Surprise Study Pilot 10 Analysis

Marjan Biria

Study description

Pilot 10: Same version of the task as pilot 9 (prediction, bigger feedback, replacing anxiety with nervous/uncomfortable) + video/audio

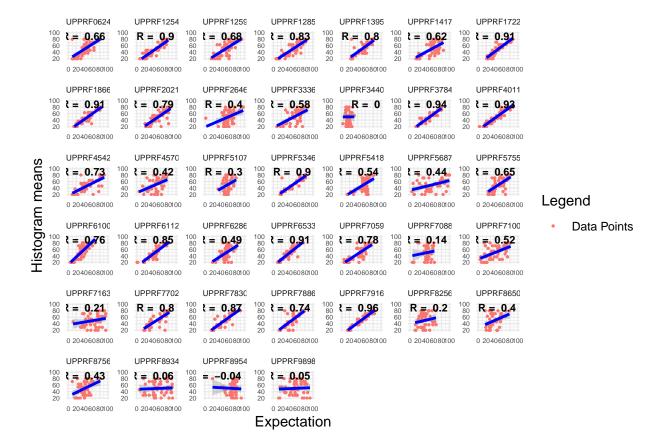
Goal: We used the task version from pilot 9, re-introducing the video. The task can be found here: https://app.gorilla.sc/admin/task/698788/editor

notes: 1) Gorilla does not allow recording audio and video files separately, so we have to always make sure "audio" is not ticked in the setting and we only choose the video. Video file will include the audio, otherwise, audio files will overwrite the video files. 2) The second issue is that video files are only 4s approximately despite the trials being 15s, Elena is in contact with the Gorilla support team to resolve this issue. Although it does not impact the current data, we do want to solve it for the actual experiment in case someone wants to analyse those video data in the future.

Histogram and prediction relationship

There were three people that ignored the histogram values and always chose similar values for prediction: "SUBPRF34408", "SUBPRF70880", "SUBPRF89542". I will do the analysis with and without them to see how they influenced the group correlations with anxiety and mood.

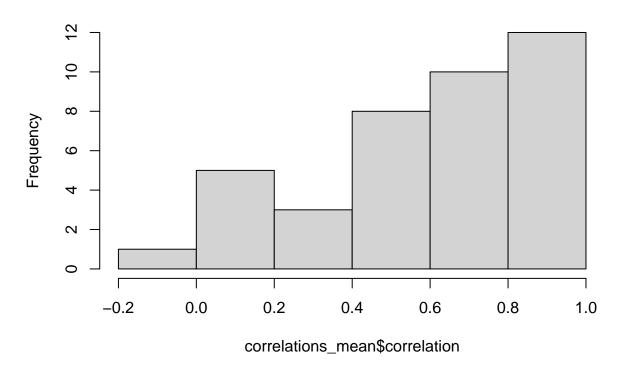
[1] "average correlation between prediction and m_hist: 0.591578772599636"



Below we can see the average correlation between hist_mean and prediction and the histogram of the individual correlations.

- ## [1] "average of correlations: 0.591578772599636"
- ## [1] "sd of correlations: 0.298916065846828"

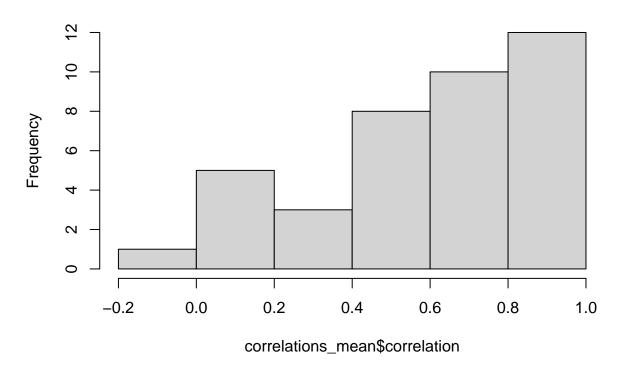
Histogram of correlations_mean\$correlation



I will remove those 3 people who ignored the histogram and gave similar predictions across all trials and repeat the correlation:

- ## [1] "average of correlations: 0.591578772599636"
- ## [1] "sd of correlations: 0.298916065846828"

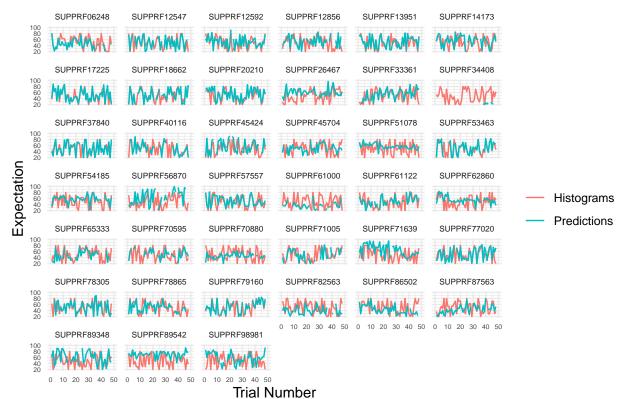
Histogram of correlations_mean\$correlation



Histogram and prediction across trials

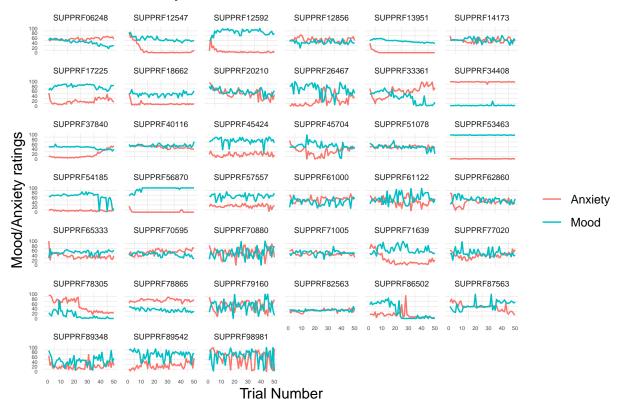
The figure below shows the histogram and expectation values over time and across trials.

Expectation across time



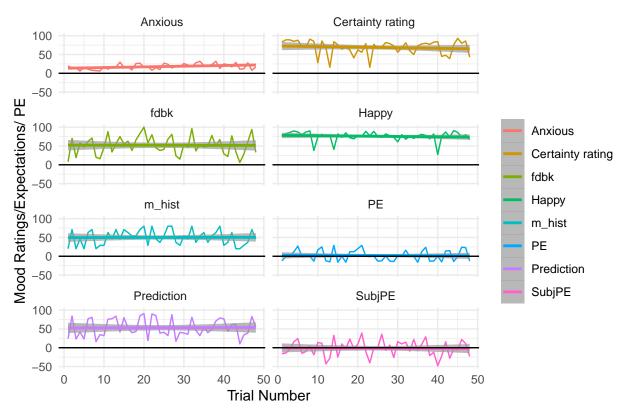
Anxiety and mood across trials

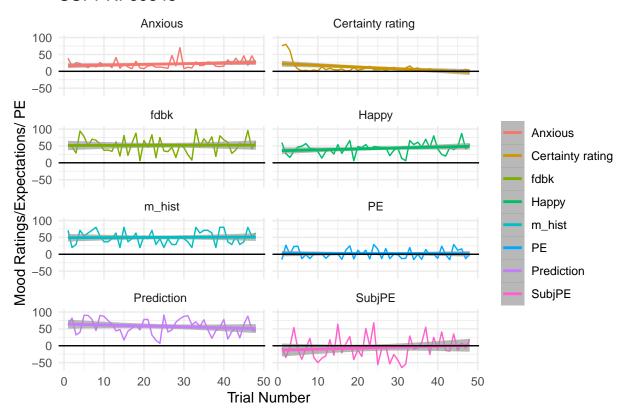
Mood and Anxiety Across Time

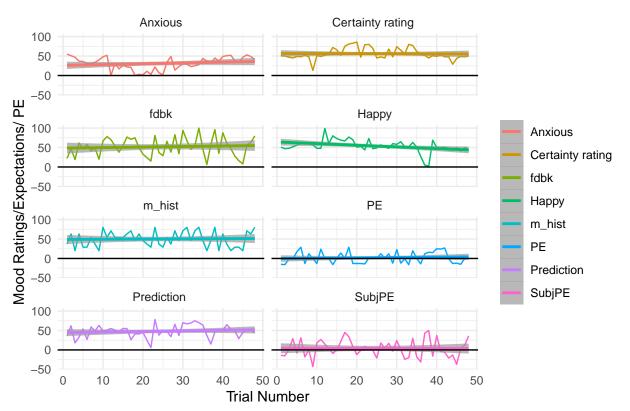


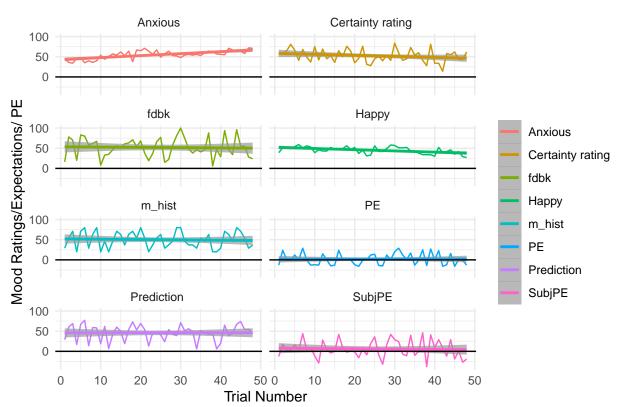
Plots for all variables per subject

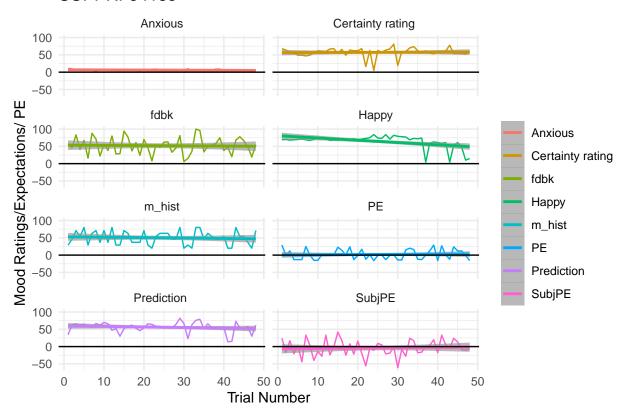
Now let's make plots for each subject that show how prediction, feedback, histogram mean, anxiety, mood, confidence rating, PE (feedback - histogram), subj_PE (feedback-prediction): one plot per subject. Two subjects had a flat 0 line for anxiety.

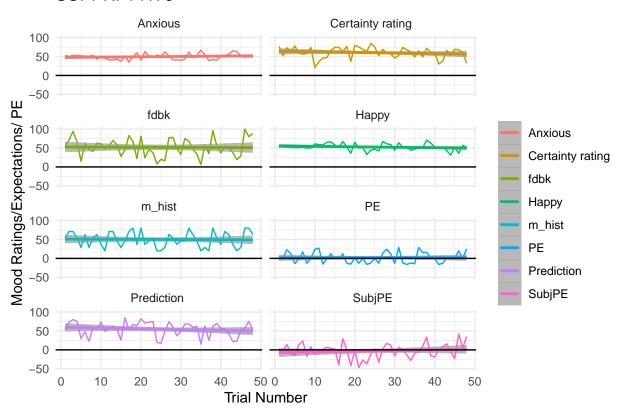


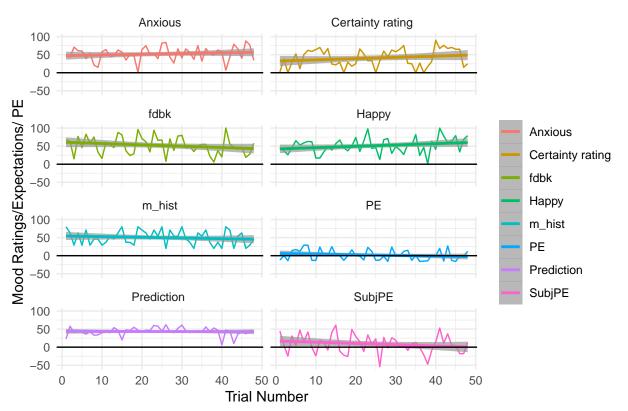


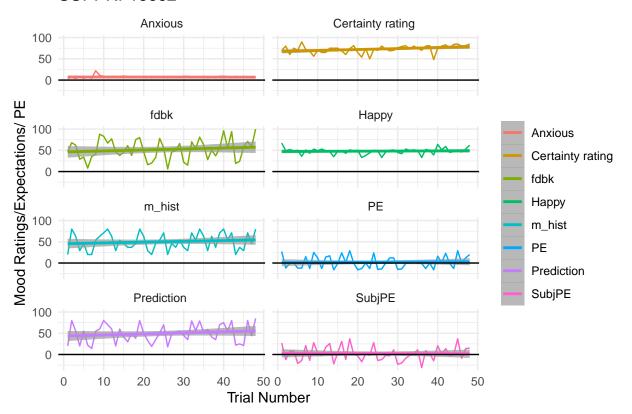


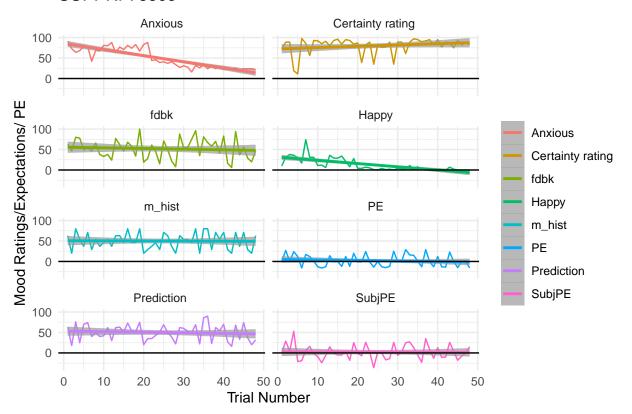


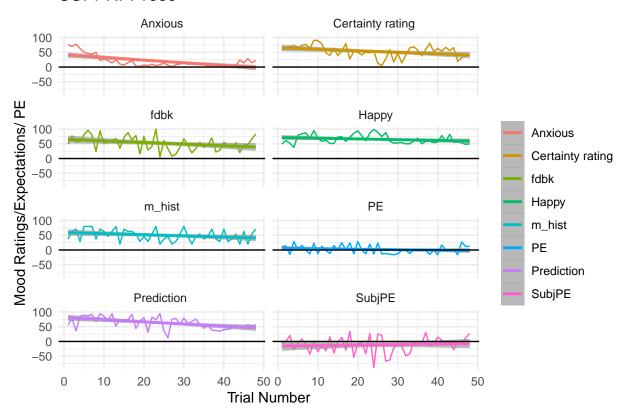


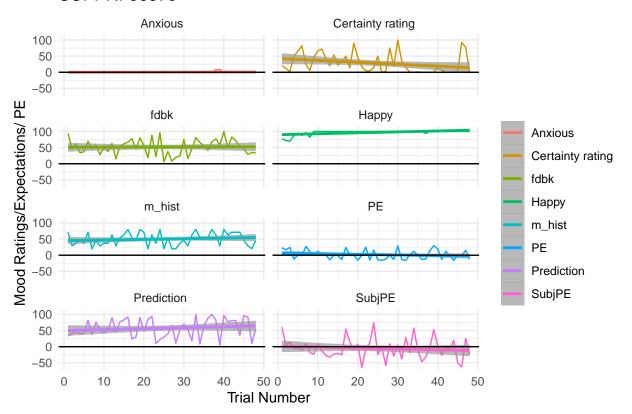


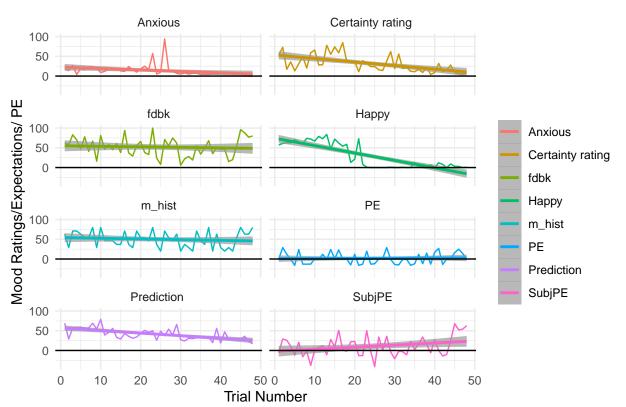


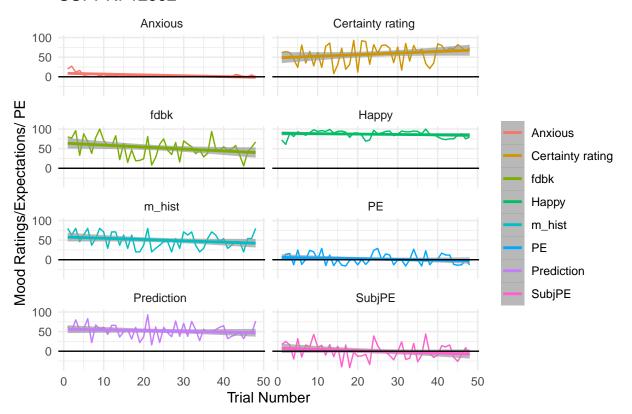


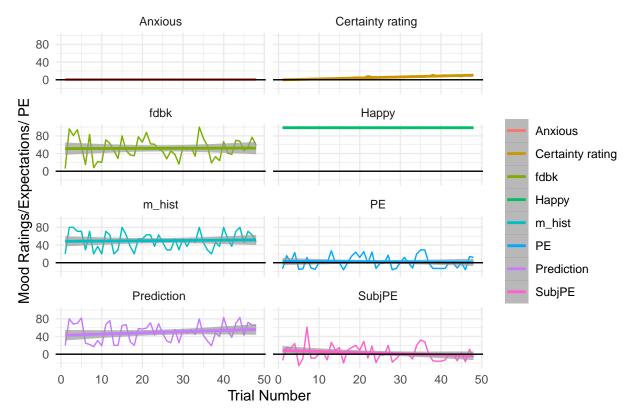


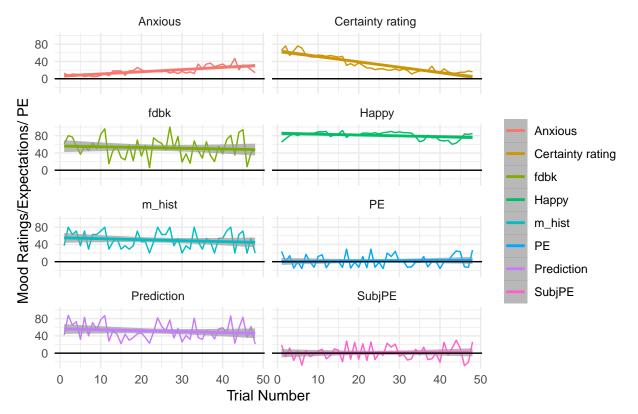


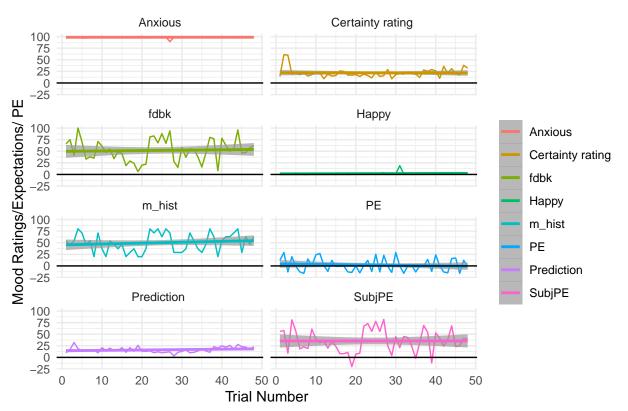


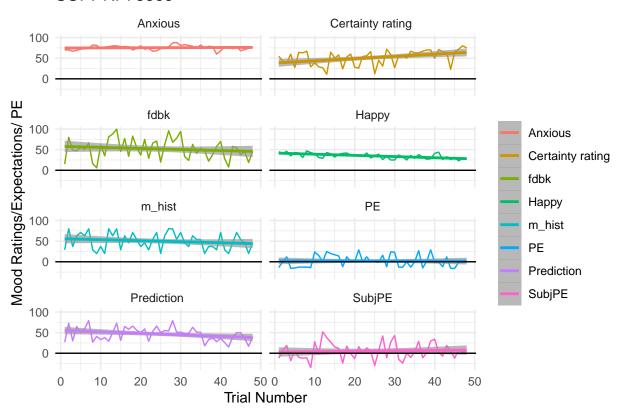


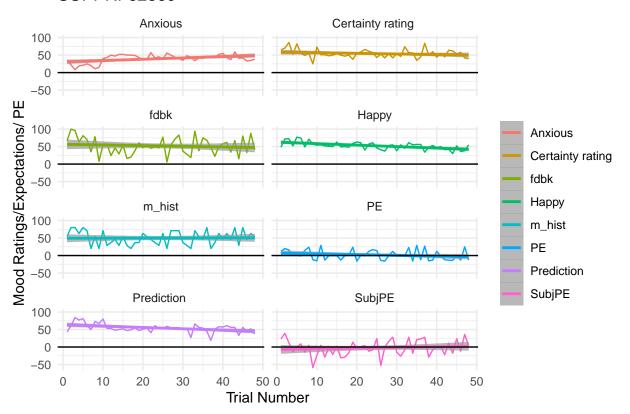


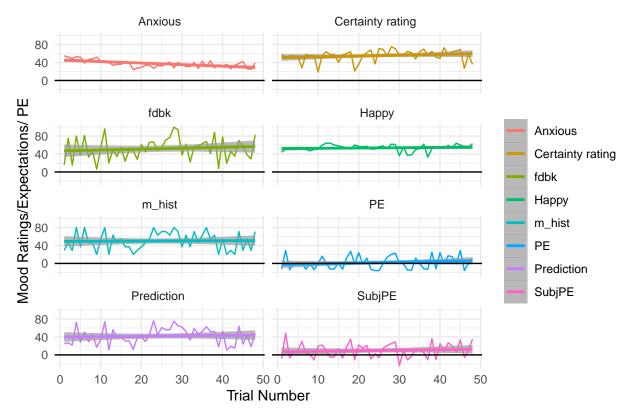


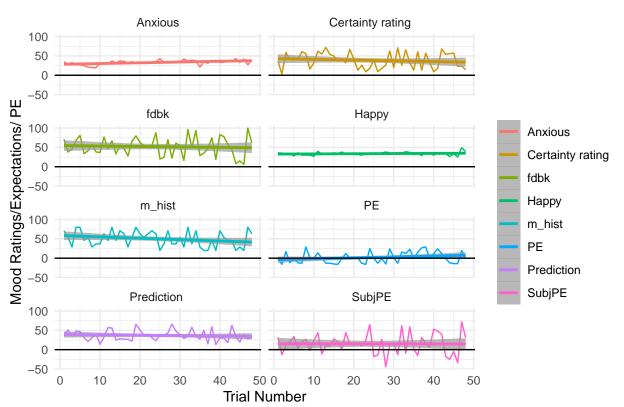


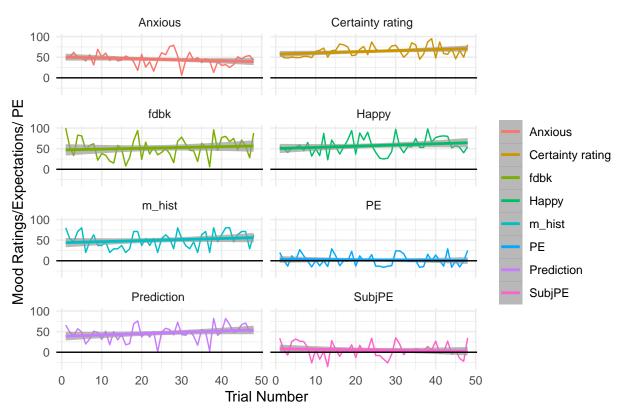


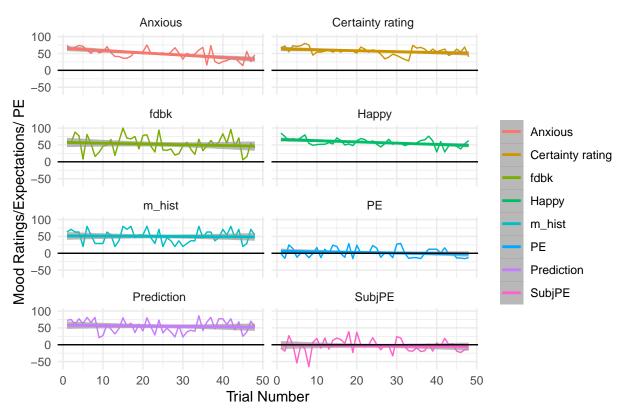


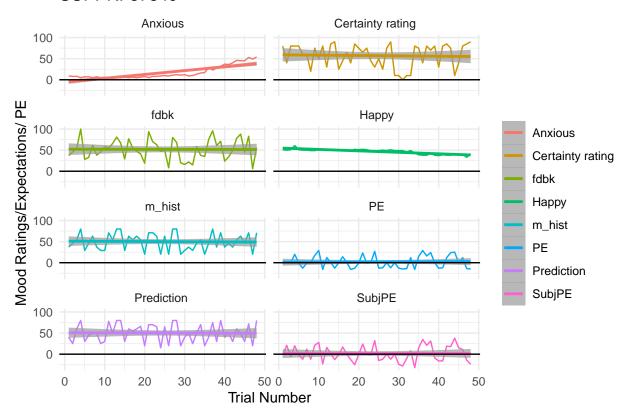


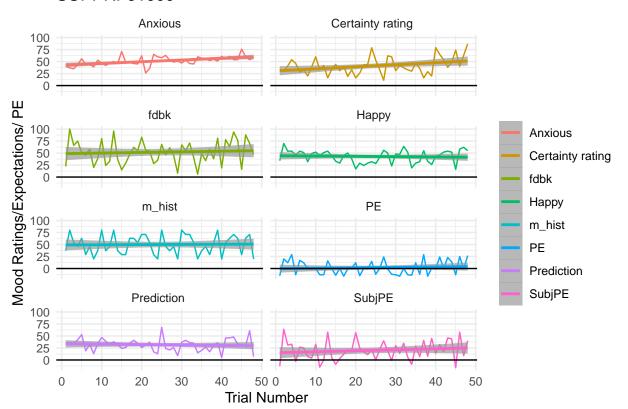


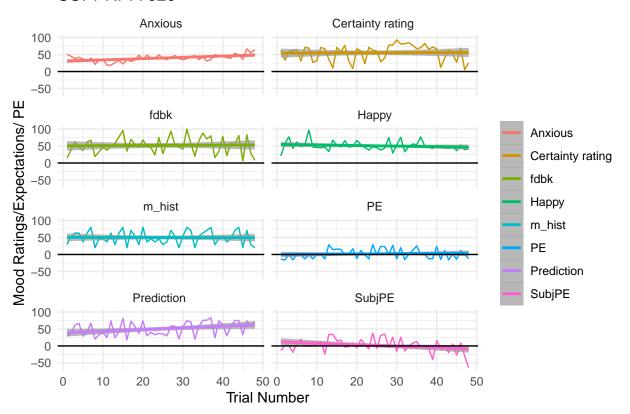


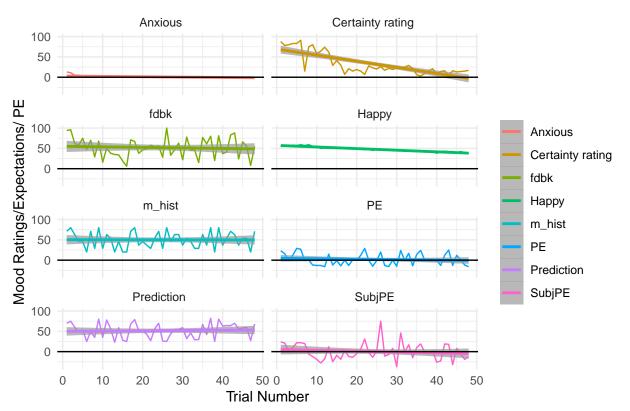


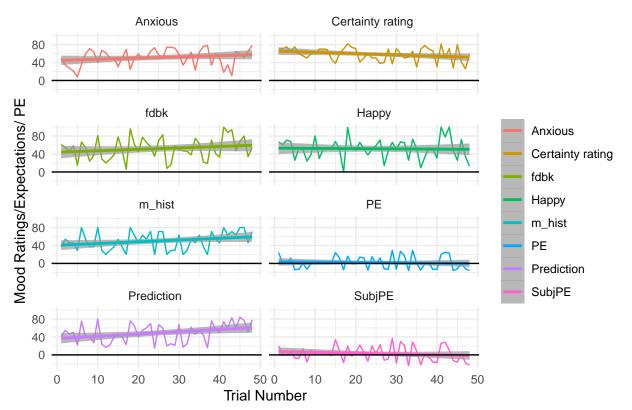


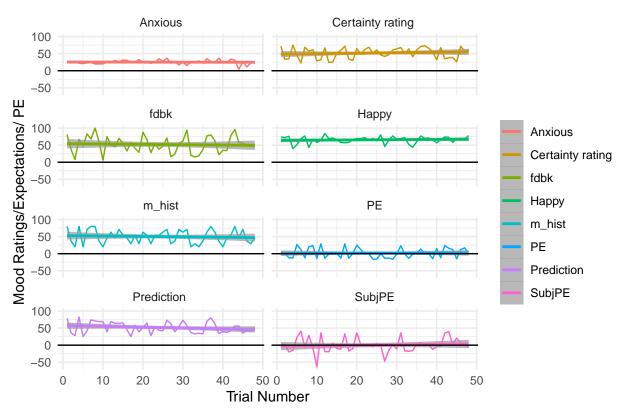


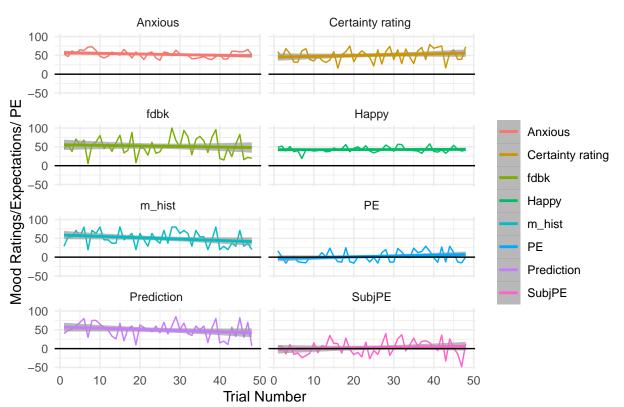


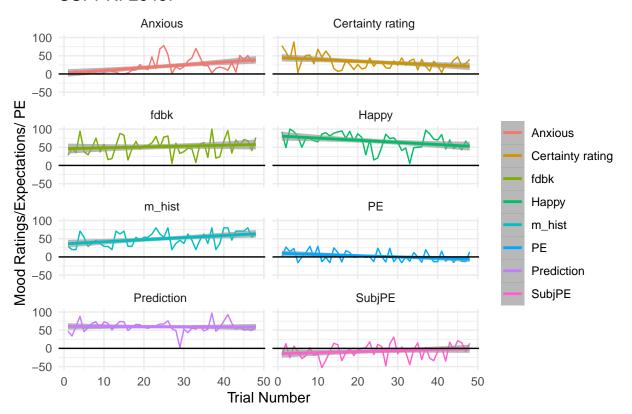


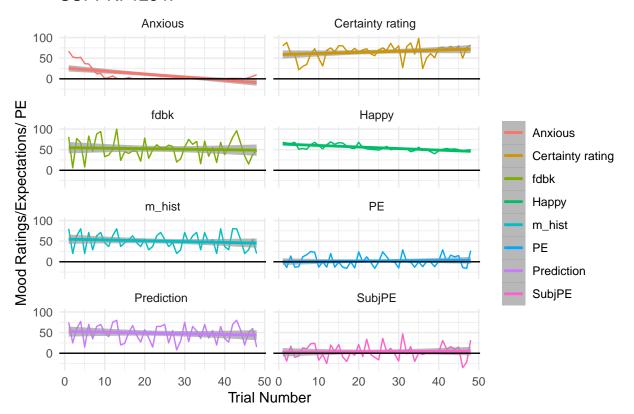


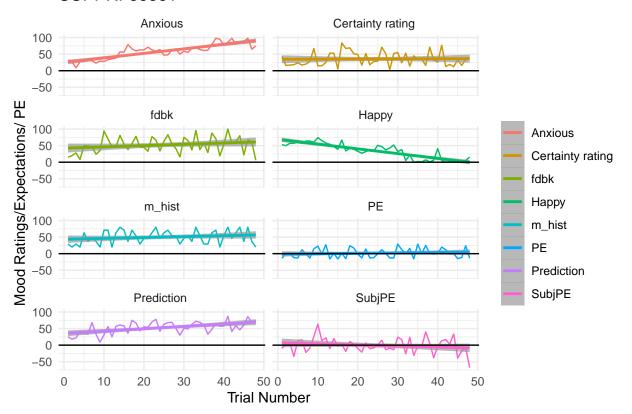


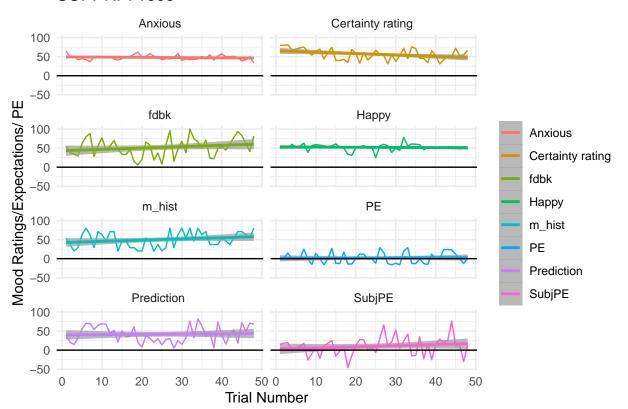


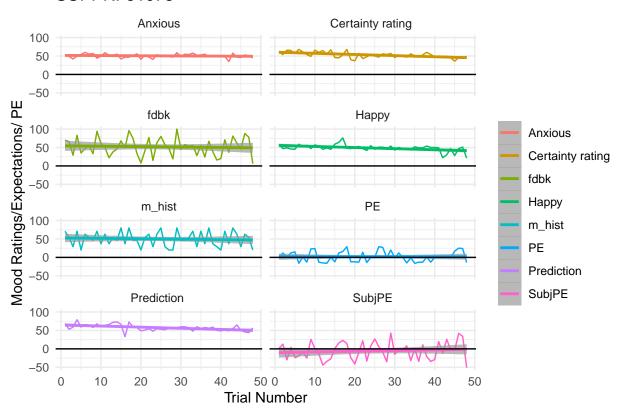


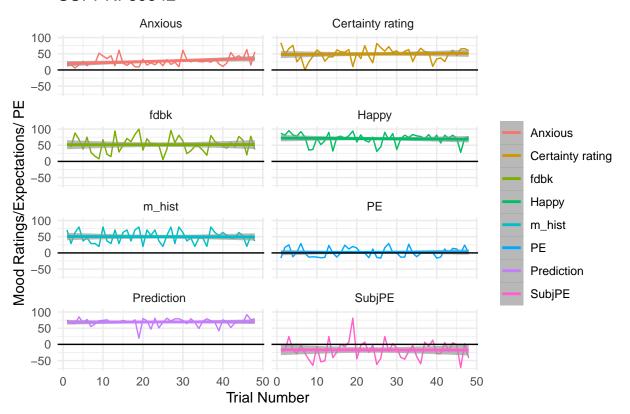


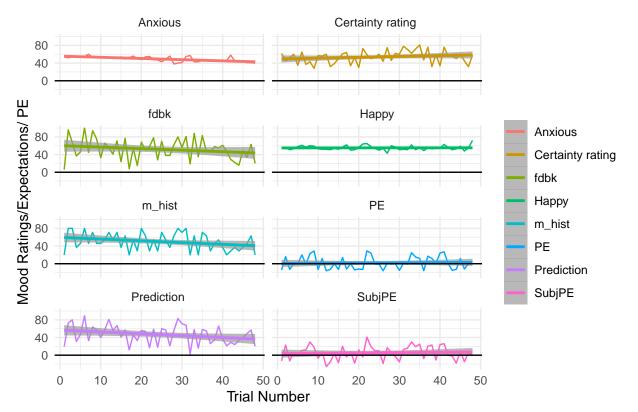


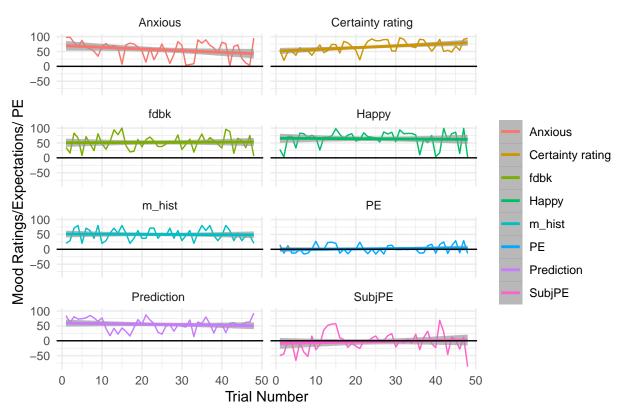


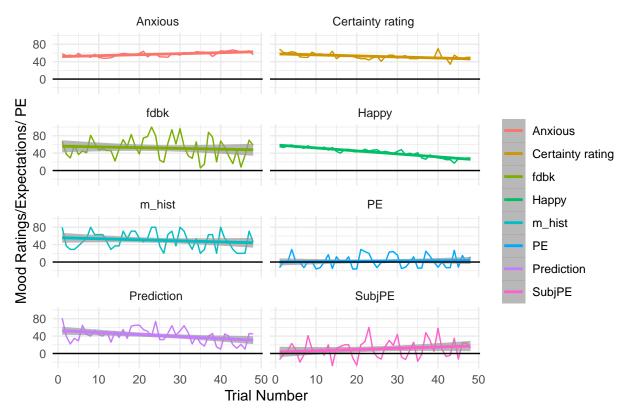


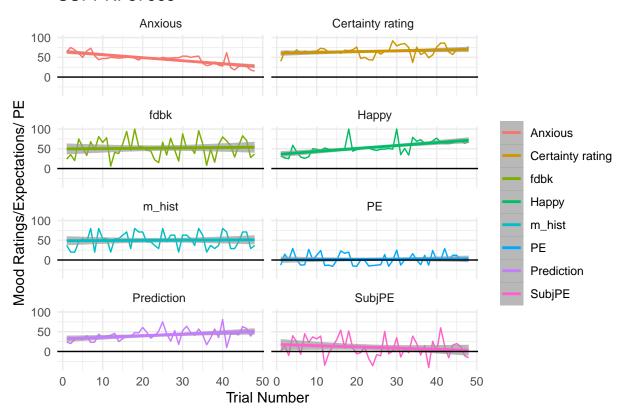








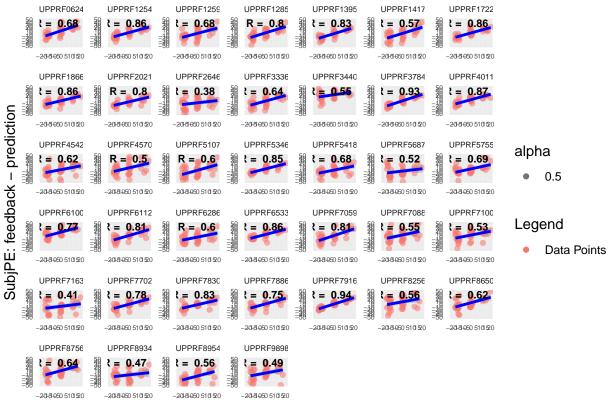




PE and SubjPE relationship

We now look at the relationship between PE (feedback - histogram_mean) and SubjPE (feedback - prediction). The correlation between SubjPE and objective PE remain similar to previous pilots.

[1] "average correlation between PE and SubjPE: 0.686357933098262"



PE: feedback - hist_mean

SubjPE and Anxiety relationship

Let's now look at the relationship between SubjPE and Anxiety ratings:

[1] "average correlation between Anxiety and SubjPE: -0.0891129101944432"



I will exclude subject "SUPPRF13951" and "SUPPRF34408", "SUPPRF53463", "SUPPRF54185", "SUPPRF56870" who rated always 0 or 100 for anxiety. After repeating the previous plot and correlations without them, the correlation becomes -0.10 from -0.09. I will add them back in.

[1] "average correlation between Anxiety and SubjPE: -0.100044677288319"



SubjPE and Anxiety

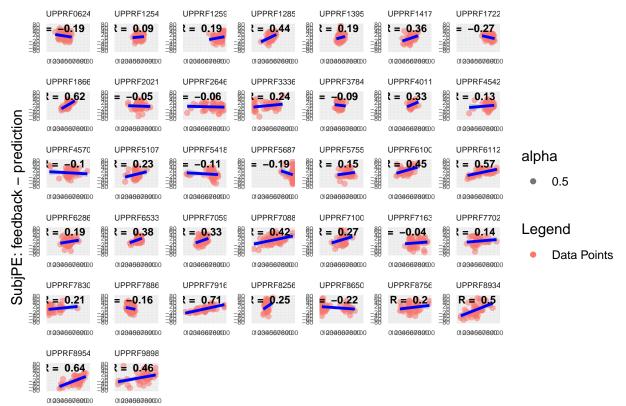
Let's look at the same relationship for mood and subjPE:

[1] "average correlation between Mood and SubjPE: 0.180251962174898"



I will remove subjects "SUPPRF34408" and "SUPPRF53463" who again rated 0 or 100. Let's look at the same relationship for mood and subjPE without these people. The correlation increases to 0.21 from 0.18.

[1] "average correlation between Mood and SubjPE: 0.194867474595861"



Mood

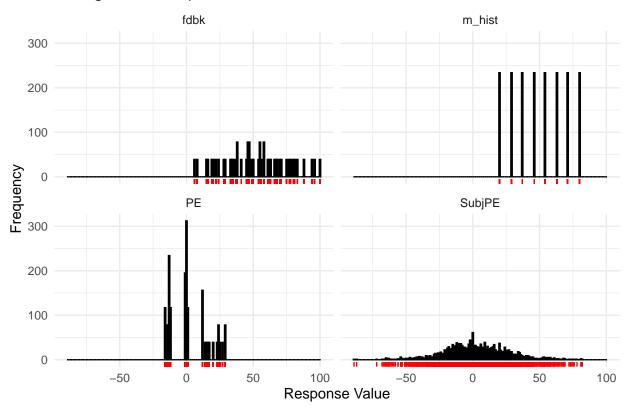
The correlation for mood is less than before (0.18/0.21 vs 0.26), and consistently higher than anxiety (which is the same as before despite changing the wording of the question). We now will look whether the average correlations are significantly different from zero for both anxiety and mood.

```
## [1] "corr Anxiety and SubjPE"
##
##
   One Sample t-test
## data: correlations_Ax_SubPE$correlation
## t = -2.8832, df = 33, p-value = 0.006879
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.17064167 -0.02944768
## sample estimates:
## mean of x
## -0.1000447
## [1] "corr happiness and SubjPE"
##
   One Sample t-test
##
## data: correlations_H_SubPE$correlation
## t = 4.4922, df = 36, p-value = 7.008e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.1068905 0.2828444
## sample estimates:
## mean of x
## 0.1948675
```

Histogram for PE, SubjPE, feedback, hist_m

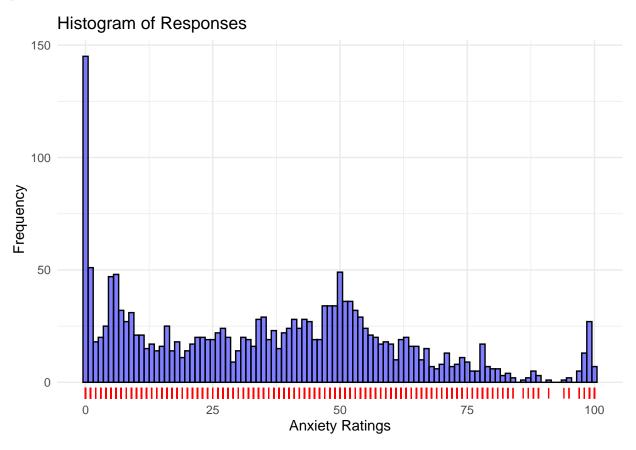
Let's have a look at histograms of SubjPE, PE, feedback and histogram means:

Histogram of Responses

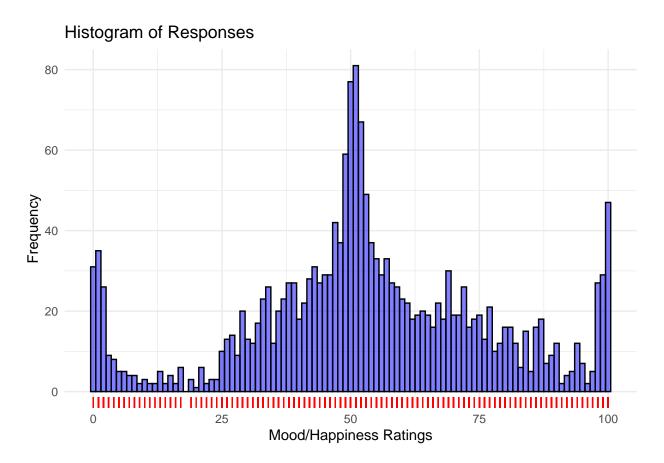


Histogram for anxiety ratings

We have again less variability in anxiety ratings, so the previous pilot was not because we changed the question!



Histogram for mood ratings



ICC and LME models for mood and anxiety

we will now look at the ICC outcome for anxiety The ICC is lower than the study without feedback (which was 0.80), it is moderate according to guidelines by Koo and Li (2016): below 0.50: poor between 0.50 and 0.75: moderate between 0.75 and 0.90: good above 0.90: excellent

```
## [1] "lmer for anxiety with just the intercept"
## [1] 0.7465389
                          97.5 %
                  2.5 %
## .sig01
               17.92998 28.10639
## .sigma
               12.62812 13.47283
## (Intercept) 28.32054 42.58438
## [1] 15120.9
## [1] 15113.18
## Data: final_df10
## Models:
## model1: Response_Ax ~ 1 + (1 | Random_ID)
## model2: Response_Ax ~ Response_PE + (1 | Random_ID)
##
                AIC BIC logLik deviance Chisq Df Pr(>Chisq)
## model1
             3 15125 15142 -7559.6
                                      15119
## model2
             4 15112 15134 -7551.9
                                      15104 15.583 1 7.895e-05 ***
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' 1
## [1] 15113.95
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + +(1 | Random_ID)
##
      Data: final df10
##
## REML criterion at convergence: 15105.9
##
## Scaled residuals:
##
       Min
                1Q Median
                                3Q
                                       Max
## -4.0822 -0.4630 -0.0345 0.3571
                                   6.2550
##
## Random effects:
   Groups
                          Variance Std.Dev.
                                   22.62
  Random_ID (Intercept) 511.7
   Residual
                          168.6
                                   12.98
## Number of obs: 1872, groups: Random_ID, 39
## Fixed effects:
##
                   Estimate Std. Error t value
                   35.59988
## (Intercept)
                               3.63472
                                         9.794
## Response_SubjPE -0.05399
                               0.01359
                                        -3.973
##
```

```
## Correlation of Fixed Effects:
##
               (Intr)
## Rspns_SbjPE -0.010
## Data: final_df10
## Models:
## model1: Response_Ax ~ 1 + (1 | Random_ID)
## model3: Response_Ax ~ Response_SubjPE + +(1 | Random_ID)
         npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)
## model1
           3 15125 15142 -7559.6
                                     15119
## model3
            4 15112 15134 -7551.8
                                     15104 15.691 1 7.456e-05 ***
## ---
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' 1
## Data: final df10
## Models:
## model2: Response_Ax ~ Response_PE + (1 | Random_ID)
## model3: Response_Ax ~ Response_SubjPE + +(1 | Random_ID)
         npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)
##
           4 15112 15134 -7551.9
                                     15104
## model3
          4 15112 15134 -7551.8
                                     15104 0.1082 0
                 2.5 %
                         97.5 %
              17.92998 28.10639
## .sig01
## .sigma
              12.62812 13.47283
## (Intercept) 28.32054 42.58438
## [1] "lmer for anxiety with just the intercept"
## [1] 0.7465389
## [1] 15121.57
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + Response_fdbk + (1 | Random_ID)
      Data: final_df10
##
## REML criterion at convergence: 15111.6
##
## Scaled residuals:
##
      Min
              1Q Median
                               3Q
                                      Max
## -4.0683 -0.4653 -0.0348 0.3573 6.2562
##
## Random effects:
## Groups
             Name
                         Variance Std.Dev.
## Random_ID (Intercept) 509.6
                                  22.57
                                  12.99
## Residual
                         168.6
## Number of obs: 1872, groups: Random_ID, 39
##
## Fixed effects:
##
                  Estimate Std. Error t value
                  36.29237
                              3.72997
                                       9.730
## (Intercept)
```

```
## Response_SubjPE -0.04371
                               0.01872 -2.335
                               0.01744 -0.797
## Response_fdbk -0.01390
##
## Correlation of Fixed Effects:
              (Intr) Rs_SPE
## Rspns_SbjPE 0.153
## Respns_fdbk -0.233 -0.688
## [1] 15113.95
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + (1 | Random_ID)
##
      Data: final_df10
##
## REML criterion at convergence: 15105.9
## Scaled residuals:
##
      Min
               1Q Median
                                30
                                       Max
## -4.0822 -0.4630 -0.0345 0.3571 6.2550
##
## Random effects:
## Groups
             Name
                          Variance Std.Dev.
## Random_ID (Intercept) 511.7
                                   22.62
                          168.6
                                   12.98
## Residual
## Number of obs: 1872, groups: Random_ID, 39
##
## Fixed effects:
##
                   Estimate Std. Error t value
## (Intercept)
                   35.59988
                               3.63472
                                        9.794
## Response_SubjPE -0.05399
                               0.01359 -3.973
## Correlation of Fixed Effects:
##
               (Intr)
## Rspns SbjPE -0.010
## Data: final_df10
## Models:
## model2: Response_Ax ~ Response_SubjPE + (1 | Random_ID)
## model1: Response_Ax ~ Response_SubjPE + Response_fdbk + (1 | Random_ID)
##
         npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)
             4 15112 15134 -7551.8
                                      15104
## model2
            5 15113 15141 -7551.5
## model1
                                      15103 0.6394 1
                                                          0.4239
## [1] 15118.89
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_fdbk + (1 | Random_ID)
##
      Data: final_df10
## REML criterion at convergence: 15110.9
##
## Scaled residuals:
##
      Min 1Q Median
                                3Q
                                       Max
```

```
## -4.0925 -0.4563 -0.0299 0.3644 6.2101
##
## Random effects:
## Groups Name
                         Variance Std.Dev.
## Random_ID (Intercept) 500.8
                                  22.38
## Residual
                         169.1
                                  13.00
## Number of obs: 1872, groups: Random_ID, 39
##
## Fixed effects:
##
                Estimate Std. Error t value
## (Intercept)
                37.62506
                          3.65551 10.293
## Response_fdbk -0.04190
                            0.01267 -3.307
## Correlation of Fixed Effects:
##
              (Intr)
## Respns_fdbk -0.180
## Data: final_df10
## Models:
## model3: Response_Ax ~ Response_fdbk + (1 | Random_ID)
## model1: Response_Ax ~ Response_SubjPE + Response_fdbk + (1 | Random_ID)
         npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)
## model3
            4 15116 15138 -7554.2
                                     15108
## model1
            5 15113 15141 -7551.5
                                     15103 5.4244 1
                                                        0.01986 *
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' ' 1
```

The ICC outcome for mood: The ICC is moderate according to guidelines by Koo and Li (2016).

```
## [1] "lmer for mood with just the intercept"
## [1] 0.6437434
##
                  2.5 %
                          97.5 %
## .sig01
               15.34013 24.09587
## .sigma
               13.81085 14.73467
## (Intercept) 47.29225 59.53788
## [1] 15437.52
## [1] 15384.58
## Data: final_df10
## Models:
## model0: Response_H ~ 1 + (1 | Random_ID)
## model1: Response_H ~ Response_SubjPE + (1 | Random_ID)
         npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)
## model0
            3 15442 15458 -7717.8
                                      15436
## model1
             4 15382 15404 -7687.1
                                      15374 61.486 1 4.459e-15 ***
## ---
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' 1
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: Response_H ~ Response_SubjPE + (Response_SubjPE | Random_ID)
##
     Data: final_df10
## Control: lmerControl(optimizer = "bobyqa")
##
##
                 BIC
                      logLik deviance df.resid
   15292.1 15325.3 -7640.0 15280.1
##
                                           1866
##
## Scaled residuals:
      Min
               1Q Median
                                3Q
                                       Max
## -5.3899 -0.4299 0.0201 0.4616 4.5245
##
## Random effects:
## Groups
              Name
                              Variance Std.Dev. Corr
##
   Random_ID (Intercept)
                              365.9256 19.1292
##
              {\tt Response\_SubjPE}
                                0.0355 0.1884
                                               -0.05
                              180.2017 13.4239
## Number of obs: 1872, groups: Random_ID, 39
## Fixed effects:
                   Estimate Std. Error t value
                   53.22371
                               3.08091 17.275
## (Intercept)
## Response_SubjPE 0.11523
                               0.03364
##
## Correlation of Fixed Effects:
##
               (Intr)
## Rspns_SbjPE -0.052
## [1] 15292.05
```

1st and last trial correlations for anxiety and mood

The correlations between first and last anxiety and mood ratings are doubled compared to the previous pilot! Do we have people who gave less reliable answers? We can see it from the ratings themselves, or predictions where some people ignored the histograms.

```
## [1] "Correlation between first and last anxiety rating"
##
   Pearson's product-moment correlation
##
## data: final_df10[final_df10$Trial.Number == 1, ]$Response_Ax and final_df10[final_df10$Trial.Number
## t = 3.9135, df = 37, p-value = 0.0003758
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.2719854 0.7316774
## sample estimates:
##
      cor
## 0.54107
## [1] "Correlation between first and last mood rating"
##
##
   Pearson's product-moment correlation
##
## data: final_df10[final_df10$Trial.Number == 1, ]$Response_H and final_df10[final_df10$Trial.Number
## t = 4.0628, df = 37, p-value = 0.0002423
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.2908819 0.7410709
## sample estimates:
##
         cor
## 0.5554232
```

Objective PE and anxiety relationship

We now will run everything again but this time using the objective PE (feedback - histogram_mean) instead of the subjective one (feedback - prediction). The correlation between subjective_PE and anxiety is very similar to objective_PE and anxiety (-0.089 vs -0.087).

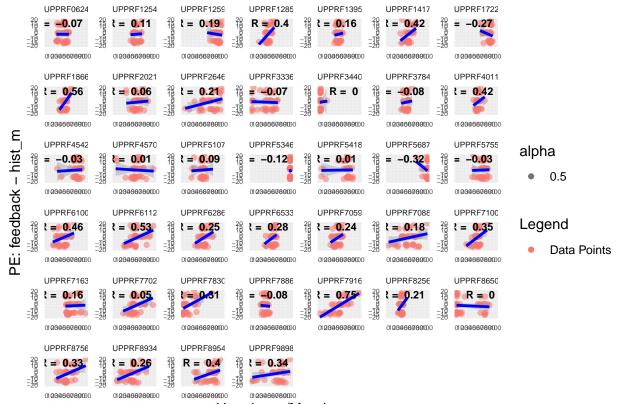
[1] "average correlation between Anxiety and PE: -0.0871134288971019"



Objective PE and mood relationship

We will repeat the same thing for the relationship between PE and mood. Same as anxiety, the relationship between the objective PE and mood is very similar to the $Subj_PE$ (0.18 vs 0.17).

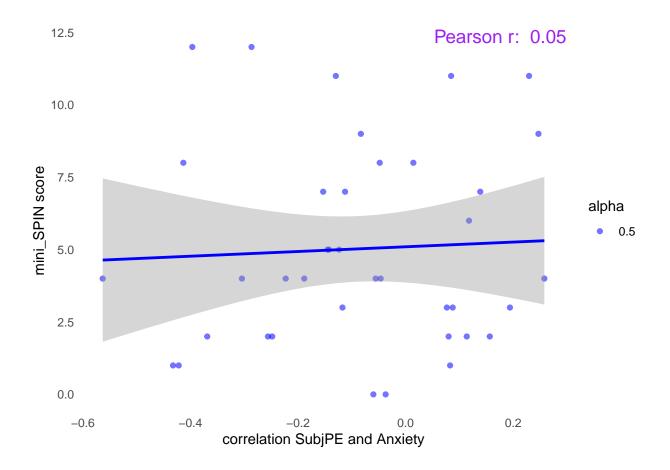
[1] "average correlation between happiness and objective PE: 0.170645092999263"



mini-SPIN count

Since these people were not screened for social anxiety, let's see how many of them had social anxiety scores higher than or equal to 6. We have 14/38 people who have high social anxiety on mini-SPIN, compared to 21/38 in the previous pilot.

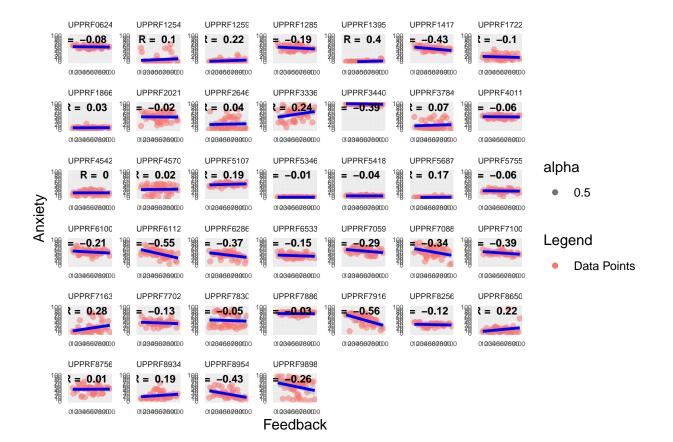
[1] "Out of 38 people, these people had a mini_SPIN total score higher or equal to 6: 14"
[1] 0.04964543



Feedback and anxiety relationship

We will now look at the relationship between feedback and anxiety (so without taking prediction or histogram into account). Again, almost the same as the same relationship with Subj_PE and PE. We can again see people who almost always rated 0 or 100 for anxiety.

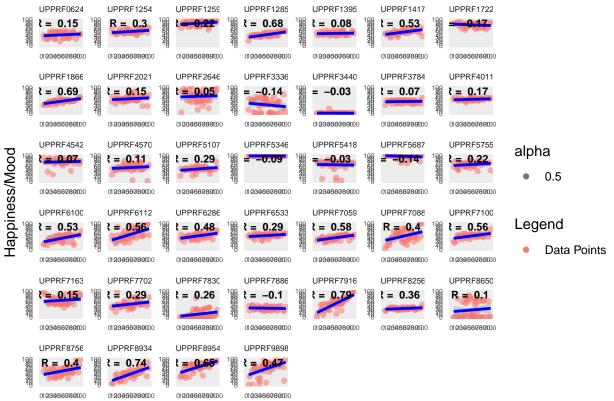
[1] "average correlation between feedback and anxiety: -0.078696284365809"



Feedback and mood relationship

The relationship between feedback and happiness: this relationship is stronger than the one with $Subj_PE$ (0.27 vs 0.17).

[1] "average correlation between feedback and happiness: 0.27407792712756"



Feedback

What would be the best next step? We had 14 people with social anxiety, maybe 1) looking these people separately, 2) looking at correlations with mini-SPIN, 3) shall we collect more data? 4) shall we collect data in people with high social anxiety (although we also don't want to use our participants for piloting considering it will be hard to recruit people for the video pilot, we may want to maybe increase the sample to 30/40, so only 15/25 more people to combine with this current pilot data?), 5) would introducing attention shift, or algorithm to produce the feedback change anything, as we do not have any effects of the PE at the moment, it seems?

LME's taking subjective PE and feedback into account

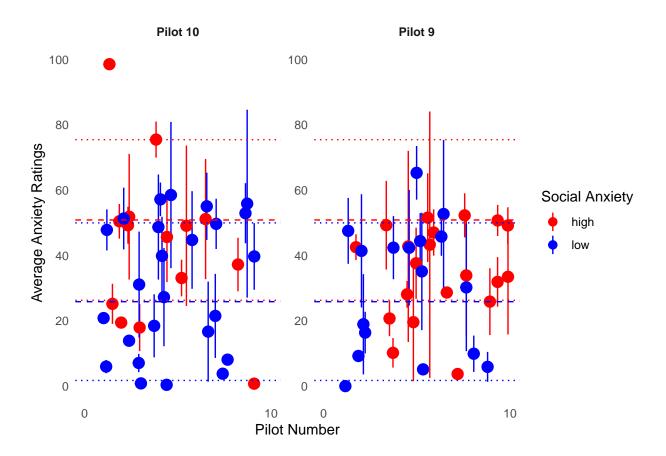
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + (1 | Random_ID)
##
     Data: final_df10
##
## REML criterion at convergence: 15105.9
## Scaled residuals:
      Min
               1Q Median
                                30
                                       Max
## -4.0822 -0.4630 -0.0345 0.3571 6.2550
##
## Random effects:
                         Variance Std.Dev.
## Groups
             Name
## Random_ID (Intercept) 511.7
                                   22.62
## Residual
                         168.6
                                   12.98
## Number of obs: 1872, groups: Random_ID, 39
## Fixed effects:
##
                   Estimate Std. Error t value
## (Intercept)
                  35.59988
                            3.63472
                                       9.794
## Response_SubjPE -0.05399
                              0.01359 - 3.973
##
## Correlation of Fixed Effects:
               (Intr)
## Rspns_SbjPE -0.010
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_fdbk + (1 | Random_ID)
     Data: final_df10
## REML criterion at convergence: 15110.9
##
## Scaled residuals:
      Min
           1Q Median
                                3Q
                                       Max
## -4.0925 -0.4563 -0.0299 0.3644 6.2101
## Random effects:
## Groups
                         Variance Std.Dev.
## Random_ID (Intercept) 500.8
                                   22.38
                          169.1
                                   13.00
## Number of obs: 1872, groups: Random_ID, 39
## Fixed effects:
                Estimate Std. Error t value
## (Intercept)
                37.62506
                            3.65551 10.293
## Response_fdbk -0.04190
                             0.01267 -3.307
## Correlation of Fixed Effects:
               (Intr)
## Respns_fdbk -0.180
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula: Response_Ax ~ Response_SubjPE + Response_fdbk + (1 | Random_ID)
      Data: final_df10
##
##
## REML criterion at convergence: 15111.6
## Scaled residuals:
       Min
              1Q Median
                                3Q
                                       Max
## -4.0683 -0.4653 -0.0348 0.3573 6.2562
##
## Random effects:
## Groups
              Name
                          Variance Std.Dev.
## Random_ID (Intercept) 509.6
                                   22.57
## Residual
                          168.6
                                   12.99
## Number of obs: 1872, groups: Random_ID, 39
##
## Fixed effects:
##
                   Estimate Std. Error t value
## (Intercept)
                  36.29237
                                        9.730
                              3.72997
## Response_SubjPE -0.04371
                               0.01872 -2.335
## Response_fdbk
                 -0.01390
                               0.01744 - 0.797
##
## Correlation of Fixed Effects:
##
               (Intr) Rs_SPE
## Rspns_SbjPE 0.153
## Respns_fdbk -0.233 -0.688
```

Anxiety plots + mini SPIN + average ratings from Eleanor's study

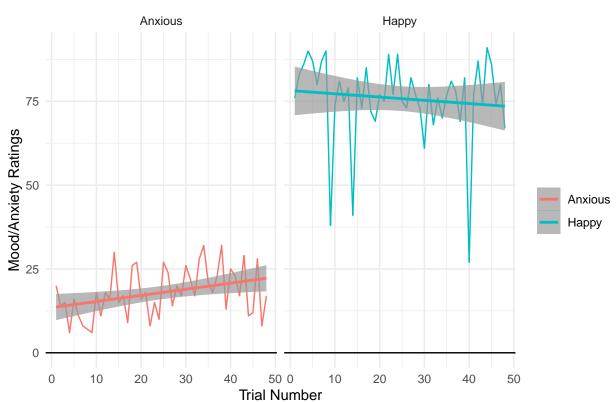
In the plot below we have anxiety ratings during pilots 9 (no video, more people with high social anxiety) and 10 (video re-introduced, less people with high social anxiety).

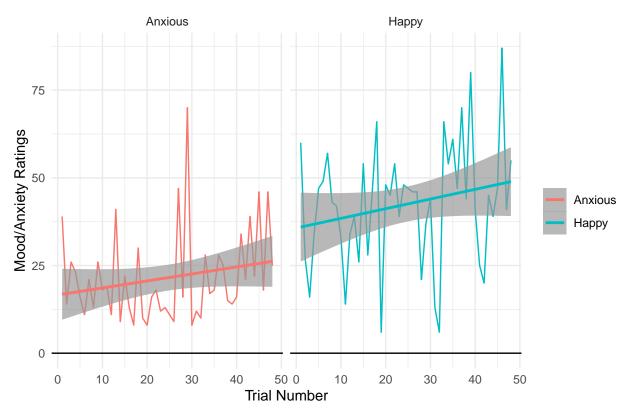
- The dots represent the average anxiety ratings per subject with SD as vertical lines.
- The red and blue dots represent people who had mini-SPIN scores that were above or below threshold of 6.
- The blue dashed line is the average anxiety after the externally focused attention in Eleanor's study and the dotted lines are \pm SD. The red ones are for the self-focused attention.
- Pilot 9 was without video where we had 21/37 subjects with high social anxiety and people had to type their answers, and pilot 10 was with the video where we had only 14/38 people with high social anxiety and people had to verbally share their answers.

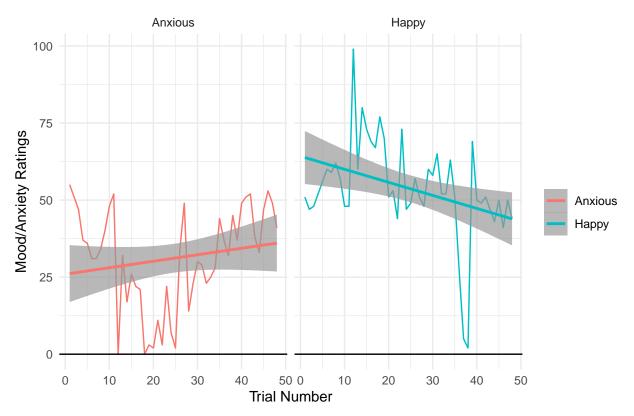


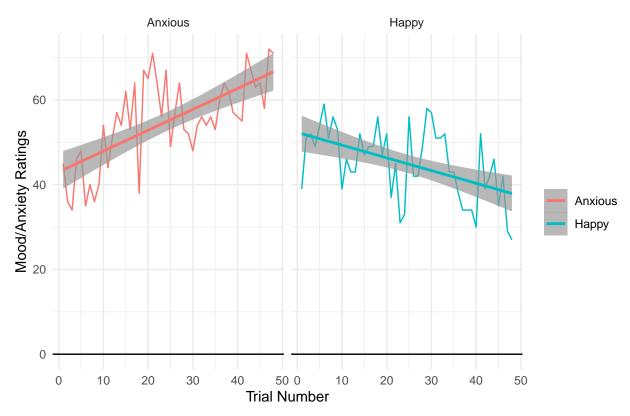
Anxiety and Mood over time within subjects + regression line

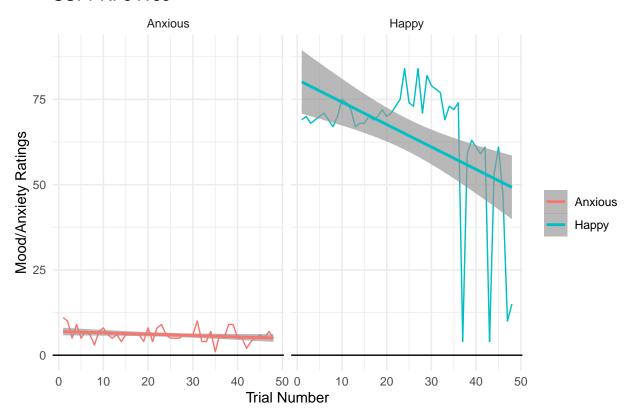
I will now repeat the same plot we had within subjects and fit a regression line for mood/anxiety ratings

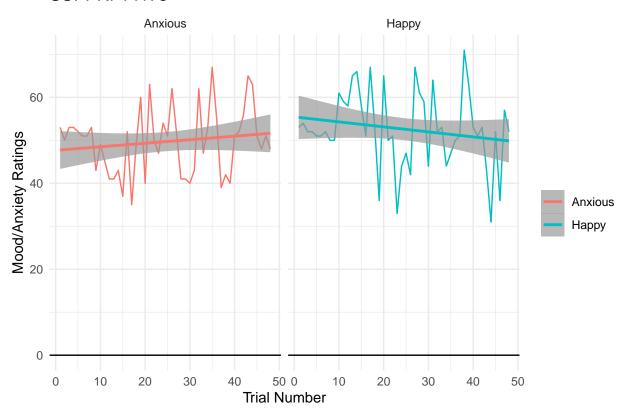


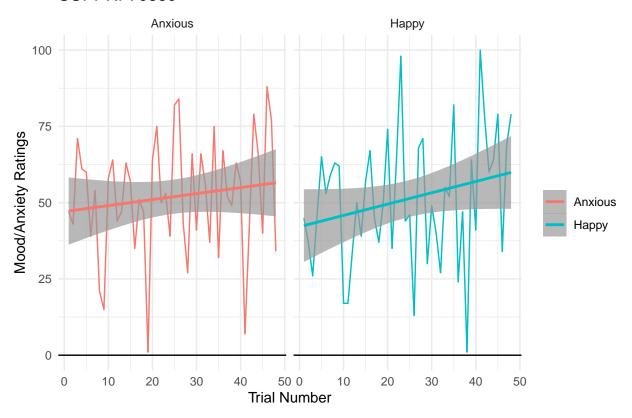


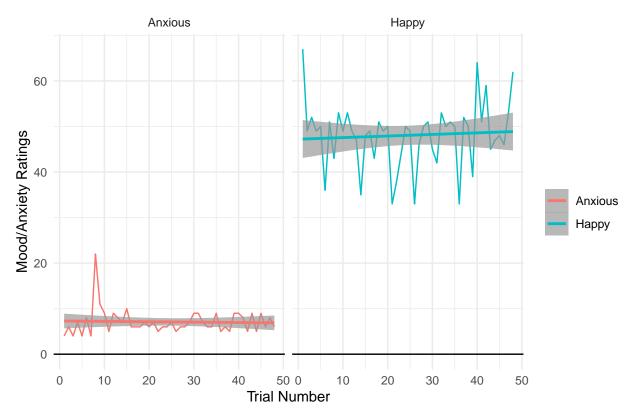


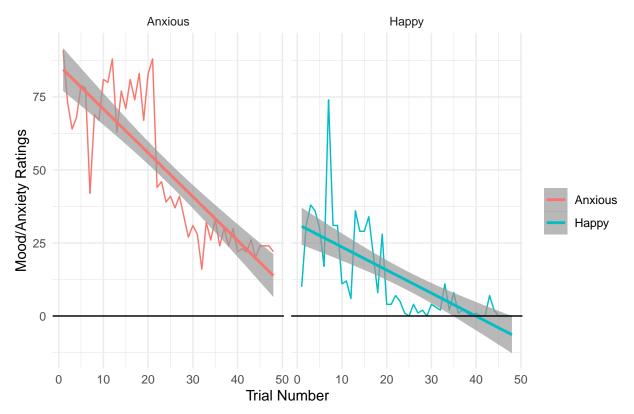


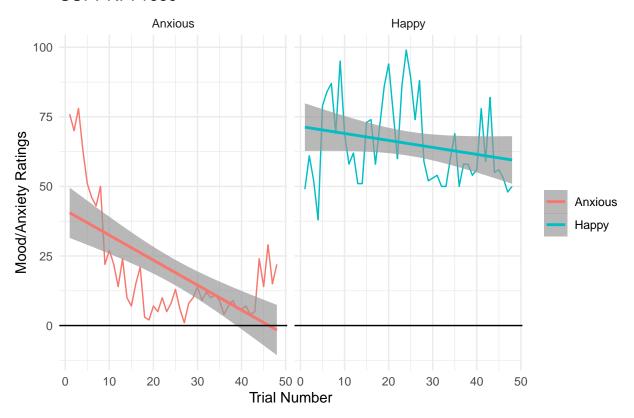


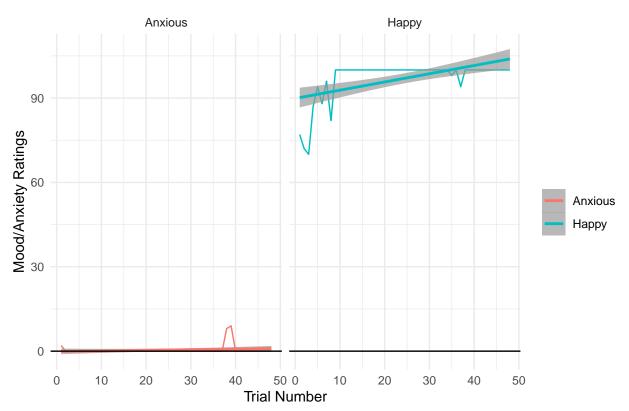


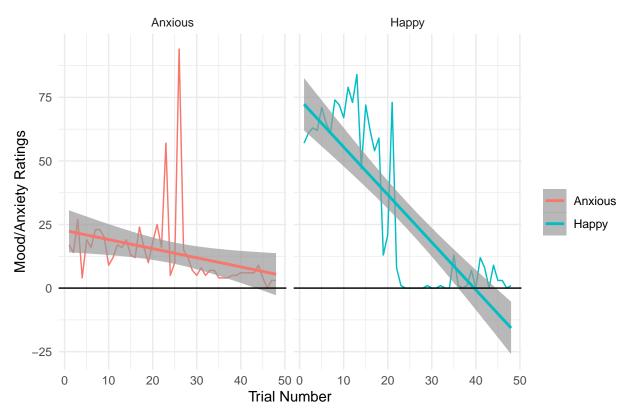


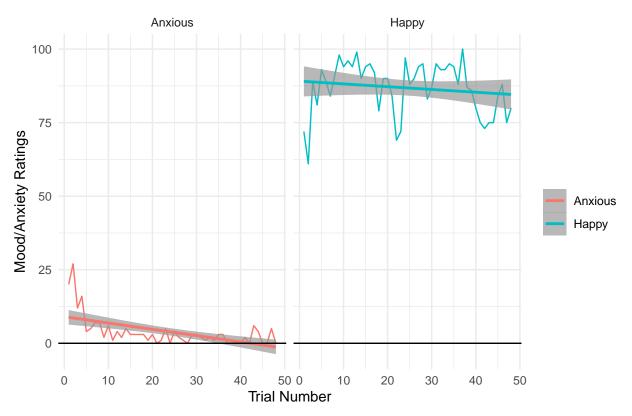


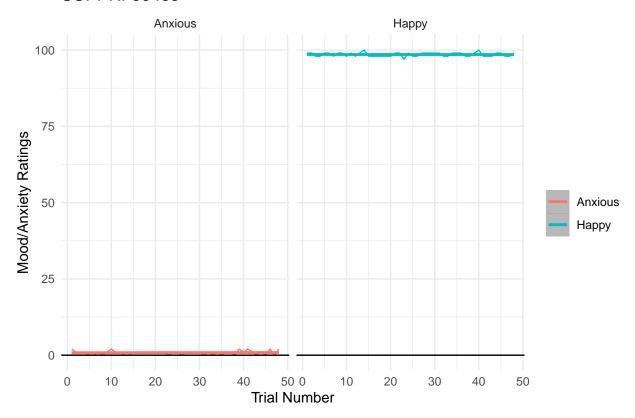


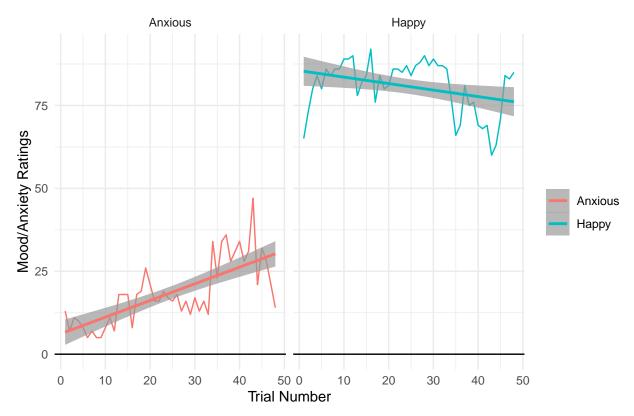


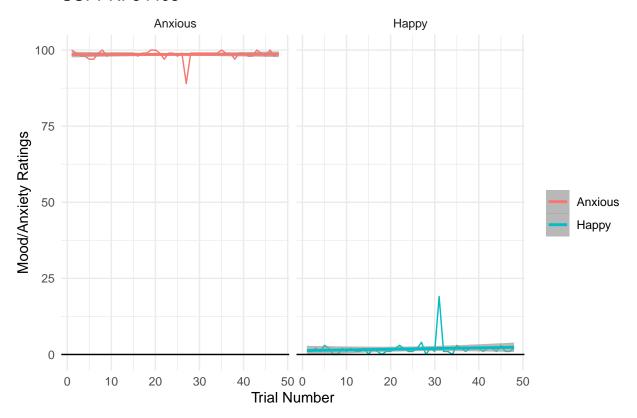


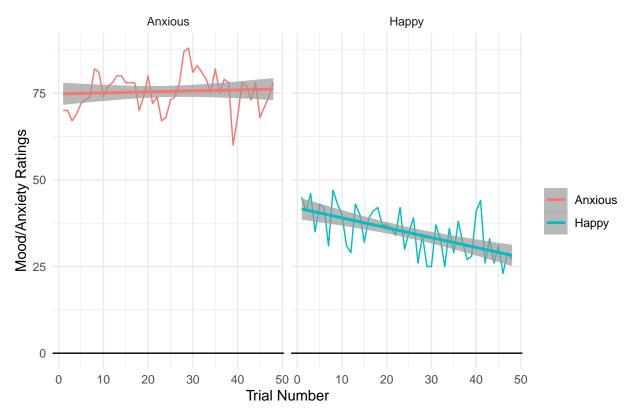


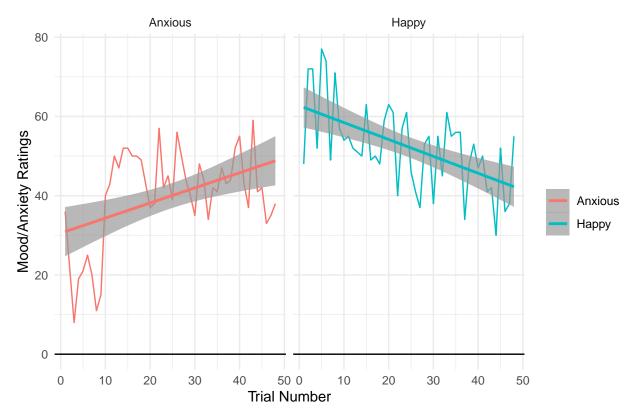


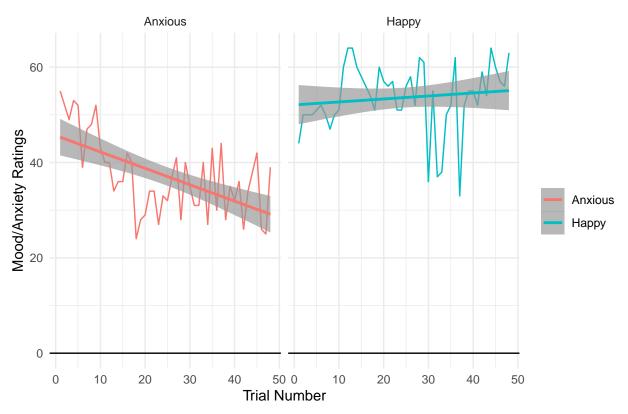


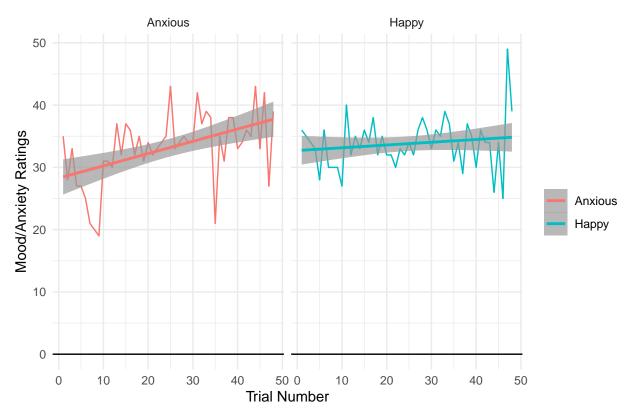


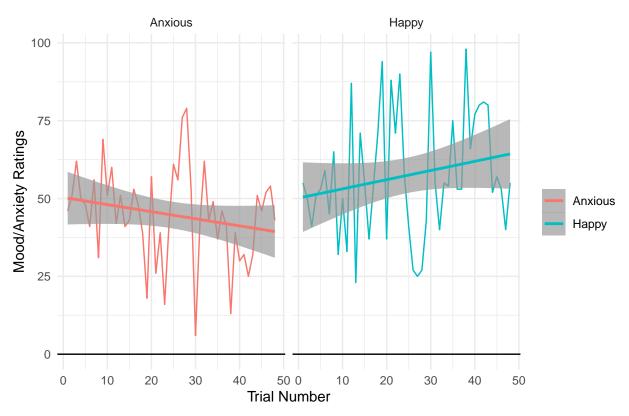


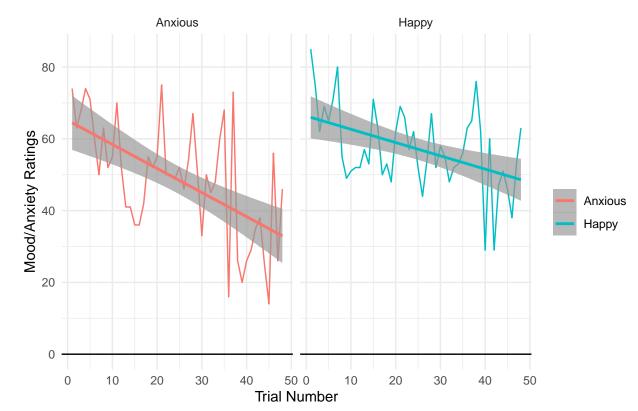


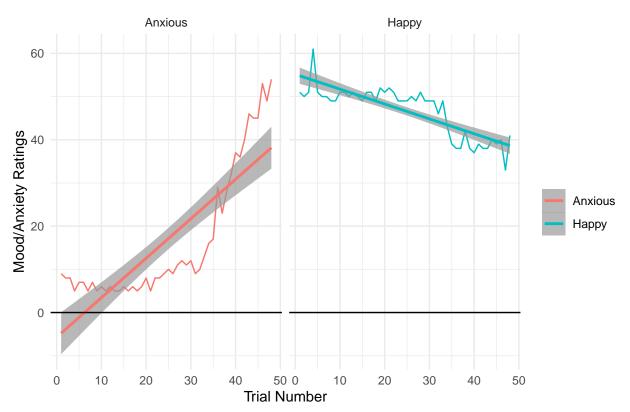


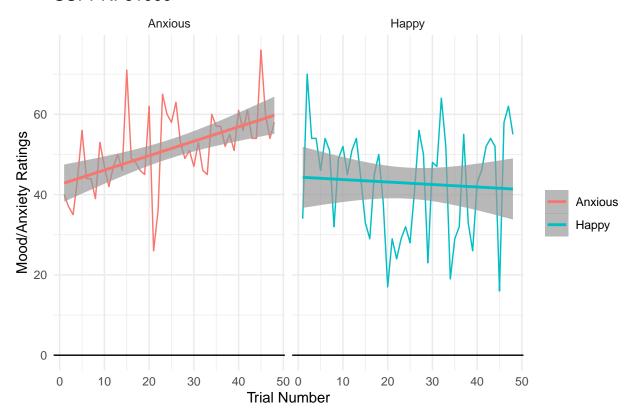


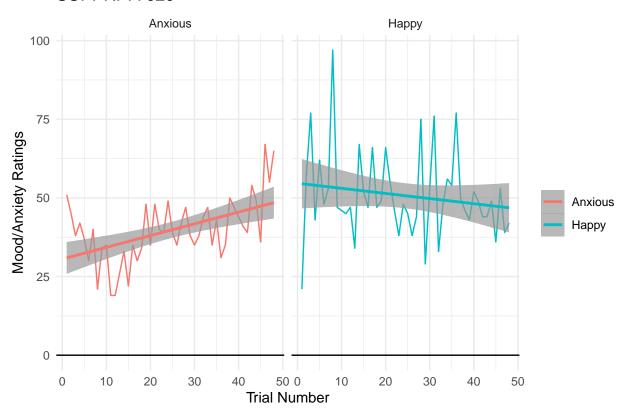


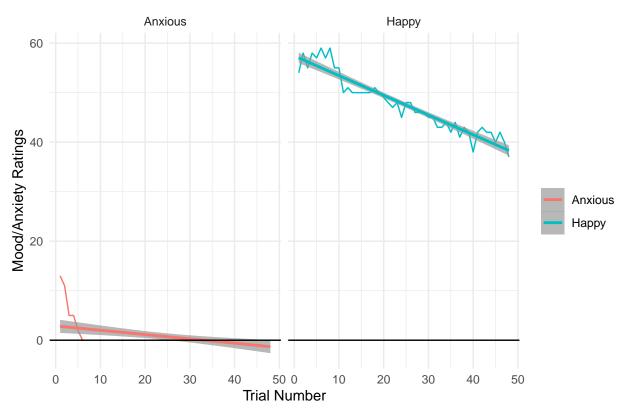


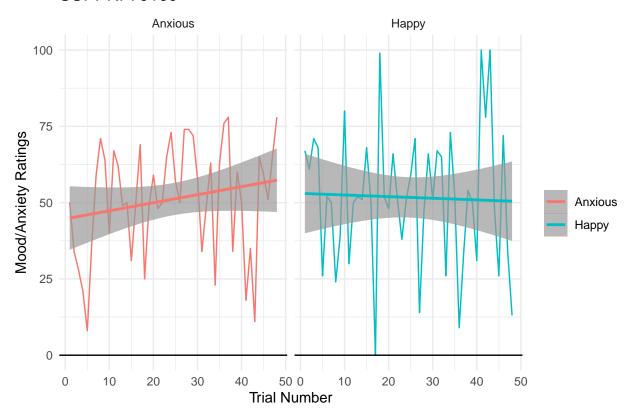


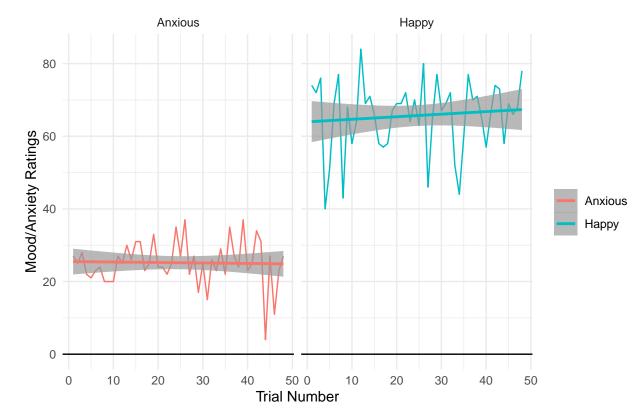


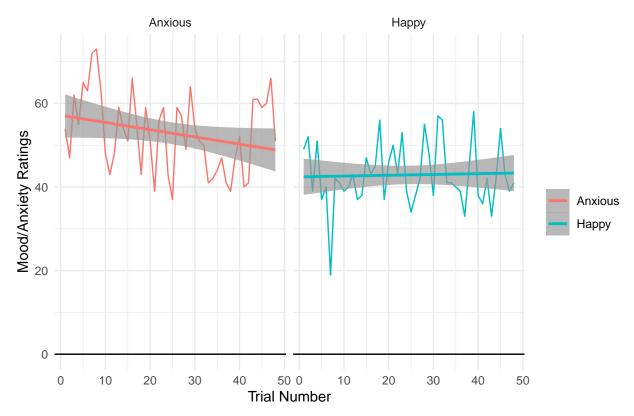


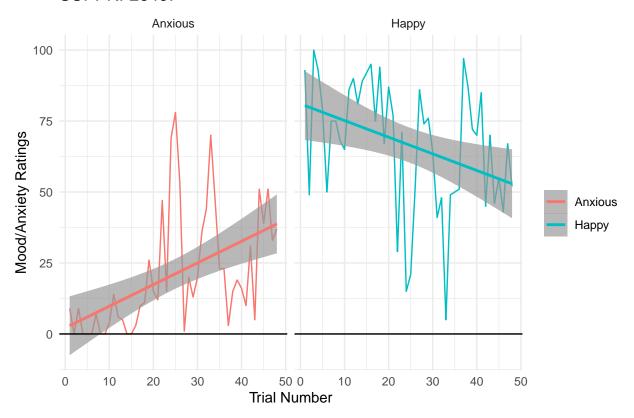


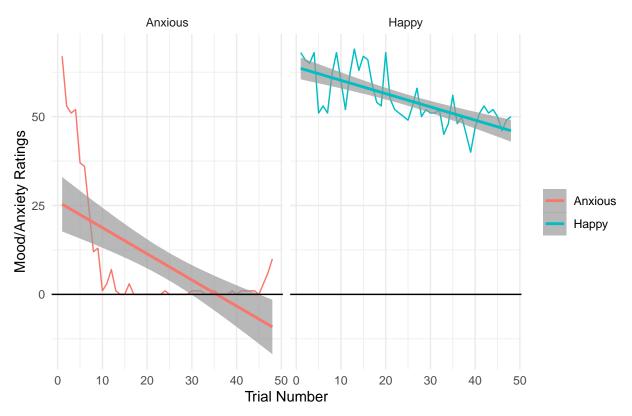


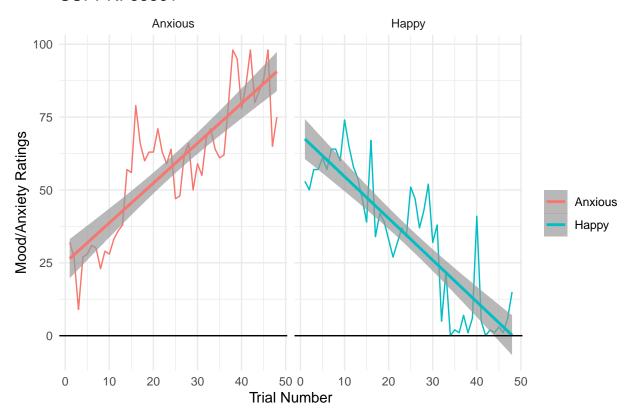


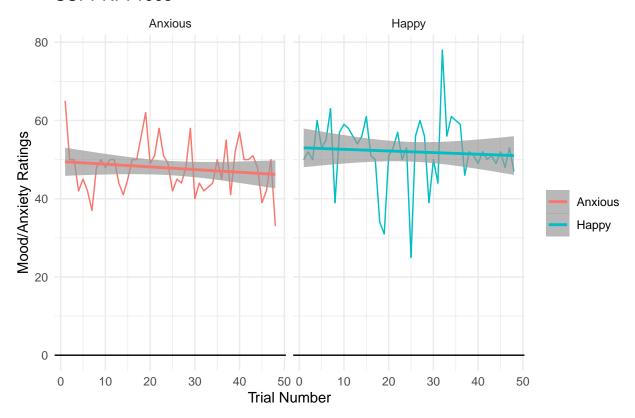


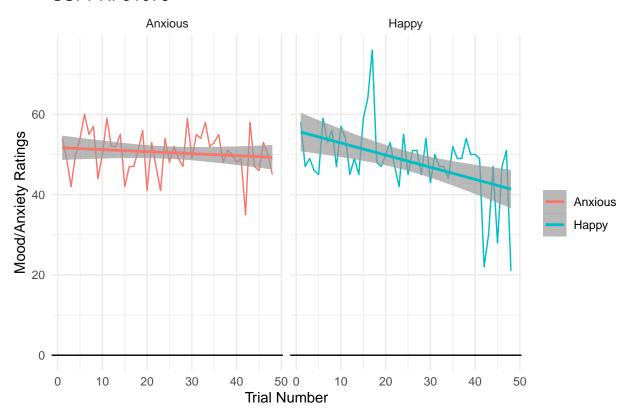


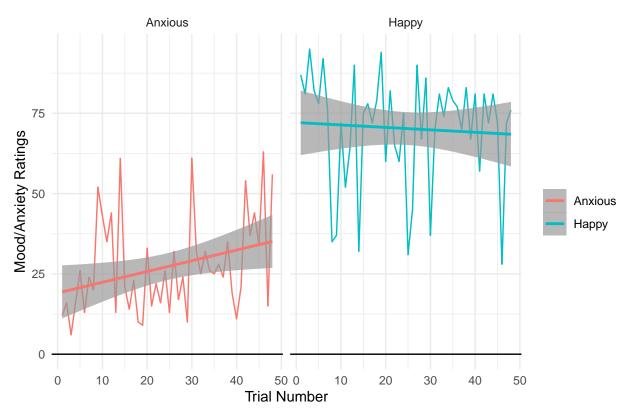


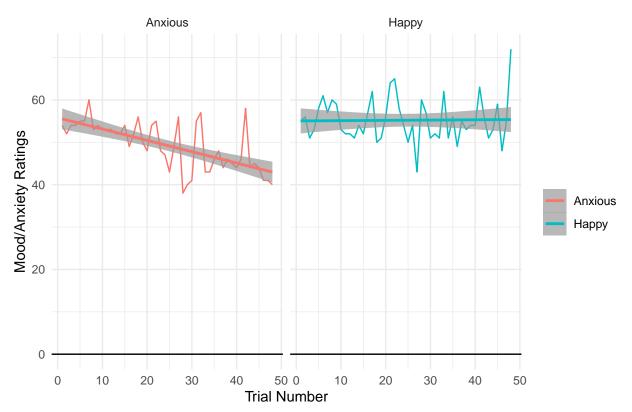


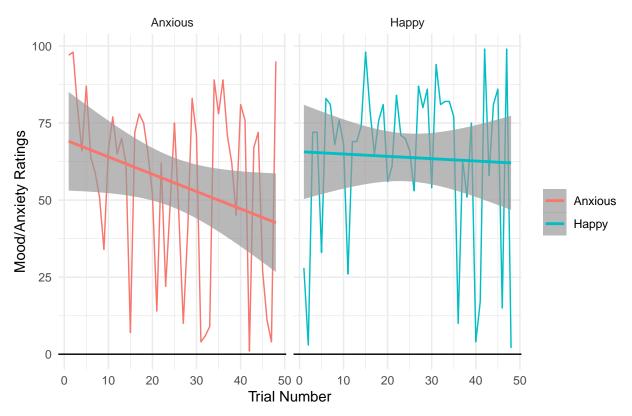


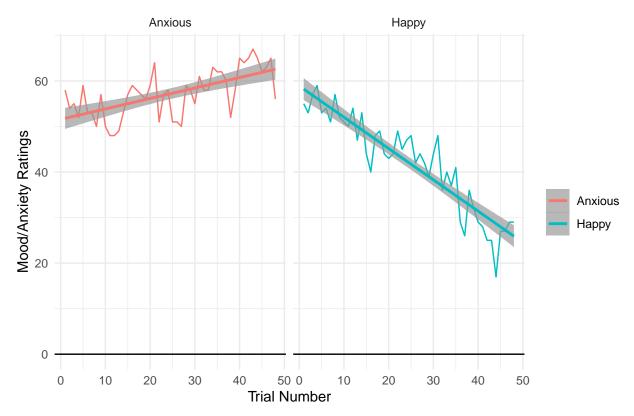


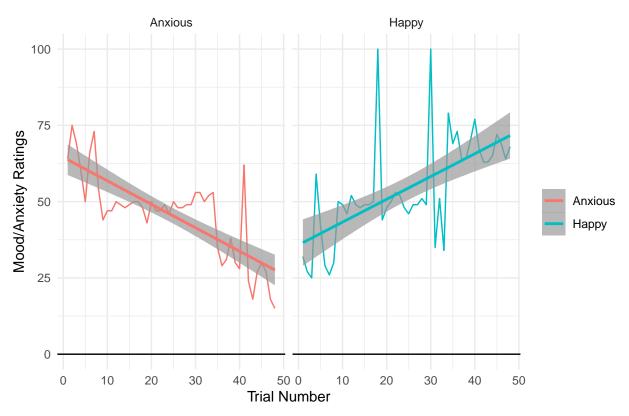












Anxiety and Mood over time- group plot

I will now repeat the same plot but for all subjects

