Surprise study pilot 14

Marjan Biria

2024-02-07

Study description

This study has the same version of PE's as pilots 10 and 11, also having video and audio only. In this pilot we 1) replaced the judge pictures with cartoon images to see whether it becomes more believable as many people believed those images were AI generated, 2) we allow participants to also choose their own avatar, again to make it more believable that the others are also real, 3) we have created a new narrative around public speaking and what is the goal of this task and added criteria they will be judged open to make them more nervous but also make it more believable if they get a rating that they cannot guess what they were judged upon at each trial (they are told the judges will receive a new criterion per trial), 4) lastly, we changed two feedback questions to inquire about believability, one of them being a scale asking "How stressful was this as a social situation?"

The Gorilla experiment is the following: https://app.gorilla.sc/admin/project/120255 The task is the following: https://app.gorilla.sc/admin/task/741126/editor?version=13

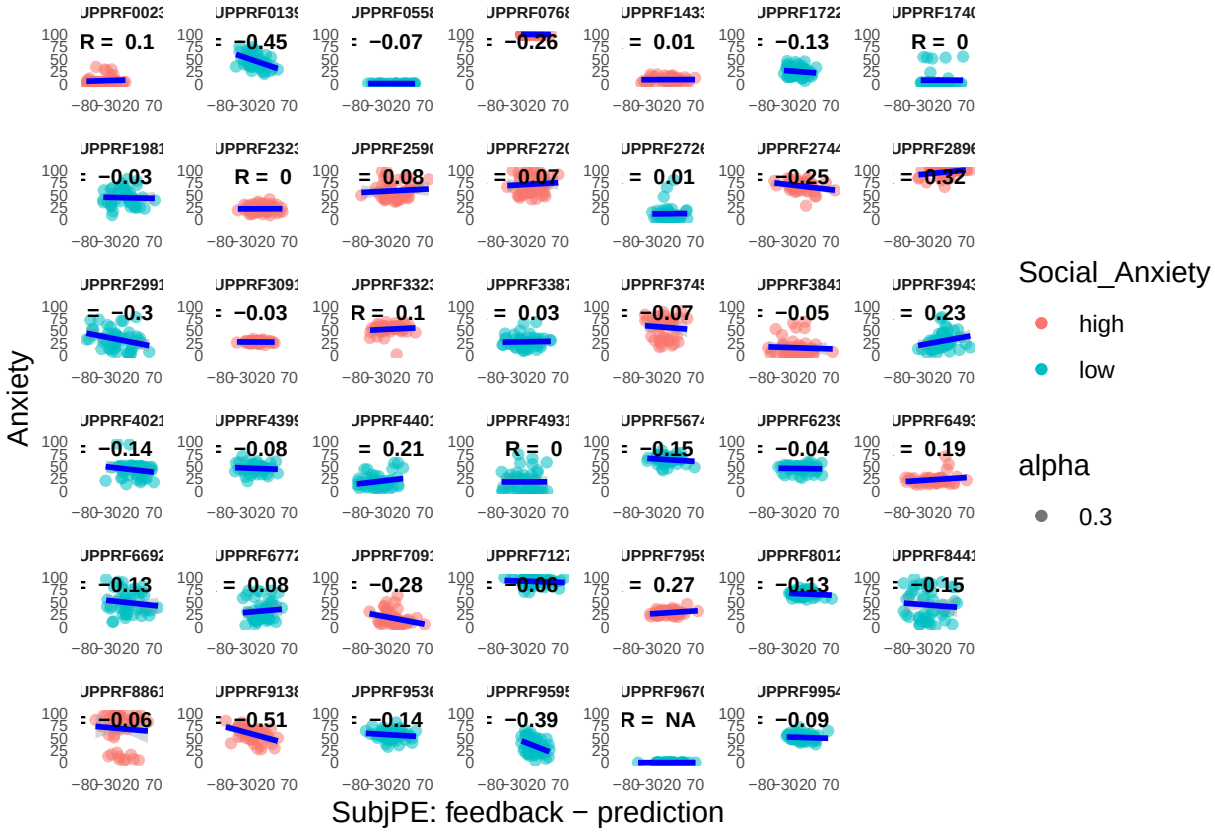
I will write this here as well for us to have, if we forget again:
Positive PE: we selected numbers from a normal distribution ranging from 12-20, added this number to the mean of the histogram.
Negative PE: we selected numbers from a normal distribution ranging from 12-20, subtracted this number to the mean of the histogram.
Big positive PE: Per judge, we added 10 to the biggest positive feedback we had generated before.
Neutral PE: we added -1, 0, 1 to the mean of the histograms

In this sample, only 17 people out of 41 had high social anxiety.

```
## # A tibble: 41 x 2
##    Random_ID   Trial.Number
##    <chr>              <int>
##  1 SUPPRF00235           48
##  2 SUPPRF01396           48
##  3 SUPPRF05583           26
##  4 SUPPRF07686           48
##  5 SUPPRF14330           48
##  6 SUPPRF17228           48
##  7 SUPPRF17407           48
##  8 SUPPRF19818           48
##  9 SUPPRF23233           48
## 10 SUPPRF25907           48
## # i 31 more rows
```

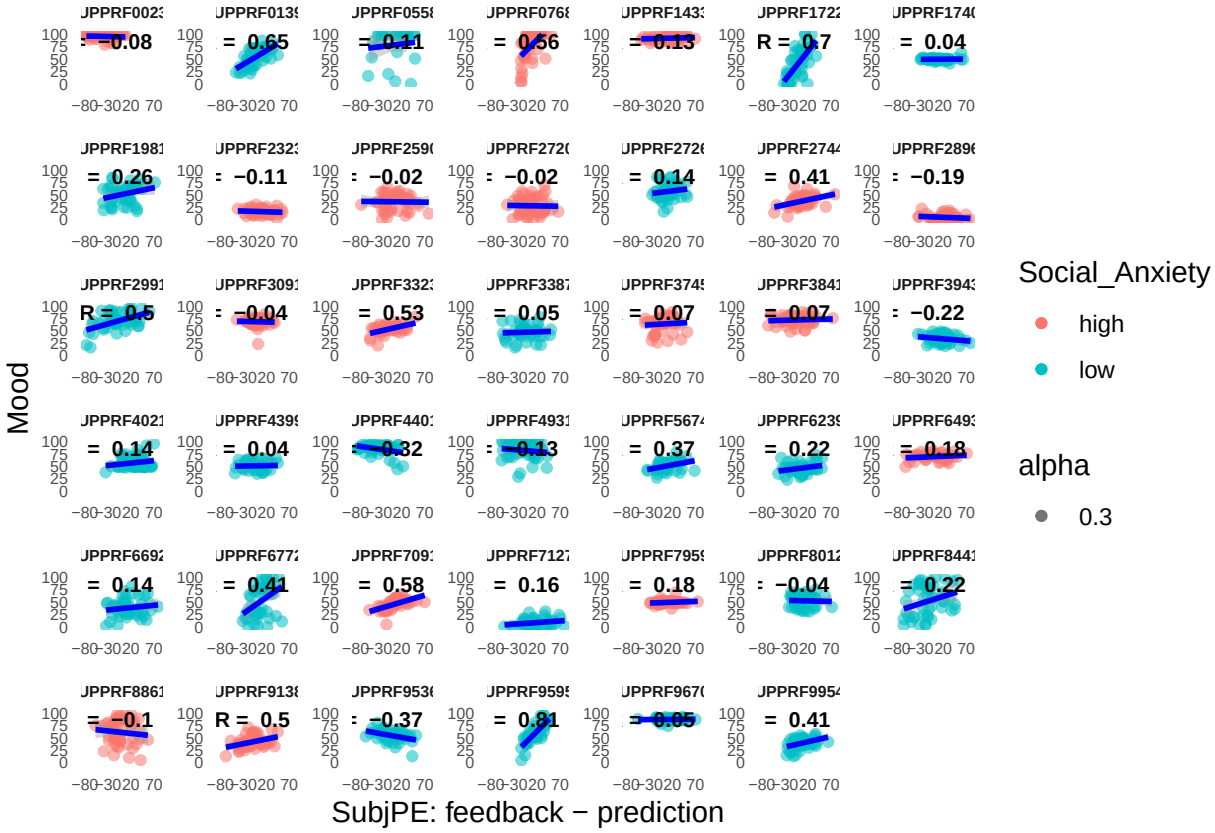
Relationship between Anxiety and SubjPE

```
## [1] "average correlation between anxiety and SubjPE: -0.0570621407466249"
```



Relationship between Mood and SubjPE

```
## [1] "average correlation between mood and SubjPE: 0.169860303615549"
```

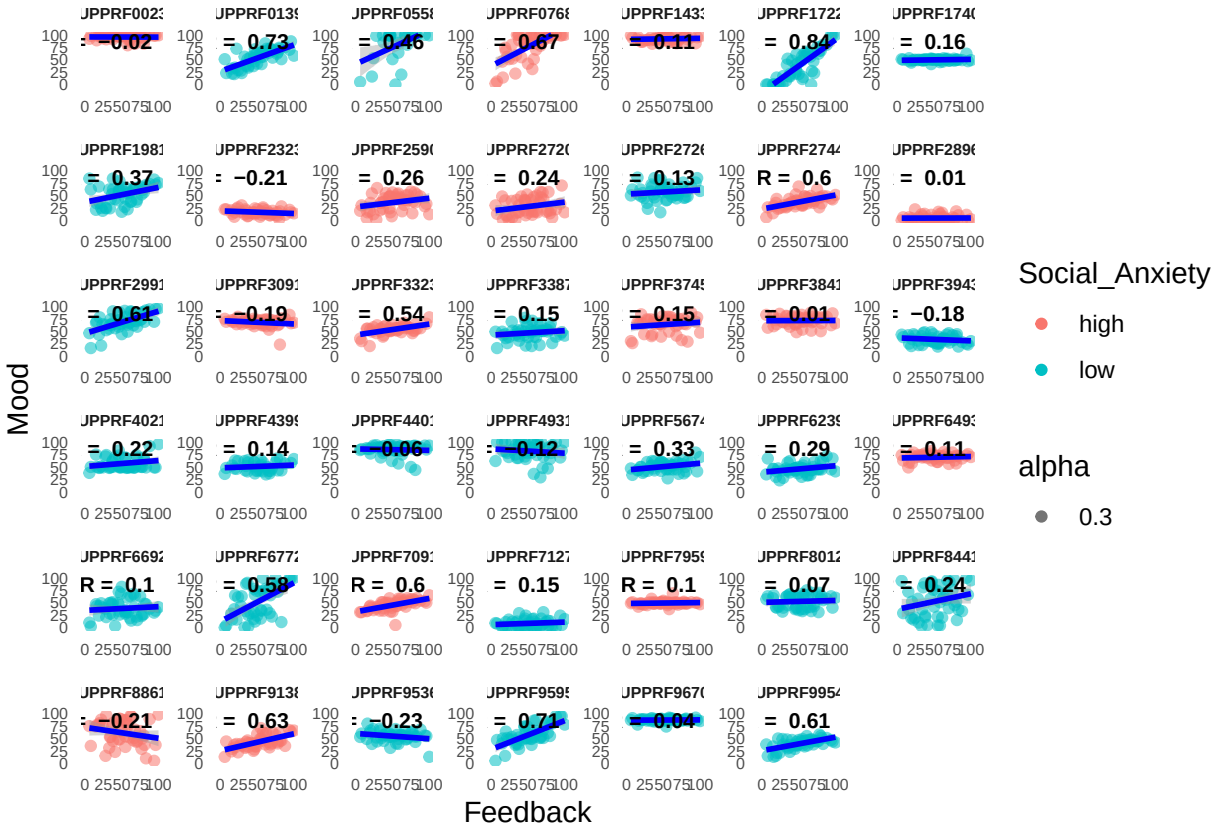


Relationship between Mood and feedback

The relationship between mood and feedback seems to be stronger than mood and subjective PE (0.24 vs 0.17), so it seems people may care more about the feedback as receiving reward or punishment, rather than social PE? The relationship between subjective PE with both anxiety and mood has been the weakest across all pilots. We need to make sure it is only because of changing the pictures of virtual players, maybe by changing the narrative they would assume that the other players are also learning how to do public speaking (less intimidating than someone who is an expert?).

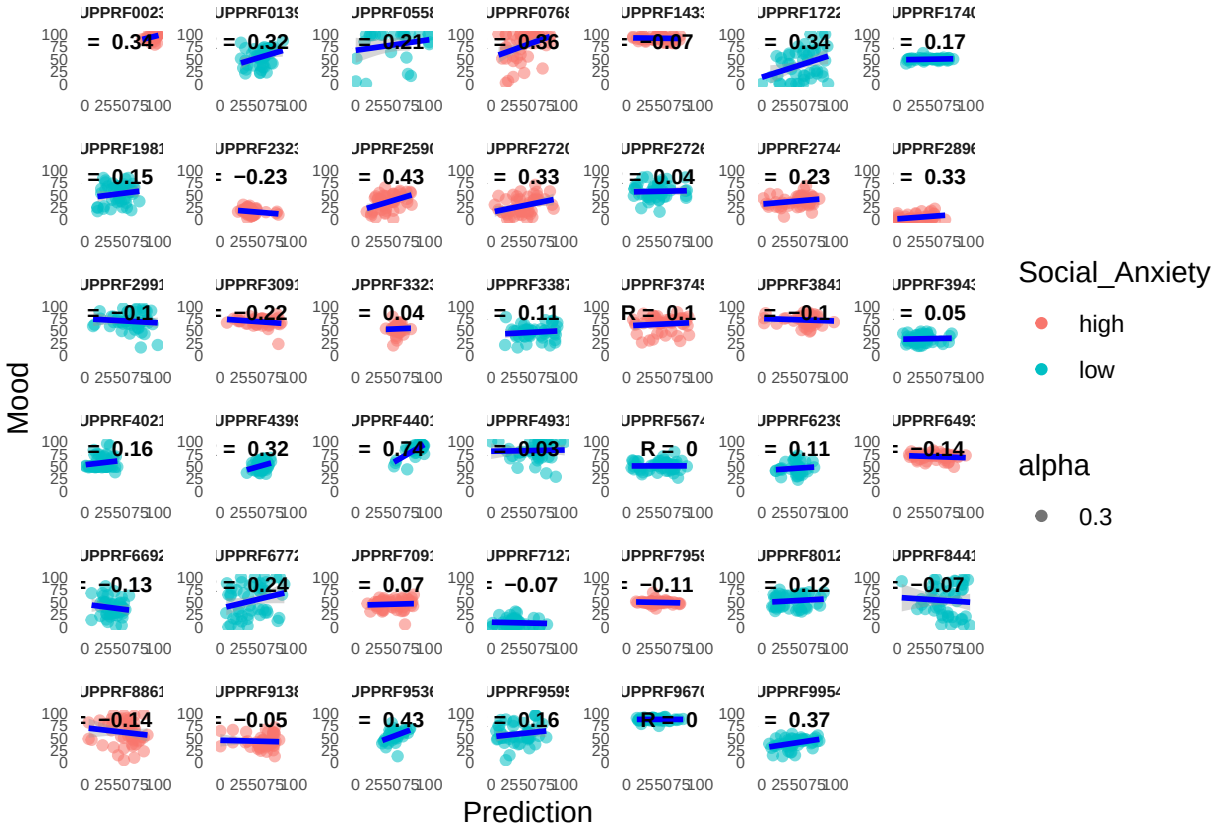
To be sure, I wonder if we want to repeat the pilot and just replace the images?

```
## [1] "average correlation between mood and feedback: 0.237770542807818"
```



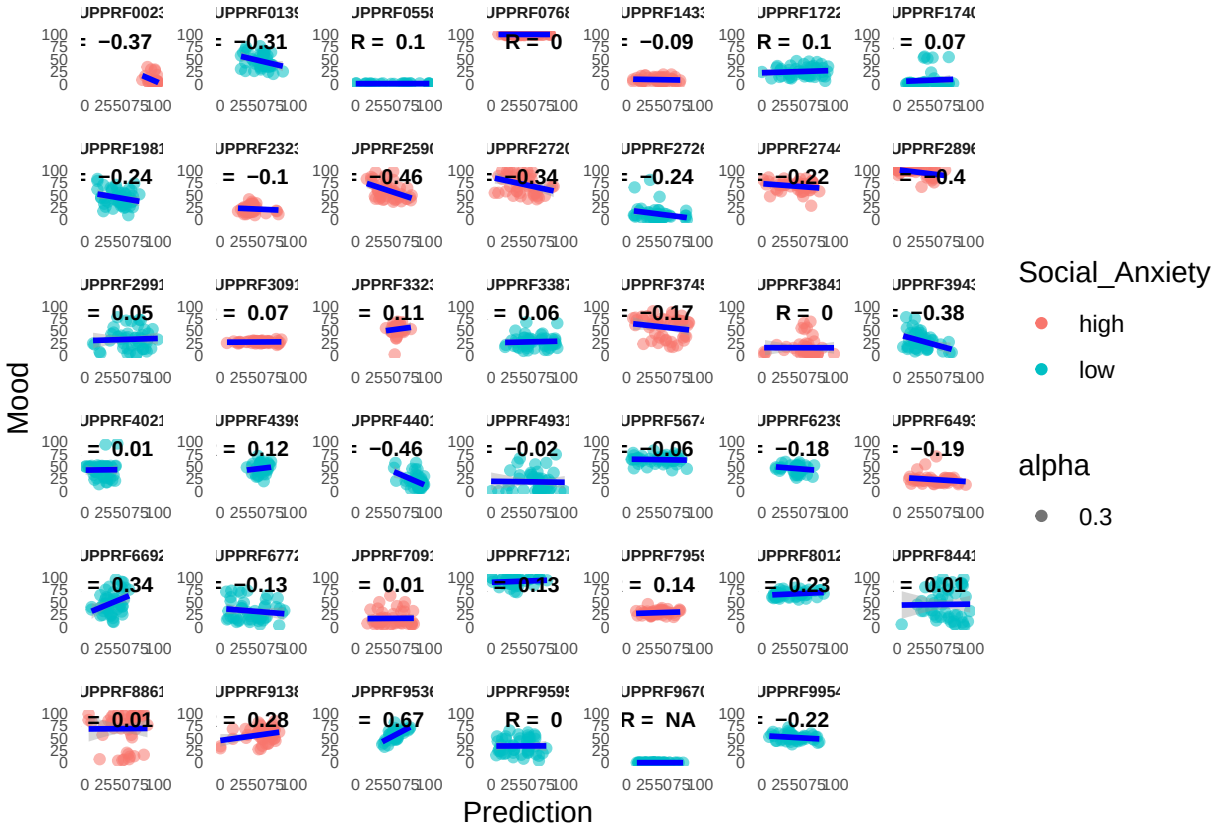
Relationship between Mood and prediction

```
## [1] "average correlation between mood and prediction: 0.119584769789578"
```



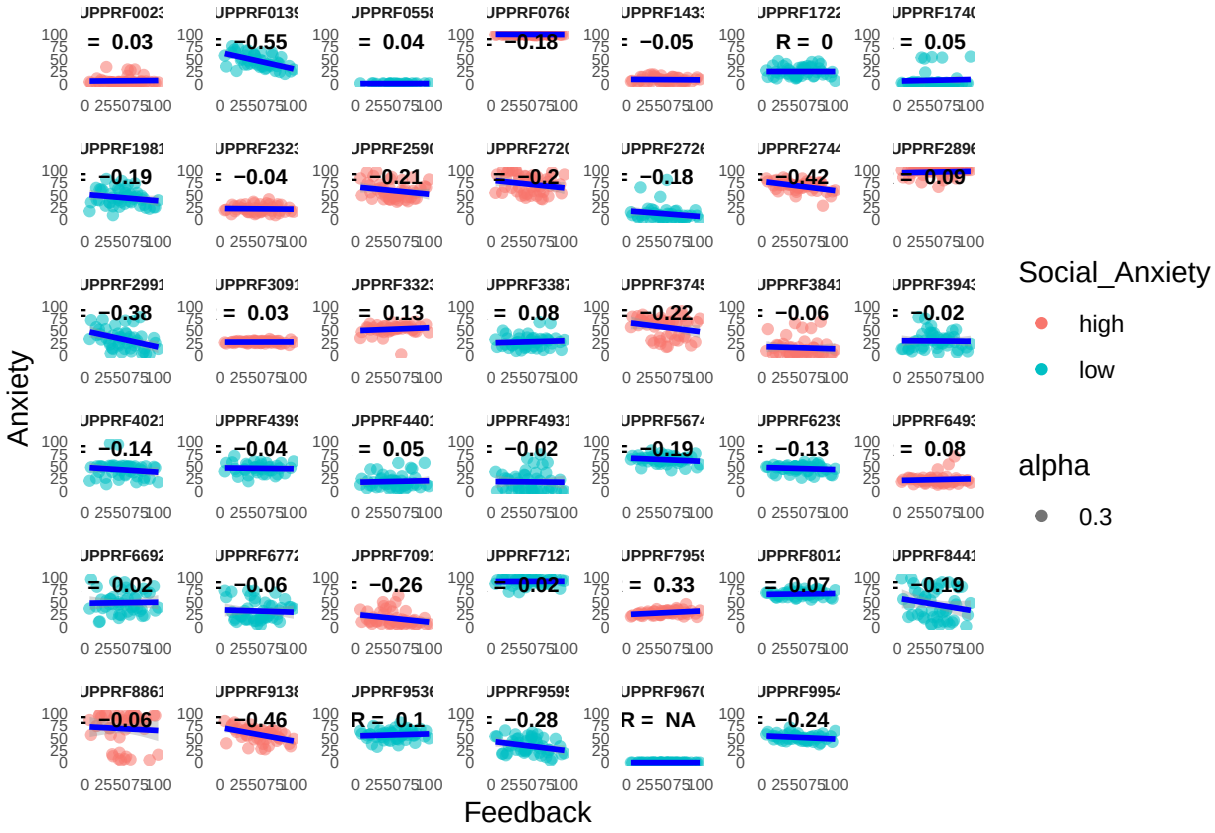
Relationship between Anxiety and prediction

```
## [1] "average correlation between anxiety and prediction: -0.0520117558366197"
```

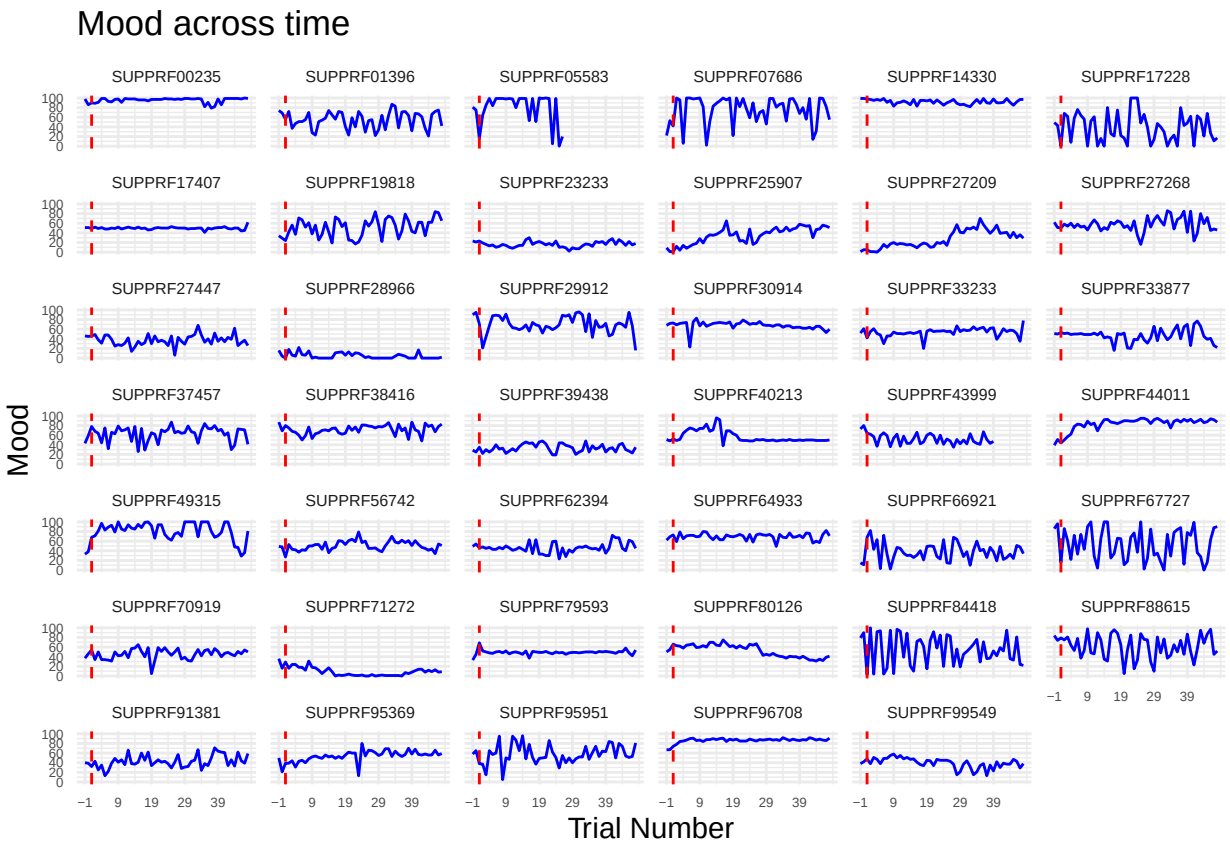


Relationship between Anxiety and feedback

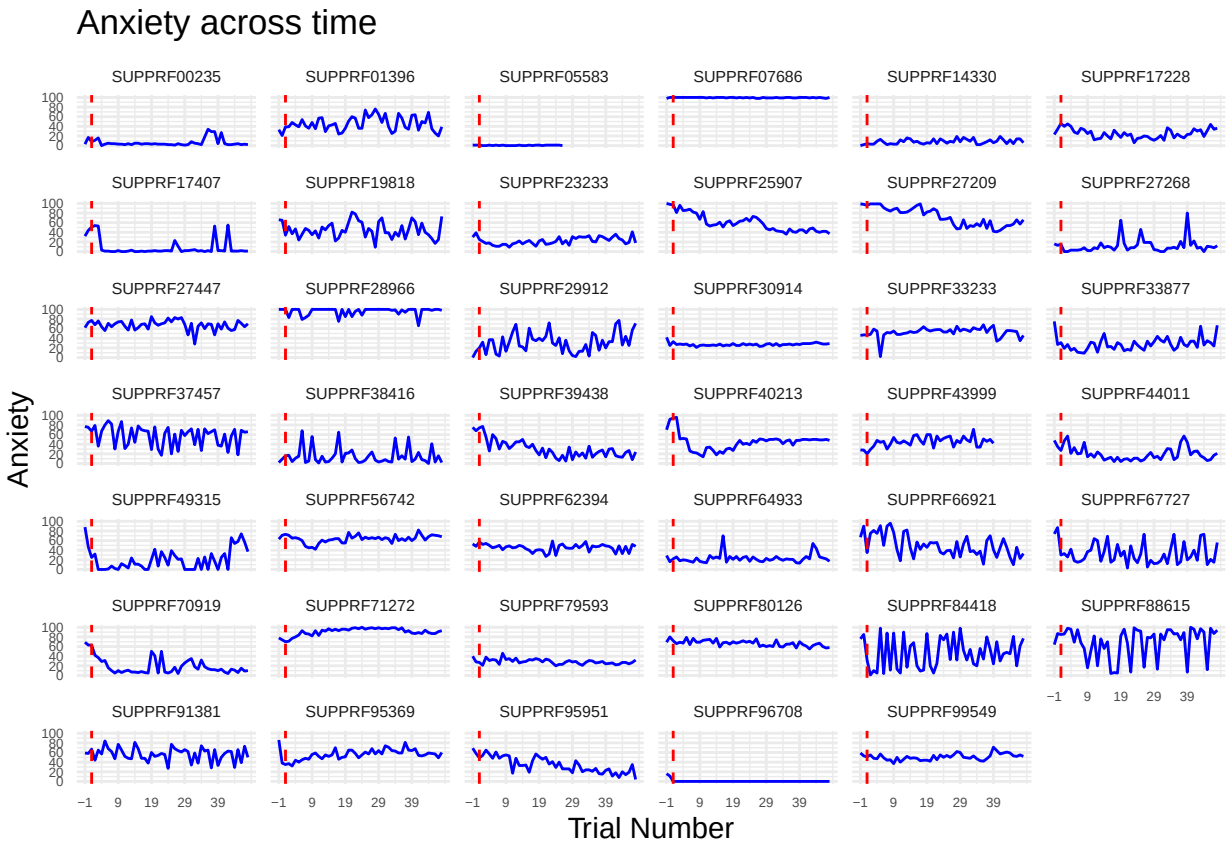
```
## [1] "average correlation between anxiety and feedback: -0.0912289783912545"
```



Mood over time



Anxiety over time



Participants' feedback

In the feedback we received from people, some people mentioned they did not like being watched and some mentioned since they did not see the other person, they did not care and did not think they were real and did not influence how they felt. We did have a few people who mentioned "the observer" or "they" when trying to descrive the virtual players. But overall, seeing how weaker the relationship is, I don't think having the cartoon images helps.

But since in this new pilot we have had several changes at once, shall we repeat the pilot with the new narrative but with the old pictures we had?

Anxiety- SubjPE correlations

Anxiety- Feedback correlations

If we wanted to run this again, we need the file “avatar_pilot_task_with_feedback.csv” which Elena has but since it has identifiable data we have not put it on Github.

Mood - SubjPE correlations

Mood-Feedback correlations

LME models for Mood and SubjPE

I will now look at the best LME models including feedback as well. But since subjective PE does include feedback, I will either only include feedback or SubjPE.

When looking at subjective PE, the best model is Mood ~ SubjPE + (SubjPE | Random_ID) with an AIC of 16192.33 When including feedback the best model is Mood ~ feedback + (feedback | Random_ID) with an AIC of 16014.14

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_SubjPE + (1 | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16325
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.2795 -0.4758  0.0368  0.5210  3.7862
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Random_ID (Intercept) 452.6    21.28
##  Residual              241.3    15.54
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)     53.48397    3.34176  16.005
## Response_SubjPE  0.11018    0.01438   7.663
##
## Correlation of Fixed Effects:
##             (Intr)
## Rspns_SbjPE -0.013
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_SubjPE + (Response_SubjPE | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16180.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.6350 -0.4948  0.0411  0.5136  3.7705
##
## Random effects:
##  Groups    Name            Variance  Std.Dev. Corr
##  Random_ID (Intercept)     407.58042 20.1886
##            Response_SubjPE   0.06984  0.2643  -0.06
##  Residual                  213.60721 14.6153
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
```

```
##                 Estimate Std. Error t value
## (Intercept)     53.14198    3.17504  16.737
## Response_SubjPE  0.14283    0.04369   3.269
##
## Correlation of Fixed Effects:
##             (Intr)
## Rspns_SbjPE -0.059
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_SubjPE * mini_SPIN_total + (Response_SubjPE |
##     Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16183.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.6484 -0.4944  0.0453  0.5159  3.7828
##
## Random effects:
##  Groups    Name            Variance  Std.Dev. Corr
##  Random_ID (Intercept)     417.22393 20.426
##            Response_SubjPE   0.06915  0.263   -0.06
##  Residual                  213.61570 14.616
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                                 Estimate Std. Error t value
## (Intercept)                     52.45925    7.07615   7.414
## Response_SubjPE                  0.24286    0.09611   2.527
## mini_SPIN_total                  0.12053    1.13898   0.106
## Response_SubjPE:mini_SPIN_total -0.01806    0.01548  -1.167
##
## Correlation of Fixed Effects:
##             (Intr) Rs_SPE m_SPIN
## Rspns_SbjPE -0.060
## mn_SPIN_ttl -0.891  0.052
## R_SPE:_SPIN  0.052 -0.892 -0.057
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_fdbk + (1 | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16239.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1198 -0.4919  0.0402  0.5467  3.4963
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Random_ID (Intercept) 418.6    20.46
```

```
##  Residual              231.1    15.20
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)   44.58492    3.30207   13.50
## Response_fdbk  0.17784    0.01459   12.19
##
## Correlation of Fixed Effects:
##             (Intr)
## Respns_fdbk -0.229

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_fdbk + (Response_fdbk | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16002.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5575 -0.4904  0.0469  0.5256  3.9285
##
## Random effects:
##  Groups    Name          Variance Std.Dev. Corr
##  Random_ID (Intercept)   608.7952 24.6738
##            Response_fdbk   0.0673  0.2594  -0.56
##  Residual                194.2163 13.9362
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)   44.30178    3.92932  11.275
## Response_fdbk  0.18269    0.04269   4.279
##
## Correlation of Fixed Effects:
##             (Intr)
## Respns_fdbk -0.578

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_H ~ Response_fdbk * mini_SPIN_total + (Response_fdbk |
##     Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16004.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5575 -0.4866  0.0457  0.5265  3.9442
##
## Random effects:
##  Groups    Name          Variance Std.Dev. Corr
##  Random_ID (Intercept)   614.304  24.7852
```

```
##            Response_fdbk   0.066   0.2569  -0.55
##  Residual                194.212  13.9360
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                               Estimate Std. Error t value
## (Intercept)                   38.02799    8.69603   4.373
## Response_fdbk                  0.29286    0.09328   3.139
## mini_SPIN_total                1.13193    1.39885   0.809
## Response_fdbk:mini_SPIN_total -0.01988    0.01500  -1.325
##
## Correlation of Fixed Effects:
##             (Intr) Rspns_ m_SPIN
## Respns_fdbk -0.568
## mn_SPIN_ttl -0.891  0.506
## Rsp_:_SPIN_  0.506 -0.891 -0.568
```

```
## [1] 16333.03
```

```
## [1] 16192.33
```

```
## [1] 16199.38
```

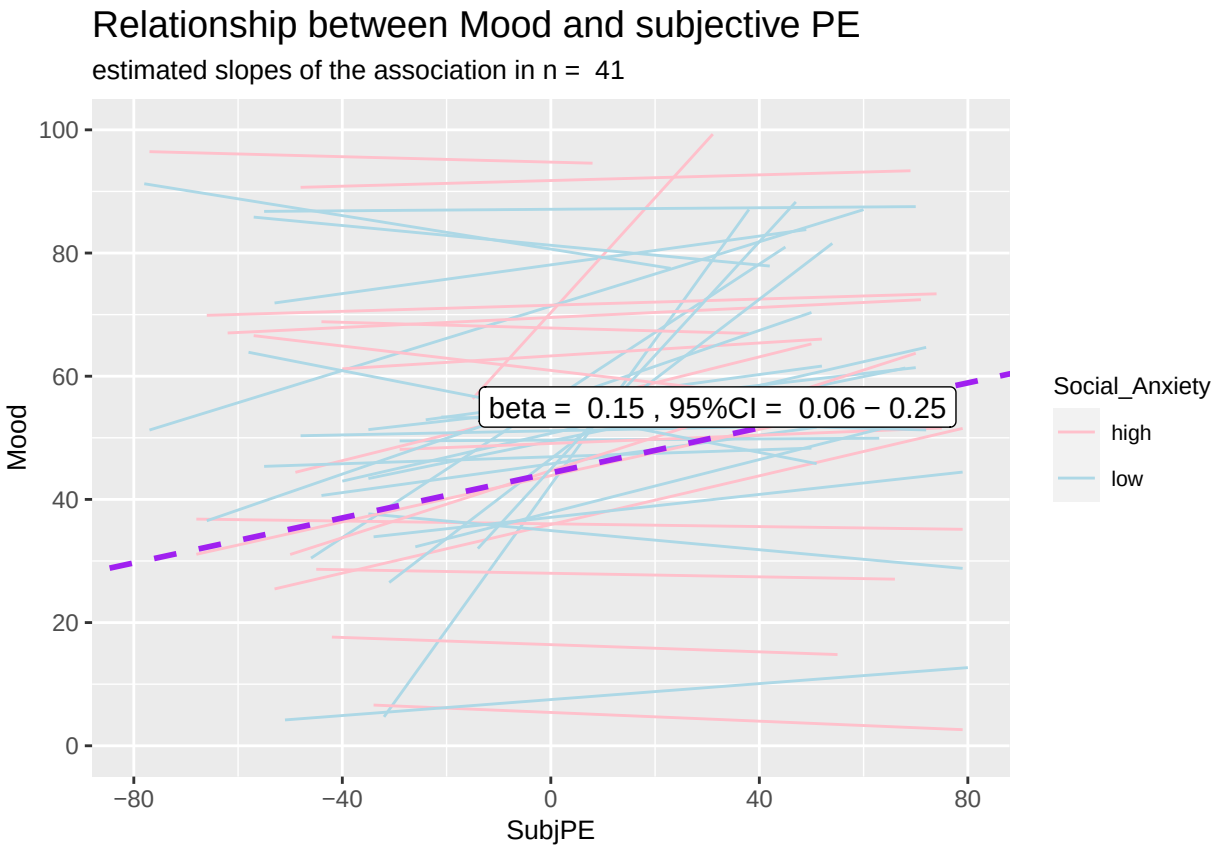
```
## [1] 16247.58
```

```
## [1] 16014.14
```

```
## [1] 16020.84
```

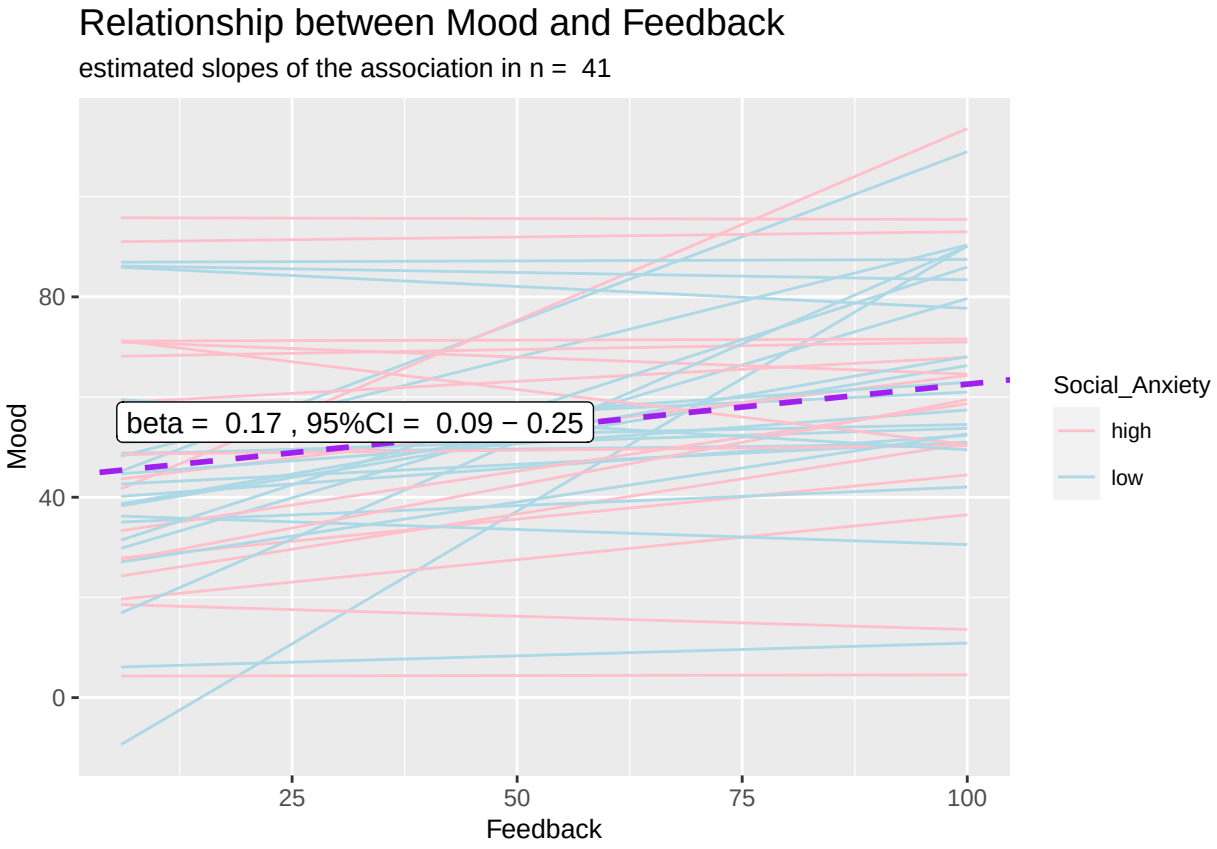
Individual plots with LME for Mood with SubjPE

When looking at subjective PE, the best model is Mood ~ SubjPE + (SubjPE | Random_ID) with an AIC of 16192.33



Individual plots with LME for Mood with feedback instead of SubjPE

When including feedback the best model is Mood ~ feedback + (feedback | Random_ID) with an AIC of 16014.14



LME models for Anxiety and SubjPE

When looking at subjective PE, the best model is Anxiety ~ SubjPE + (SubjPE | Random_ID) with an AIC of 16097.78 When including feedback the best model is Anxiety ~ feedback + (Random_ID) with an AIC of 16084.09

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + (1 | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16090
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4122 -0.4978 -0.0214  0.4007  4.6974
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Random_ID (Intercept) 641.0    25.32
##  Residual              211.7    14.55
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)     40.60638    3.96800  10.233
## Response_SubjPE -0.03483    0.01347  -2.585
##
## Correlation of Fixed Effects:
##             (Intr)
## Rspns_SbjPE -0.010

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE + (Response_SubjPE | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16085.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4263 -0.4982 -0.0235  0.3904  4.7073
##
## Random effects:
##  Groups    Name            Variance  Std.Dev. Corr
##  Random_ID (Intercept)     6.358e+02 25.21456
##            Response_SubjPE 3.349e-03  0.05787 -0.14
##  Residual                  2.097e+02 14.47930
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)     40.59479    3.95308  10.269
```

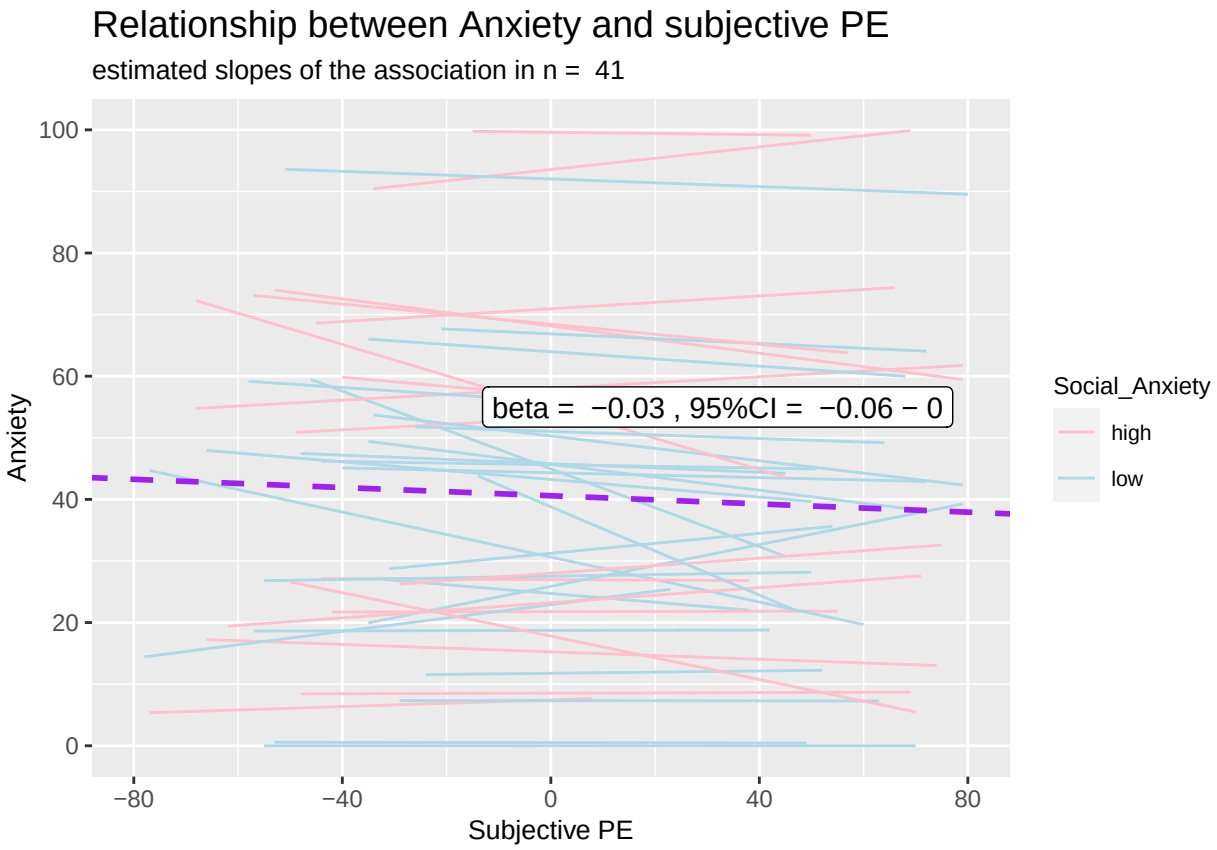
```
## Response_SubjPE -0.03332    0.01638  -2.034
##
## Correlation of Fixed Effects:
##             (Intr)
## Rspns_SbjPE -0.087
## optimizer (bobyqa) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_SubjPE * mini_SPIN_total + (Response_SubjPE |
##     Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16088
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4369 -0.4909 -0.0227  0.3936  4.7108
##
## Random effects:
##  Groups    Name            Variance  Std.Dev. Corr
##  Random_ID (Intercept)     6.066e+02 24.62886
##            Response_SubjPE 3.448e-03  0.05872 -0.21
##  Residual                  2.097e+02 14.47985
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                     28.030647   8.508686   3.294
## Response_SubjPE                 -0.059781   0.036886  -1.621
## mini_SPIN_total                  2.277306   1.369265   1.663
## Response_SubjPE:mini_SPIN_total  0.004817   0.005937   0.811
##
## Correlation of Fixed Effects:
##             (Intr) Rs_SPE m_SPIN
## Rspns_SbjPE -0.130
## mn_SPIN_ttl -0.891  0.113
## R_SPE:_SPIN  0.113 -0.895 -0.122
## optimizer (bobyqa) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_fdbk + (1 | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16076.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4537 -0.5042 -0.0314  0.4204  4.6990
```

```
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Random_ID (Intercept) 633.2    25.16
##  Residual              210.2    14.50
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)   43.78428    4.00926  10.921
## Response_fdbk -0.06324    0.01392  -4.545
##
## Correlation of Fixed Effects:
##             (Intr)
## Respns_fdbk -0.180

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_fdbk + (Response_fdbk | Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16072.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4765 -0.4987 -0.0323  0.4086  4.7191
##
## Random effects:
##  Groups    Name          Variance  Std.Dev. Corr
##  Random_ID (Intercept)   6.807e+02 26.0911
##            Response_fdbk 3.284e-03  0.0573  -0.36
##  Residual                2.084e+02 14.4348
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)   43.75779    4.15102  10.541
## Response_fdbk -0.06277    0.01651  -3.802
##
## Correlation of Fixed Effects:
##             (Intr)
## Respns_fdbk -0.339
## optimizer (bobyqa) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?

## Linear mixed model fit by REML ['lmerMod']
## Formula: Response_Ax ~ Response_fdbk * mini_SPIN_total + (Response_fdbk |
##     Random_ID)
##    Data: final_df14
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 16075.8
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4812 -0.4974 -0.0344  0.4075  4.7205
##
## Random effects:
##  Groups    Name          Variance  Std.Dev. Corr
##  Random_ID (Intercept)   650.93437 25.5134
##            Response_fdbk   0.00354  0.0595  -0.35
##  Residual                208.36408 14.4348
## Number of obs: 1939, groups:  Random_ID, 41
##
## Fixed effects:
##                                Estimate Std. Error t value
## (Intercept)                   30.482216   8.951108   3.405
## Response_fdbk                 -0.052512   0.036875  -1.424
## mini_SPIN_total                2.397379   1.440078   1.665
## Response_fdbk:mini_SPIN_total -0.001848   0.005906  -0.313
##
## Correlation of Fixed Effects:
##             (Intr) Rspns_ m_SPIN
## Respns_fdbk -0.338
## mn_SPIN_ttl -0.891  0.301
## Rsp_:_SPIN_  0.302 -0.892 -0.338
## optimizer (bobyqa) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?

## [1] 16098.03

## [1] 16097.78

## [1] 16104.05

## [1] 16084.09

## [1] 16084.61

## [1] 16091.8
```

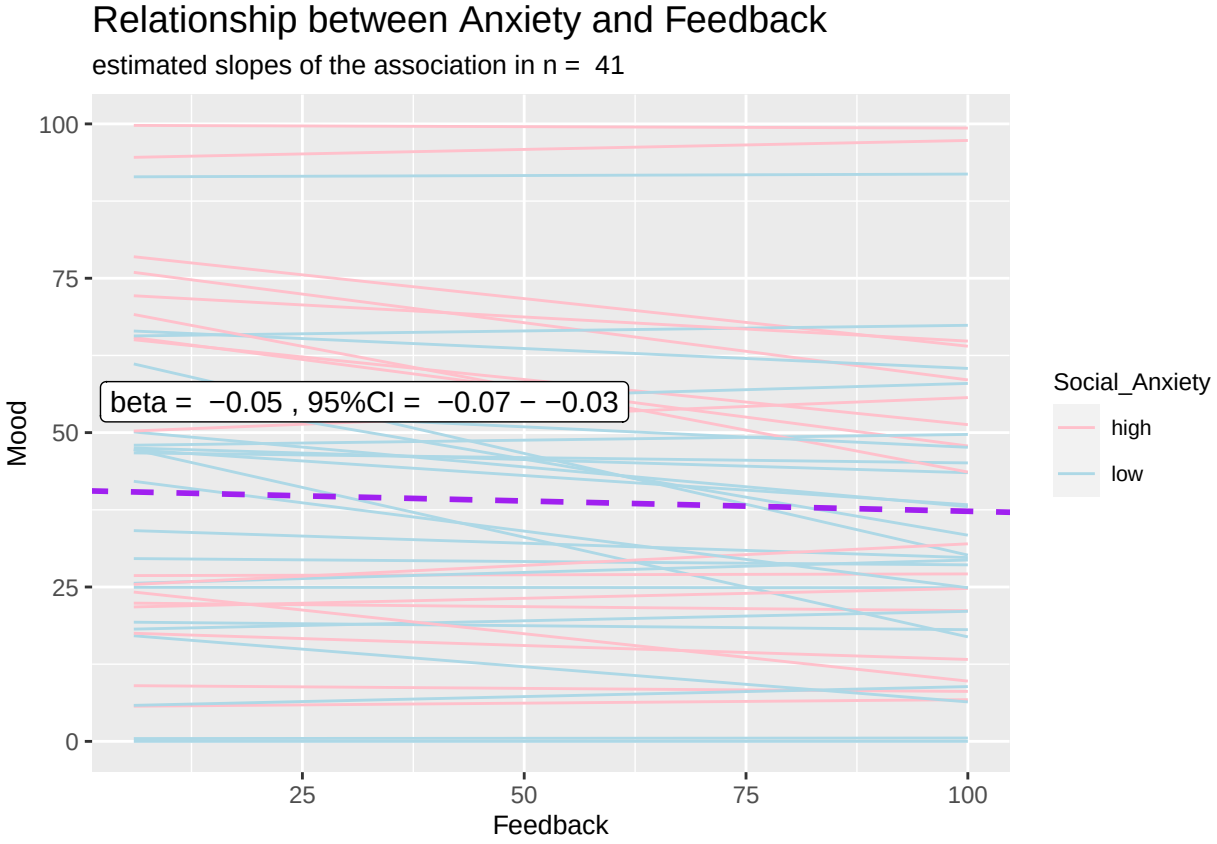
Individual plots with LME for Anxiety with SubjPE

When looking at subjective PE, the best model is Anxiety ~ SubjPE + (SubjPE | Random_ID) with an AIC of 16097.78



Individual plots with LME for Anxiety with feedback instead of SubjPE

When including feedback the best model is Anxiety ~ feedback + (Random_ID) with an AIC of 16084.09



ICC for anxiety

we will now look at the ICC outcome for anxiety The ICC for anxiety is 0.75, which is moderate/good 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] 0.7489496
```

```
##                 2.5 %   97.5 %
## .sig01      20.27662 31.43409
## .sigma      14.12142 15.04917
## (Intercept) 32.68154 48.32104
```

ICC for mood

The ICC for mood is 0.42, which is lower than anxiety and is actually within the poor category, 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 mood with just the intercept"
```

```
## [1] 0.6271569
```

```
##                 2.5 %   97.5 %
## .sig01      16.46445 25.58591
## .sigma      15.29259 16.29728
## (Intercept) 47.43910 60.19157
```