Models for Perceptual Difficulty Paper

Exp.2

Mixed effects logistic regression predicting redundant adjective use from fixed effects of redundant property, with random by-subject and by-item intercepts

going from high difficulty-material redundant(0) to low difficulty-color redundant(1) & no redundancy(0) to redundancy(1) –> should be positive

```
##
## high_difficulty  low_difficulty
##             357             335
```

```
##
##   0   1
## 519 173
```

```
##
##                     0   1
##   high_difficulty 333  24
##   low_difficulty  186 149
```

```
##
##          boot_leather_green          bottle_glass_green
##                          42                          44
##        bottle_plastic_green           chair_metal_green
##                          38                          48
##          chair_metal_purple           cup_plastic_green
##                          44                          39
##         jacket_denim_purple          pitcher_metal_blue
##                          40                          45
##            plate_paper_blue plate_plastic_blue_original
##                          88                          46
##            spoon_wood_green            table_metal_blue
##                          45                          86
##           table_metal_green table_metal_silver_original
##                          44                          43
```

```
##
##    boot  bottle   chair     cup  jacket pitcher   plate   spoon   table
##      42      82      92      39      40      45     134      45     173
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: redundant ~ ctrialType + (1 + ctrialType | gameid) + (1 | targetName)
##    Data: targets
##
##      AIC      BIC   logLik deviance df.resid
##    509.5    536.8   -248.8    497.5      686
```

```
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.92499 -0.27108 -0.01346  0.05618  2.73976
##
## Random effects:
##  Groups     Name        Variance Std.Dev. Corr
##  gameid     (Intercept) 12.622   3.553
##             ctrialType  38.280   6.187    -0.91
##  targetName (Intercept)  3.375   1.837
## Number of obs: 692, groups:  gameid, 51; targetName, 14
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.946      1.049  -4.717 2.40e-06 ***
## ctrialType     7.569      1.779   4.255 2.09e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr)
## ctrialType -0.778
```

Exp.3

Mixed effects linear regression predicting logRT to redundant adjective from fixed effects of redundant property –> to replicate the effect from Exp1

going from high difficulty(0) to low difficulty = material to color adjectives –> logRT decreases = should be negative

```
##
##    color material
##     6338     6310
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: logRT ~ cFeatureQuestion + (1 + cFeatureQuestion | targetname) +
##     (1 + cFeatureQuestion | workerid)
##    Data: pd_all
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+06))
##
##      AIC      BIC   logLik deviance df.resid
##   5026.5   5093.5  -2504.3   5008.5    12639
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.2595 -0.5602 -0.0824  0.4742 11.9512
##
## Random effects:
##  Groups     Name             Variance  Std.Dev. Corr
##  workerid   (Intercept)      0.0586192 0.24211
##             cFeatureQuestion 0.0062700 0.07918  0.25
##  targetname (Intercept)      0.0007589 0.02755
```

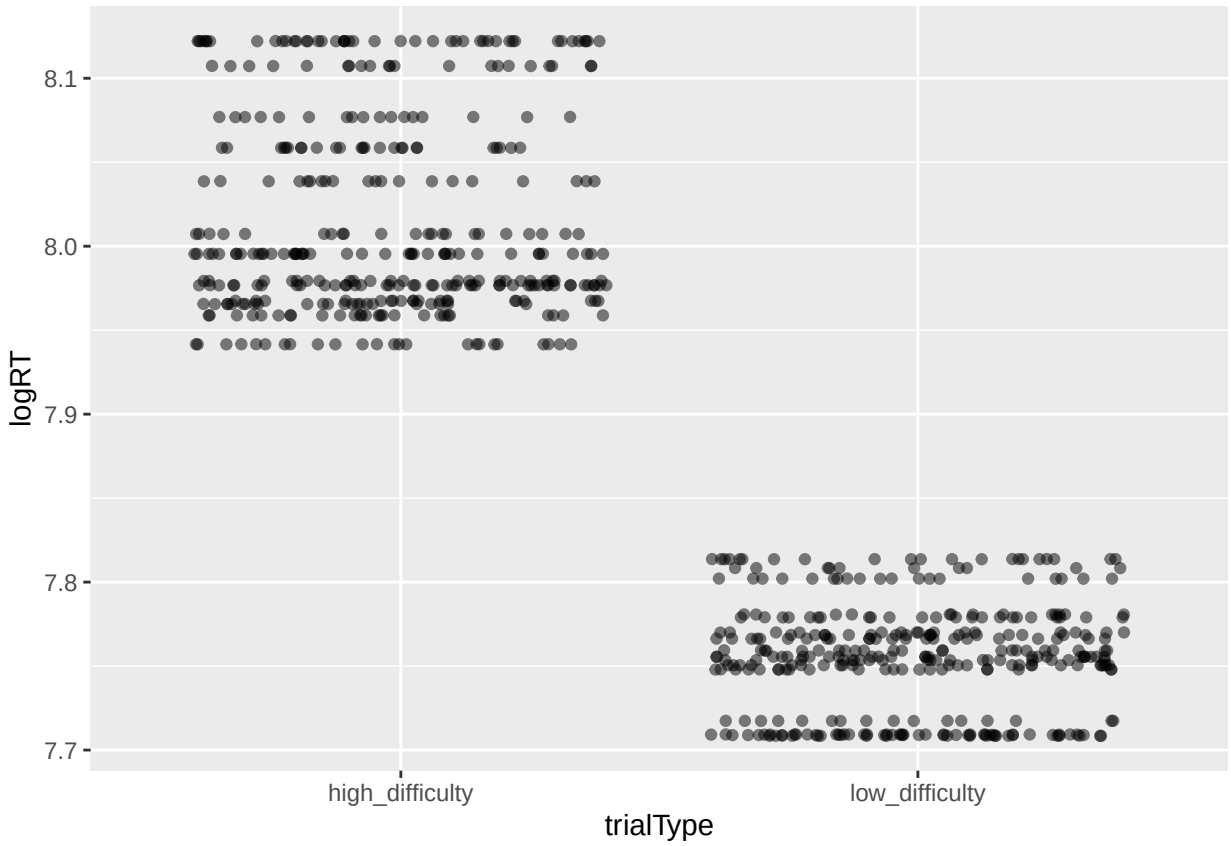
```
##             cFeatureQuestion 0.0037598 0.06132  0.50
##  Residual                    0.0769377 0.27738
## Number of obs: 12648, groups:  workerid, 407; targetname, 14
##
## Fixed effects:
##                   Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)        7.95484    0.01431 127.77628   555.7  < 2e-16 ***
## cFeatureQuestion   0.24159    0.01763  15.83905    13.7 3.37e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## cFeaturQstn 0.285
```

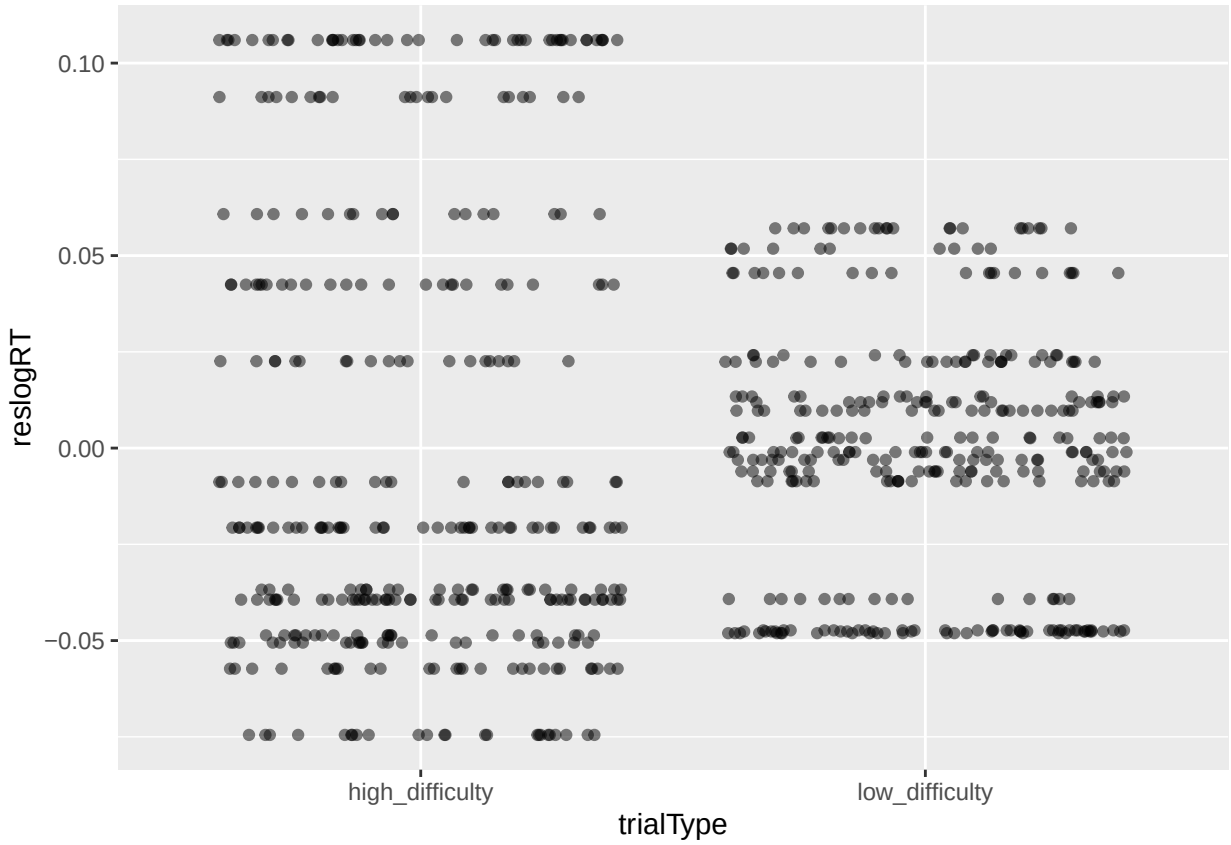
Exp.2 & Exp. 3

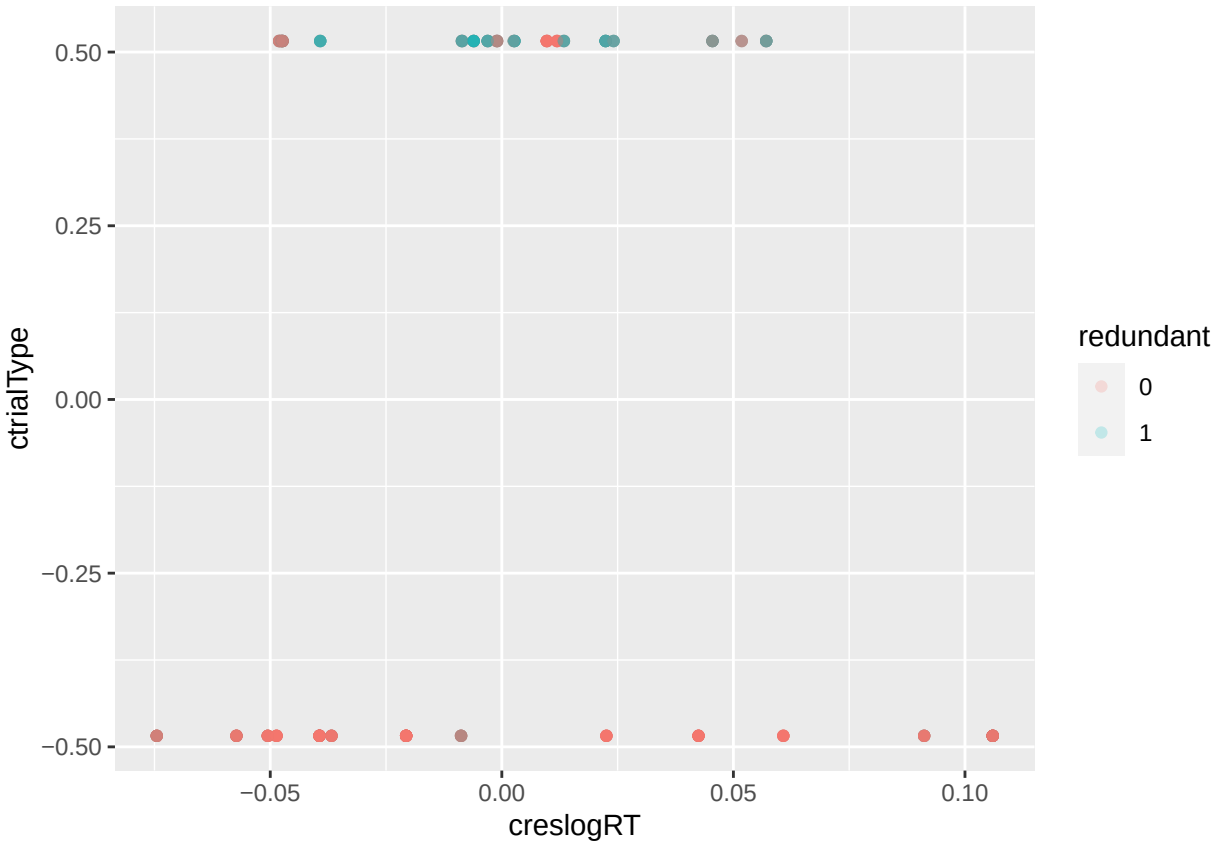
Mixed effects logistic regression predicting redundant adjective use from redundant property (color or material), RT to redundant adjective in context and their interaction

bigger RT = more perceptually difficulty = less redundant adjective use

```
##
## Call:
## lm(formula = logRT ~ ctrialType, data = tomodel)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.074505 -0.039379 -0.006076  0.022572  0.105989
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.890476   0.001821 4333.79   <2e-16 ***
## ctrialType  -0.259528   0.003643  -71.24   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04789 on 690 degrees of freedom
## Multiple R-squared:  0.8803, Adjusted R-squared:  0.8801
## F-statistic:  5075 on 1 and 690 DF,  p-value: < 2.2e-16
```


8.1
8.0
7.9
7.8
7.7
logRT
high_difficulty
low_difficulty
trialType


0.10
0.05
0.00
−0.05
reslogRT
high_difficulty
low_difficulty
trialType




```
## Warning in (function (fn, par, lower = rep.int(-Inf, n), upper = rep.int(Inf, :
## failure to converge in 10000 evaluations
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0609965 (tol = 0.002, component 1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: redundant ~ ctrialType * creslogRT + (1 + ctrialType * creslogRT |
##     gameid) + (1 | targetName)
##    Data: tomodel
##
##      AIC      BIC   logLik deviance df.resid
##    526.1    594.2   -248.1    496.1      677
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.90878 -0.23938 -0.01235  0.04826  2.72875
##
## Random effects:
##  Groups     Name                 Variance Std.Dev. Corr
##  gameid     (Intercept)           12.439   3.527
##             ctrialType            36.522   6.043   -0.90
##             creslogRT             31.000   5.568    0.70 -0.90
##             ctrialType:creslogRT 265.162  16.284   -0.89  0.97 -0.77
##  targetName (Intercept)            3.698   1.923
## Number of obs: 692, groups:  gameid, 51; targetName, 14
```

```
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)             -5.007      1.292  -3.876 0.000106 ***
## ctrialType               7.427      2.275   3.264 0.001097 **
## creslogRT              -11.303     16.411  -0.689 0.490978
## ctrialType:creslogRT    13.572     31.873   0.426 0.670236
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) ctrlTy crslRT
## ctrialType  -0.827
## creslogRT    0.324 -0.358
## ctrlTyp:cRT -0.332  0.394 -0.761
## convergence code: 0
## Model failed to converge with max|grad| = 0.0609965 (tol = 0.002, component 1)
## failure to converge in 10000 evaluations
```

```
## # Check for Multicollinearity
##
## Low Correlation
##
##              Parameter  VIF Increased SE
##             ctrialType 1.20         1.09
##              creslogRT 2.40         1.55
##  ctrialType:creslogRT 2.48         1.57
```

```
agr = tomodel %>%
  mutate(redundant = as.numeric(as.character(redundant)), binreslogRT=cut_interval(reslogRT,10)) %>%
  group_by(trialType,binreslogRT) %>%
  summarise(Proportion = mean(redundant), CILow=ci.low(redundant),CIHigh=ci.high(redundant)) %>%
  ungroup() %>%
  mutate(YMin=Proportion-CILow,YMax=Proportion+CIHigh)
```

```
## `summarise()` regrouping output by 'trialType' (override with `.groups` argument)
```

```
# no effect of RT above and beyond trial type:
ggplot(agr, aes(x=binreslogRT,y=Proportion,color=trialType,group=trialType)) +
  geom_point() +
  geom_smooth() +
  geom_errorbar(aes(ymin=YMin,ymax=YMax),width=.25) +
  theme(axis.text.x=element_text(angle=45,hjust=1,vjust=1))
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 4
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : Chernobyl! trL>n 6
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : Chernobyl! trL>n 6
```

```
## Warning in sqrt(sum.squares/one.delta): NaNs produced
```

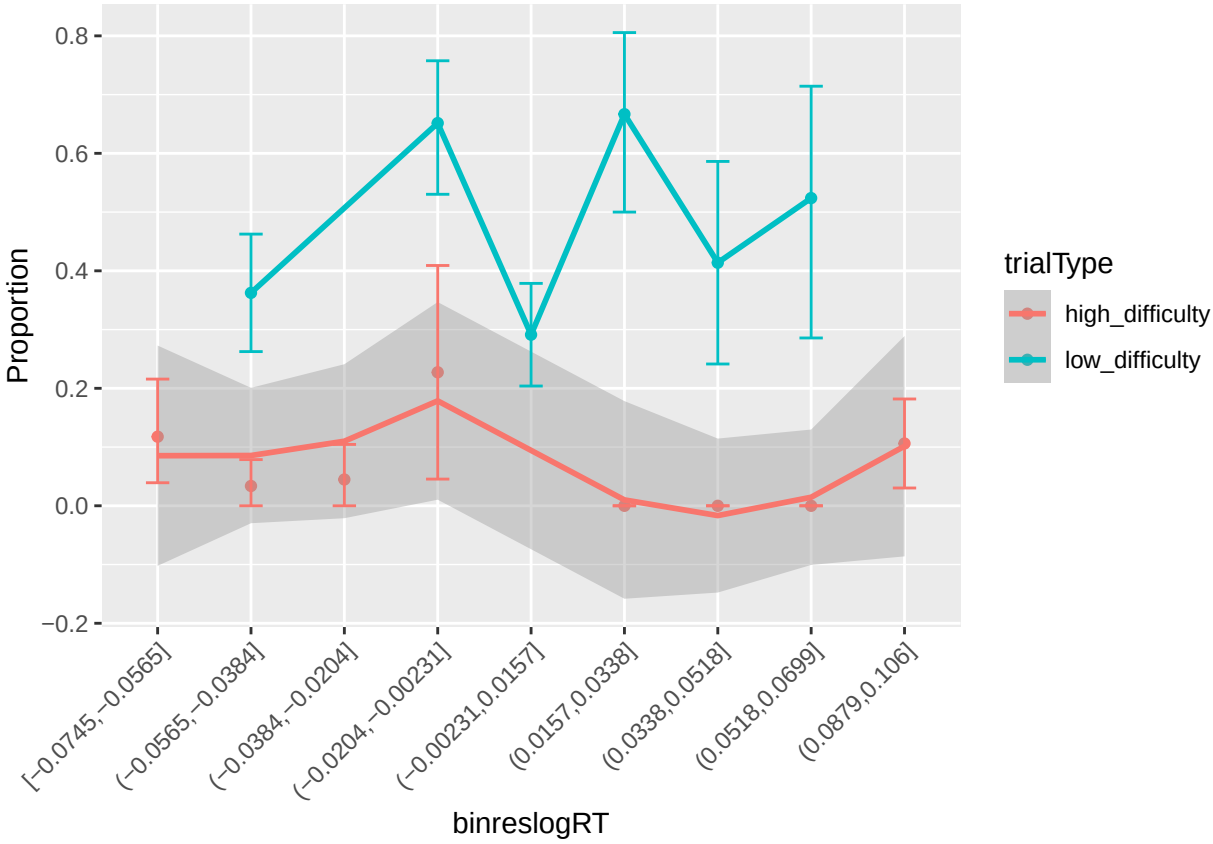
```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used at 4
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius 2
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal condition
## number 0
```

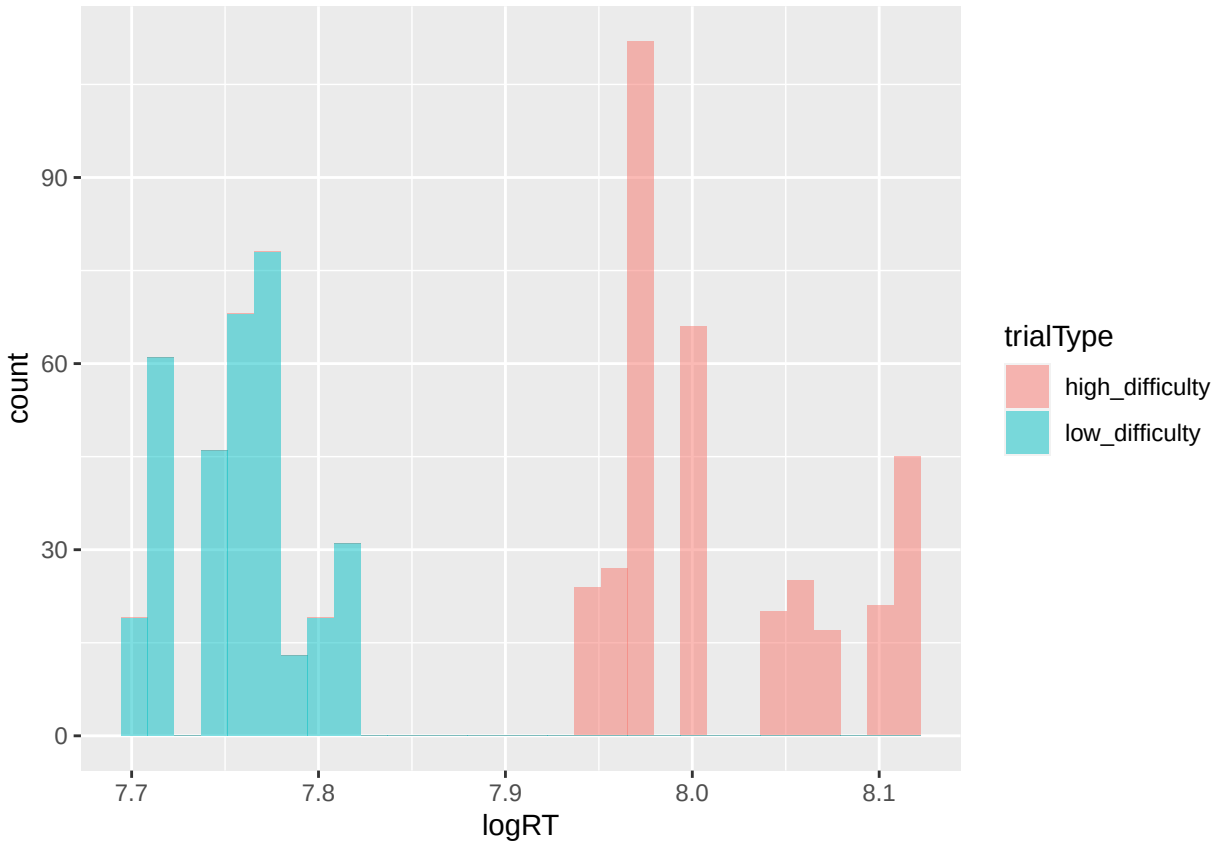
```
## Warning in stats::qt(level/2 + 0.5, pred$df): NaNs produced
```

```
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning -
## Inf
```



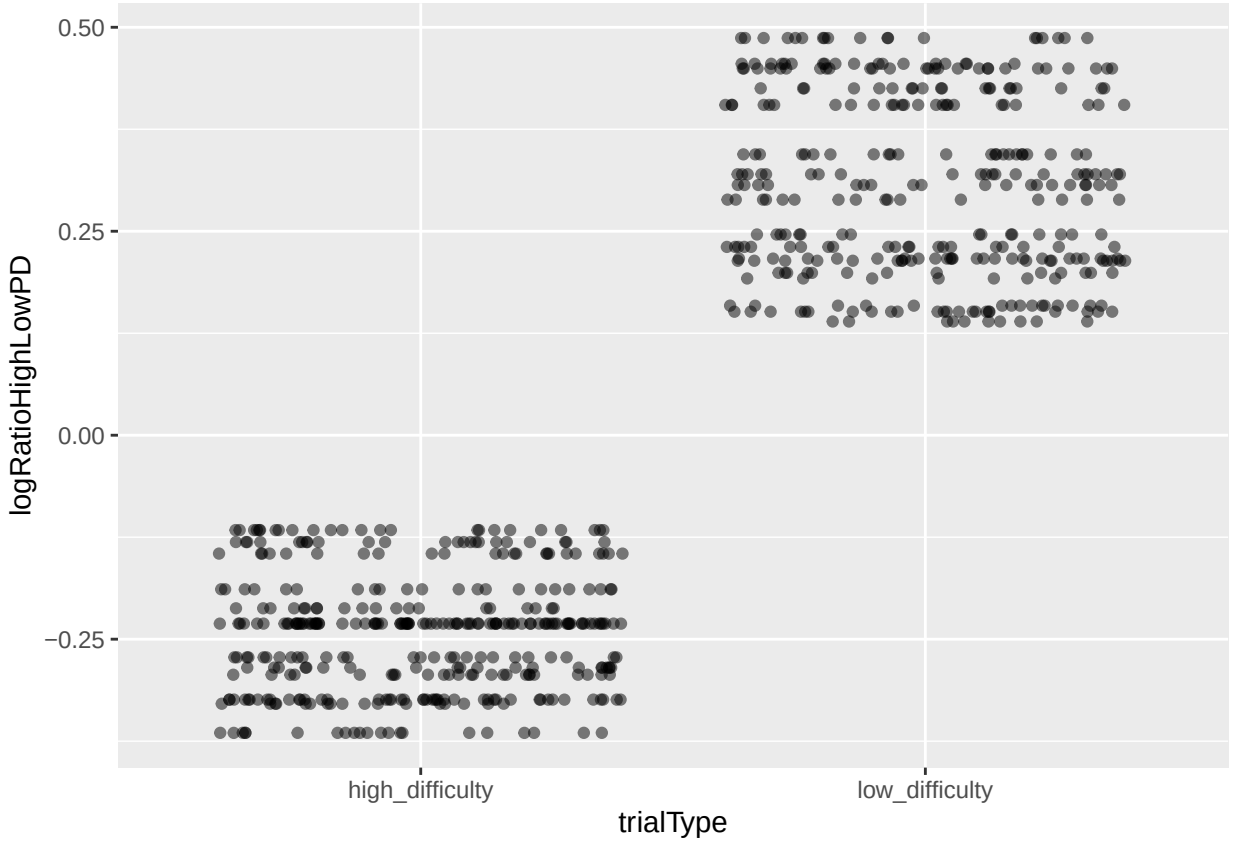

```
# histograms of mean logRTs by trial type completely disjoint. so: trial type explains difference in re
ggplot(tomodel, aes(x=logRT,fill=trialType)) +
  geom_histogram(alpha=.5)
```

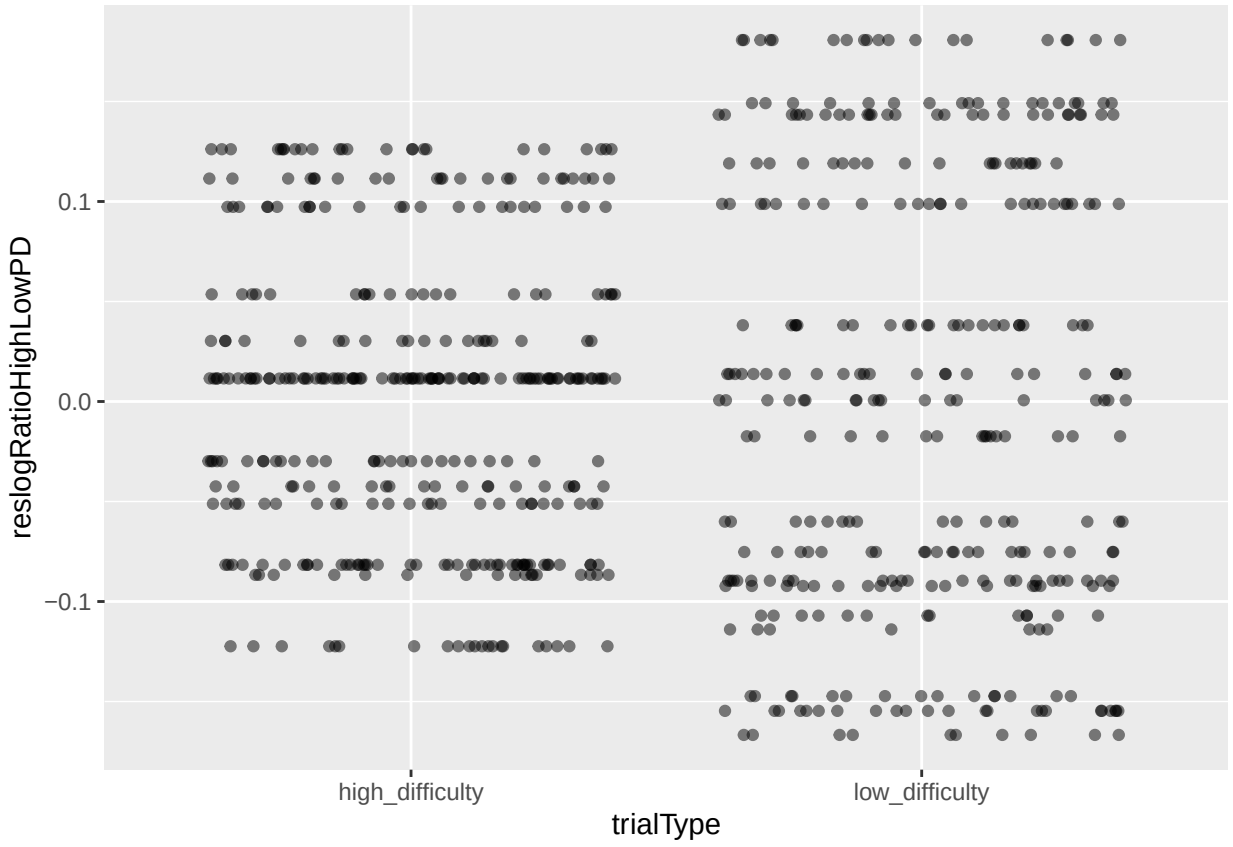
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

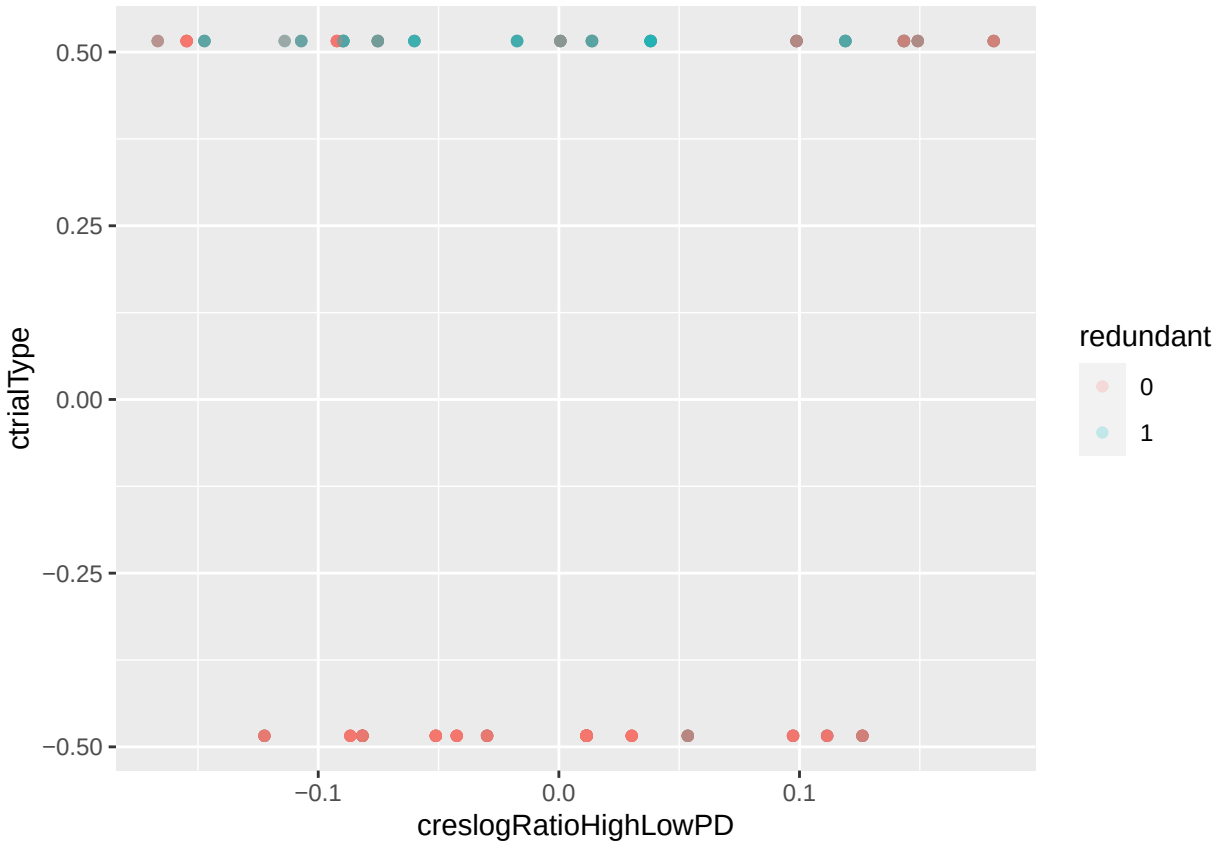


Same model with perceptual difficulty difference score for each context (difference between RTs to target's sufficient and redundant feature)

```
##
## Call:
## lm(formula = logRatioHighLowPD ~ ctrialType, data = tomodel)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16670 -0.08165  0.01146  0.09732  0.18072
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.023109   0.003513   6.578 9.44e-11 ***
## ctrialType  0.548513   0.007030  78.024  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09242 on 690 degrees of freedom
## Multiple R-squared:  0.8982, Adjusted R-squared:  0.898
## F-statistic:  6088 on 1 and 690 DF,  p-value: < 2.2e-16
```


0.50
0.25
0.00
−0.25
logRatioHighLowPD
high_difficulty
low_difficulty
trialType


0.1
0.0
−0.1
reslogRatioHighLowPD
high_difficulty
low_difficulty
trialType




```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: redundant ~ creslogRatioHighLowPD * ctrialType + (1 + ctrialType |
##     gameid) + (1 | targetName)
##    Data: tomodel
##
##      AIC      BIC   logLik deviance df.resid
##    510.7    547.1   -247.4    494.7      684
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5636 -0.2424 -0.0103  0.0577  3.0475
##
## Random effects:
##  Groups     Name        Variance Std.Dev. Corr
##  gameid     (Intercept) 14.317   3.784
##             ctrialType  44.518   6.672    -0.92
##  targetName (Intercept)  3.886   1.971
## Number of obs: 692, groups:  gameid, 51; targetName, 14
##
## Fixed effects:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                       -5.264      1.065  -4.945 7.62e-07 ***
## creslogRatioHighLowPD              5.309      3.868   1.373    0.170
## ctrialType                         7.719      1.726   4.472 7.76e-06 ***
## creslogRatioHighLowPD:ctrialType  -12.501      8.857  -1.411    0.158
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) crRHLPD ctrlTy
## crslgRtHLPD -0.254
## ctrialType  -0.720  0.067
## crslRHLPD:T  0.266 -0.566  -0.162
```

```
## # Check for Multicollinearity
##
## Low Correlation
##
##                          Parameter  VIF Increased SE
##             creslogRatioHighLowPD 1.47         1.21
##                         ctrialType 1.03         1.01
##  creslogRatioHighLowPD:ctrialType 1.51         1.23
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0338011 (tol = 0.002, component 1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## redundant ~ creslogRatioHighLowPD + ctrialType + (1 + creslogRatioHighLowPD +
##     ctrialType | gameid) + (1 | targetName)
##    Data: tomodel
##
##      AIC      BIC   logLik deviance df.resid
##    512.3    557.7   -246.1    492.3      682
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.91234 -0.19988 -0.01199  0.02512  2.74248
##
## Random effects:
##  Groups     Name                   Variance Std.Dev. Corr
##  gameid     (Intercept)            13.485   3.672
##             creslogRatioHighLowPD 24.109   4.910     0.38
##             ctrialType            41.603   6.450    -0.90  0.06
##  targetName (Intercept)             3.593   1.896
## Number of obs: 692, groups:  gameid, 51; targetName, 14
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            -5.0711     1.0460  -4.848 1.25e-06 ***
## creslogRatioHighLowPD   0.9378     3.3684   0.278    0.781
## ctrialType              7.4773     1.7358   4.308 1.65e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) cRHLPD
## crslgRtHLPD -0.043
```

```
## ctrialType  -0.731 -0.050
## convergence code: 0
## Model failed to converge with max|grad| = 0.0338011 (tol = 0.002, component 1)
```

```
## # Check for Multicollinearity
##
## Low Correlation
##
##              Parameter  VIF Increased SE
##  creslogRatioHighLowPD 1.00         1.00
##             ctrialType 1.00         1.00
```

```
## Data: tomodel
## Models:
## m4: redundant ~ creslogRatioHighLowPD * ctrialType + (1 + ctrialType |
## m4:     gameid) + (1 | targetName)
## m5: redundant ~ creslogRatioHighLowPD + ctrialType + (1 + creslogRatioHighLowPD +
## m5:     ctrialType | gameid) + (1 | targetName)
##    npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m4    8 510.74 547.06 -247.37   494.74
## m5   10 512.28 557.68 -246.14   492.28 2.4572  2     0.2927
```