Size: 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 entries with missing data due to wrong answers, we have 100 subjects and 1171 observations left.

We consider the following independent variables: - LEAYRS a numerical variable with values ranging from 1 to 16. Learning in years. - POST1 a percentage scaled to lay between 0 and 10. It indicates the post test proficency. - PRE a percentage scaled to lay between 0 and 10. It indicates the previous knowledge. - 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.

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

lmm <- lmer(log(Time) ~ type + LEAYRS + (PRE + POST1)^2 + AMSP + HRSD + POST1\*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 + (PRE + POST1)^2 + AMSP + HRSD + POST1 \*   
## HRSD + (1 | List:Name) + (1 | List:type:Order)  
## Data: DLM  
##   
## AIC BIC logLik deviance df.resid   
## 1417.8 1478.6 -696.9 1393.8 1159   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.295 -0.665 -0.104 0.589 4.411   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0607 0.246   
## List:type:Order (Intercept) 0.0228 0.151   
## Residual 0.1591 0.399   
## Number of obs: 1171, groups: List:Name, 100; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 7.87084 1.52052 5.18  
## typeMISMATCH 0.10843 0.05870 1.85  
## LEAYRS 0.01140 0.00815 1.40  
## PRE -0.36933 0.22244 -1.66  
## POST1 -0.02448 0.15508 -0.16  
## AMSP -0.16706 0.05258 -3.18  
## HRSD 2.03105 0.76228 2.66  
## PRE:POST1 0.03700 0.02260 1.64  
## POST1:HRSD -0.20866 0.07719 -2.70  
##   
## Correlation of Fixed Effects:  
## (Intr) tMISMA LEAYRS PRE POST1 AMSP HRSD PRE:PO  
## typMISMATCH -0.012   
## LEAYRS 0.009 -0.005   
## PRE -0.913 -0.007 -0.022   
## POST1 -0.990 -0.007 0.003 0.896   
## AMSP -0.078 -0.005 -0.234 0.094 -0.041   
## HRSD -0.483 -0.004 0.095 0.226 0.508 -0.232   
## PRE:POST1 0.915 0.007 0.019 -0.998 -0.901 -0.106 -0.222   
## POST1:HRSD 0.483 0.004 -0.092 -0.224 -0.508 0.223 -0.999 0.220

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 1425 1445 -708 1417   
## lmm.1 5 1423 1449 -707 1413 3.2 1 0.073 .  
## ---  
## 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 1425 1445 -708 1417   
## lmm.2 5 1426 1452 -708 1416 0.22 1 0.64

lmm.3 <- update(lmm.0, .~. + POST1)   
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) + POST1  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1425 1445 -708 1417   
## lmm.3 5 1426 1452 -708 1416 0.06 1 0.8

lmm.3 <- update(lmm.0, .~. + PRE)   
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) + PRE  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1425 1445 -708 1417   
## lmm.3 5 1426 1451 -708 1416 1.01 1 0.32

lmm.3 <- update(lmm.0, .~. + POST1)   
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) + POST1  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1425 1445 -708 1417   
## lmm.3 5 1426 1452 -708 1416 0.06 1 0.8

lmm.3 <- update(lmm.0, .~. + PRE\*POST1)   
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) + PRE + POST1 +   
## lmm.3: PRE:POST1  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.0 4 1425 1445 -708 1417   
## lmm.3 7 1427 1462 -706 1413 4.03 3 0.26

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 1425 1445 -708 1417   
## lmm.4 5 1422 1447 -706 1412 4.79 1 0.029 \*  
## ---  
## 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 1425 1445 -708 1417   
## lmm.5 5 1425 1450 -707 1415 1.57 1 0.21

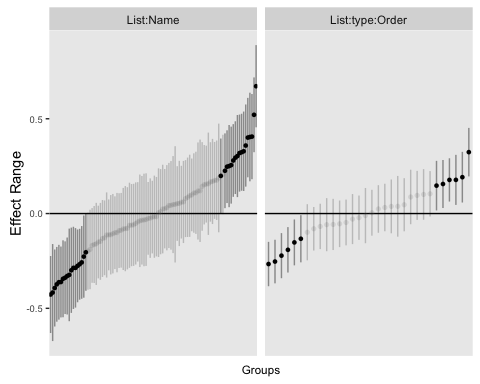
We conclude that only AMSP has a significant effect at a 0.05 level while the type has a significant effect at a level of 0.1.

An automatic, step-backward procedure from the package lmerTest starting from a model that includes all second order fixed factor intereactions lead to a different conclusion: the factors AMSP, POST1, HRSD and the interaction between these last two are significant.

Finally, Analysis of Variance Table of type III with Satterthwaite approximation for degrees of freedom gives produces yet another different analysis. It is reported at the end of this document.

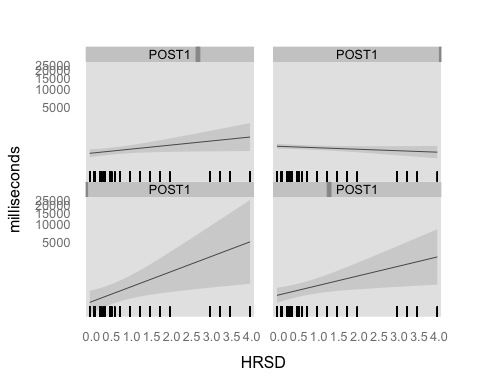
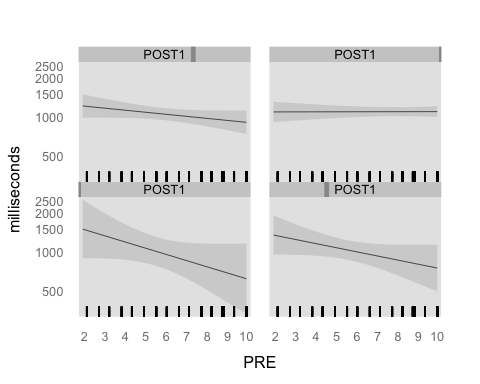
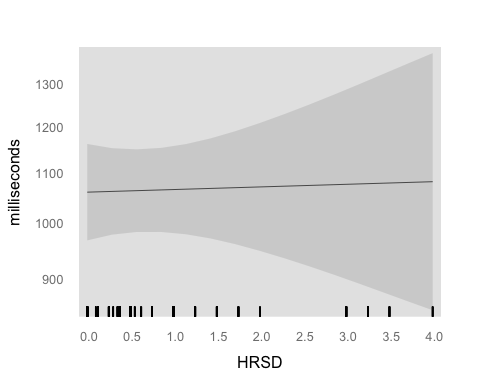
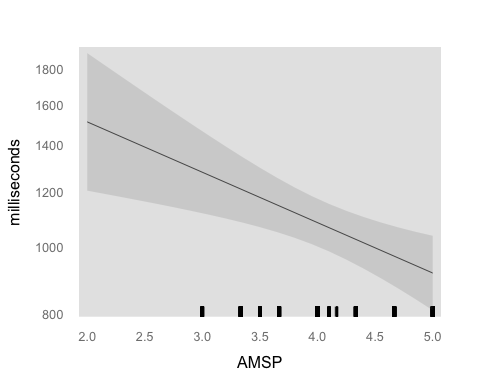
All three analysis indicate however that AMSP is significant.

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. The HRSD values do not seem to have a considerable impact while PRE and POST1 interactions are relevant but only among subjects with low POST1. Finally, HRSD and POST1 interactions are depicted.



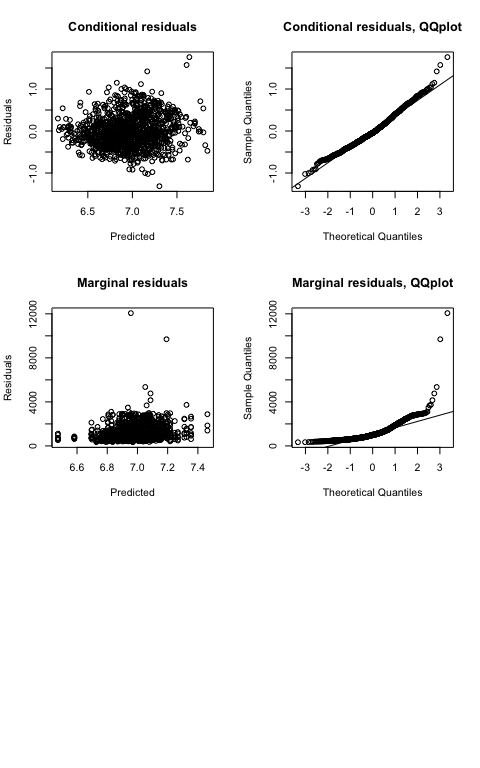
## Loading required package: effects

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



## Diagnostic plots

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 + LEAYRS + (PRE + POST1)^2 + AMSP + HRSD + POST1\*HRSD ), 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 0.543 0.543 1 32.7 3.41 0.0738 .   
## LEAYRS 0.311 0.311 1 99.2 1.96 0.1651   
## PRE 0.438 0.438 1 98.3 2.76 0.1000   
## POST1 0.004 0.004 1 99.7 0.02 0.8749   
## AMSP 1.606 1.606 1 103.6 10.09 0.0020 \*\*  
## HRSD 1.129 1.129 1 105.4 7.10 0.0089 \*\*  
## PRE:POST1 0.426 0.426 1 98.4 2.68 0.1048   
## POST1:HRSD 1.162 1.162 1 105.3 7.31 0.0080 \*\*  
## ---  
## 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   
## 1417.8 1478.6 -696.9 1393.8 1159   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.295 -0.665 -0.104 0.589 4.411   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0607 0.246   
## List:type:Order (Intercept) 0.0228 0.151   
## Residual 0.1591 0.399   
## Number of obs: 1171, groups: List:Name, 100; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 7.87084 1.52052 99.70000 5.18 1.2e-06 \*\*\*  
## typeMISMATCH 0.10843 0.05870 32.70000 1.85 0.0738 .   
## LEAYRS 0.01140 0.00815 99.20000 1.40 0.1651   
## PRE -0.36933 0.22244 98.30000 -1.66 0.1000   
## POST1 -0.02448 0.15508 99.70000 -0.16 0.8749   
## AMSP -0.16706 0.05258 103.60000 -3.18 0.0020 \*\*   
## HRSD 2.03105 0.76228 105.40000 2.66 0.0089 \*\*   
## PRE:POST1 0.03700 0.02260 98.40000 1.64 0.1048   
## POST1:HRSD -0.20866 0.07719 105.30000 -2.70 0.0080 \*\*   
## ---  
## 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 -0.913 -0.007 -0.022   
## POST1 -0.990 -0.007 0.003 0.896   
## AMSP -0.078 -0.005 -0.234 0.094 -0.041   
## HRSD -0.483 -0.004 0.095 0.226 0.508 -0.232   
## PRE:POST1 0.915 0.007 0.019 -0.998 -0.901 -0.106 -0.222   
## POST1:HRSD 0.483 0.004 -0.092 -0.224 -0.508 0.223 -0.999 0.220