were several subjects that showed very consistent averages through all five of their training sessions. Some subjects showed mixed signs of improvements in their stress levels but overall these improvements were eventually loss in later training sessions. Please see "Suturing Session Average for Subject 3" as an example of this observation.

5. Performance Data

(a) Time

In order to analyze the time taken by the subjects to perform cutting and suturing tasks, we have drawn the time barplots for all 15 subjects. We have used Green color for Cutting and Red for Suturing.

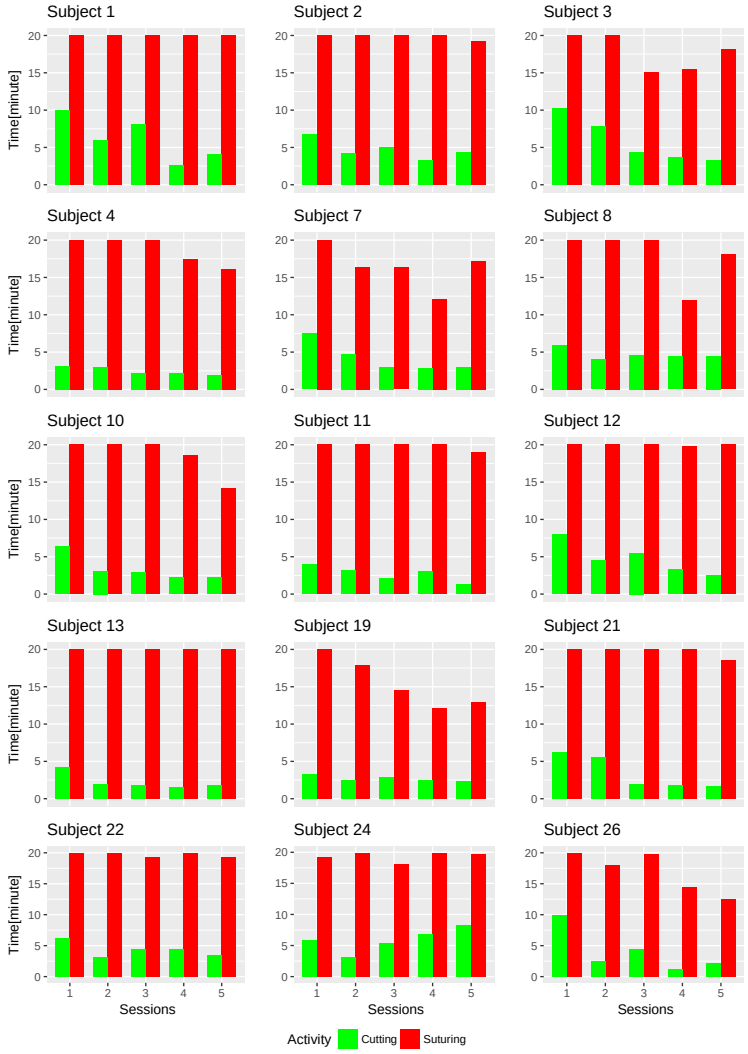


Figure 26: Cutting and Suturing Time for 15 subjects for 5 sessions

(i) Suturing-:

From Figure 31 and Figure 32, we can say that for almost all the subjects, the time taken by the subjects to perform suturing is constantly same and high from session 1 to session 5. Subject 4, 10 and 26 seem to improve their performance after session 3.

With the exception of subject 19, the time taken for suturing decreases with progress in all 5 sessions. Thus, we can say that this subject was getting better at the task of suturing and was therefore able to finish the job earlier.

(ii) Cutting-:

From Figure 39 and Figure 40, we can say that for almost half of the subjects the cutting time remains almost the same throughout all 5 sessions. For the other half there does not seem to be much change in cutting time.It remains almost the same all throughout the sessions. For subject 3,7 and 21 the cutting time decreased s they moved from session1 to 5. With the exception of subject 24 the time taken for cutting increases with progress in sessions. Thus, we can say that she was not getting better at the task of cutting.
Overall we don't see much improvement in performance time of all the subjects as they move from session 1 to 5.

(iii) Comparison of cutting ad suturing time
For all the subjects the time taken to perform suturing is way more than than time taken for cutting.
Thus we can say that the subjects found cutting easier in comparison with suturing, as they did cutting faster than suturing.

(b) Accuracy

In order to check the performance of the subjects we have also checked the accuracy of the subjects. We have considered two accuracy variables: Score and Number of Sutures.

(b.1) Number of Sutures

For all the subjects the allotted time for suturing was 20 minutes.
We have used barplots to analyze the number of sutures by each subject. These bar plots are shown in the Figure 41 and Figure 42-:

Figure 27: Number of Sutures for 15 subjects for 5 sessions

From Figure we can say that for almost all the subjects, the number of sutures are either increasing as they are progressing from session 1 to session 5 or remaining constantly good as they progress from session 1. We can say that performance of the subjects got better after their first session. With the exception of subject 13, his performance decreased after the 3rd session. Also, for subject 24, the performance was consistently good from session 1-3, but reduced in session-4 and increased again in session 5.
Overall mostly, we can see an upward trend in number of sutures for almost all the subjects as they move from session 1 to 5. Thus, we can say that with time and experience the subjects got confident and did well in their tasks of suturing.

(b.2) Score

Each subject was given scores for their tasks performed in all the sessions. These scores were given by 2 scorers.
We have use barplots to analyze the score received by each subject for the 2 tasks from both the scorers. These plots are shown in the Figure 43 and Figure 44-:

(i) Cutting

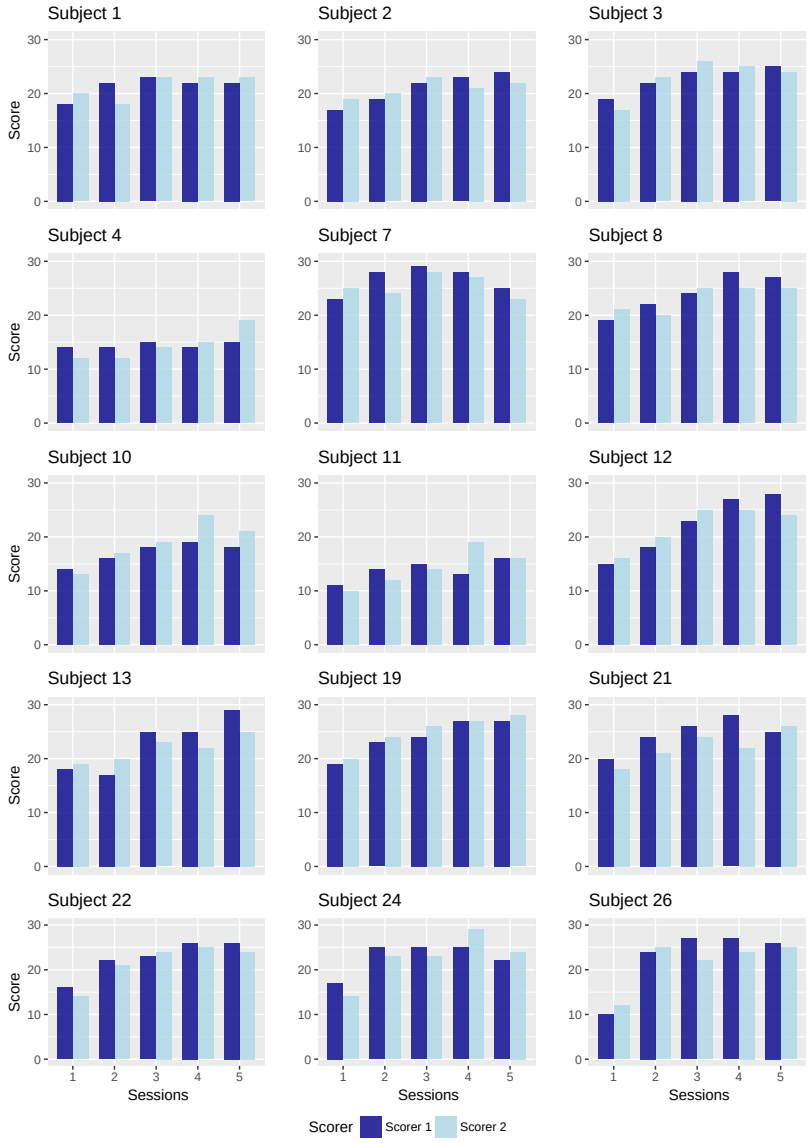


Figure 28: Cutting Score for 15 subjects for 5 sessions

From this plot we can say that the for almost all the subjects the scores given by both the scorers vary slightly. They do not seem to agree on the performance of the subjects.

Also, we can say that for almost all the subjects the scorer-1 seems to give slightly more more score than scorer-2.

Also, we can say that for almost all the subjects as they moved from session 1 to session 5, the performance either improved or remained consistently well towards session 3,4 and 5. We can say this from their increase in scores with progress in sessions.
With the exception of subject 4, he didn't seem to improve much as his scores remain almost the same for all 5 sessions. Also, for subject 11, his performance decreased after 3rd session. Overall mostly, we can see an upward trend in scores for almost all the subjects as they move from session 1 to 5. Thus, we can say that with time and experience the subjects got confident and did well in their tasks of cutting.

(ii) Suturing

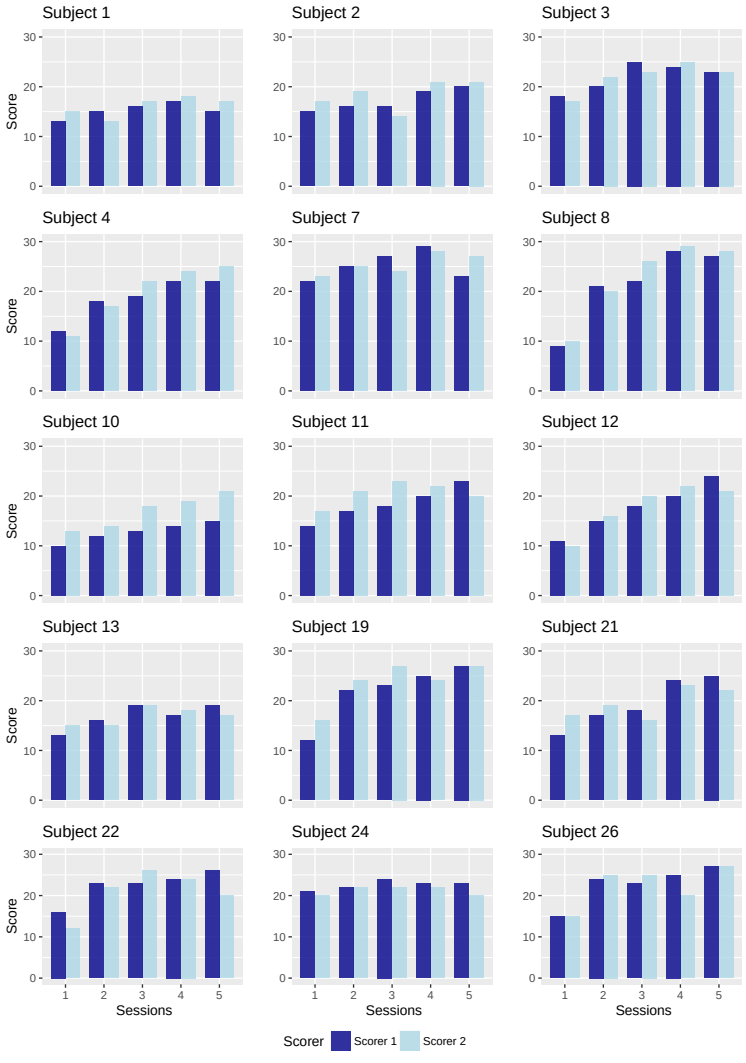


Figure 29: Suturing Score for 15 subjects for 5 sessions

From this plot we can say that the for almost all the subjects the scores given by both

the scorers vary slightly. They do not seem to agree on the performance of the subjects. Also, we can say that for almost all the subjects the scorer-2 seems to give slightly more more score than scorer-1. With the exception og subject 24 where for all 5 sessions, scorer-2 slighly less marks than scorer-1.

Also, we can say that for almost all the subjects as they moved from session 1 to session 5, the performance either improved or remained consistently well towards session 3,4 and 5. We can say this from their increase in scores with progress in sessions.
With the exception of subject 1 and 24, they don't seem to improve much as their scores remain almost the same for all 5 sessions. The score of 1st subject decreased after session 4. The score of 24th subject decreased after session 4. Also, for subject 7, his performance decreased after 3rd session.
Overall mostly, we can see an upward trend in scores for almost all the subjects as they move from session 1 to 5. Thus, we can say that with time and experience the subjects got confident and did well in their tasks of suturing.