Mixed Model APCPS

Data

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

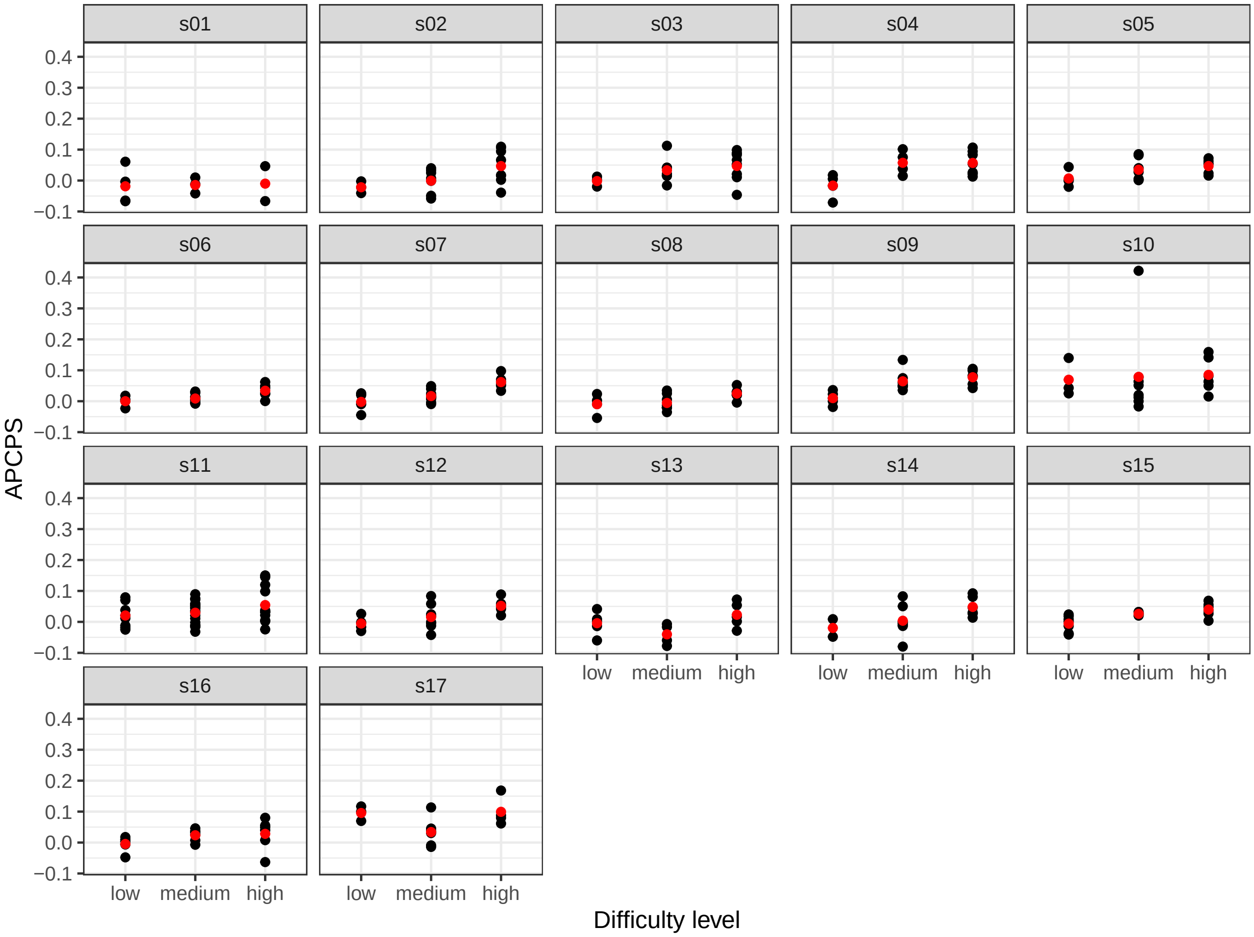
Subject	Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
s01	low	-0.0036	-0.0668	-0.0648	0.0608															
s01	medium	-0.0115	-0.0416	0.0099																
s01	high	0.0464	-0.0665																	
s02	low	-0.0406	-0.0026	-0.0220																
s02	medium	0.0330	-0.0010	-0.0582	0.0071	0.0250	0.0398	-0.0492												
s02	high	0.1093	0.0944	0.0171	0.0166	0.1070	0.0662	0.0024	-0.0391											
s03	low	0.0127	0.0019	-0.0196																
s03	medium	-0.0159	0.0158	0.0213	0.0152	0.1124	0.0414	0.0419												
s03	high	-0.0462	0.0838	0.0986	0.0530	0.0664	0.0208	0.0107	0.0890											
s04	low	0.0175	-0.0169	-0.0712	0.0053															
s04	medium	0.1013	0.0753	0.0150	0.0379															
s04	high	0.0127	0.1067	0.0941	0.0539	0.0263	0.0216	0.0832												
s05	low	-0.0009	0.0439	-0.0203	0.0041															
s05	medium	0.0053	0.0011	0.0401	0.0030	0.0276	0.0852	0.0817												
s05	high	0.0233	0.0526	0.0626	0.0551	0.0721	0.0159													
s06	low	0.0177	0.0051	-0.0231	0.0003															
s06	medium	0.0146	0.0034	0.0074	0.0270	-0.0035	0.0126	0.0313	-0.0060	-0.0079										
s06	high	0.0425	0.0620	0.0004	0.0236	0.0247	0.0494													
s07	low	-0.0447	-0.0096	0.0256	0.0197															
s07	medium	0.0113	-0.0044	0.0144	0.0489	0.0136	0.0190	-0.0093	0.0400	0.0078	0.0242									
s07	high	0.0975	0.0554	0.0629	0.0510	0.0699	0.0336													
s08	low	-0.0542	0.0234	0.0023	-0.0095															
s08	medium	0.0346	-0.0354	0.0252	-0.0208	-0.0211	0.0055	-0.0204	-0.0133											
s08	high	0.0524	0.0202	0.0298	0.0196	-0.0044	0.0311													
s09	low	-0.0187	-0.0001	0.0227	0.0360	0.0097														
s09	medium	0.1335	0.0747	0.0501	0.0356	0.0466	0.0593	0.0510												
s09	high	0.0988	0.0555	0.0425	0.0826	0.0808	0.1048													
s10	low	0.0252	0.1397	0.0430																
s10	medium	0.0124	0.0200	0.0778	-0.0171	0.0514	0.0000	0.4215	0.0636											
s10	high	0.0821	0.1413	0.0635	0.0152	0.1591	0.0501													
s11	low	0.0117	0.0794	0.0382	-0.0186	-0.0117	-0.0251	0.0208	0.0704											
s11	medium	0.0895	0.0318	0.0595	-0.0018	0.0721	0.0205	0.0452	0.0539	0.0452	0.0361	-0.0151	-0.0112	-0.0319	0.0759	0.0571	0.0109	-0.0112	0.0519	-0.0158
s11	high	0.1446	0.0321	0.0985	0.0324	0.0215	0.0018	0.0374	-0.0243	0.1502	0.0327	0.1198	0.0051							
s12	low	0.0260	-0.0165	-0.0296	-0.0014															
s12	medium	0.0584	0.0185	0.0244	0.0836	-0.0043	-0.0007	-0.0423	-0.0123											
s12	high	0.0887	0.0541	0.0587	0.0459	0.0419	0.0208													
s13	low	-0.0600	-0.0137	0.0022	0.0415	0.0088														
s13	medium	-0.0165	-0.0597	-0.0070	-0.0779															
s13	high	0.0016	0.0724	0.0537	0.0160	-0.0287														
s14	low	-0.0481	0.0089																	
s14	medium	-0.0100	-0.0048	0.0827	0.0005	-0.0133	-0.0799	0.0505												
s14	high	0.0464	0.0268	0.0923	0.0810	0.0136	0.0303													
s15	low	-0.0134	0.0241	-0.0410	0.0084	-0.0374	0.0180	0.0015												
s15	medium	0.0322	0.0279	0.0201	0.0239															
s15	high	0.0683	0.0369	0.0268	0.0480	0.0551	0.0036													
s16	low	-0.0063	0.0178	-0.0478	0.0100	0.0034														
s16	medium	0.0460	-0.0071	0.0371	0.0073	0.0327	0.0253													
s16	high	0.0553	0.0077	0.0399	0.0801	0.0489	-0.0631													
s17	low	0.1168	0.0698	0.1002																
s17	medium	0.1135	-0.0141	0.0452	0.0371	0.0307	-0.0092													
s17	high	0.0876	0.1682	0.0800	0.0615															

Level	n	MD	SD
low	72	0.0046	0.0399
medium	124	0.0237	0.0523
high	106	0.0498	0.0433

Summary by group

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

```
(q <-Datos %>% ggplot(aes(x=Level,y=APCPS)) +
    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 (APCPS) 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

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

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.

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

APCPS_mixed_lme4 <- lme4::lmer(APCPS ~ Level + (1+Level|Subject),data=Datos,REML=F)

anova(APCPS_mixed_reducido,APCPS_mixed_lme4)
```

```
## Data: Datos
## Models:
## APCPS_mixed_reducido: APCPS ~ 1 + (1 + Level | Subject)
## APCPS_mixed_lme4: APCPS ~ Level + (1 + Level | Subject)
##                      npar   AIC  BIC logLik deviance Chisq Df Pr(>Chisq)
## APCPS_mixed_reducido    8  -978 -949    497     -994
## APCPS_mixed_lme4       10 -1003 -966    512    -1023    29  2 0.00000051 ***
## ---
## 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.

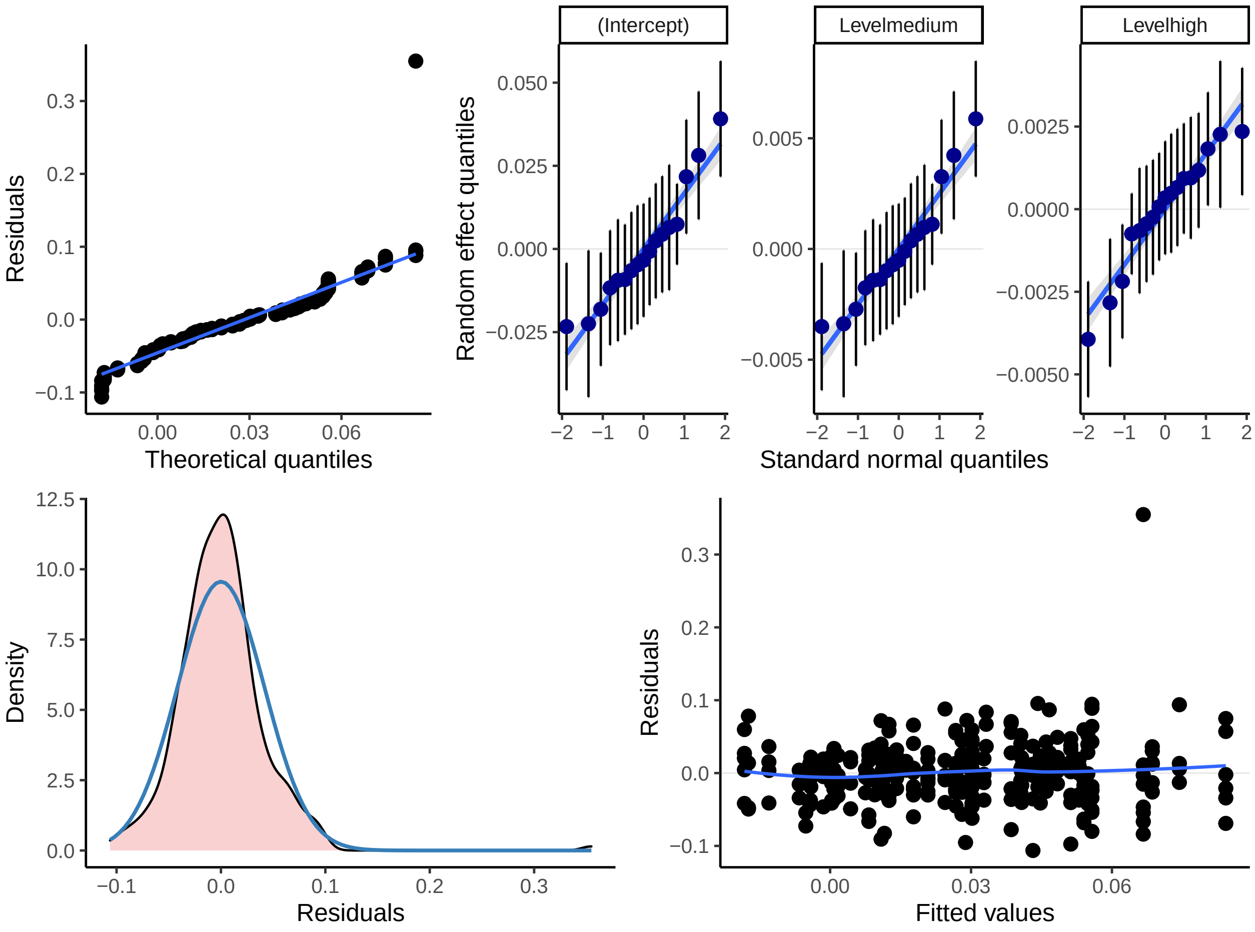
```r
APCPS_mixed_lme4 <- lme4::lmer(APCPS ~ Level + (1+Level|Subject),data=Datos)
summary(APCPS_mixed_lme4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: APCPS ~ Level + (1 + Level | Subject)
##    Data: Datos
```

```
##
## REML criterion at convergence: -997.8
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -2.485 -0.546 -0.028  0.450  8.297
##
## Random effects:
##  Groups   Name        Variance   Std.Dev. Corr
##  Subject  (Intercept) 0.00037402 0.01934
##           Levelmedium 0.00000844 0.00290   1.00
##           Levelhigh   0.00000379 0.00195  -1.00 -1.00
##  Residual             0.00182904 0.04277
## Number of obs: 302, groups:  Subject, 17
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.00509    0.00692    0.74
## Levelmedium  0.01657    0.00647    2.56
## Levelhigh    0.04396    0.00659    6.67
##
## Correlation of Fixed Effects:
##             (Intr) Lvlmdm
## Levelmedium -0.501
## Levelhigh   -0.613  0.603
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
p<-plot_model(APCPS_mixed_lme4, type = "diag")

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



```
contr <- glht(APCPS_mixed_lme4,linfct=mcp(Level="Tukey"))
summary(contr, test = adjusted("holm"))
```

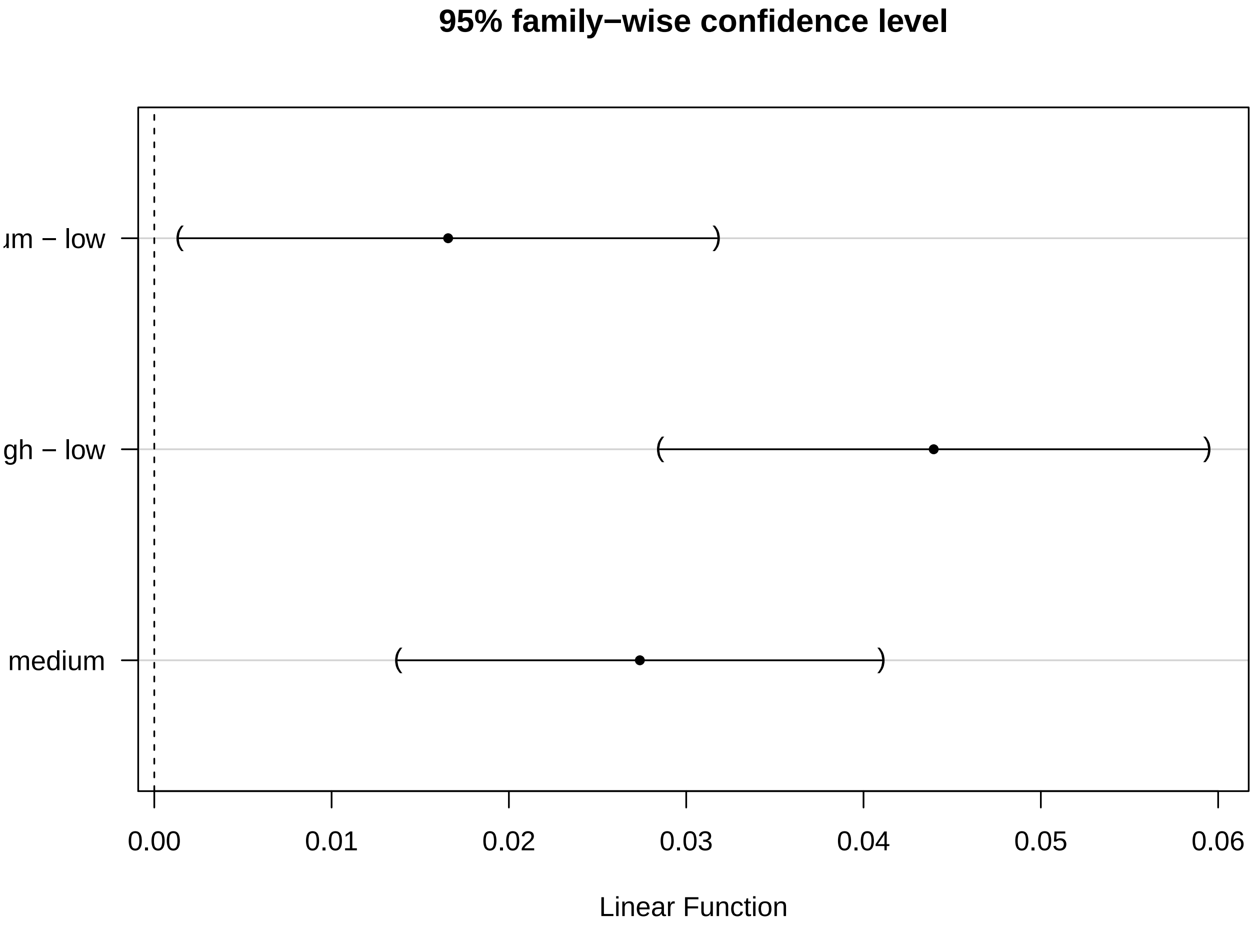
```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = APCPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Linear Hypotheses:
##                      Estimate Std. Error z value       Pr(>|z|)
## medium - low == 0     0.01657    0.00647    2.56           0.01 *
## high - low == 0       0.04396    0.00659    6.67 0.000000000079 ***
## high - medium == 0    0.02738    0.00582    4.70 0.000005160206 ***
## ---
## 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 = APCPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Quantile = 2.342
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##                      Estimate lwr     upr
## medium - low == 0    0.01657  0.00141 0.03173
## high - low == 0      0.04396  0.02851 0.05940
## high - medium == 0   0.02738  0.01374 0.04103
```

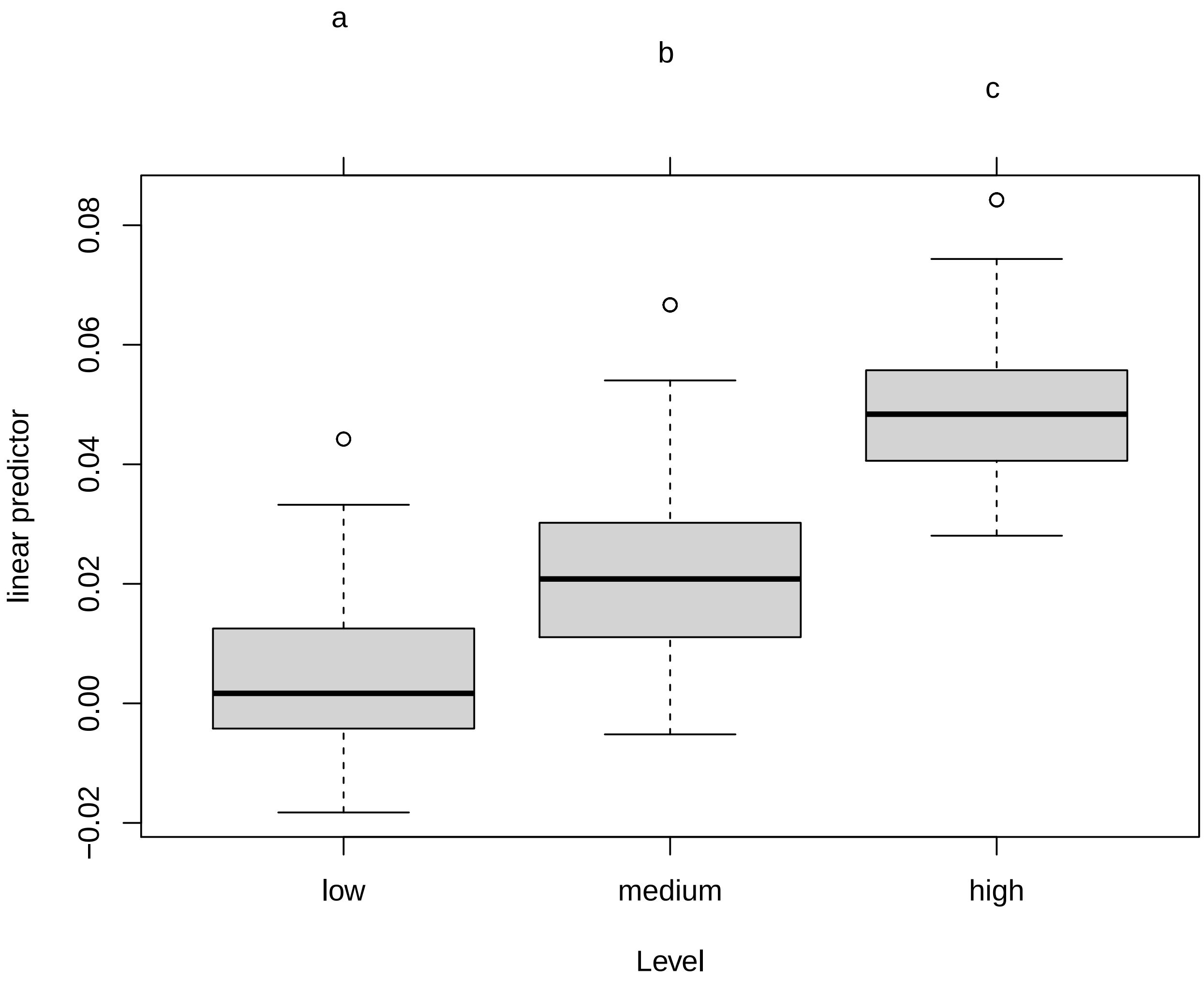
```
plot(confint(contr))
```



```
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
s10	SequenceMemory_r24	FALSE	3	2.744	1.1568	0.4215	1.6368	-1.974	7176806	0	r24	24	14	medium	0.3548	0.0667
s10	SequenceMemory_r05	FALSE	1	3.083	0.4308	0.1397	0.8144	-1.820	8153300	0	r05	5	5	low	0.0955	0.0442
s11	SequenceMemory_r22	FALSE	6	3.611	0.5424	0.1502	0.8607	-2.056	6139927	0	r22	22		high	0.0945	0.0557
s17	SequenceMemory_r18	FALSE	6	3.191	0.5366	0.1682	1.0147	-1.900	8073634	0	r18	18	7	high	0.0938	0.0744
s11	SequenceMemory_r10	FALSE	6	3.505	0.5068	0.1446	0.9528	-2.008	7335892	0	r10	10	11	high	0.0889	0.0557
s03	SequenceMemory_r21	FALSE	3	3.775	0.4244	0.1124	0.6202	-2.071	3588223	0	r21	21	12	medium	0.0880	0.0245

Subject	Name	Training	Nivel	BLPS	MPDC	APCPS	PD	Entropy	TTP	PDS	SequenceMemory	SMN	id	Level	res	fit
s04	SequenceMemory_r27	FALSE	1	3.449	-0.2455	-0.0712	-0.0211	-1.681	3807563	0	r27	27	13	low	-0.0827	0.0116
s10	SequenceMemory_r17	FALSE	3	4.511	-0.0772	-0.0171	0.2468	-2.151	6219667	0	r17	17	10	medium	-0.0838	0.0667
s14	SequenceMemory_r25	FALSE	3	3.697	-0.2953	-0.0799	0.0867	-1.767	6080251	0	r25	25	12	medium	-0.0907	0.0108
s01	SequenceMemory_r17	FALSE	6	4.023	-0.2674	-0.0665	0.1099	-1.911	2292544	0	r17	17	7	high	-0.0953	0.0288
s03	SequenceMemory_r03	FALSE	6	4.688	-0.2167	-0.0462	0.2240	-2.162	8432543	0	r03	3	1	high	-0.0975	0.0512
s16	SequenceMemory_r28	FALSE	6	4.849	-0.3060	-0.0631	-0.0414	-2.184	14432973	0	r28	28	15	high	-0.1063	0.0432



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

```
Datos2 = Datos
Datos2$res = residuals(APCPS_mixed_lme4,type="pearson")
Datos2$fit = fitted(APCPS_mixed_lme4,type="pearson")

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

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

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

```
##
##  Shapiro-Wilk normality test
##
## data:  Datos2$res
## W = 0.88, p-value = 0.00000000000003
```

```
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 = 0.000000000000000000290330359278416, sd =
##  0.0417227812574007
##  Parameters assumed to have been estimated from data
##
## data:  Datos2$res
## Anmax = Inf, p-value = 0.0006
```

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

```
## # A tibble: 1 x 4
##     df1   df2 statistic     p
##   <int> <int>     <dbl> <dbl>
## 1     2   299     0.395 0.674
```

The same model without the outlier

We repeat the analysis without the outlier

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

```
APCPS_mixed_lme4 <- lme4::lmer(APCPS ~ Level + (1+Level|Subject),data=Datos)
summary(APCPS_mixed_lme4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: APCPS ~ Level + (1 + Level | Subject)
##    Data: Datos
##
## REML criterion at convergence: -1080
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8691 -0.6134 -0.0069  0.5478  2.6454
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Subject  (Intercept) 0.000522 0.02285
##           Levelmedium 0.000347 0.01862  -0.61
##           Levelhigh   0.000080 0.00895  -0.61  1.00
##  Residual             0.001326 0.03641
## Number of obs: 301, groups:  Subject, 17
```

```
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.00544    0.00707    0.77
## Levelmedium  0.01353    0.00717    1.89
## Levelhigh    0.04356    0.00605    7.20
##
## Correlation of Fixed Effects:
##             (Intr) Lvlmdm
## Levelmedium -0.680
## Levelhigh   -0.620  0.674
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
anova(APCPS_mixed_lme4)
```

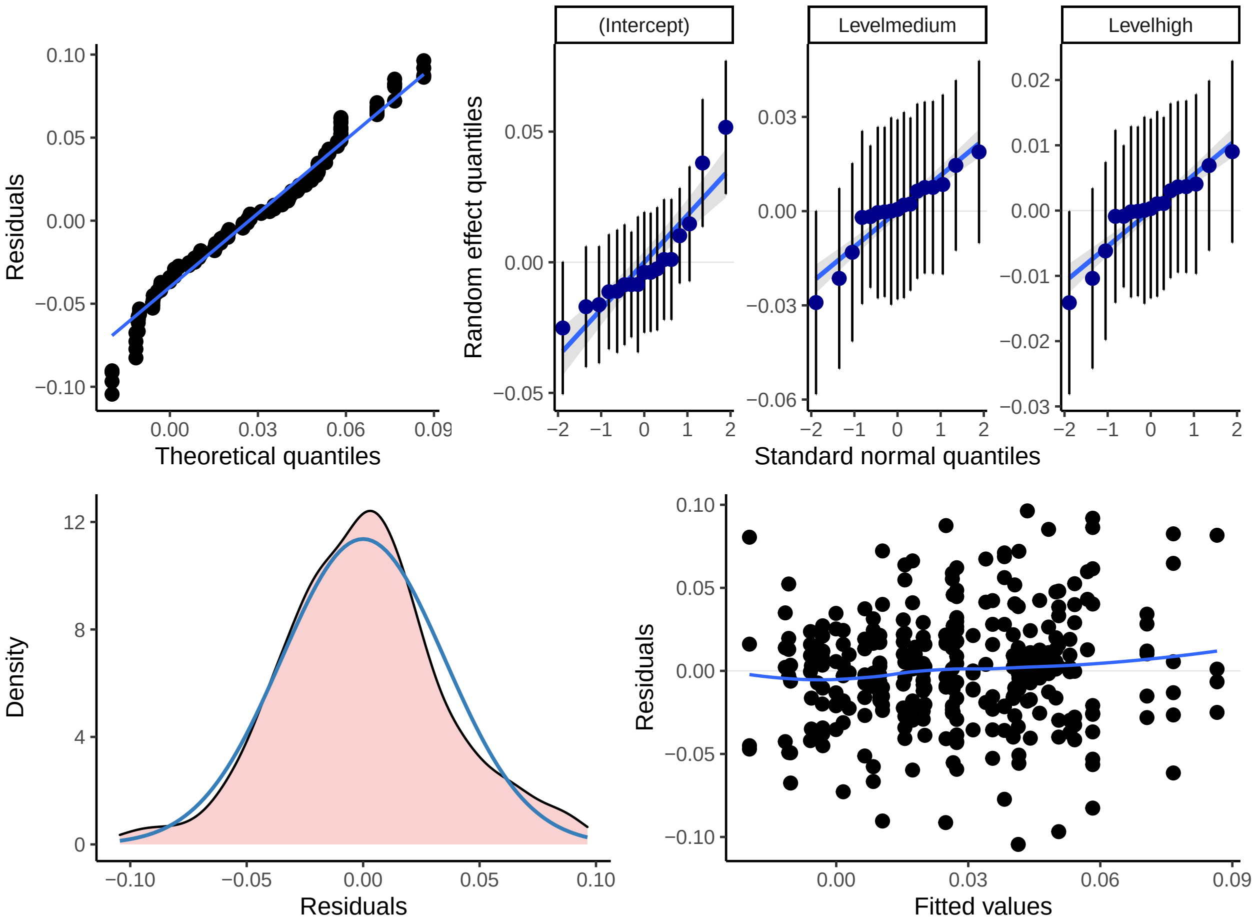
```
## Analysis of Variance Table
##       npar Sum Sq Mean Sq F value
## Level    2 0.0902  0.0451      34
```

```
coef(APCPS_mixed_lme4)
```

```
## $Subject
##     (Intercept) Levelmedium Levelhigh
## s01   -0.019717   0.0154606   0.04459
## s02   -0.005651   0.0140741   0.04387
## s03    0.002935   0.0220085   0.04762
## s04    0.001596   0.0323892   0.05258
## s05    0.006502   0.0198880   0.04659
## s06   -0.003157   0.0130612   0.04337
## s07    0.006475   0.0132871   0.04344
## s08   -0.011580   0.0115234   0.04267
## s09    0.020146   0.0281071   0.05046
## s10    0.043427  -0.0078922   0.03317
## s11    0.015594   0.0117806   0.04268
## s12    0.001579   0.0157691   0.04464
## s13   -0.010796   0.0004295   0.03737
## s14   -0.003039   0.0135542   0.04361
## s15   -0.003079   0.0210372   0.04718
## s16   -0.005858   0.0210960   0.04722
## s17    0.057081  -0.0155793   0.02945
##
## attr(,"class")
## [1] "coef.mer"
```

```
p<-plot_model(APCPS_mixed_lme4, type = "diag")

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



```
# muy importante Tukey para lme4.
contr <- glht(APCPS_mixed_lme4,linfct=mcp(Level="Tukey"))
summary(contr, test = adjusted("holm"))
```

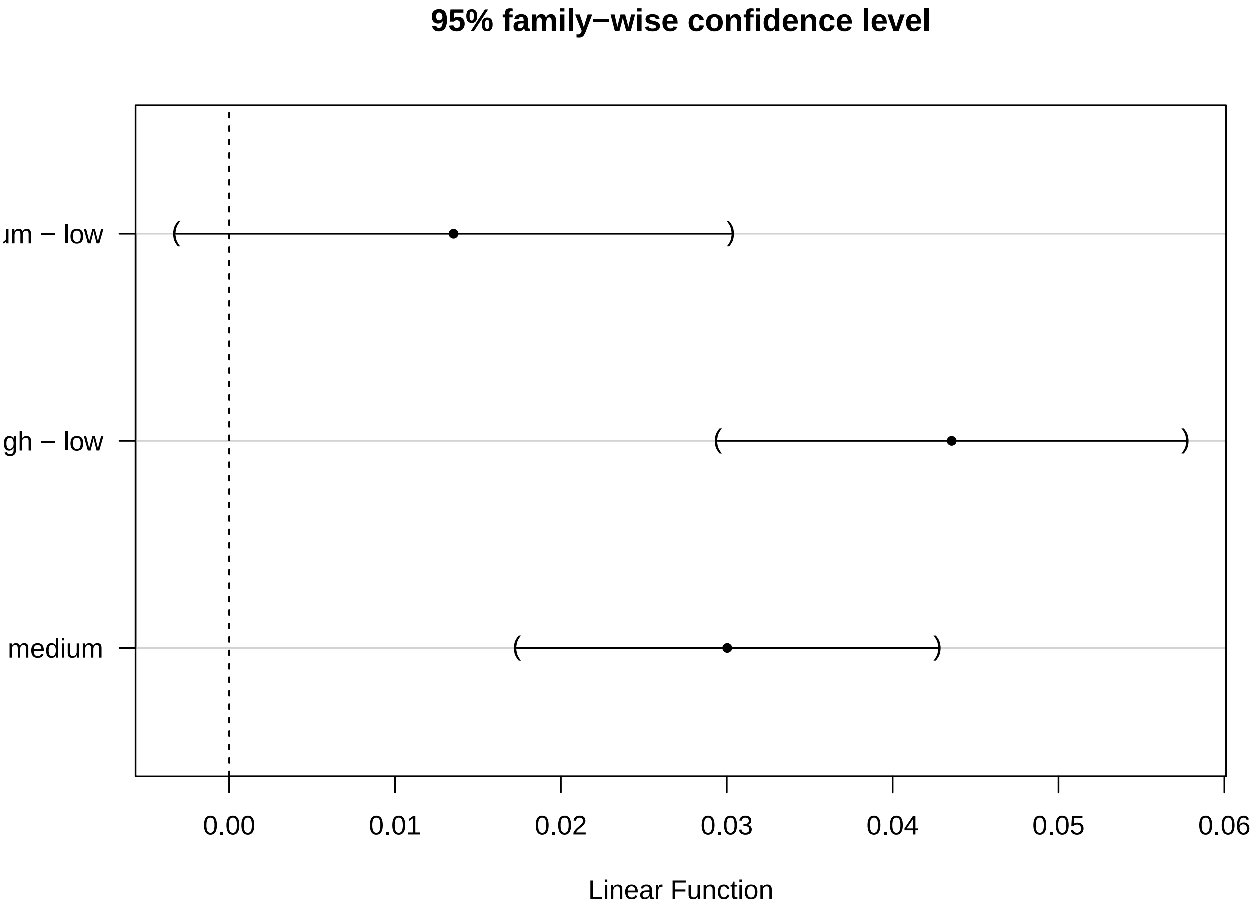
```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = APCPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Linear Hypotheses:
##                      Estimate Std. Error z value        Pr(>|z|)
## medium - low == 0     0.01353    0.00717    1.89           0.059 .
## high - low == 0       0.04356    0.00605    7.20 0.0000000000017 ***
## high - medium == 0    0.03003    0.00544    5.52 0.0000000669154 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- holm method)
```

```
confint(contr)
```

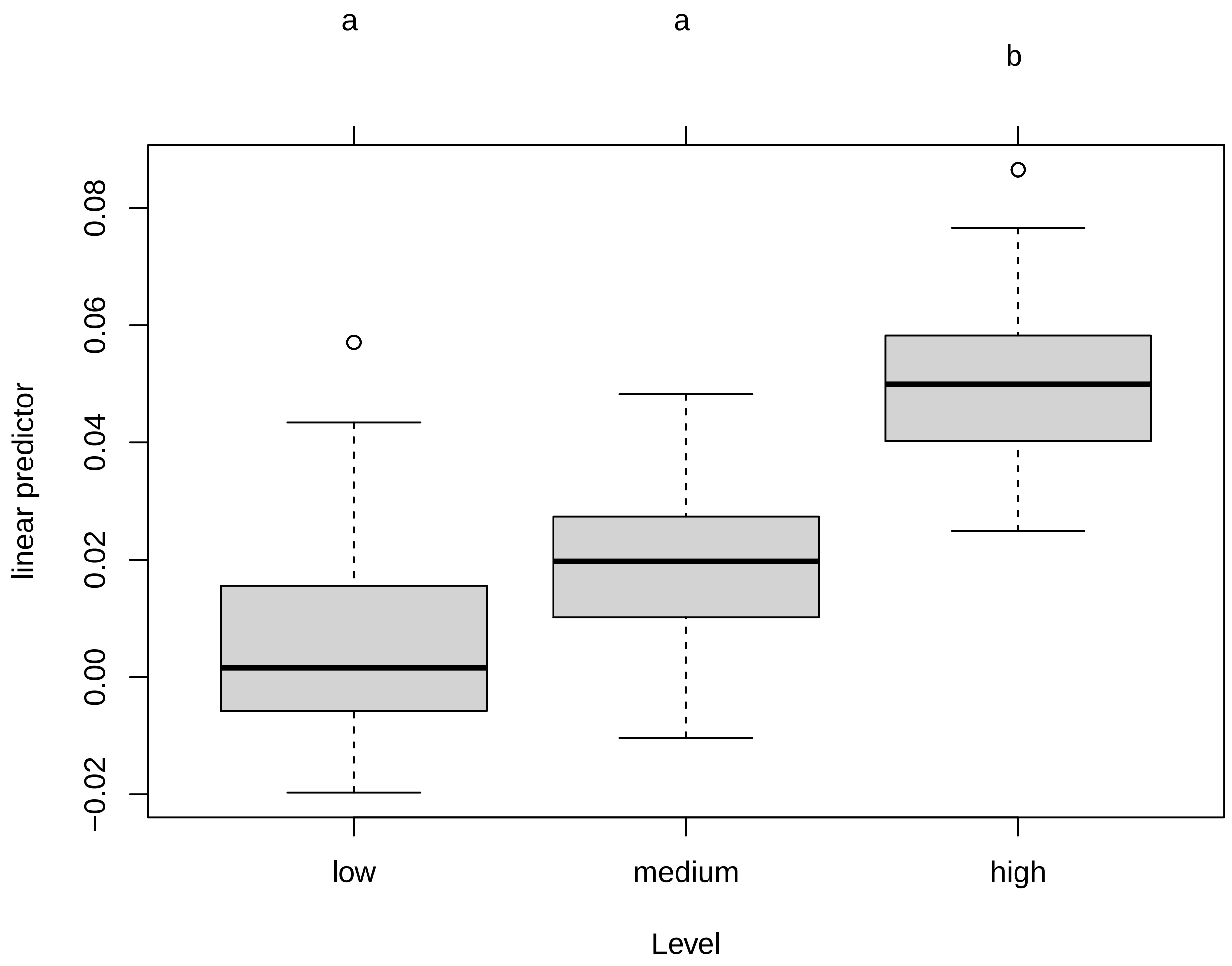
```
##
##   Simultaneous Confidence Intervals
##
```

```
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme4::lmer(formula = APCPS ~ Level + (1 + Level | Subject), data = Datos)
##
## Quantile = 2.334
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##                       Estimate lwr      upr
## medium - low == 0    0.01353 -0.00321  0.03027
## high - low == 0      0.04356  0.02945  0.05767
## high - medium == 0   0.03003  0.01734  0.04272
```

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



```
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)
```



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

```
Datos2=Datos
Datos2$res = residuals(APCPS_mixed_lme4,type="pearson")
Datos2$fit = fitted(APCPS_mixed_lme4,type="pearson")
shapiro.test(Datos2$res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Datos2$res
## W = 0.99, p-value = 0.1
```

```
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 = 0.0000000000000000206445398530857, sd =
##  0.0350986619728449
##  Parameters assumed to have been estimated from data
##
## data:  Datos2$res
## Anmax = 2.1, p-value = 0.8
```

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

```
## # A tibble: 1 x 4
##     df1   df2 statistic     p
##   <int> <int>     <dbl> <dbl>
## 1     2   298      1.89 0.153
```

Non parametric tests

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

```
##
##  Kruskal-Wallis rank sum test
##
## data:  APCPS by Level
## Kruskal-Wallis chi-squared = 57, df = 2, p-value = 0.0000000000003
```

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

```
##
##  Pairwise comparisons using Tukey and Kramer (Nemenyi) test
##                    with Tukey-Dist approximation for independent samples
##
## data:  APCPS by Level
##
##        low              medium
## medium 0.012            -
## high   0.00000000000092 0.00000052607408
##
## P value adjustment method: none
```

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

```
##        low             medium
## medium 0.024           -
## high   0.0000000000083 0.0000002818669
```