Size: 75% POST1 FILTERED L2

## Data analysis

We load the data and remove the cases with NA values.

We have 100 subjects for a total of 1600 observations.

After removal of subjects with POST1 and POST2 <75% and entries with missing data due to wrong answers, we have 86 subjects and 1012 observations left.

We consider the following independent variables: - LEAYRS a numerical variable with values ranging from 1 to 16. Learning in years. - LEA a percentage scaled to lay between 0 and 10. It indicates the novelty of target construction, measured by subtracting pretest score from posttestscore (POST1-PRE). This value indicates the previous knowledge of construction. - AMSP a numerical variable with values ranging from 1 to 5. - HRSD a numerical variable with values ranging from 0 to 11.5. L2 use in hours per day.

We rescaled some of these variables to be on similar scales. We treat these as fixed factors under study. In addition we have the random factors described earlier. The same modelling set up applies here.

Our base model is the following lmm <- lmer(log(Time) ~ type + LEAYRS + LEA + AMSP + HRSD + (1 | List:Name) + (1 | List:type:Order), data = DLM)

lmm <- lmer(log(Time) ~ type + LEAYRS + LEA + AMSP + HRSD + (1 | List:Name) + (1 | List:type:Order), data = DLM, REML=FALSE)  
summary(lmm)

## Linear mixed model fit by maximum likelihood ['lmerMod']  
## Formula:   
## log(Time) ~ type + LEAYRS + LEA + AMSP + HRSD + (1 | List:Name) +   
## (1 | List:type:Order)  
## Data: DLM  
##   
## AIC BIC logLik deviance df.resid   
## 1247.5 1291.8 -614.7 1229.5 1003   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.197 -0.649 -0.113 0.579 4.290   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0666 0.258   
## List:type:Order (Intercept) 0.0236 0.154   
## Residual 0.1616 0.402   
## Number of obs: 1012, groups: List:Name, 86; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 7.577386 0.230699 32.8  
## typeMISMATCH 0.105459 0.060372 1.7  
## LEAYRS 0.009678 0.008524 1.1  
## LEA 0.000822 0.015486 0.1  
## AMSP -0.164402 0.054752 -3.0  
## HRSD -0.031589 0.035696 -0.9  
##   
## Correlation of Fixed Effects:  
## (Intr) tMISMA LEAYRS LEA AMSP   
## typMISMATCH -0.122   
## LEAYRS 0.023 -0.006   
## LEA -0.397 -0.011 0.072   
## AMSP -0.945 -0.003 -0.205 0.270   
## HRSD 0.079 0.002 0.060 0.033 -0.216

We use a manual, step-forward procedure with likelihood ratio test to see which of the fixed effects are significant.

lmm.0 <- lmer(log(Time) ~ (1 | List:Name) + (1 | List:type:Order), data = DLM)  
  
lmm.1 <- update(lmm.0, .~. + type)  
anova(lmm.0,lmm.1)

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

## Data: DLM  
## Models:  
## lmm.0: log(Time) ~ (1 | List:Name) + (1 | List:type:Order)  
## lmm.1: log(Time) ~ (1 | List:Name) + (1 | List:type:Order) + type  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## lmm.0 4 1252 1272 -622 1244   
## lmm.1 5 1251 1276 -621 1241 2.93 1 0.087 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

lmm.2 <- update(lmm.0, .~.+LEAYRS)  
anova(lmm.0,lmm.2)

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

## Data: DLM  
## Models:  
## lmm.0: log(Time) ~ (1 | List:Name) + (1 | List:type:Order)  
## lmm.2: log(Time) ~ (1 | List:Name) + (1 | List:type:Order) + LEAYRS  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1252 1272 -622 1244   
## lmm.2 5 1254 1279 -622 1244 0.16 1 0.69

lmm.3 <- update(lmm.0, .~. + LEA)   
anova(lmm.0,lmm.3)

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

## Data: DLM  
## Models:  
## lmm.0: log(Time) ~ (1 | List:Name) + (1 | List:type:Order)  
## lmm.3: log(Time) ~ (1 | List:Name) + (1 | List:type:Order) + LEA  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1252 1272 -622 1244   
## lmm.3 5 1254 1278 -622 1244 0.89 1 0.35

lmm.4 <- update(lmm.0, .~. + AMSP)  
anova(lmm.0,lmm.4)

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

## Data: DLM  
## Models:  
## lmm.0: log(Time) ~ (1 | List:Name) + (1 | List:type:Order)  
## lmm.4: log(Time) ~ (1 | List:Name) + (1 | List:type:Order) + AMSP  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## lmm.0 4 1252 1272 -622 1244   
## lmm.4 5 1245 1269 -617 1235 9.8 1 0.0017 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

lmm.5 <- update(lmm.0, .~. + HRSD)  
anova(lmm.0,lmm.5)

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

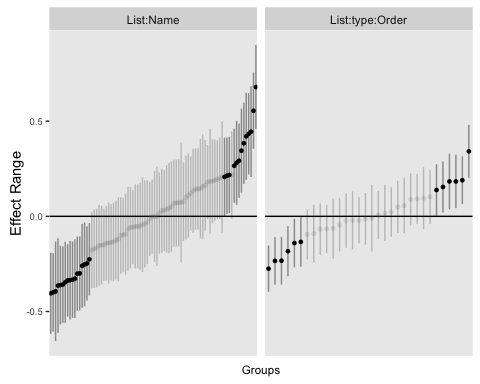
## Data: DLM  
## Models:  
## lmm.0: log(Time) ~ (1 | List:Name) + (1 | List:type:Order)  
## lmm.5: log(Time) ~ (1 | List:Name) + (1 | List:type:Order) + HRSD  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1252 1272 -622 1244   
## lmm.5 5 1252 1277 -621 1242 2.48 1 0.12

We conclude that only AMSP has a significant effect at a 0.05 level while type is significant at a 0.1 level.

An automatic, step-backward procedure from the package lmerTest starting from a model that includes all second order fixed factor intereactions lead to a similar conclusion: only the factor AMSP is kept in the model.

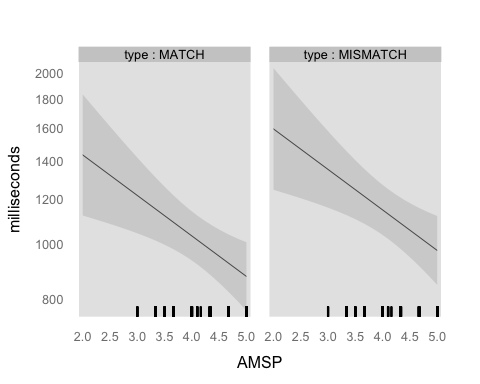
Finally, Analysis of Variance Table of type III with Satterthwaite approximation for degrees of freedom leads to the same conclusions as the likelihood ratio test procedure. It is reported at the end of this document.

The figures show random and the fixed effects. Again the subject effects are significant. We back transformed the fixed effects in linear scale and added 0.95-confidence level bands. It is evident the reduction in reaction time as AMSP increases. Although type has an effect it is not large.



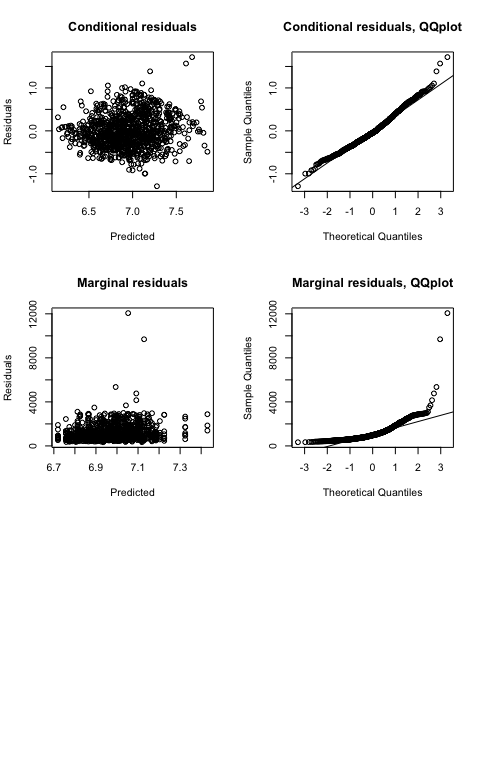
## Loading required package: effects

## Warning: package 'effects' was built under R version 3.2.4



## Diagnostic plots

lmm<- lmer(log(Time) ~ (type + HRSD + LEAYRS + AMSP) + (1 | Name)+ (1 | type:List:Order), data = DLM)  
par(mfrow=c(3,2))  
# plot(lm4,which=1:4)  
  
plot(fitted(lmm, type = "response"), residuals(lmm, type = "response"),  
 main = "Conditional residuals", xlab = "Predicted", ylab = "Residuals")  
  
res <- residuals(lmm, type = "response")  
qqnorm(res, main = "Conditional residuals, QQplot")  
qqline(res)  
  
lm.0 <- lm(log(Time) ~ (type + HRSD + LEAYRS + AMSP), data = DLM)  
x <- model.matrix(lm.0)  
pred <- x %\*% fixef(lmm)  
res <- DLM$Time - pred  
plot(pred, res, main = "Marginal residuals", xlab = "Predicted", ylab = "Residuals")  
qqnorm(res, main = "Marginal residuals, QQplot")  
qqline(res)



The joint qqplot looks normal. The marginal looks less nice.

## Anova Table with Satterwhite

require(lmerTest)

## Loading required package: lmerTest

##   
## Attaching package: 'lmerTest'

## Det følgende objekt er maskeret fra 'package:lme4':  
##   
## lmer

## Det følgende objekt er maskeret fra 'package:stats':  
##   
## step

lmm <- lmer(log(Time) ~ type + LEAYRS + (PRE + POST1)^2 + AMSP + HRSD + POST1\*HRSD + (1 | List:Name) + (1 | List:type:Order), data = DLM, REML=FALSE)  
anova(lmm)

## Analysis of Variance Table of type III with Satterthwaite   
## approximation for degrees of freedom  
## Sum Sq Mean Sq NumDF DenDF F.value Pr(>F)   
## type 0.503 0.503 1 32.7 3.12 0.0869 .   
## LEAYRS 0.314 0.314 1 85.3 1.94 0.1672   
## PRE 0.408 0.408 1 85.0 2.53 0.1158   
## POST1 0.099 0.099 1 84.9 0.61 0.4369   
## AMSP 1.780 1.780 1 90.1 11.01 0.0013 \*\*  
## HRSD 0.458 0.458 1 85.5 2.83 0.0961 .   
## PRE:POST1 0.406 0.406 1 85.0 2.51 0.1167   
## POST1:HRSD 0.470 0.470 1 85.5 2.91 0.0917 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(lmm)

## Linear mixed model fit by maximum likelihood t-tests use  
## Satterthwaite approximations to degrees of freedom [lmerMod]  
## Formula:   
## log(Time) ~ type + LEAYRS + (PRE + POST1)^2 + AMSP + HRSD + POST1 \*   
## HRSD + (1 | List:Name) + (1 | List:type:Order)  
## Data: DLM  
##   
## AIC BIC logLik deviance df.resid   
## 1249.1 1308.2 -612.6 1225.1 1000   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.191 -0.659 -0.115 0.586 4.333   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0628 0.251   
## List:type:Order (Intercept) 0.0234 0.153   
## Residual 0.1616 0.402   
## Number of obs: 1012, groups: List:Name, 86; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 9.06971 1.81012 84.80000 5.01 2.9e-06 \*\*\*  
## typeMISMATCH 0.10624 0.06019 32.70000 1.76 0.0869 .   
## LEAYRS 0.01169 0.00839 85.30000 1.39 0.1672   
## PRE -0.45310 0.28514 85.00000 -1.59 0.1158   
## POST1 -0.14449 0.18498 84.90000 -0.78 0.4369   
## AMSP -0.18419 0.05550 90.10000 -3.32 0.0013 \*\*   
## HRSD 2.41534 1.43525 85.50000 1.68 0.0961 .   
## PRE:POST1 0.04589 0.02895 85.00000 1.58 0.1167   
## POST1:HRSD -0.24658 0.14458 85.50000 -1.71 0.0917 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) tMISMA LEAYRS PRE POST1 AMSP HRSD PRE:PO  
## typMISMATCH -0.012   
## LEAYRS 0.009 -0.005   
## PRE -0.840 -0.008 -0.074   
## POST1 -0.992 -0.004 0.003 0.825   
## AMSP -0.090 -0.006 -0.226 0.104 -0.011   
## HRSD -0.112 0.003 0.102 -0.233 0.124 -0.125   
## PRE:POST1 0.845 0.008 0.070 -0.999 -0.832 -0.117 0.231   
## POST1:HRSD 0.113 -0.003 -0.101 0.233 -0.125 0.121 -1.000 -0.231