Mixed Model Entropy

Data

```
Datos %>%
  ungroup() %>%
  dplyr::select(Subject,Level,Entropy) %>%
  group_by(Subject,Level) %>%
  mutate(mid = 1:n()) %>%
  pivot_wider(names_from=mid,values_from=Entropy) %>%
  arrange(Subject,Level) %>%
  kable("latex", booktabs = T) %>%
  kable_styling(latex_options = c("striped", "scale_down"))
```

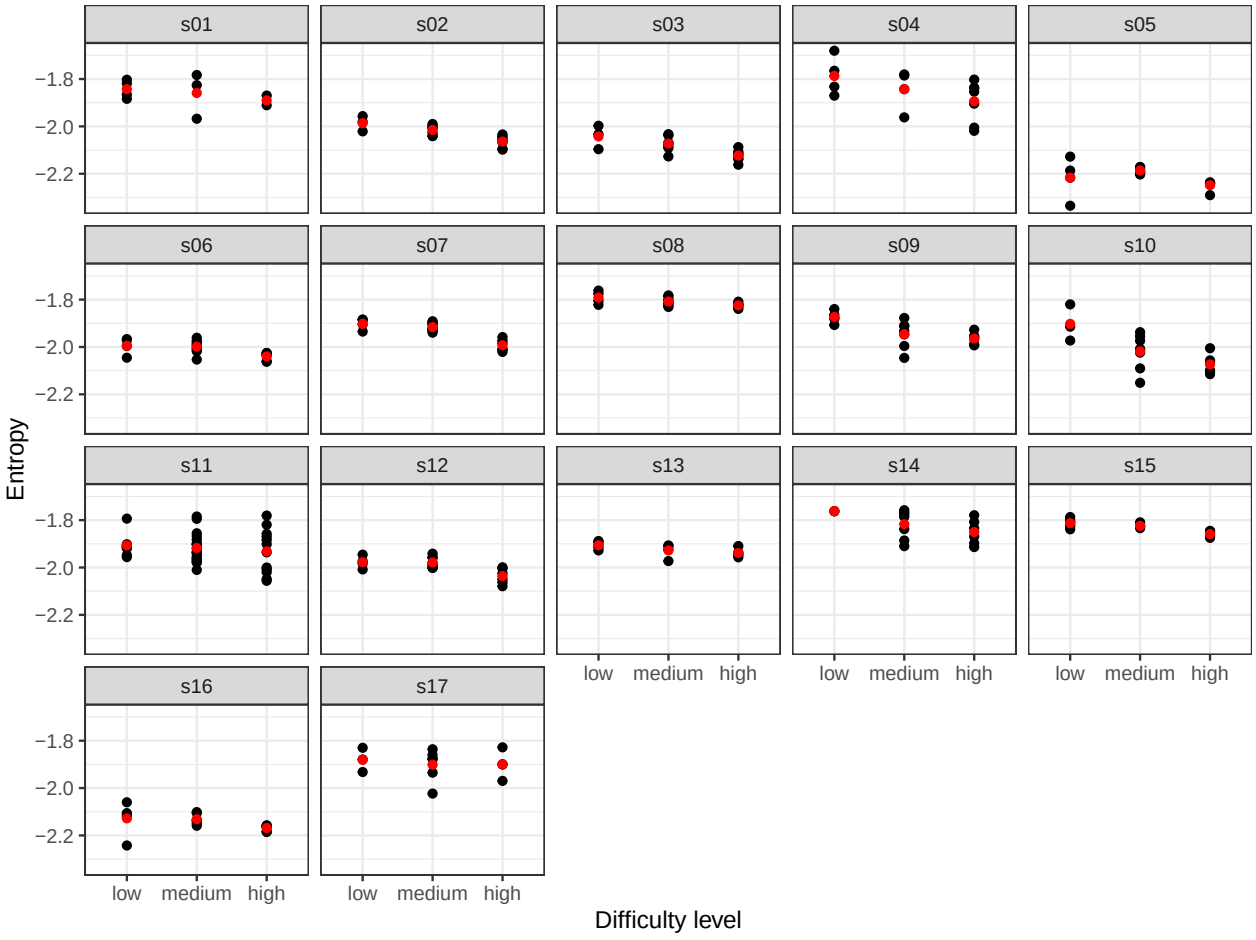
Subject	Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
s01	low	-1.865	-1.884	-1.819	-1.803															
s01	medium	-1.826	-1.968	-1.783																
s01	high	-1.870	-1.911																	
s02	low	-1.981	-1.956	-2.021																
s02	medium	-2.041	-2.027	-2.005	-1.998	-2.007	-1.990	-2.040												
s02	high	-2.095	-2.098	-2.062	-2.062	-2.051	-2.042	-2.034	-2.067											
s03	low	-2.033	-1.997	-2.096																
s03	medium	-2.038	-2.033	-2.127	-2.069	-2.071	-2.080	-2.091												
s03	high	-2.162	-2.137	-2.136	-2.087	-2.119	-2.123	-2.112	-2.108											
s04	low	-1.870	-1.832	-1.681	-1.765															
s04	medium	-1.962	-1.843	-1.785	-1.781															
s04	high	-2.019	-2.005	-1.904	-1.837	-1.837	-1.853	-1.802												
s05	low	-2.335	-2.216	-2.127	-2.187															
s05	medium	-2.171	-2.203	-2.190	-2.180	-2.196	-2.172	-2.191												
s05	high	-2.245	-2.239	-2.236	-2.239	-2.290	-2.239													
s06	low	-2.045	-1.996	-1.971	-1.967															
s06	medium	-2.053	-2.018	-1.973	-2.012	-1.982	-1.989	-1.996	-1.961	-2.006										
s06	high	-2.060	-2.025	-2.034	-2.026	-2.032	-2.063													
s07	low	-1.885	-1.885	-1.905	-1.935															
s07	medium	-1.927	-1.927	-1.896	-1.932	-1.903	-1.905	-1.921	-1.940	-1.891	-1.923									
s07	high	-2.009	-2.021	-1.973	-2.011	-1.984	-1.958													
s08	low	-1.822	-1.805	-1.776	-1.762															
s08	medium	-1.831	-1.815	-1.808	-1.782	-1.816	-1.801	-1.787	-1.826											
s08	high	-1.822	-1.819	-1.839	-1.825	-1.809	-1.832													
s09	low	-1.907	-1.877	-1.864	-1.881	-1.840														
s09	medium	-2.046	-1.996	-1.947	-1.912	-1.933	-1.878	-1.910												
s09	high	-1.966	-1.988	-1.954	-1.927	-1.993	-1.957													
s10	low	-1.973	-1.820	-1.914																
s10	medium	-2.090	-1.938	-1.956	-2.151	-2.010	-2.024	-1.974	-2.009											
s10	high	-2.062	-2.057	-2.005	-2.115	-2.103	-2.099													
s11	low	-1.913	-1.948	-1.907	-1.902	-1.919	-1.794	-1.956	-1.914											
s11	medium	-1.970	-2.010	-1.933	-1.977	-1.938	-1.903	-1.953	-1.881	-1.904	-1.855	-1.963	-1.795	-1.785	-1.92	-1.94	-1.979	-1.867	-1.962	-1.897
s11	high	-2.008	-1.901	-2.000	-1.870	-1.936	-1.819	-2.049	-1.781	-2.056	-1.858	-2.020	-1.884							
s12	low	-1.946	-2.008	-1.984	-1.973															
s12	medium	-2.000	-2.000	-1.990	-2.002	-1.957	-1.942	-1.958	-1.981											
s12	high	-2.024	-2.049	-2.079	-2.062	-2.004	-1.999													
s13	low	-1.915	-1.906	-1.893	-1.888	-1.928														
s13	medium	-1.908	-1.923	-1.907	-1.973															
s13	high	-1.946	-1.936	-1.909	-1.944	-1.957														
s14	low	-1.762	-1.763																	
s14	medium	-1.910	-1.886	-1.786	-1.838	-1.758	-1.767	-1.775												
s14	high	-1.779	-1.897	-1.869	-1.808	-1.913	-1.835													
s15	low	-1.821	-1.839	-1.799	-1.826	-1.787	-1.808	-1.810												
s15	medium	-1.809	-1.834	-1.826	-1.826															
s15	high	-1.869	-1.845	-1.848	-1.875	-1.855	-1.861													
s16	low	-2.242	-2.106	-2.060	-2.120	-2.111														
s16	medium	-2.139	-2.159	-2.102	-2.149	-2.135	-2.105													
s16	high	-2.157	-2.185	-2.159	-2.162	-2.166	-2.184													
s17	low	-1.932	-1.830	-1.879																
s17	medium	-2.023	-1.935	-1.860	-1.836	-1.877	-1.877													
s17	high	-1.970	-1.900	-1.900	-1.828															

Level	n	MD	SD
low	72	-1.923	0.1242
medium	124	-1.955	0.1101
high	106	-1.994	0.1236

Summary by group

```
Datos %>%
  group_by(Level) %>%
  summarise(n=n(),MD=mean(Entropy),SD=sd(Entropy)) %>%
    kable() %>%   kable_styling(latex_options = c("striped"))
```

```
(q <-Datos %>% ggplot(aes(x=Level,y=Entropy)) +
    geom_point() + facet_wrap(~ Subject)+
    labs(x="Difficulty level")+theme_bw()+
    stat_summary(fun="mean", geom="point",color="red"))
```



Random Intercept and Slope Model

The following model is used to investigate whether there are significant differences between the study variables:

$$y_{ij} = \mu + l_k + s_j + (sl)_{jk} + \epsilon_{ij}, \tag{1}$$

where y_{ij} is the response variable (Entropy) for the i-th observation from the j-th subject, μ is the intercept, l_k is the k-th difficulty level, s_j is the jth subject effect, $(sl)_{jk}$ is the subject-level effect, i.e., the k-th level effect at the j-th subject, ϵ_{ij} is the error term (residual) for the ith observation from the jth subject.

We called *level l* a fixed effect, and ϵ is our *error term* that represent deviations from our predictions due to *random* factors that we cannot control experimentally. However, several measurements were taken for each subject at each difficulty level and that violates the assumption of independence of a linear model. On the other hand, each individual has a different cognitive load capacity, and this will be a characteristic factor that will affect all the responses of the same subject, which will make these responses interdependent instead of independent, see figure **??**. The way we approaches this situation is adding a random effect to the subject and to the subject-level interaction. This allows us to solve this lack of independence by assuming a different intercept and slope for each subject. And finally, we assume that the residual, subject and subject-level effects are all relations of separate distributions, all with zero means:

$$\begin{aligned} \epsilon_{ij} &\sim N(0, \sigma^2), \\ s_j &\sim N(0, \sigma_s^2), \\ (sl)_{jk} &\sim N(0, \sigma_{sl}^2). \end{aligned}$$

Hence, s_j and $(sl)_{jk}$ are now random effects, and μ and l_k are fixed effects.

Using the **R** notation the model is

$$Entropy = (b_0 + u_{Subject}) + b_{Level} Level + \epsilon$$

In order to evaluate if there is an effect due to the difficulty level we will use the likelihood ratio test of the model with the *Level* effect against the model without the *Level* effect.

```
Entropy_mixed_reducido <- lme4::lmer(Entropy ~ 1 + (1+Level|Subject),data=Datos,REML=F)

Entropy_mixed_lme4 <- lme4::lmer(Entropy ~ Level + (1+Level|Subject),data=Datos,REML=F)

anova(Entropy_mixed_reducido,Entropy_mixed_lme4)
```

```
## Data: Datos
## Models:
## Entropy_mixed_reducido: Entropy ~ 1 + (1 + Level | Subject)
## Entropy_mixed_lme4: Entropy ~ Level + (1 + Level | Subject)
##                         npar  AIC  BIC logLik deviance Chisq Df Pr(>Chisq)
## Entropy_mixed_reducido    8 -872 -842    444     -888
## Entropy_mixed_lme4       10 -894 -857    457     -914  25.9  2  0.0000024 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value of the ratio test is significant at a level of 0.001.

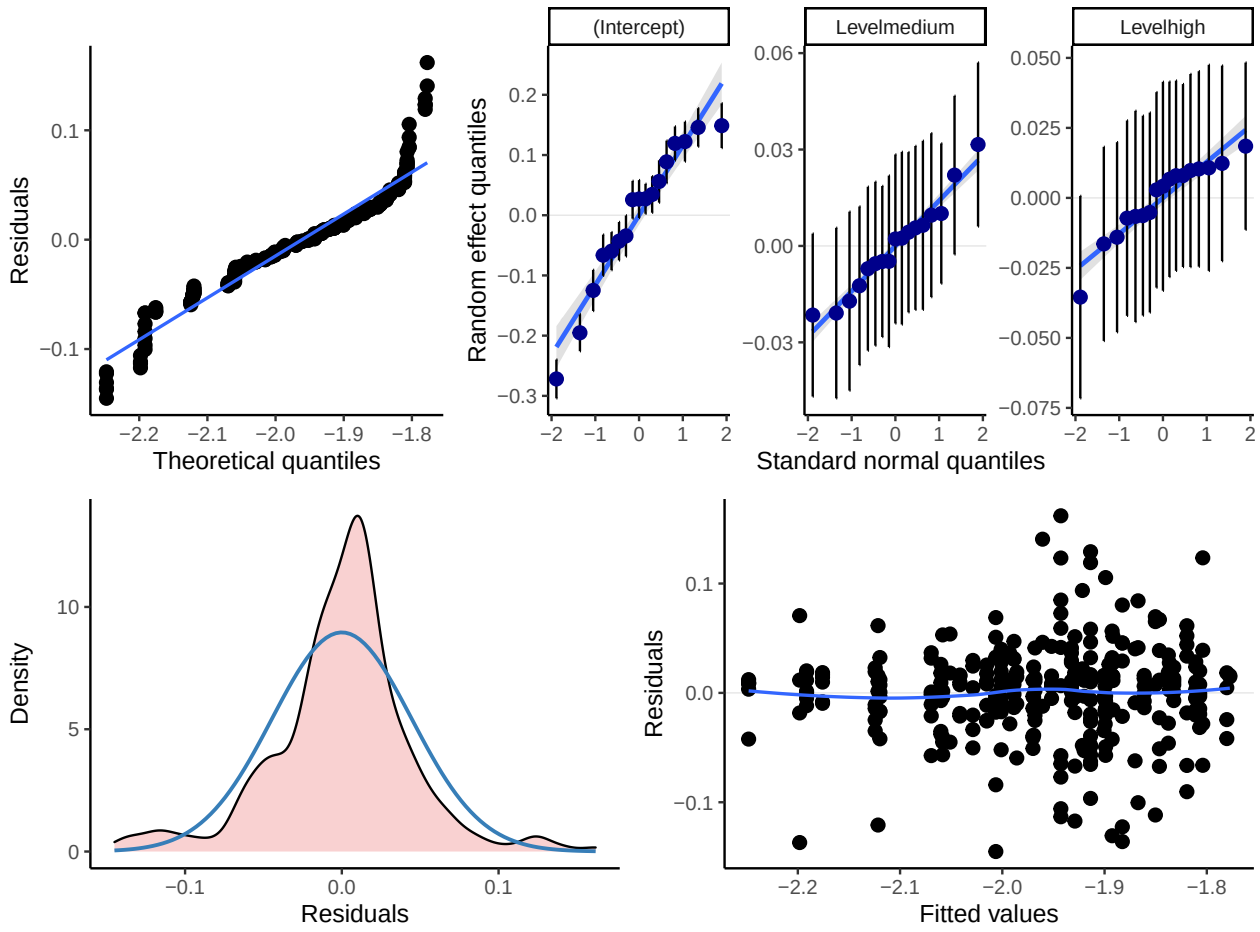
```
Entropy_mixed_lme4 <- lme4::lmer(Entropy ~ Level + (1+Level|Subject),data=Datos)
summary(Entropy_mixed_lme4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Entropy ~ Level + (1 + Level | Subject)
##    Data: Datos
##
## REML criterion at convergence: -892.3
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
```

```
## -3.127 -0.420  0.102  0.419  3.491
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Subject  (Intercept) 0.014371 0.1199
##           Levelmedium 0.000373 0.0193   -0.64
##           Levelhigh   0.000499 0.0223   -0.14  0.85
##  Residual             0.002152 0.0464
## Number of obs: 302, groups:  Subject, 17
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) -1.92627    0.02961  -65.05
## Levelmedium -0.02482    0.00847   -2.93
## Levelhigh   -0.06225    0.00907   -6.87
##
## Correlation of Fixed Effects:
##             (Intr) Lvlmdm
## Levelmedium -0.471
## Levelhigh   -0.197  0.699
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
p<-plot_model(Entropy_mixed_lme4, type = "diag")

({p[[1]]+theme(plot.title=element_blank(),plot.subtitle=element_blank())+scale_x_continuous(name="Theor
```




```
contr <- glht(Entropy_mixed_lme4,linfct=mcp(Level="Tukey"))
summary(contr, test = adjusted("holm"))
```

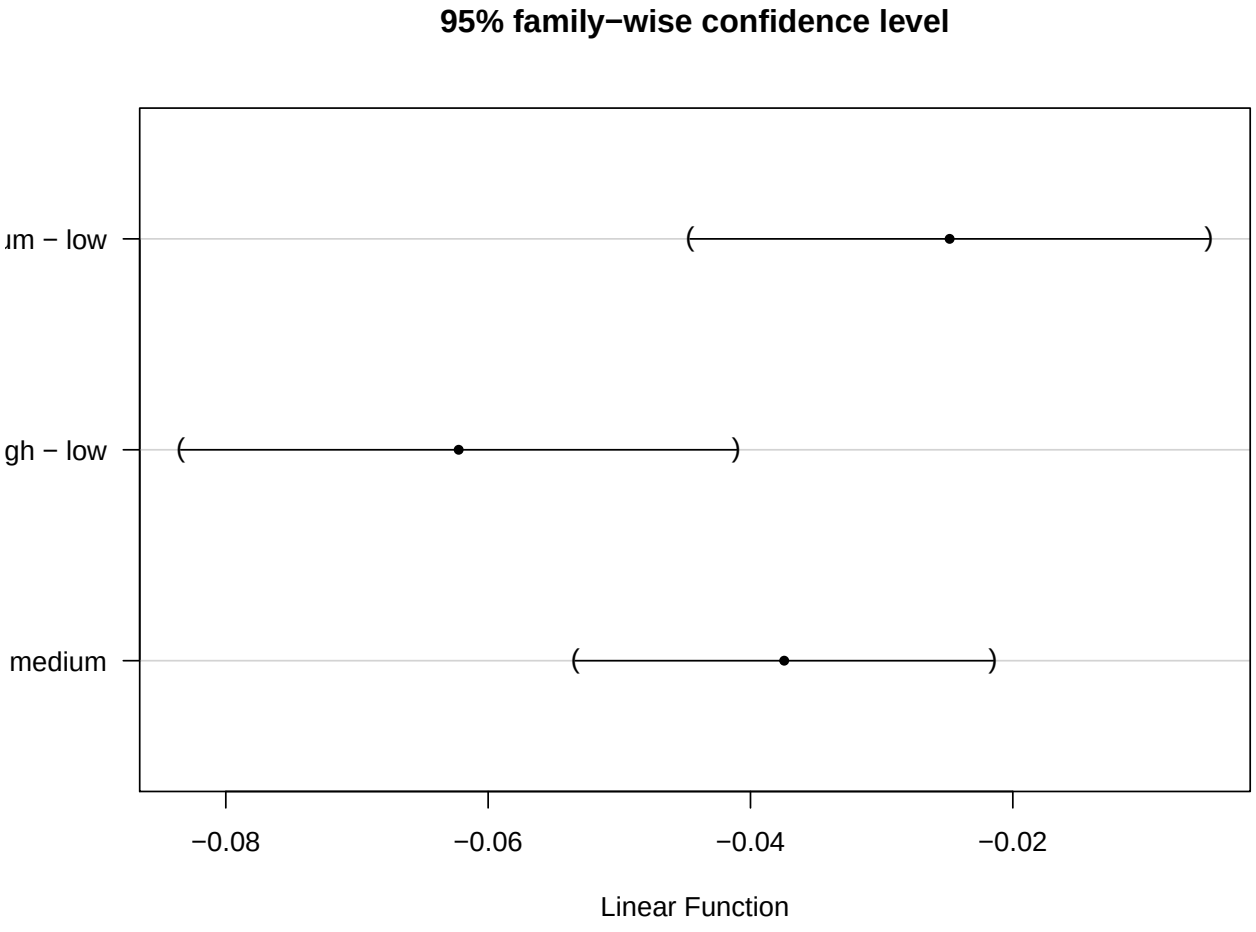
```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = Entropy ~ Level + (1 + Level | Subject),
##     data = Datos)
##
## Linear Hypotheses:
##                     Estimate Std. Error z value      Pr(>|z|)
## medium - low == 0  -0.02482    0.00847   -2.93        0.0034 **
## high - low == 0    -0.06225    0.00907   -6.87 0.00000000002 ***
## high - medium == 0 -0.03743    0.00682   -5.49 0.00000008060 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- holm method)
```

```
confint(contr)
```

```
##
##   Simultaneous Confidence Intervals
```

```
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = Entropy ~ Level + (1 + Level | Subject),
##     data = Datos)
##
## Quantile = 2.336
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##                     Estimate lwr      upr
## medium - low == 0  -0.02482 -0.04460 -0.00504
## high - low == 0    -0.06225 -0.08343 -0.04107
## high - medium == 0 -0.03743 -0.05336 -0.02150
```

```
plot(confint(contr))
```




```
contr.cld <- cld(contr)
old.par <- par(mai=c(1,1,1.25,1), no.readonly = TRUE)
plot(contr.cld)
```

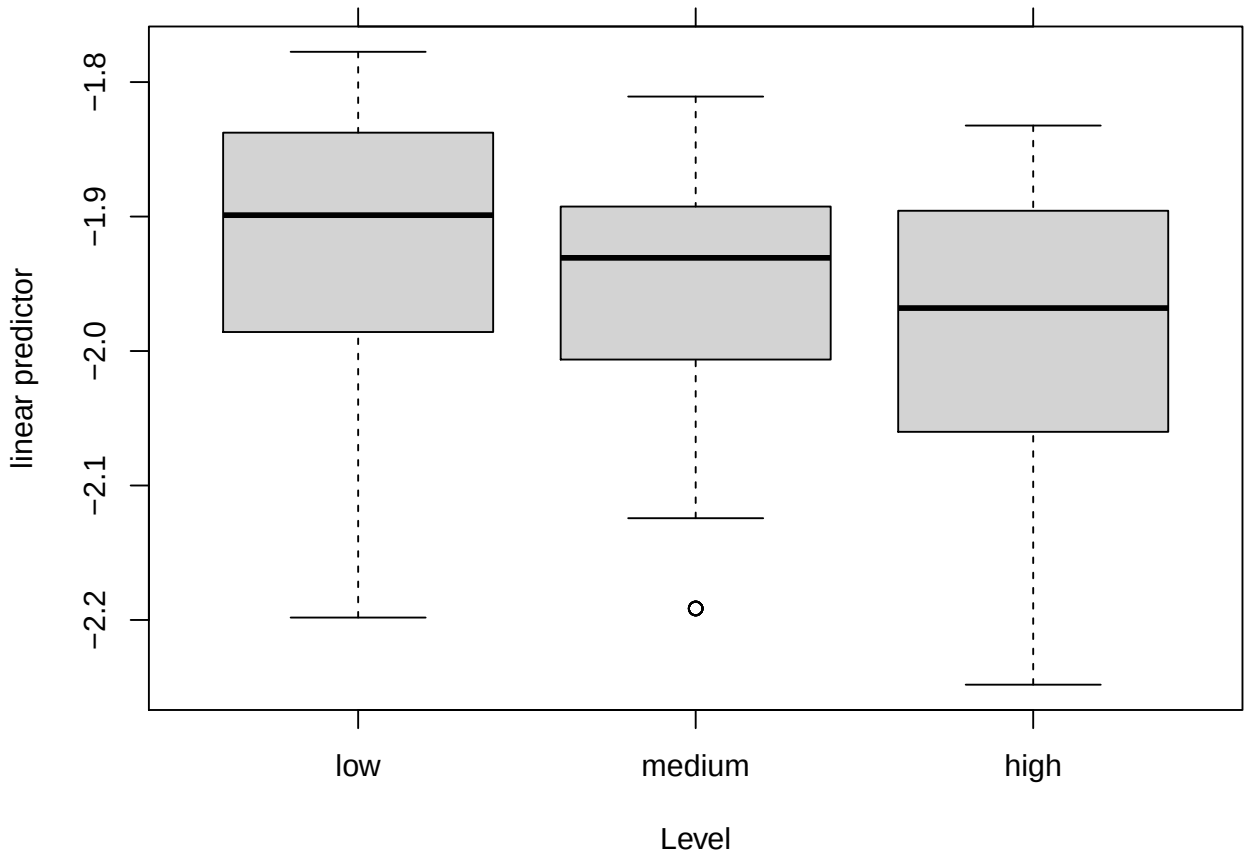
Subject	Name	Training	Nivel	BLPS	MPDC	APCPS	PD	Entropy	TTP	PDS	SequenceMemory	SMN	id	Level	res	fit
s11	SequenceMemory_r18	FALSE	6	3.521	-0.0856	-0.0243	0.0779	-1.781	7136712	0	r18	18		high	0.1619	-1.943
s10	SequenceMemory_r05	FALSE	1	3.083	0.4308	0.1397	0.8144	-1.820	8153300	0	r05	5	5	low	0.1406	-1.961
s11	SequenceMemory_r27	FALSE	3	3.558	-0.1134	-0.0319	0.0068	-1.785	4086712	0	r27	27		medium	0.1291	-1.914
s04	SequenceMemory_r27	FALSE	1	3.449	-0.2455	-0.0712	-0.0211	-1.681	3807563	0	r27	27	13	low	0.1235	-1.804
s11	SequenceMemory_r15	FALSE	6	3.522	0.0065	0.0018	0.1540	-1.819	6498755	0	r15	15	20	high	0.1233	-1.943
s11	SequenceMemory_r20	FALSE	3	3.507	-0.0394	-0.0112	0.1822	-1.795	6120082	0	r20	20		medium	0.1191	-1.914

Subject	Name	Training	Nivel	BLPS	MPDC	APCPS	PD	Entropy	TTP	PDS	SequenceMemory	SMN	id	Level	res	fit
s16	SequenceMemory_r01	FALSE	1	4.759	-0.0301	-0.0063	0.2297	-2.242	7535552	0	r01	1	1	low	-0.1208	-2.122
s04	SequenceMemory_r05	FALSE	6	3.622	0.3864	0.1067	0.8141	-2.005	8352831	0	r05	5	4	high	-0.1225	-1.883
s17	SequenceMemory_r01	FALSE	3	3.647	0.4141	0.1135	0.8695	-2.023	6279569	0	r01	1	1	medium	-0.1306	-1.893
s04	SequenceMemory_r01	FALSE	6	3.992	0.0507	0.0127	0.6136	-2.019	8253083	0	r01	1	1	high	-0.1360	-1.883
s05	SequenceMemory_r01	FALSE	1	5.046	-0.0044	-0.0009	0.4557	-2.335	4385784	0	r01	1	1	low	-0.1368	-2.198
s10	SequenceMemory_r17	FALSE	3	4.511	-0.0772	-0.0171	0.2468	-2.151	6219667	0	r17	17	10	medium	-0.1450	-2.006

a

b

c




```
par(old.par)
```

```
Datos2 = Datos
Datos2$res = residuals(Entropy_mixed_lme4,type="pearson")
Datos2$fit = fitted(Entropy_mixed_lme4,type="pearson")

Datos2 %>% arrange(desc(res)) %>% head() %>%   kable() %>%   kable_styling(latex_options = c("striped",
```

```
Datos2 %>% arrange(desc(res)) %>% tail() %>%   kable() %>%   kable_styling(latex_options = c("striped",
```

```
shapiro.test(Datos2$res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Datos2$res
## W = 0.96, p-value = 0.00000007
```

```
goftest::ad.test(Datos2$res,null="pnorm",mean=mean(Datos2$res), sd=sd(Datos2$res), estimated=TRUE)
```

```
##
##  Anderson-Darling test of goodness-of-fit
##  Braun's adjustment using 17 groups
##  Null hypothesis: Normal distribution
##  with parameters mean = 0.000000000000000307333440625072, sd =
##  0.0445282614355351
##  Parameters assumed to have been estimated from data
##
## data:  Datos2$res
## Anmax = 2.9, p-value = 0.4
```

```
rstatix::levene_test(data=ungroup(Datos2),res~Level)
```

```
## # A tibble: 1 x 4
##     df1   df2 statistic     p
##   <int> <int>     <dbl> <dbl>
## 1     2   299    0.0508 0.950
```

Non parametric tests

```
kruskal.test(Entropy ~ Level, data=Datos)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  Entropy by Level
## Kruskal-Wallis chi-squared = 16, df = 2, p-value = 0.0003
```

```
PMCMR::posthoc.kruskal.nemenyi.test(data=Datos,Entropy~Level, dist="Tukey")
```

```
##
##  Pairwise comparisons using Tukey and Kramer (Nemenyi) test
##                     with Tukey-Dist approximation for independent samples
##
## data:  Entropy by Level
##
##        low     medium
## medium 0.08026 -
## high   0.00016 0.06226
##
## P value adjustment method: none
```

```
PMCMRplus::tukeyTest(data=Datos,Entropy~Level)
```

```
##        low     medium
## medium 0.16760 -
## high   0.00035 0.03730
```