Mixed Model BLPS

Data

```
Datos %>%
  ungroup() %>%
  dplyr::select(Subject,Level,BLPS) %>%
  group_by(Subject,Level) %>%
  mutate(mid = 1:n()) %>%
  pivot_wider(names_from=mid,values_from=BLPS) %>%
  arrange(Subject,Level) %>%
  kable()
```

Subject	Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
s01	low	3.653	3.950	3.768	3.283															
s01	medium	3.584	4.069	3.404																
s01	high	3.489	4.023																	
s02	low	4.113	3.889	4.149																
s02	medium	3.980	4.076	4.256	3.963	3.919	3.818	4.322												
s02	high	3.849	3.908	4.104	4.104	3.740	3.860	4.083	4.359											
s03	low	4.040	3.984	4.359																
s03	medium	4.171	4.024	4.275	4.130	3.775	4.059	4.087												
s03	high	4.688	4.056	3.998	4.034	4.068	4.266	4.276	3.954											
s04	low	3.588	3.621	3.449	3.380															
s04	medium	3.531	3.332	3.392	3.309															
s04	high	3.992	3.622	3.418	3.387	3.478	3.532	3.217												
s05	low	5.046	4.450	4.460	4.532															
s05	medium	4.478	4.598	4.383	4.516	4.455	4.152	4.220												
s05	high	4.630	4.484	4.430	4.472	4.560	4.645													
s06	low	4.053	3.967	4.010	3.907															
s06	medium	4.088	4.035	3.897	3.924	3.961	3.918	3.865	3.916	4.046										
s06	high	3.999	3.830	4.092	3.978	3.990	3.978													
s07	low	3.862	3.726	3.649	3.748															
s07	medium	3.755	3.811	3.665	3.635	3.685	3.671	3.822	3.685	3.680	3.700									
s07	high	3.654	3.839	3.688	3.829	3.692	3.752													
s08	low	3.738	3.413	3.416	3.424															
s08	medium	3.438	3.648	3.414	3.513	3.596	3.464	3.522	3.593											
s08	high	3.358	3.458	3.474	3.474	3.518	3.451													
s09	low	3.819	3.670	3.559	3.553	3.545														
s09	medium	3.637	3.704	3.668	3.632	3.644	3.468	3.574												
s09	high	3.549	3.751	3.712	3.509	3.680	3.512													
s10	low	3.828	3.083	3.612																
s10	medium	4.201	3.740	3.586	4.511	3.829	4.066	2.744	3.782											
s10	high	3.857	3.644	3.773	4.260	3.702	4.077													
s11	low	3.722	3.574	3.610	3.806	3.826	3.555	3.800	3.512											
s11	medium	3.594	3.903	3.603	3.942	3.570	3.664	3.701	3.495	3.578	3.489	3.956	3.507	3.558	3.514	3.628	3.899	3.688	3.7	3.781
s11	high	3.505	3.617	3.635	3.535	3.744	3.522	3.986	3.521	3.611	3.505	3.618	3.665							
s12	low	3.754	4.087	4.075	3.932															

Subject	Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
s12	medium	3.779	3.927	3.876	3.694	3.900	3.844	4.056	3.994											
s12	high	3.733	3.922	3.987	3.989	3.848	3.913													
s13	low	4.008	3.798	3.706	3.553	3.771														
s13	medium	3.815	4.030	3.776	4.255															
s13	high	3.846	3.567	3.564	3.784	3.995														
s14	low	3.561	3.363																	
s14	medium	3.794	3.712	3.183	3.571	3.427	3.697	3.256												
s14	high	3.278	3.627	3.341	3.237	3.715	3.457													
s15	low	3.582	3.492	3.627	3.516	3.583	3.437	3.501												
s15	medium	3.393	3.467	3.474	3.463															
s15	high	3.417	3.464	3.506	3.498	3.426	3.620													
s16	low	4.759	4.227	4.375	4.302	4.300														
s16	medium	4.205	4.496	4.136	4.399	4.250	4.192													
s16	high	4.223	4.511	4.291	4.138	4.272	4.849													
s17	low	3.414	3.322	3.340																
s17	medium	3.647	3.875	3.470	3.439	3.561	3.706													
s17	high	3.597	3.191	3.450	3.329															

Summary by group

```
Datos %>%
  group_by(Level) %>%
  summarise(n=n(),MD=mean(BLPS),SD=sd(BLPS)) %>%
  kable()
```

Level	n	MD	SD
low	72	3.793	0.3651
medium	124	3.803	0.3191
high	106	3.810	0.3589

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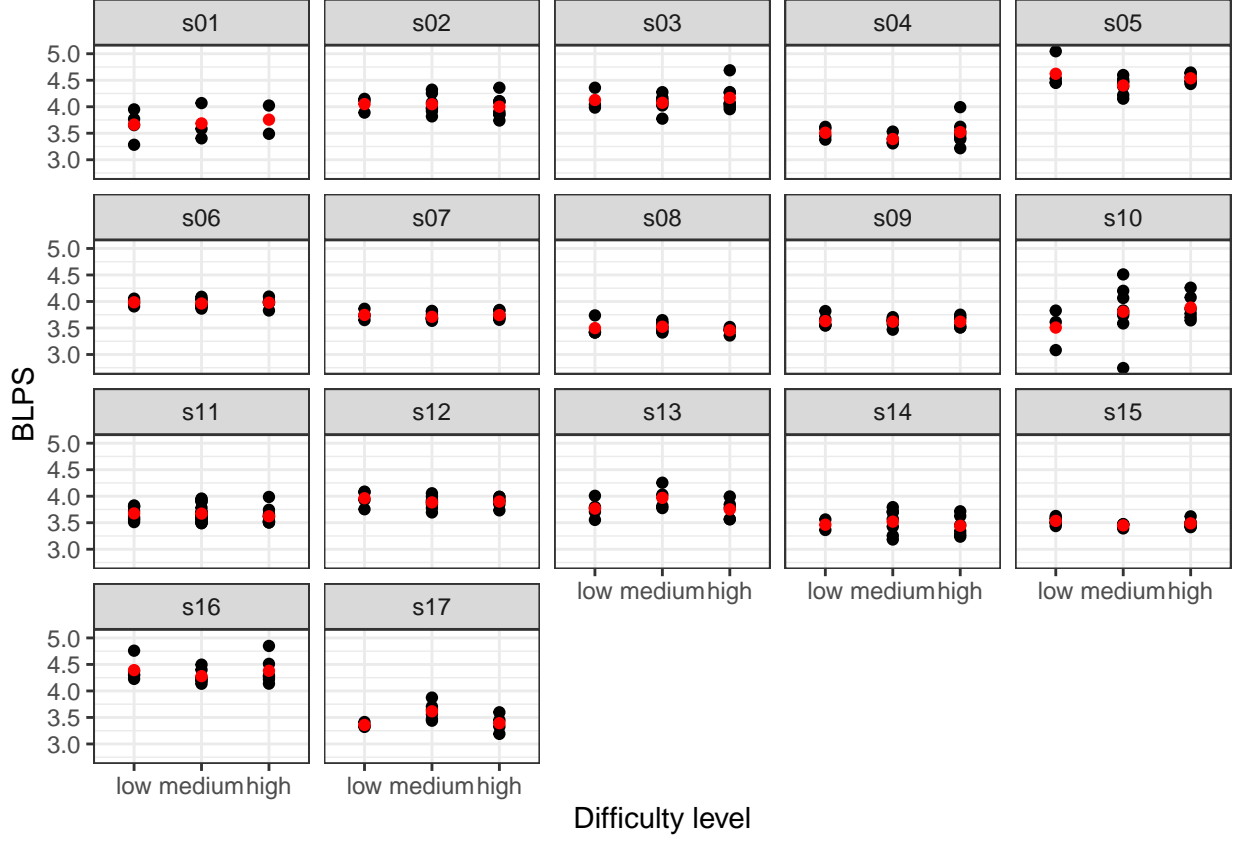
Random Intercept and Slope Model

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 (BLPS) 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:

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

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

3

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

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.

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

BLPS_mixed_lme4 <- lme4::lmer(BLPS ~ Level + (1+Level|Subject),data=Datos,REML=F)

anova(BLPS_mixed_reducido,BLPS_mixed_lme4)
```

```
## Data: Datos
## Models:
## BLPS_mixed_reducido: BLPS ~ 1 + (1 + Level | Subject)
## BLPS_mixed_lme4: BLPS ~ Level + (1 + Level | Subject)
##                     npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## BLPS_mixed_reducido    8 -98.7 -69.0   57.3     -115
## BLPS_mixed_lme4       10 -94.7 -57.6   57.3     -115  0.01  2          1
```

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

```
BLPS_mixed_lme4 <- lme4::lmer(BLPS ~ Level + (1+Level|Subject),data=Datos)
summary(BLPS_mixed_lme4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: BLPS ~ Level + (1 + Level | Subject)
##    Data: Datos
##
## REML criterion at convergence: -100.1
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -5.762 -0.516 -0.107  0.467  4.046
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  Subject  (Intercept) 0.1116959 0.3342
##           Levelmedium 0.0036876 0.0607   -1.00
##           Levelhigh   0.0000672 0.0082   -1.00  1.00
##  Residual             0.0324263 0.1801
## Number of obs: 302, groups:  Subject, 17
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  3.802370   0.083862   45.34
## Levelmedium -0.000169   0.030833   -0.01
## Levelhigh   -0.001887   0.027905   -0.07
##
## Correlation of Fixed Effects:
##             (Intr) Lvlmdm
## Levelmedium -0.640
## Levelhigh   -0.265  0.577
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
p<-plot_model(BLPS_mixed_lme4, type = "diag")
```

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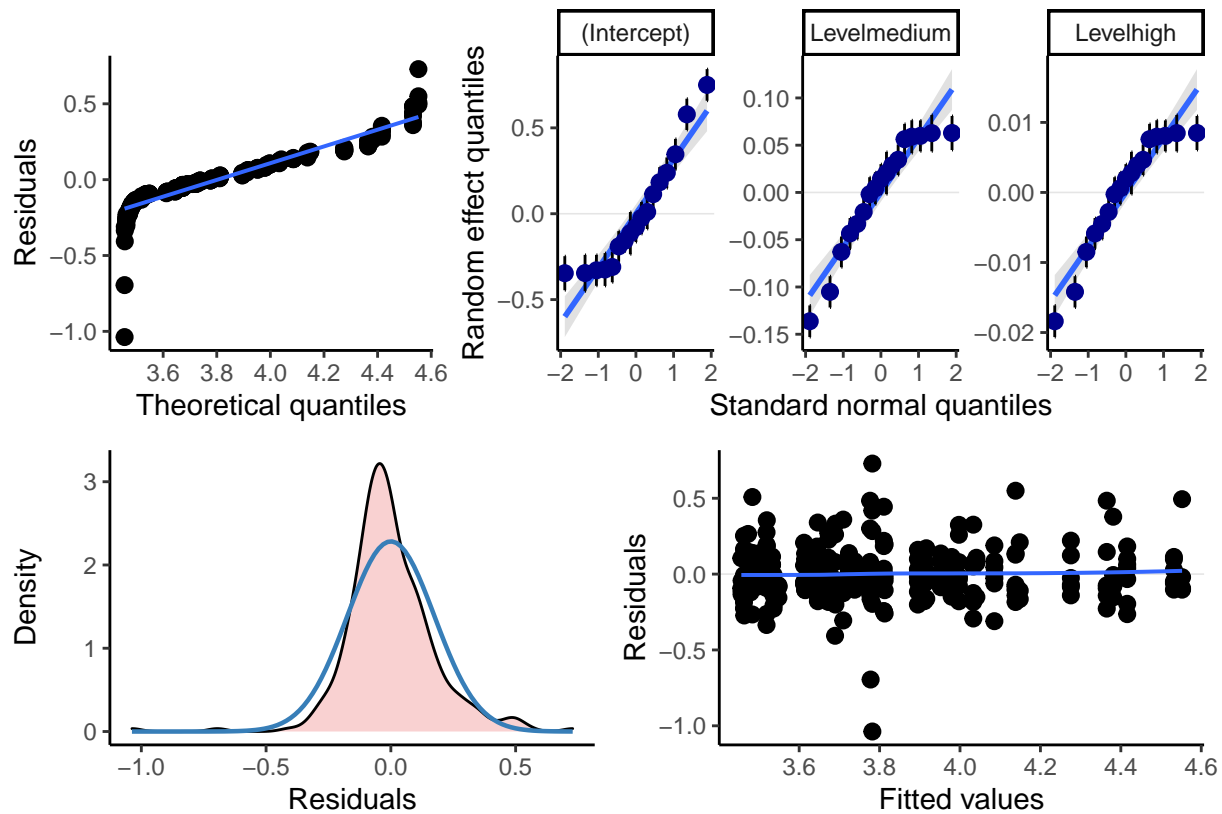


```
contr <- glht(BLPS_mixed_lme4,linfct=mcp(Level="Tukey"))
summary(contr, test = adjusted("holm"))
```

```
##
##    Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = BLPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Linear Hypotheses:
##                    Estimate Std. Error z value Pr(>|z|)
## medium - low == 0  -0.000169   0.030833   -0.01        1
## high - low == 0    -0.001887   0.027905   -0.07        1
## high - medium == 0 -0.001718   0.027152   -0.06        1
## (Adjusted p values reported -- holm method)
```

```
confint(contr)
```

```
##
##    Simultaneous Confidence Intervals
##
## Multiple Comparisons of Means: Tukey Contrasts
##
```

5

```
## 
## Fit: lme4::lmer(formula = BLPS ~ Level + (1 + Level | Subject), data = Datos)
## 
## Quantile = 2.342
## 95% family-wise confidence level
## 
## 
## Linear Hypotheses:
##                    Estimate  lwr       upr
## medium - low == 0  -0.000169 -0.072379  0.072040
## high - low == 0    -0.001887 -0.067240  0.063466
## high - medium == 0 -0.001718 -0.065307  0.061872
```

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


95% family−wise confidence level
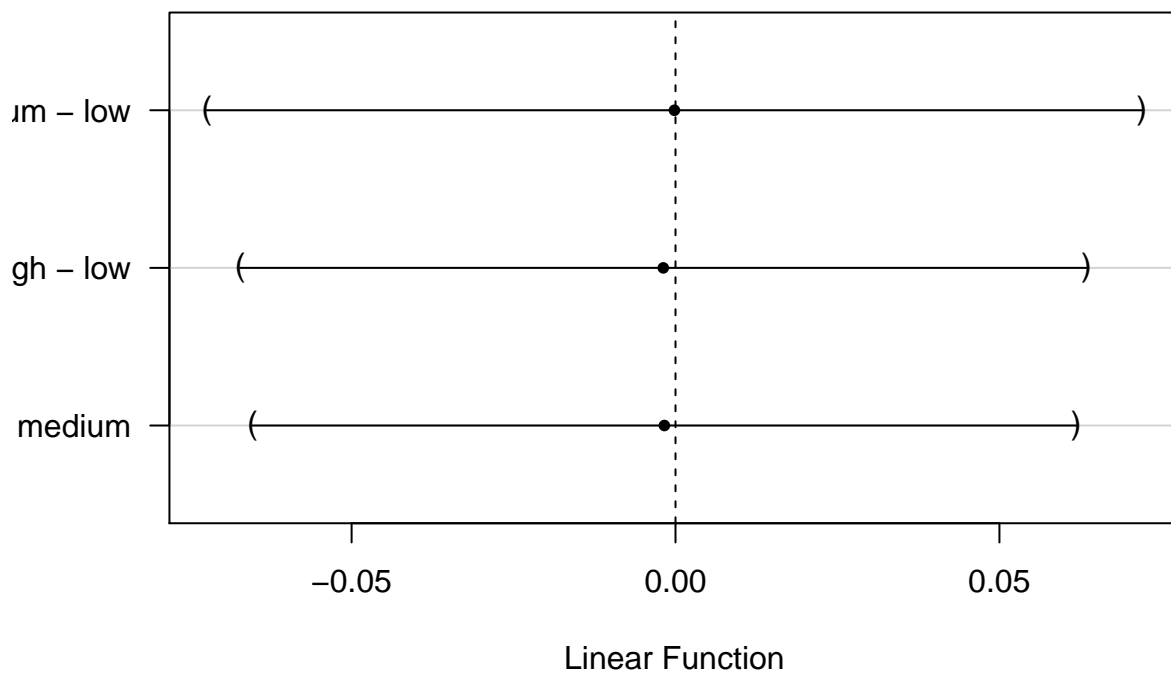
```
contr.cld <- cld(contr)
old.par <- par(mai=c(1,1,1.25,1), no.readonly = TRUE)
plot(contr.cld)
```
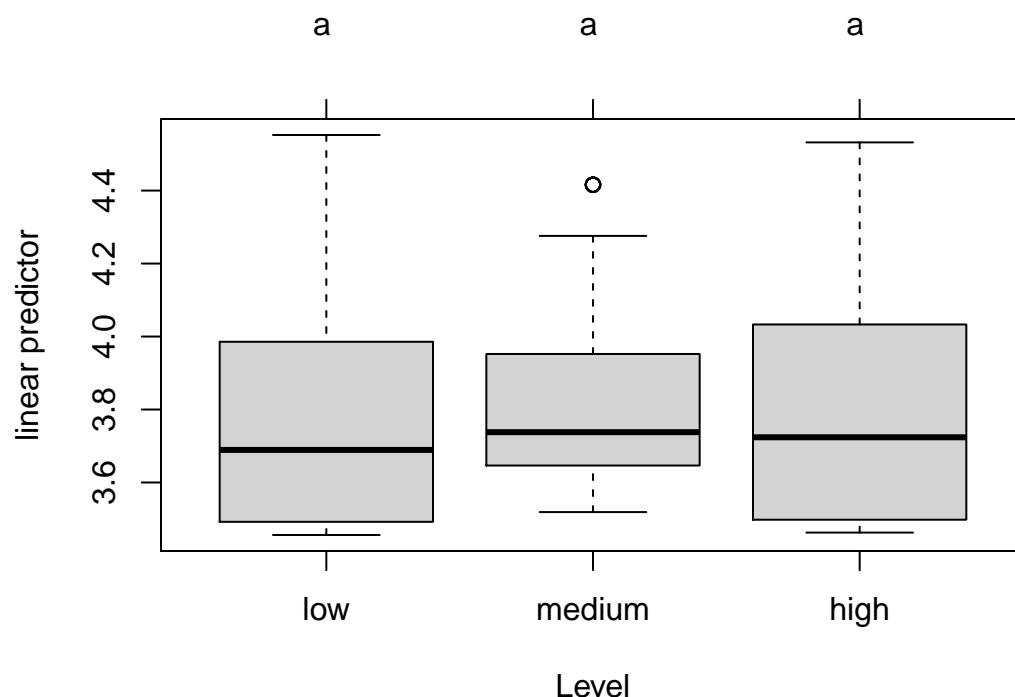
```
par(old.par)
```

```
Datos2 = Datos
Datos2$res = residuals(BLPS_mixed_lme4,type="pearson")
Datos2$fit = fitted(BLPS_mixed_lme4,type="pearson")

Datos2 %>% arrange(desc(res)) %>% head() %>% kable()
```

Subject	Name	Training	Nivel	BLPS	MPD	CAPC	PSPD	Entropy	TTP	PDS	Sequence	Memory	SdNn	Level	res	fit
s10	SequenceMemory	FALSE	Er17	4.511	- 0.0772	- 0.0171	0.2468	- 2.151	6219	6670	r17	17	10	medium	0.7287	3.782
s03	SequenceMemory	FALSE	Er03	4.688	- 0.2167	- 0.0462	0.2240	- 2.162	8432	5430	r03	3	1	high	0.5496	4.138
s04	SequenceMemory	FALSE	Er01	3.992	0.0507	0.0127	0.6136	- 2.019	8253	0830	r01	1	1	high	0.5081	3.484
s05	SequenceMemory	FALSE	Er01	5.046	- 0.0044	- 0.0009	0.4557	- 2.335	4385	7840	r01	1	1	low	0.4940	4.552
s16	SequenceMemory	FALSE	Er28	4.849	- 0.3060	- 0.0631	- 0.0414	- 2.184	1443	2978	r28	28	15	high	0.4838	4.365
s10	SequenceMemory	FALSE	Er18	4.260	0.0646	0.0152	0.3601	- 2.115	2091	1750	r18	18	11	high	0.4837	3.776

```
Datos2 %>% arrange(desc(res)) %>% tail() %>% kable()
```

Subject	Name	Training	Nivel	BLPS	MPD	CAPC	PSPD	Entropy	TTP	PDS	Sequence	Memory	SdNn	Level	res	fit
s01	SequenceMemory	FALSE	Er12	3.404	0.0336	0.0099	0.3407	- 1.783	11024	0480	r12	12	6	medium	- 0.3056	3.710
s03	SequenceMemory	FALSE	Er21	3.775	0.4244	0.1124	0.6202	- 2.071	3588	2230	r21	21	12	medium	- 0.3104	4.085
s14	SequenceMemory	FALSE	Er15	3.183	0.2633	0.0827	0.4829	- 1.786	5920	7900	r15	15	7	medium	- 0.3356	3.519

Subject	Name	Training	Nivel	BLPS	MPD	CAPCP	SD	Entropy	FTP	PDS	Sequence	SMN	MNr	Level	res	fit
s01	SequenceMemory	FALSE	r23	3.283	0.1995	0.0608	0.5783	-1.803	7136	7480	r23	23	9	low	-0.4066	3.689
s10	SequenceMemory	FALSE	r05	3.083	0.4308	0.1397	0.8144	-1.820	8153	3000	r05	5	5	low	-0.6951	3.778
s10	SequenceMemory	FALSE	r24	2.744	1.1568	0.4215	1.6368	-1.974	7176	8060	r24	24	14	medium	-1.0376	3.782

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

```
## 
##  Shapiro-Wilk normality test
## 
## data:  Datos2$res
## W = 0.92, p-value = 0.00000000002
```

```
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 = -0.0000000000000000183810143510862, sd =
##  0.174706133134124
##  Parameters assumed to have been estimated from data
## 
## data:  Datos2$res
## Anmax = 3, 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     0.157 0.855
```

The same model without the outlier

We repeat the analysis without the outlier

```
# we exclude the outlier
Datos <- Datos %>% filter(!(Subject=="s10"&SMN==24),
                          !(Subject=="s10" & SMN==5))
```

```
BLPS_mixed_lme4 <- lme4::lmer(BLPS ~ Level + (1+Level|Subject),data=Datos)
summary(BLPS_mixed_lme4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: BLPS ~ Level + (1 + Level | Subject)
##    Data: Datos
## 
## REML criterion at convergence: -156.8
## 
## Scaled residuals:
##    Min     1Q Median     3Q    Max
```

```
## -2.530 -0.591 -0.137  0.509  3.844
##
## Random effects:
##  Groups    Name        Variance  Std.Dev. Corr
##  Subject  (Intercept)  0.1110231 0.33320
##           Levelmedium  0.0032405 0.05692  -1.00
##           Levelhigh    0.0000274 0.00524  -1.00  1.00
##  Residual              0.0263929 0.16246
## Number of obs: 300, groups:  Subject, 17
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   3.81602    0.08315   45.89
## Levelmedium  -0.00528    0.02822   -0.19
## Levelhigh    -0.01547    0.02528   -0.61
##
## Correlation of Fixed Effects:
##             (Intr) Lvlmdm
## Levelmedium -0.638
## Levelhigh   -0.230  0.567
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
anova(BLPS_mixed_lme4)
```

```
## Analysis of Variance Table
##       npar Sum Sq Mean Sq F value
## Level    2 0.0109 0.00544    0.21
```

```
coef(BLPS_mixed_lme4)
```

```
## $Subject
##     (Intercept) Levelmedium Levelhigh
## s01       3.694    0.015628 -0.013544
## s02       4.048   -0.044840 -0.019109
## s03       4.155   -0.063163 -0.020795
## s04       3.486    0.051111 -0.010279
## s05       4.556   -0.131649 -0.027097
## s06       3.991   -0.035201 -0.018222
## s07       3.730    0.009352 -0.014122
## s08       3.481    0.052004 -0.010197
## s09       3.619    0.028384 -0.012370
## s10       3.907   -0.020828 -0.016899
## s11       3.652    0.022822 -0.012882
## s12       3.922   -0.023386 -0.017134
## s13       3.819   -0.005801 -0.015516
## s14       3.465    0.054620 -0.009956
## s15       3.498    0.048972 -0.010476
## s16       4.385   -0.102559 -0.024420
## s17       3.465    0.054767 -0.009942
##
## attr(,"class")
## [1] "coef.mer"
```

```
p<-plot_model(BLPS_mixed_lme4, type = "diag")
```

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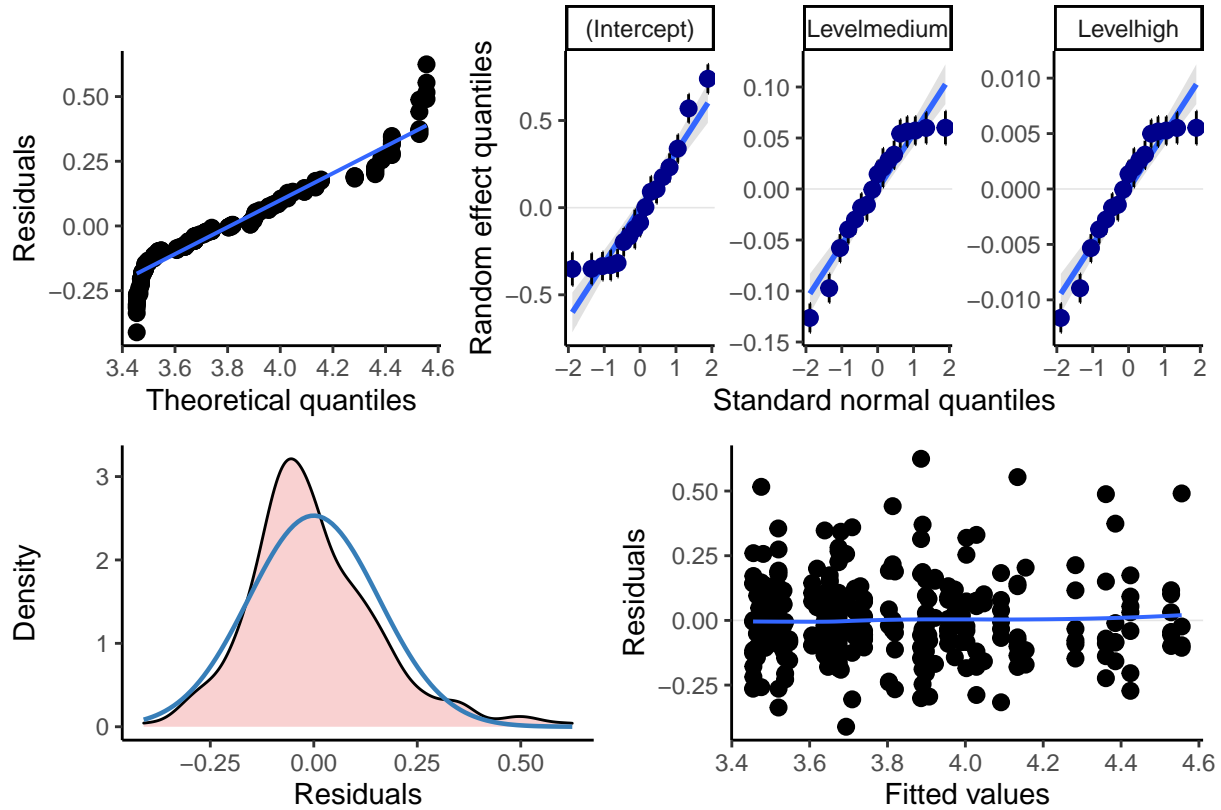


```
# muy importante Tukey para lme4.
contr <- glht(BLPS_mixed_lme4,linfct=mcp(Level="Tukey"))
summary(contr, test = adjusted("holm"))
```

```
##
##      Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = BLPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Linear Hypotheses:
##                    Estimate Std. Error z value Pr(>|z|)
## medium - low == 0  -0.00528    0.02822   -0.19        1
## high - low == 0     -0.01547    0.02528   -0.61        1
## high - medium == 0 -0.01019    0.02504   -0.41        1
## (Adjusted p values reported -- holm method)
```

```
confint(contr)
```

```
##
##      Simultaneous Confidence Intervals
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
```

```
## Fit: lme4::lmer(formula = BLPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Quantile = 2.341
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##                   Estimate lwr      upr
## medium - low == 0  -0.00528 -0.07134  0.06078
## high - low == 0    -0.01547 -0.07465  0.04372
## high - medium == 0 -0.01019 -0.06881  0.04843
```

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

95% family–wise confidence level
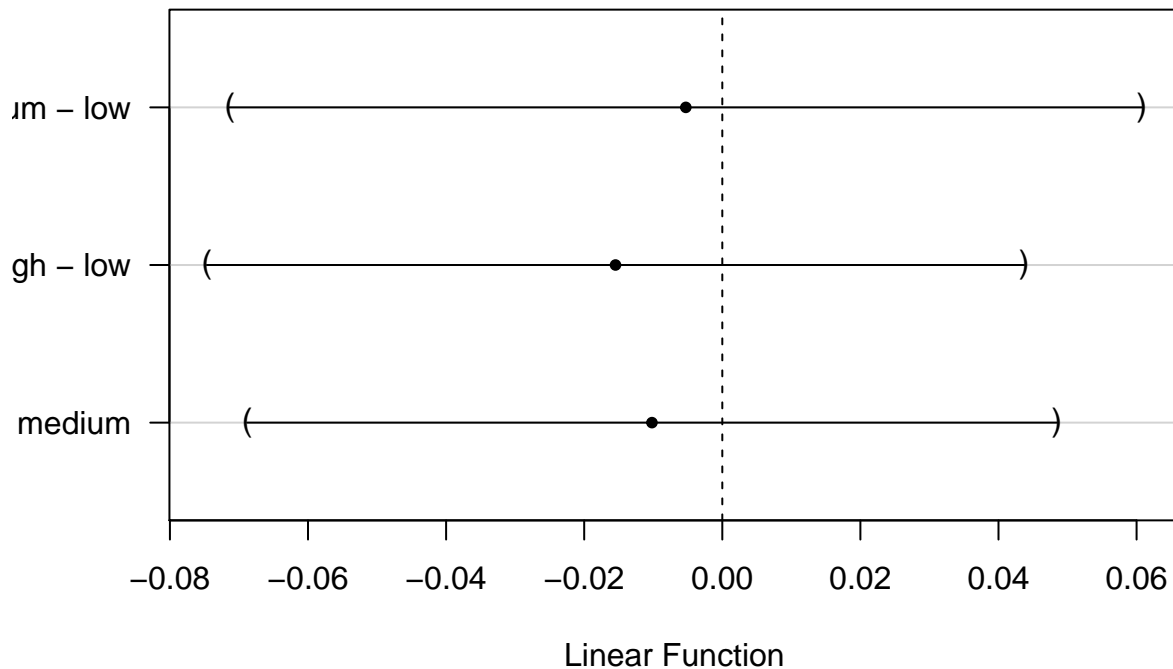

```
contr.cld <- cld(contr)
### use sufficiently large upper margin
old.par <- par(mai=c(1,1,1.25,1), no.readonly = TRUE)
### plot
plot(contr.cld)
```
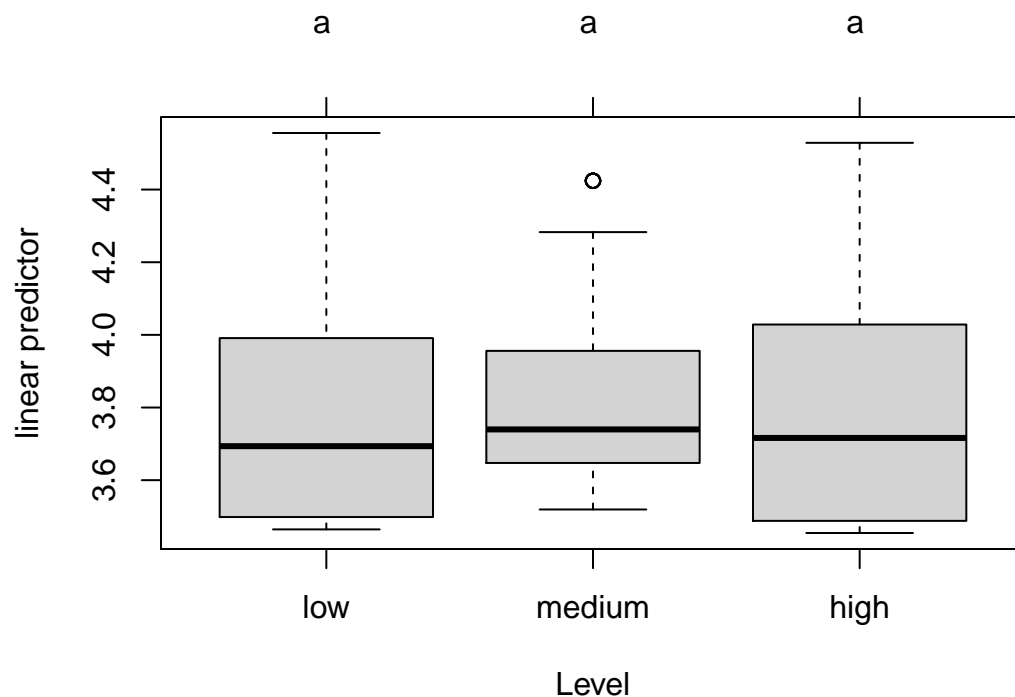
```r
par(old.par)
```

```r
Datos2=Datos
Datos2$res = residuals(BLPS_mixed_lme4,type="pearson")
Datos2$fit = fitted(BLPS_mixed_lme4,type="pearson")
shapiro.test(Datos2$res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Datos2$res
## W = 0.96, p-value = 0.0000003
```

```r
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 = 0.00000000000000000592471312172808, sd =
##  0.157567596860321
##  Parameters assumed to have been estimated from data
##
## data:  Datos2$res
## Anmax = 2.5, p-value = 0.6
```

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

```
## # A tibble: 1 x 4
##     df1   df2 statistic     p
##   <int> <int>     <dbl> <dbl>
## 1     2   297     0.132 0.876
```

Non parametric tests

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

```
##
##  Kruskal-Wallis rank sum test
##
## data:  BLPS by Level
## Kruskal-Wallis chi-squared = 0.46, df = 2, p-value = 0.8
```

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

```
##
##  Pairwise comparisons using Tukey and Kramer (Nemenyi) test
##                    with Tukey-Dist approximation for independent samples
##
## data:  BLPS by Level
##
##        low  medium
## medium 0.79 -
## high   0.96 0.90
##
## P value adjustment method: none
```

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

```
##        low  medium
## medium 0.98 -
## high   0.99 1.00
```