BOI_Regression

Stan West

2024-10-07

This document describes the relationship between scores on the ADC and performance on a semantic decision task. Independent samples t-tests show that groups overall do not differ in accuracy or response latency $t(603.71) = -0.56$, $p = 0.58$ & $t(597.39) = 1.49$, $p = 0.14$, respectively. Distributions of response latency and accuracy across groups and BOI are shown in the violin plots below. Also plotted is the distribution of accuracy and resposne latency across a range of ADC scores. Each point refers to the average latency or average accuracy for a participant in each of the BOI levels. This is separated out by group. Finally, there are nested linear models showing the relationship between each of our DVs (resposne latency & accuracy) and group, group and BOI level, and group, BOI level, and ADC score. Nested model comparisons show that the best fitting model for both of the DVs includes group & BOI level, but not ADC scores.

Table of means and violin plots

Table 1: Descriptives Between Groups

Group	Mean Duration	SD Duration	Average Percent Correct	SD Percent Correct
ASD	977.7054	184.7615	0.8499320	0.1245079
NA	1001.5459	208.4481	0.8440951	0.1324285

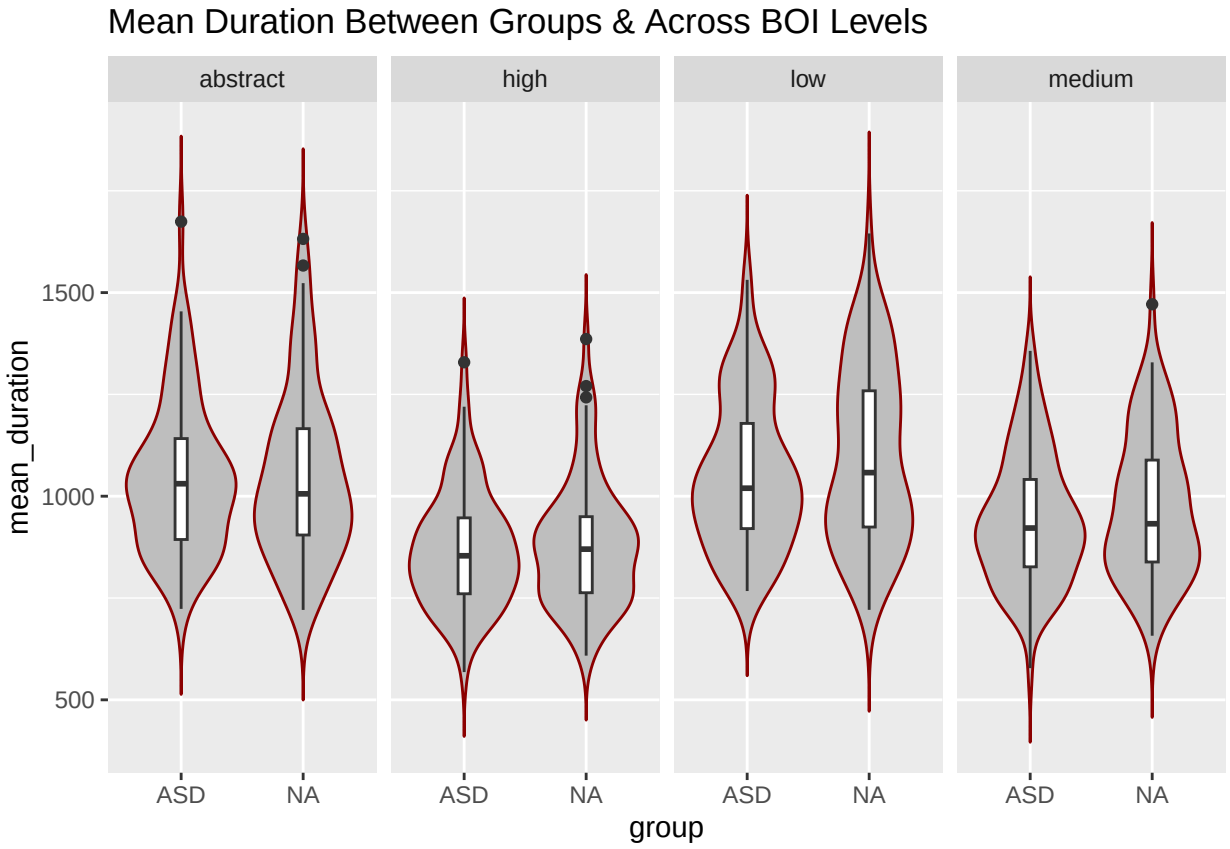
Table 2: Descriptives Between Groups and BOI Levels

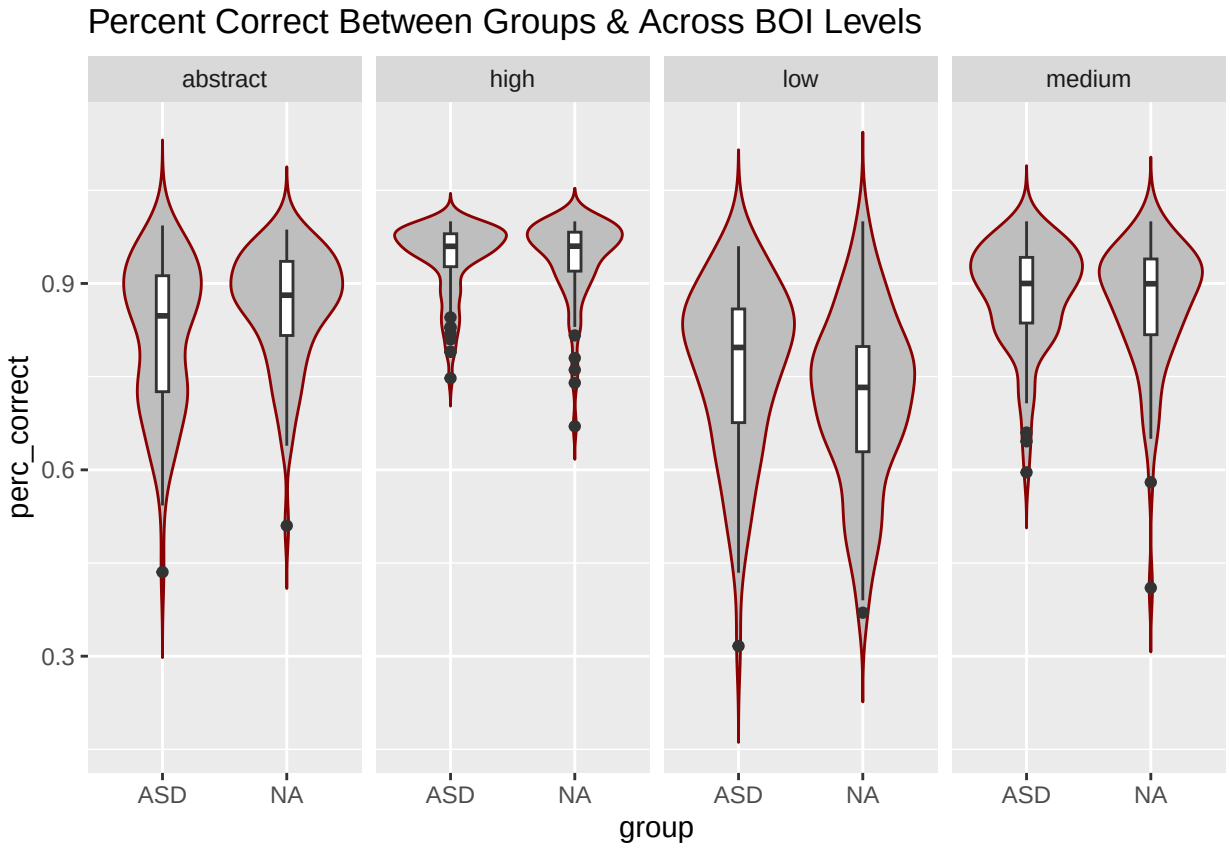
Group	Level	Mean Duration	SD Duration	Average Percent Correct	SD Percent Correct
ASD	abstract	1045.2303	186.0486	0.8152527	0.1211803
ASD	high	870.3469	144.4500	0.9401900	0.0589838
ASD	low	1052.1727	183.1366	0.7633212	0.1382650
ASD	medium	943.0715	160.4966	0.8809642	0.0854218
NA	abstract	1048.6056	210.4664	0.8611279	0.0934746
NA	high	883.5153	159.3386	0.9394869	0.0667352
NA	low	1098.1592	219.4395	0.7136195	0.1401492
NA	medium	975.9036	176.6799	0.8621461	0.1045079

Table 3: ADC Descriptives Between Groups

Group	Mean ADC	SD ADC
ASD	38.56579	15.29411

Group	Mean ADC	SD ADC
NA	23.36842	14.14665





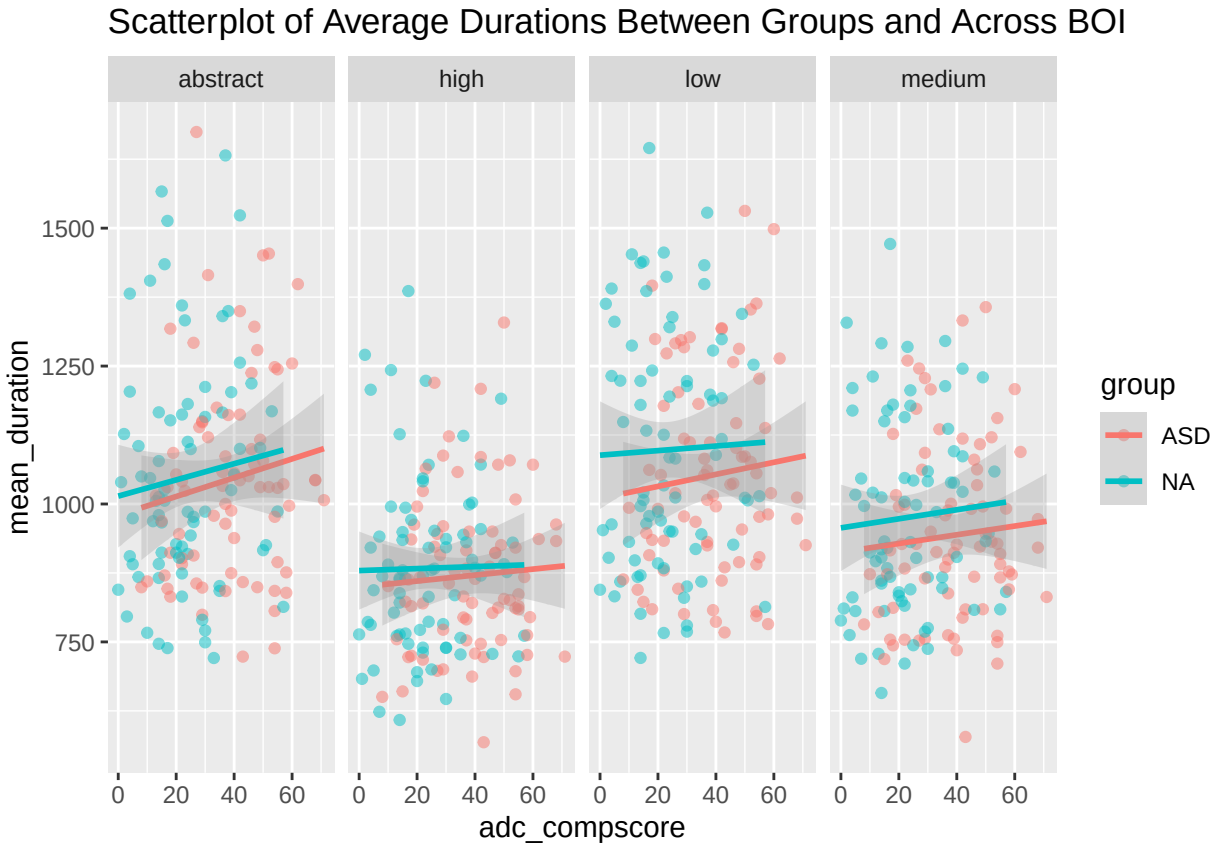
T-tests of overall response latency and accuracy. Groups do not differ in overall response latency or accuracy.

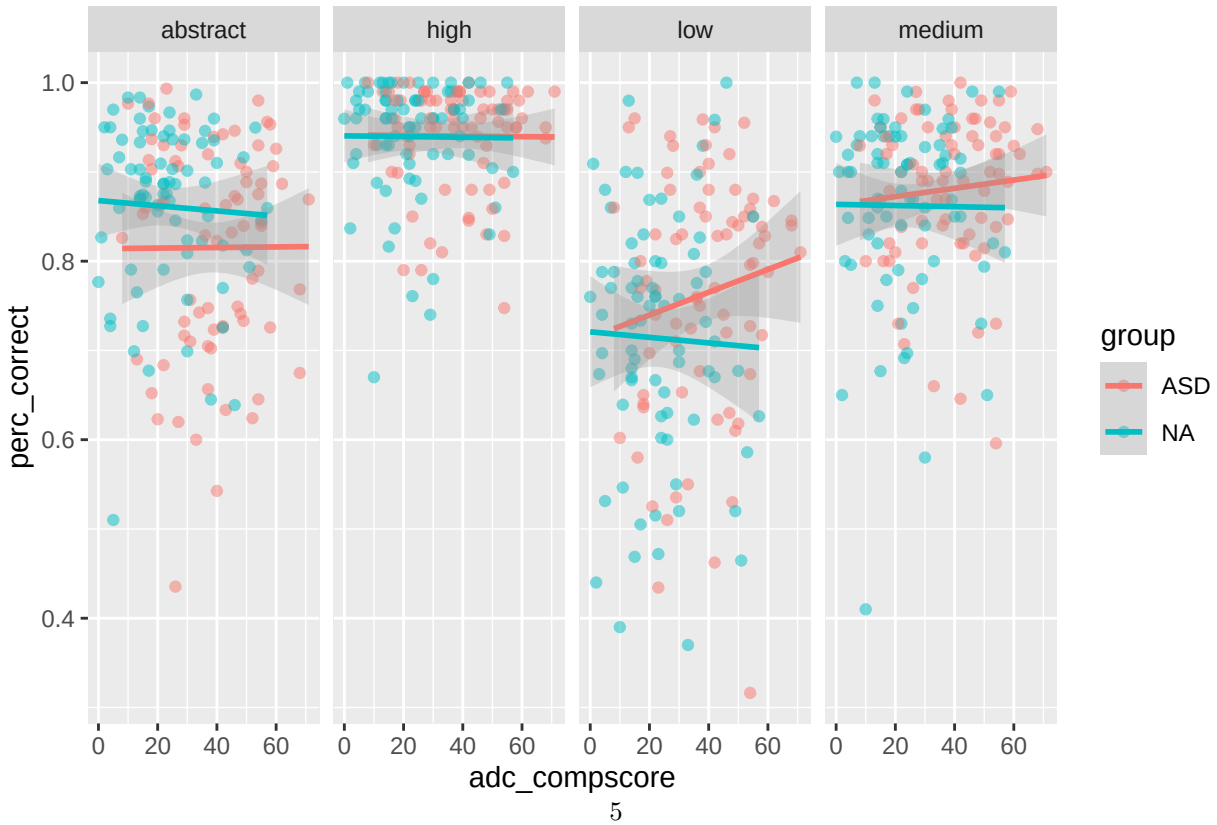
```
##
##  Welch Two Sample t-test
##
## data:  mean_duration by group
## t = -1.4923, df = 597.39, p-value = 0.1361
## alternative hypothesis: true difference in means between group ASD and group NA is not equal to 0
## 95 percent confidence interval:
##  -55.215863   7.534715
## sample estimates:
## mean in group ASD  mean in group NA
##          977.7054         1001.5459

##
##  Welch Two Sample t-test
##
## data:  perc_correct by group
## t = 0.55989, df = 603.71, p-value = 0.5758
## alternative hypothesis: true difference in means between group ASD and group NA is not equal to 0
## 95 percent confidence interval:
##  -0.01463697  0.02631075
## sample estimates:
## mean in group ASD  mean in group NA
```

```
##           0.8499320          0.8440951
```

Plotting duration and percent correct over ADC scores.





Nested model comparisons (group, group + level, & group + level + ADC, with interactions)

```
m1_duration <- lm(mean_duration ~ group, data = df_descrip)
m2_duration <- lm(mean_duration ~ group * level, data = df_descrip)
m3_duration <- lm(mean_duration ~ group * level * adc_compscore, data = df_descrip)
list_mods_dur <- list(model_1 = m1_duration,model_2 = m2_duration,model_3 = m3_duration)
lapply(list_mods_dur, summary)
```

```
## $model_1
##
## Call:
## lm(formula = mean_duration ~ group, data = df_descrip)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -409.57 -141.51  -33.56  116.38  696.59
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   977.71      11.30  86.549   <2e-16 ***
## groupNA        23.84      15.98   1.492    0.136
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 197 on 606 degrees of freedom
## Multiple R-squared:  0.003661,   Adjusted R-squared:  0.002017
## F-statistic: 2.227 on 1 and 606 DF,  p-value: 0.1361
##
##
## $model_2
##
## Call:
## lm(formula = mean_duration ~ group * level, data = df_descrip)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -377.17 -137.33  -19.78  109.96  629.07
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1045.230     20.831  50.176  < 2e-16 ***
## groupNA                3.375     29.460   0.115 0.908822
## levelhigh           -174.883     29.460  -5.936 4.93e-09 ***
## levellow               6.942     29.460   0.236 0.813780
## levelmedium         -102.159     29.460  -3.468 0.000562 ***
## groupNA:levelhigh      9.793     41.662   0.235 0.814244
## groupNA:levellow      42.611     41.662   1.023 0.306824
## groupNA:levelmedium   29.457     41.662   0.707 0.479817
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 181.6 on 600 degrees of freedom
## Multiple R-squared:  0.1614, Adjusted R-squared:  0.1516
## F-statistic:  16.5 on 7 and 600 DF,  p-value: < 2.2e-16
##
##
## $model_3
##
## Call:
## lm(formula = mean_duration ~ group * level * adc_compscore, data = df_descrip)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -373.3 -134.7  -24.8  110.9  648.8
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       979.5777    56.7751  17.254   <2e-16 ***
## groupNA                            34.7183    69.6853   0.498    0.619
## levelhigh                        -130.1352    80.2921  -1.621    0.106
## levellow                           30.5363    80.2921   0.380    0.704
## levelmedium                       -66.9746    80.2921  -0.834    0.405
## adc_compscore                       1.7024     1.3688   1.244    0.214
## groupNA:levelhigh                  -4.8776    98.5499  -0.049    0.961
## groupNA:levellow                   43.6429    98.5499   0.443    0.658
## groupNA:levelmedium                 9.4033    98.5499   0.095    0.924
## groupNA:adc_compscore              -0.2341     2.0158  -0.116    0.908
## levelhigh:adc_compscore            -1.1603     1.9358  -0.599    0.549
## levellow:adc_compscore             -0.6118     1.9358  -0.316    0.752
## levelmedium:adc_compscore          -0.9123     1.9358  -0.471    0.638
## groupNA:levelhigh:adc_compscore    -0.1268     2.8508  -0.044    0.965
## groupNA:levellow:adc_compscore     -0.4420     2.8508  -0.155    0.877
## groupNA:levelmedium:adc_compscore   0.2648     2.8508   0.093    0.926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 182.2 on 592 degrees of freedom
## Multiple R-squared:  0.1671, Adjusted R-squared:  0.146
## F-statistic: 7.918 on 15 and 592 DF,  p-value: < 2.2e-16
```

```
anova( m1_duration,m2_duration,m3_duration)
```

```
## Analysis of Variance Table
##
## Model 1: mean_duration ~ group
## Model 2: mean_duration ~ group * level
## Model 3: mean_duration ~ group * level * adc_compscore
##   Res.Df      RSS Df Sum of Sq       F Pr(>F)
## 1    606 23508989
## 2    600 19787433  6   3721556 18.6841 <2e-16 ***
## 3    592 19652755  8    134678  0.5071 0.8513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
m1_correct <- lm(perc_correct ~ group, data = df_descrip)
m2_correct <- lm(perc_correct ~ group * level, data = df_descrip)
m3_correct <- lm(perc_correct ~ group * level * adc_compscore, data = df_descrip)
list_mods_acc <- list(model_1 = m1_correct, model_2 = m2_correct, model_3 = m3_correct)
lapply(list_mods_acc, summary)
```

```
## $model_1
##
## Call:
## lm(formula = perc_correct ~ group, data = df_descrip)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53361 -0.06882  0.03027  0.10007  0.15590
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.849932   0.007372  115.30   <2e-16 ***
## groupNA     -0.005837   0.010425   -0.56    0.576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1285 on 606 degrees of freedom
## Multiple R-squared:  0.000517,   Adjusted R-squared:  -0.001132
## F-statistic: 0.3135 on 1 and 606 DF,  p-value: 0.5758
##
##
## $model_2
##
## Call:
## lm(formula = perc_correct ~ group * level, data = df_descrip)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45215 -0.05217  0.02051  0.06074  0.28638
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.81525    0.01205  67.637  < 2e-16 ***
## groupNA              0.04588    0.01705   2.691 0.007317 **
## levelhigh            0.12494    0.01705   7.329 7.51e-13 ***
## levellow            -0.05193    0.01705  -3.047 0.002417 **
## levelmedium          0.06571    0.01705   3.855 0.000128 ***
## groupNA:levelhigh   -0.04658    0.02411  -1.932 0.053810 .
## groupNA:levellow    -0.09558    0.02411  -3.965 8.23e-05 ***
## groupNA:levelmedium -0.06469    0.02411  -2.684 0.007484 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1051 on 600 degrees of freedom
## Multiple R-squared:  0.3386, Adjusted R-squared:  0.3309
## F-statistic: 43.88 on 7 and 600 DF,  p-value: < 2.2e-16
##
```

```
##
## $model_3
##
## Call:
## lm(formula = perc_correct ~ group * level * adc_compscore, data = df_descrip)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46671 -0.05201  0.02056  0.06084  0.29342
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        8.140e-01  3.288e-02  24.760  < 2e-16 ***
## groupNA                            5.391e-02  4.035e-02   1.336  0.18201
## levelhigh                          1.274e-01  4.649e-02   2.741  0.00631 **
## levellow                          -9.993e-02  4.649e-02  -2.149  0.03200 *
## levelmedium                        4.904e-02  4.649e-02   1.055  0.29194
## adc_compscore                      3.253e-05  7.926e-04   0.041  0.96728
## groupNA:levelhigh                 -5.492e-02  5.706e-02  -0.962  0.33625
## groupNA:levellow                  -4.709e-02  5.706e-02  -0.825  0.40959
## groupNA:levelmedium               -5.329e-02  5.706e-02  -0.934  0.35079
## groupNA:adc_compscore             -3.229e-04  1.167e-03  -0.277  0.78216
## levelhigh:adc_compscore           -6.456e-05  1.121e-03  -0.058  0.95409
## levellow:adc_compscore             1.245e-03  1.121e-03   1.110  0.26725
## levelmedium:adc_compscore          4.323e-04  1.121e-03   0.386  0.69990
## groupNA:levelhigh:adc_compscore    3.149e-04  1.651e-03   0.191  0.84877
## groupNA:levellow:adc_compscore    -1.265e-03  1.651e-03  -0.767  0.44363
## groupNA:levelmedium:adc_compscore -2.070e-04  1.651e-03  -0.125  0.90025
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1055 on 592 degrees of freedom
## Multiple R-squared:  0.3421, Adjusted R-squared:  0.3255
## F-statistic: 20.52 on 15 and 592 DF,  p-value: < 2.2e-16
```

```
anova(m1_correct,m2_correct,m3_correct)
```

```
## Analysis of Variance Table
##
## Model 1: perc_correct ~ group
## Model 2: perc_correct ~ group * level
## Model 3: perc_correct ~ group * level * adc_compscore
##   Res.Df     RSS Df Sum of Sq       F Pr(>F)
## 1    606 10.0110
## 2    600  6.6249  6    3.3860 50.7013 <2e-16 ***
## 3    592  6.5893  8    0.0356  0.3998 0.9208
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nested model comparisons no abstract category (group, group + level, & group + level + ADC, with interactions)

```
d_descrip_no_abstract <- df_descrip %>%
  filter(!level == "abstract")

m1_duration <- lm(mean_duration ~ group, data = d_descrip_no_abstract)
m2_duration <- lm(mean_duration ~ group * level, data = d_descrip_no_abstract)
m3_duration <- lm(mean_duration ~ group * level * adc_compscore, data = d_descrip_no_abstract)
list_mods_dur <- list(model_1 = m1_duration,model_2 = m2_duration,model_3 = m3_duration)
lapply(list_mods_dur, summary)
```

```
## $model_1
##
## Call:
## lm(formula = mean_duration ~ group, data = d_descrip_no_abstract)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -387.07 -143.59  -34.32  121.87  659.39
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   955.20      12.78  74.746   <2e-16 ***
## groupNA        30.66      18.07   1.697   0.0905 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 193 on 454 degrees of freedom
## Multiple R-squared:  0.0063, Adjusted R-squared:  0.004112
## F-statistic: 2.879 on 1 and 454 DF,  p-value: 0.09045
##
##
## $model_2
##
## Call:
## lm(formula = mean_duration ~ group * level, data = d_descrip_no_abstract)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -377.17 -133.89  -18.52  109.96  547.09
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           870.35      20.14  43.220  < 2e-16 ***
## groupNA                13.17      28.48   0.462    0.644
## levellow              181.83      28.48   6.385 4.29e-10 ***
## levelmedium            72.72      28.48   2.554    0.011 *
## groupNA:levellow       32.82      40.28   0.815    0.416
## groupNA:levelmedium    19.66      40.28   0.488    0.626
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 175.6 on 450 degrees of freedom
## Multiple R-squared:  0.1847, Adjusted R-squared:  0.1757
## F-statistic: 20.39 on 5 and 450 DF,  p-value: < 2.2e-16
##
##
## $model_3
##
## Call:
## lm(formula = mean_duration ~ group * level * adc_compscore, data = d_descrip_no_abstract)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -373.28 -131.65  -20.49  103.93  549.73
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       849.4426    54.9726  15.452   <2e-16 ***
## groupNA                            29.8407    67.4730   0.442   0.6585
## levellow                          160.6714    77.7430   2.067   0.0393 *
## levelmedium                        63.1606    77.7430   0.812   0.4170
## adc_compscore                       0.5420     1.3253   0.409   0.6827
## groupNA:levellow                   48.5205    95.4212   0.508   0.6114
## groupNA:levelmedium                14.2809    95.4212   0.150   0.8811
## groupNA:adc_compscore              -0.3609     1.9518  -0.185   0.8534
## levellow:adc_compscore              0.5485     1.8743   0.293   0.7699
## levelmedium:adc_compscore           0.2480     1.8743   0.132   0.8948
## groupNA:levellow:adc_compscore     -0.3152     2.7603  -0.114   0.9091
## groupNA:levelmedium:adc_compscore   0.3916     2.7603   0.142   0.8872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 176.4 on 444 degrees of freedom
## Multiple R-squared:  0.1877, Adjusted R-squared:  0.1676
## F-statistic: 9.326 on 11 and 444 DF,  p-value: 3.899e-15
```

```
anova( m1_duration,m2_duration,m3_duration)
```

```
## Analysis of Variance Table
##
## Model 1: mean_duration ~ group
## Model 2: mean_duration ~ group * level
## Model 3: mean_duration ~ group * level * adc_compscore
##   Res.Df      RSS Df Sum of Sq       F Pr(>F)
## 1    454 16904268
## 2    450 13869168  4   3035100 24.3800 <2e-16 ***
## 3    444 13818518  6     50651  0.2712 0.9503
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
m1_correct <- lm(perc_correct ~ group, data = d_descrip_no_abstract)
m2_correct <- lm(perc_correct ~ group * level, data = d_descrip_no_abstract)
m3_correct <- lm(perc_correct ~ group * level * adc_compscore, data = d_descrip_no_abstract)
```

```r
list_mods_acc <- list(model_1 = m1_correct, model_2 = m2_correct, model_3 = m3_correct)
lapply(list_mods_acc, summary)
```

```
## $model_1
##
## Call:
## lm(formula = perc_correct ~ group, data = d_descrip_no_abstract)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.54517 -0.06306  0.03647  0.10158  0.16158
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.861492   0.008848  97.369   <2e-16 ***
## groupNA     -0.023074   0.012512  -1.844   0.0658 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1336 on 454 degrees of freedom
## Multiple R-squared:  0.007435,   Adjusted R-squared:  0.005249
## F-statistic: 3.401 on 1 and 454 DF,  p-value: 0.06582
##
##
## $model_2
##
## Call:
## lm(formula = perc_correct ~ group * level, data = d_descrip_no_abstract)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45215 -0.04425  0.02031  0.06051  0.28638
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.940190   0.011931  78.803  < 2e-16 ***
## groupNA             -0.000703   0.016873  -0.042 0.966783
## levellow            -0.176869   0.016873 -10.482  < 2e-16 ***
## levelmedium         -0.059226   0.016873  -3.510 0.000493 ***
## groupNA:levellow    -0.048999   0.023862  -2.053 0.040609 *
## groupNA:levelmedium -0.018115   0.023862  -0.759 0.448153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.104 on 450 degrees of freedom
## Multiple R-squared:  0.4037, Adjusted R-squared:  0.397
## F-statistic: 60.92 on 5 and 450 DF,  p-value: < 2.2e-16
##
##
## $model_3
##
## Call:
## lm(formula = perc_correct ~ group * level * adc_compscore, data = d_descrip_no_abstract)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46671 -0.04715  0.02018  0.05928  0.29342
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       9.414e-01  3.251e-02  28.955  < 2e-16 ***
## groupNA                          -1.003e-03  3.991e-02  -0.025   0.9800
## levellow                         -2.274e-01  4.598e-02  -4.945 1.08e-06 ***
## levelmedium                      -7.839e-02  4.598e-02  -1.705   0.0889 .
## adc_compscore                    -3.203e-05  7.839e-04  -0.041   0.9674
## groupNA:levellow                  7.828e-03  5.644e-02   0.139   0.8897
## groupNA:levelmedium               1.631e-03  5.644e-02   0.029   0.9770
## groupNA:adc_compscore            -7.982e-06  1.154e-03  -0.007   0.9945
## levellow:adc_compscore            1.309e-03  1.109e-03   1.181   0.2382
## levelmedium:adc_compscore         4.968e-04  1.109e-03   0.448   0.6543
## groupNA:levellow:adc_compscore   -1.580e-03  1.633e-03  -0.968   0.3336
## groupNA:levelmedium:adc_compscore -5.219e-04  1.633e-03  -0.320   0.7494
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1043 on 444 degrees of freedom
## Multiple R-squared:  0.4079, Adjusted R-squared:  0.3932
## F-statistic:  27.8 on 11 and 444 DF,  p-value: < 2.2e-16
```

```
anova(m1_correct,m2_correct,m3_correct)
```

```
## Analysis of Variance Table
##
## Model 1: perc_correct ~ group
## Model 2: perc_correct ~ group * level
## Model 3: perc_correct ~ group * level * adc_compscore
##   Res.Df    RSS Df Sum of Sq       F Pr(>F)
## 1    454 8.1030
## 2    450 4.8683  4    3.2348 74.2781 <2e-16 ***
## 3    444 4.8340  6    0.0343  0.5252 0.7893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```