

E1_report

Gabriel Strain

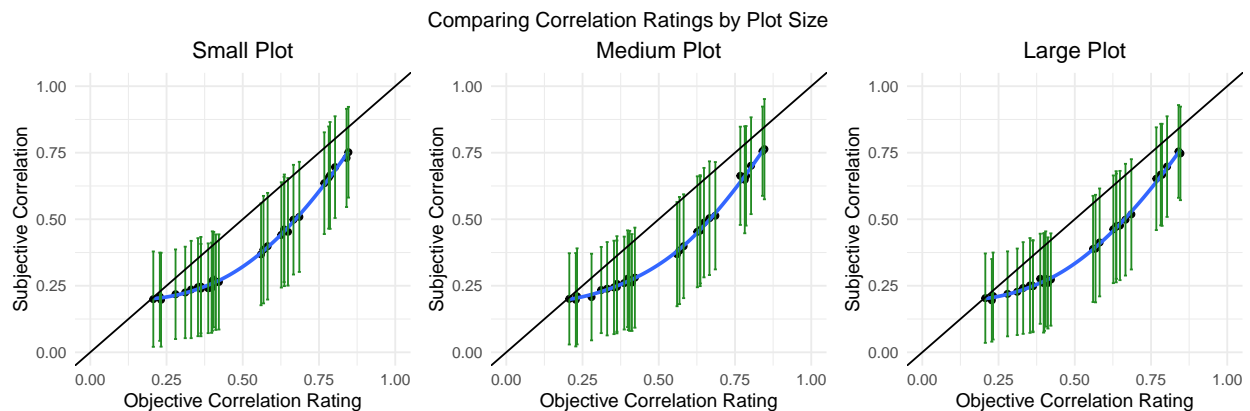
21/03/2022

Performance on Correlation Estimation is Better on Larger Plots. In addition, the Presence of a Novel Contrast Encoding also Improves Performance.

Plotting the Results

The plots below show subjective correlation performance for Small, Medium, and Large plots. These sizes corresponded to 63%, 100% and 252% size ratios. We hypothesised that performance at estimating correlation would be better when the plot was larger. N = 118, 180 experimental trials per participant (fully repeated measures: all participants saw all versions of each plot).

```
## 'geom_smooth()' using formula 'y ~ x'  
## 'geom_smooth()' using formula 'y ~ x'  
## 'geom_smooth()' using formula 'y ~ x'
```

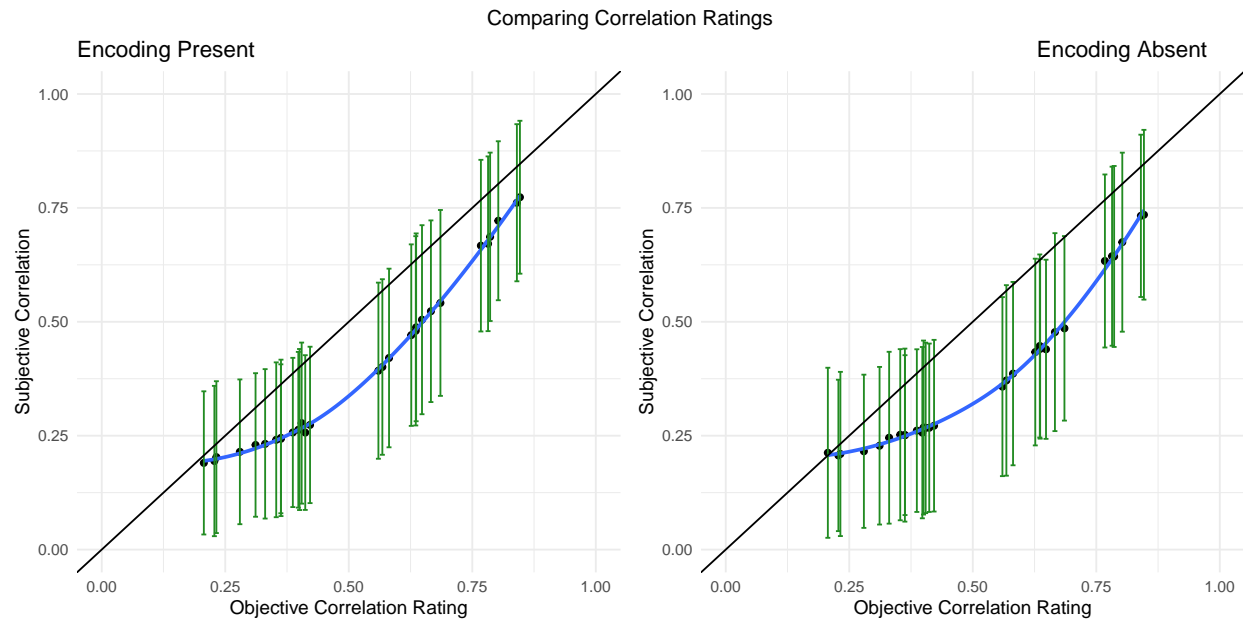


The summary statistics for the factor of size are as follows.

```
## # A tibble: 3 x 3  
##   size    mean    sd  
##   <fct> <dbl> <dbl>  
## 1 L      0.120 0.190  
## 2 M      0.122 0.193  
## 3 S      0.128 0.193
```

We utilised a novel contrast encoding whereby the contrast of a certain point was linearly related to the size of the residual of that point. The presence of this contrast encoding was associated with better performance on correlation estimation. Summary statistics are below the plots.

```
## 'geom_smooth()' using formula 'y ~ x'
## 'geom_smooth()' using formula 'y ~ x'
```



```
## # A tibble: 2 x 3
##   present mean    sd
##   <fct>   <dbl> <dbl>
## 1 N       0.132 0.198
## 2 Y       0.115 0.185
```

Modelling

Modelling Main Effects

A Linear Mixed Model was fitted with the **buildmer** package, which begins with a maximal model and drops terms until it converges and no errors are returned. The model built had the following syntax;

```
model <- lmer(difference ~ size * present + (1 | participant) + (1 | item))
```

This model has random intercepts for participants and items.

Checking Presence of Interaction

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: only_those_passed
## Models:
## model_no_interact: difference ~ size + present + (1 | participant) + (1 | item)
## model: difference ~ size * present + (1 | participant) + (1 | item)
##               npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## model_no_interact    7 -20797 -20741  10406   -20811
## model                9 -20796 -20724  10407   -20814  2.881  2    0.2368
```

There was no significant difference between models with and without the interaction.

An ANOVA comparing models with and without our interventions revealed a significant effect $\chi^2(5) = 91.91, p < .0001$. This effect was driven ratings being significantly higher for Large vs Small plots ($p = .001$) and Medium vs Small plots ($p = .025$), and being higher when the contrast encoding was employed ($p < .0001$).

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: difference ~ size * present + (1 | participant) + (1 | item)
## Data: only_those_passed
##
## REML criterion at convergence: -20749.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -5.8513 -0.6092  0.0180  0.6162  5.9993
##
## Random effects:
##   Groups      Name      Variance Std.Dev.
## participant (Intercept) 0.013378 0.11566
## item        (Intercept) 0.002309 0.04805
## Residual                0.021210 0.14564
## Number of obs: 21168, groups: participant, 118; item, 30
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   1.232e-01  1.383e-02 1.139e+02  8.907 9.68e-15 ***
## size1         -3.625e-03  1.416e-03 2.102e+04 -2.561  0.0105 *
## size2         -1.381e-03  1.416e-03 2.102e+04 -0.975  0.3294
## present1       8.721e-03  1.001e-03 2.102e+04  8.712 < 2e-16 ***
## size1:present1  1.944e-03  1.416e-03 2.102e+04  1.373  0.1698
## size2:present1  2.506e-04  1.416e-03 2.102e+04  0.177  0.8595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) size1  size2  prsnt1  sz1:p1
## size1          0.000
## size2          0.000 -0.500
## present1       0.000  0.000  0.000
## siz1:prsnt1    0.000  0.000  0.000  0.000
## siz2:prsnt1    0.000  0.000  0.000  0.001 -0.500
```

There was no effect of an interaction between the two factors. Estimated marginal means and contrasts for the two factors can be seen below:

```

## $emmeans
##   size emmean      SE df asymp.LCL asymp.UCL
##   L      0.120 0.0139 Inf    0.0923    0.147
##   M      0.122 0.0139 Inf    0.0946    0.149
##   S      0.128 0.0139 Inf    0.1010    0.155
##
## Results are averaged over the levels of: present
## Degrees-of-freedom method: asymptotic
## Confidence level used: 0.95
##
## $contrasts
##   contrast estimate      SE df z.ratio p.value
##   L - M      -0.00224 0.00245 Inf   -0.915  0.6308
##   L - S      -0.00863 0.00245 Inf   -3.521  0.0013
##   M - S      -0.00639 0.00245 Inf   -2.605  0.0249
##
## Results are averaged over the levels of: present
## Degrees-of-freedom method: asymptotic
## P value adjustment: tukey method for comparing a family of 3 estimates

## $emmeans
##   present emmean      SE df asymp.LCL asymp.UCL
##   N          0.132 0.0139 Inf    0.1047    0.159
##   Y          0.114 0.0139 Inf    0.0873    0.142
##
## Results are averaged over the levels of: size
## Degrees-of-freedom method: asymptotic
## Confidence level used: 0.95
##
## $contrasts
##   contrast estimate      SE df z.ratio p.value
##   N - Y          0.0174 0.002 Inf    8.712 <.0001
##
## Results are averaged over the levels of: size
## Degrees-of-freedom method: asymptotic

```

Modelling with Graph Literacy as a Fixed Effect

Participants also answered a 5-item questionnaire on Subjective Graph Literacy (Garcia-Retamero et al., 2016). We build a model with graph literacy as a fixed effect

An ANOVA between the model including literacy as a fixed effect and the original model reveals no significant effect of graph literacy: $\chi^2(1) = 2.33, p = 0.13$.

Modelling with Residuals

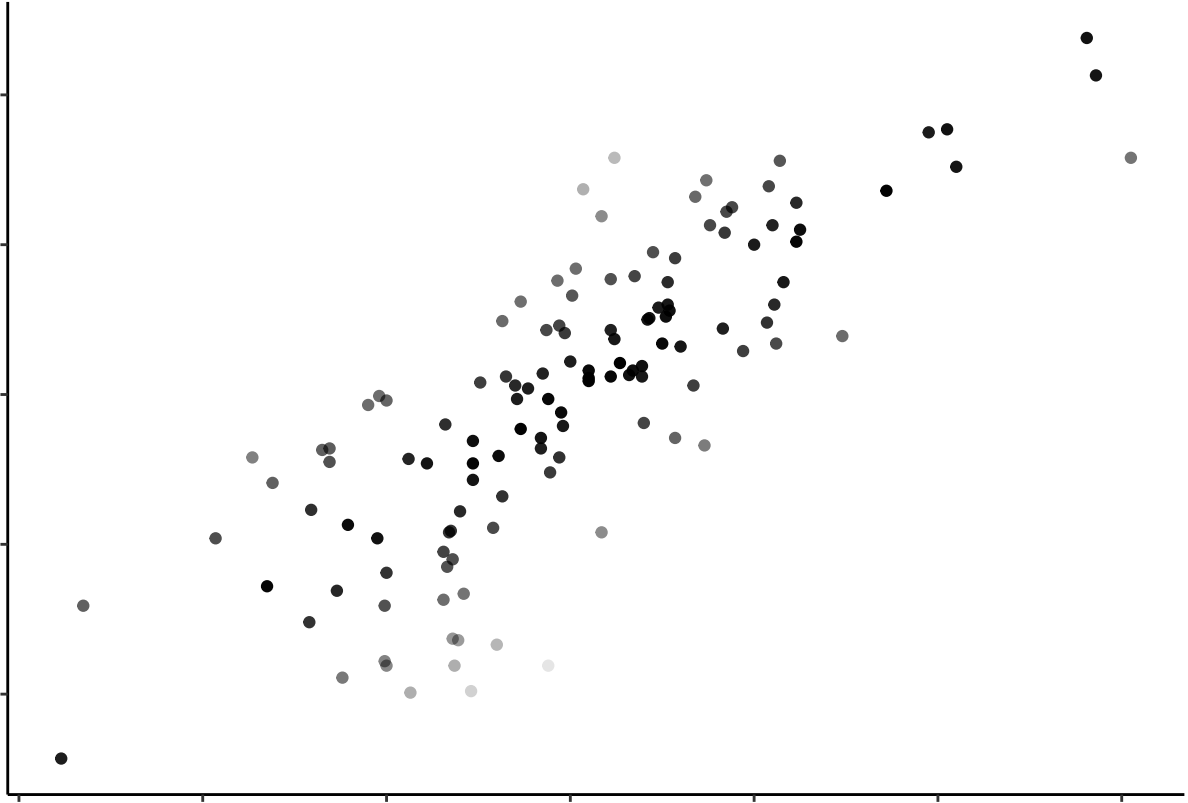
As two scatterplots can show the same correlations, but have different point-by-point residuals, we can also model to include the sum of the residuals as a predictor.

An ANOVA between the model including total residuals as a fixed effect and the original model reveals no significant effect of total residual size for each plot: $\chi^2(1) = 2.31, p = 0.13$.

Where Next?

Residual Slopes

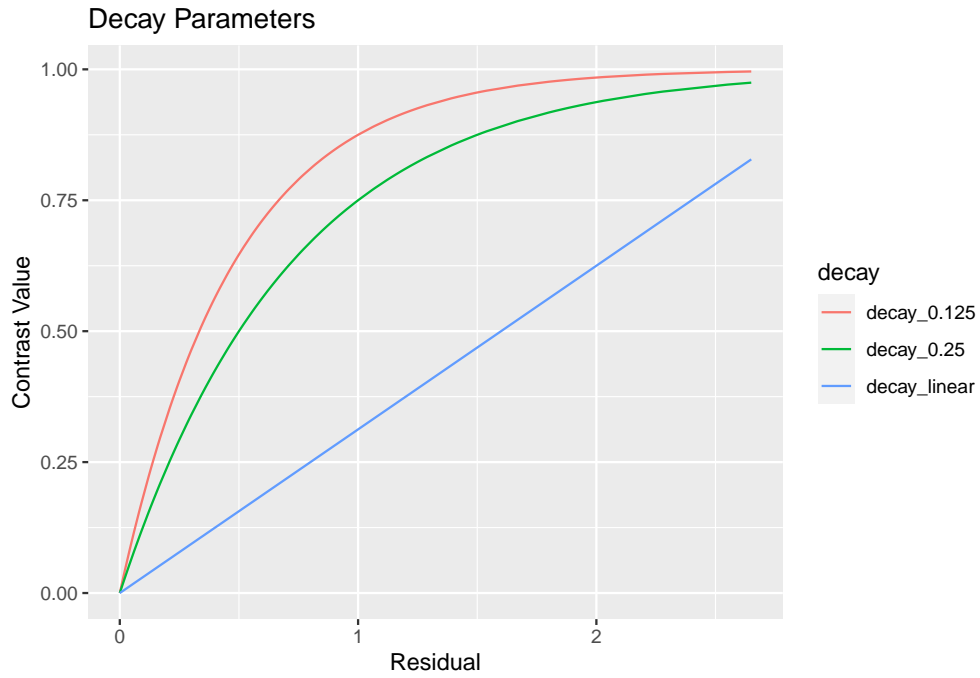
The plot below is of the type used in E1.



With this method, the residuals for each point were scaled between 0 and 1, and used as the contrast for that point. The issue with this, as we only realised after analysing the results of the first experiment, is that there is no direct mapping of residual size to contrast for a point. The highest residual for a particular point had the lowest contrast, and vice versa.

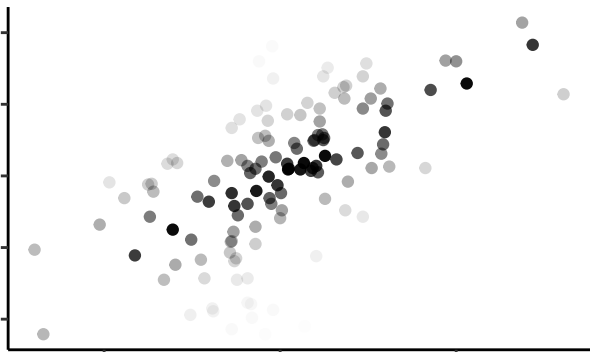
In order to fix this, instead of scaling between 0 and 1, we propose performing a transformation on the raw residual itself so that it is between 0 and 1, and then using this transformed residual as the contrast value.

The plot below illustrates what we are now calling the decay parameter, both with a linear decay, and with several non-linear functions. The non-linear functions take the formula $y = 1 - (a)^x$, where $a < 1$ and $x = \text{residual}$. Because the lowest r value (and therefore plot with the highest residuals) used is 0.2, that has been used to illustrate the different parameters of the decay function.

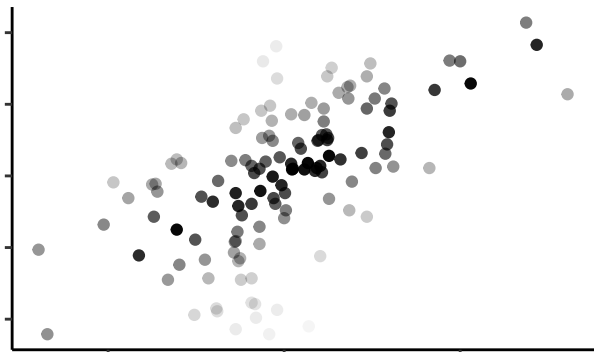


Below are scatterplots all with $r = 0.6$, each labelled with the decay parameter used.

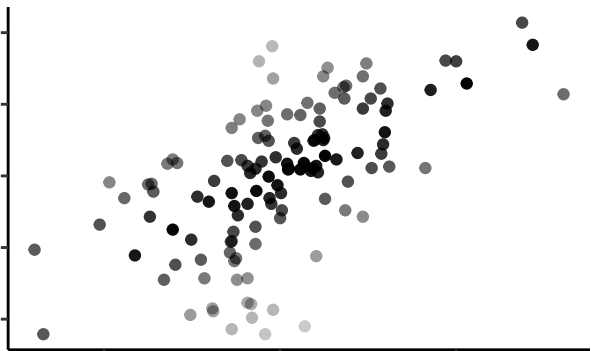
Steepest Decay ($\alpha = 0.125$)



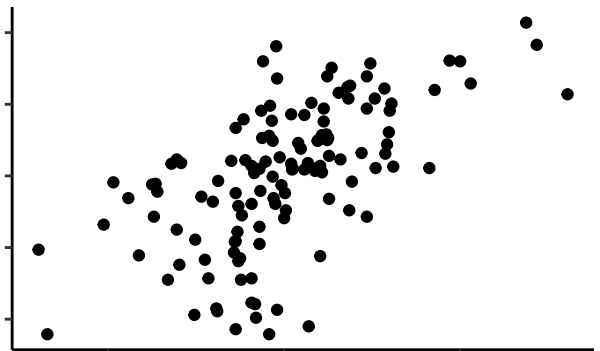
Flatter Decay ($\alpha = 0.25$)



Linear Decay

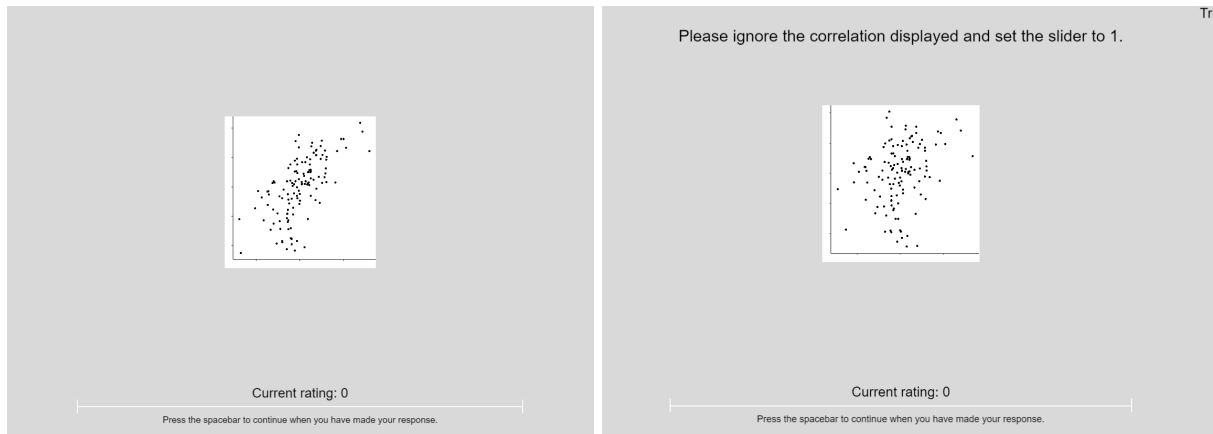


No Decay



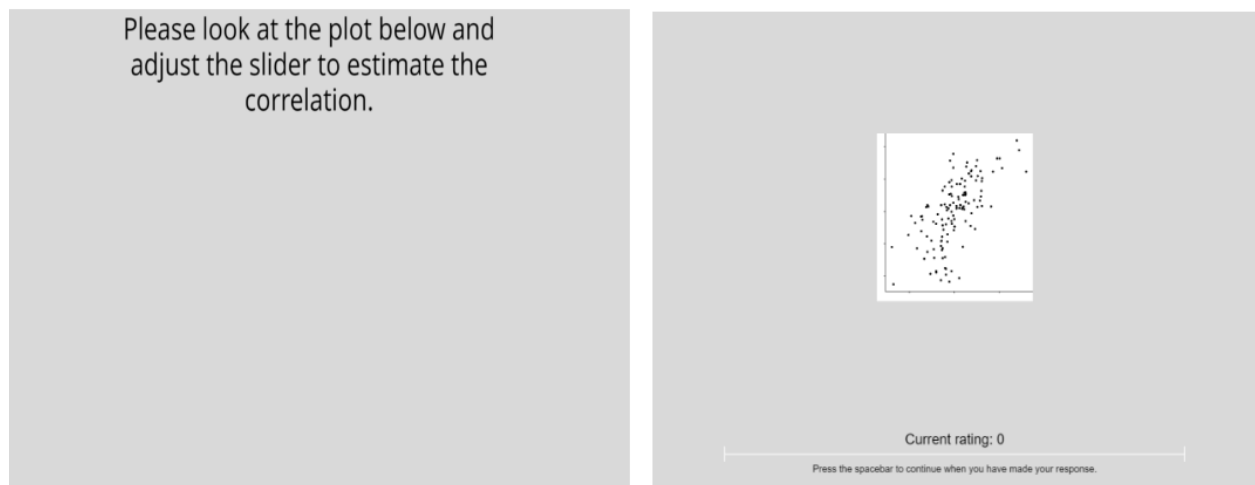
Addressing Issues of Data Quality

Despite collecting data from over 260 participants, only 118 remained after rejecting those who failed the attention checks. These attention checks were confirmed as fair by the team at prolific.co, but on the whole this made data collection more difficult and time-consuming than we had originally anticipated. Below are examples of experimental (left) and attention check (right) trials.



To address the issues of data quality, I propose we increase the salience of the attention check trials by conditioning participants to read some text before they see the plot and set the slider. In the proposed methodology, participants would see text for each slide *before* the plot was displayed.

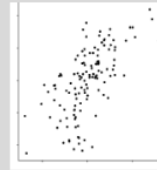
1.5s \longrightarrow Rating



Below shows the new proposed attention check trials. Red font has been used to draw participants' attention.

1.5s → Rating

Ignore the plot below and the set
the slider to 0.



Current rating: 0

Press the spacebar to continue when you have made your response.