Assignment 2

Clara Baudry & Gaja Nenadović & Hsuan Lee

06 marec, 2022

1. Convert the wide data file into a long format. Check the data and recode if necessary.

Table 1

Descriptive statistics (continued below)

	vars	n	mean	sd	median	trimmed	mad	min
momage	1	884	25.6	1.869	26	25.64	1.483	21
homecog	2	884	9.1	2.451	9	9.22	2.965	3
time	3	884	2.5	1.119	2.5	2.5	1.483	1
anti	4	884	1.819	1.823	1	1.555	1.483	0
read	5	884	4.345	1.619	4.4	4.314	1.853	0.7

	max	range	skew	kurtosis	se
momage	29	8	-0.142	-0.8477	0.06286
homecog	14	11	-0.3733	-0.4163	0.08243
time	4	3	0	-1.364	0.03762
anti	10	10	1.119	1.049	0.06132
read	8.4	7.7	0.1058	-0.7705	0.05447

We see that the *time* variable ranges from 1 to 4, which means that the first occasion is coded as 1. This means that the intercept cannot be interpreted meaningfully (the time variable will never be equal to 0). Therefore, we will recode the time variable and set its range from 0 to 3, so that the intercept is the predicted outcome for the first occasion when all other predictors are equal to 0. Looking at the ranges of the other predictors, we see that 0 is not a plausible value for any of them. This again means that the intercept will not be meaningful and we will have to extrapolate to interpret it. We will center the three predictors (*momage, read, homecog*), so that the intercept can be interpreted as the predicted outcome at the first occasion for a person with average values on all the other predictors.

Table 2

*Descriptive statistics for centered and scaled predictors (continued below)*

	vars	n	mean	sd	median	trimmed	mad
momage	1	884	-1.432e-15	1.869	0.4027	0.04113	1.483
homecog	2	884	3.544e-16	2.451	-0.09955	0.1208	2.965
time	3	884	1.5	1.119	1.5	1.5	1.483
anti	4	884	1.819	1.823	1	1.555	1.483
read	5	884	-1.129e-16	1.619	0.05464	-0.0311	1.853

	min	max	range	skew	kurtosis	se
momage	-4.597	3.403	8	-0.142	-0.8477	0.06286
homecog	-6.1	4.9	11	-0.3733	-0.4163	0.08243
time	0	3	3	0	-1.364	0.03762
anti	0	10	10	1.119	1.049	0.06132
read	-3.645	4.055	7.7	0.1058	-0.7705	0.05447

As we see from the table, 0 is now a plausible (and meaningful) value for every predictor.

## Check the linearity assumption, report and include plots.

We will check if: 1) there is a linear relationship between each of the first level predictors and the outcome, that is between *time* and *antisocial behavior* (outcome variable) and *read* and *antisocial behavior*; 2) there is a linear relationship between the level-2 predictors and the aggregated outcome, that is between *cognitive stimulation*, *antisocial behavior*, *mother's age* and *antisocial behavior*.
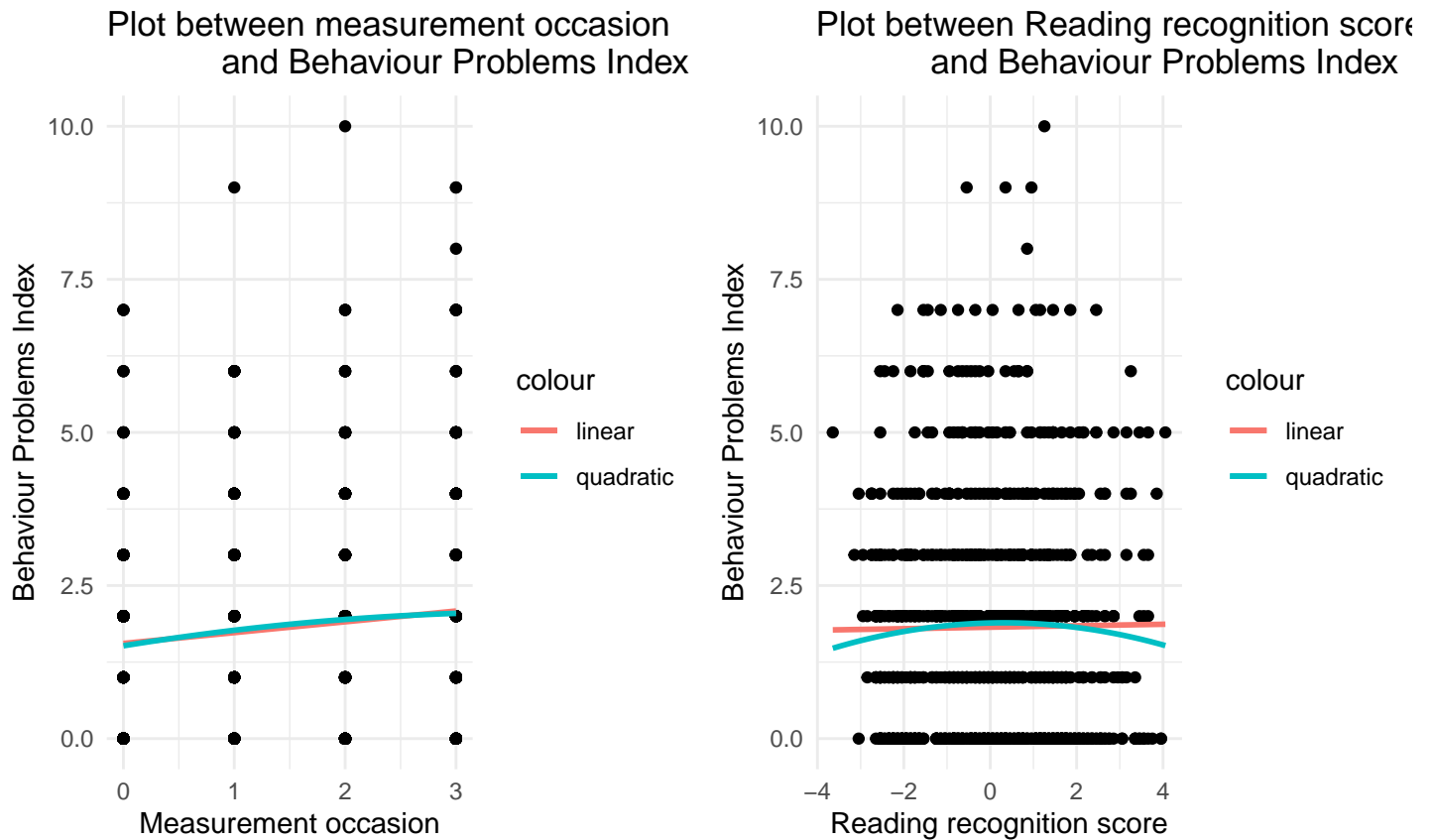

The plot of the relationship between *time* and *antisocial behavior* reveals that the quadratic trend line is practically the same as the linear trend line. It shows that the explained variance is almost the same for these two trends and the relation appears to be approximately linear.

As we can see from the *Plot between Reading recognition score and Behaviour Problems Index*, the linear and quadratic lines for the relationship between the *reading recognition skills* and *antisocial behavior* are now slightly different. The quadratic trend seems to fit somewhat better (and explain more variance in the outcome) than the linear one, albeit to a rather small degree. The slope of the linear trend line seems to be close to 0, indicating that there is no linear relationship between these two variables. There appears to be neither linear nor quadratic relationship between the two variables, but we will include *read* in the analysis to further evaluate this.

Scatterplot 1

Scatterplot 2

Scatterplot 3 and Scatterplot 4

Behaviour Problems Index vs Reading recognition
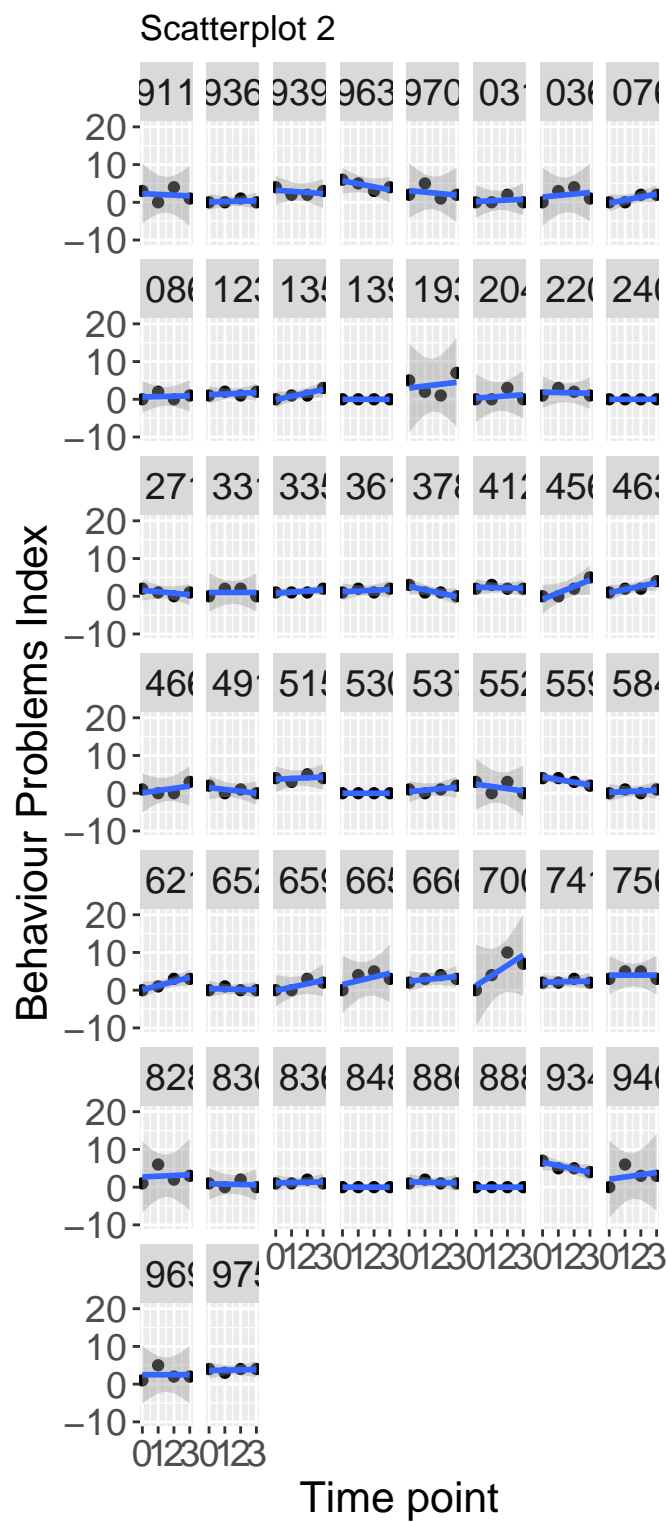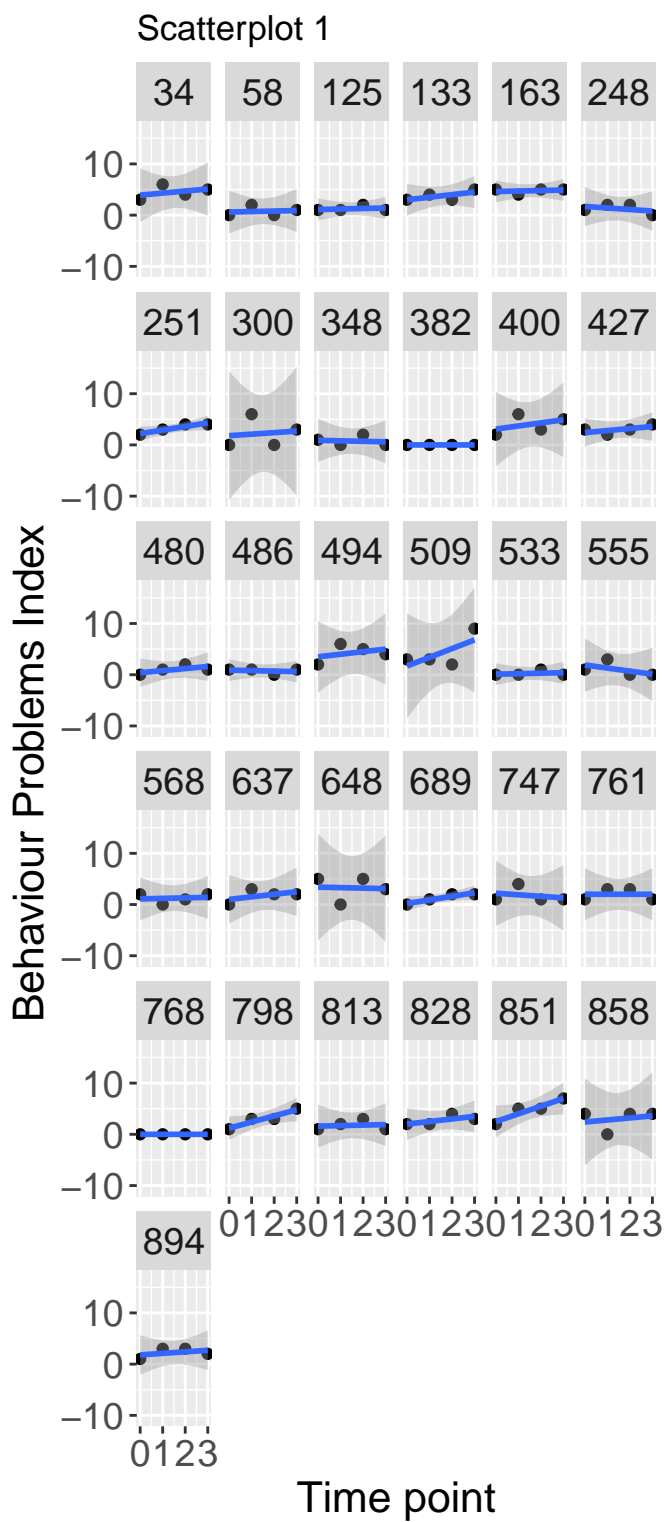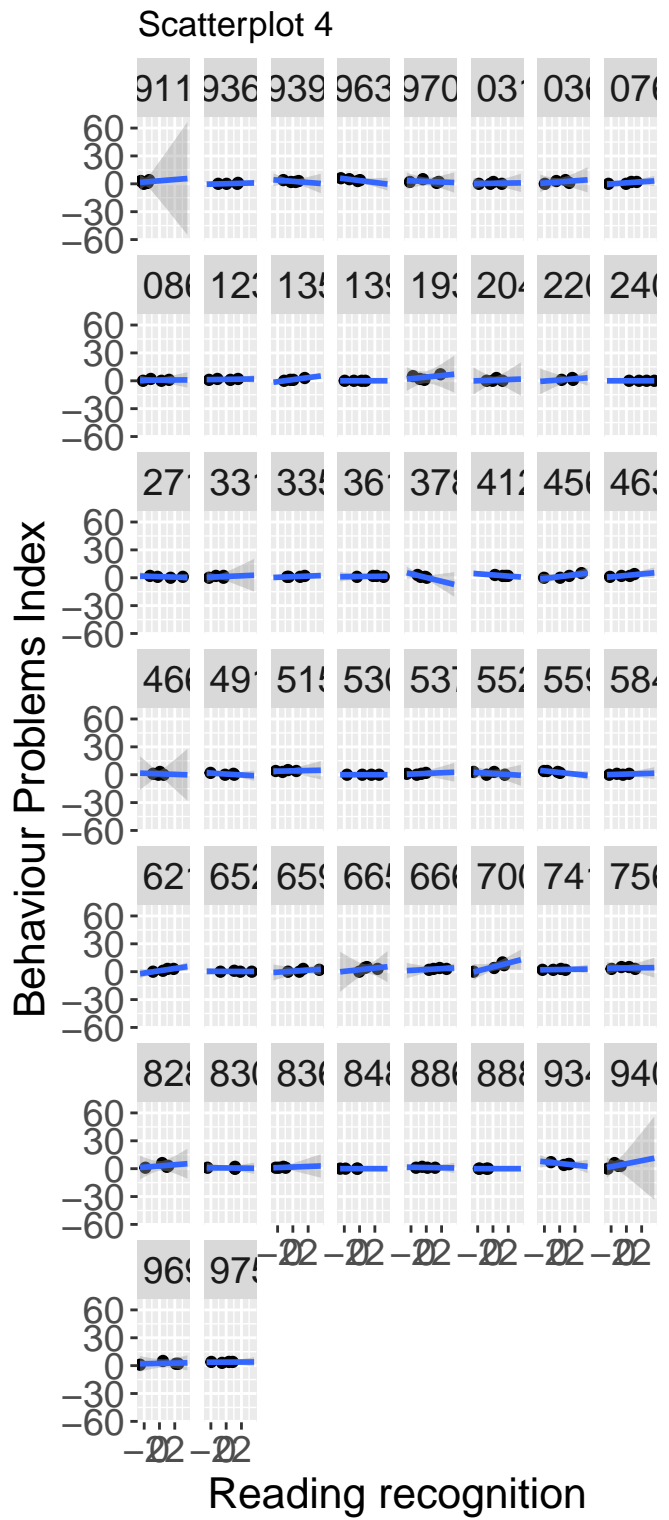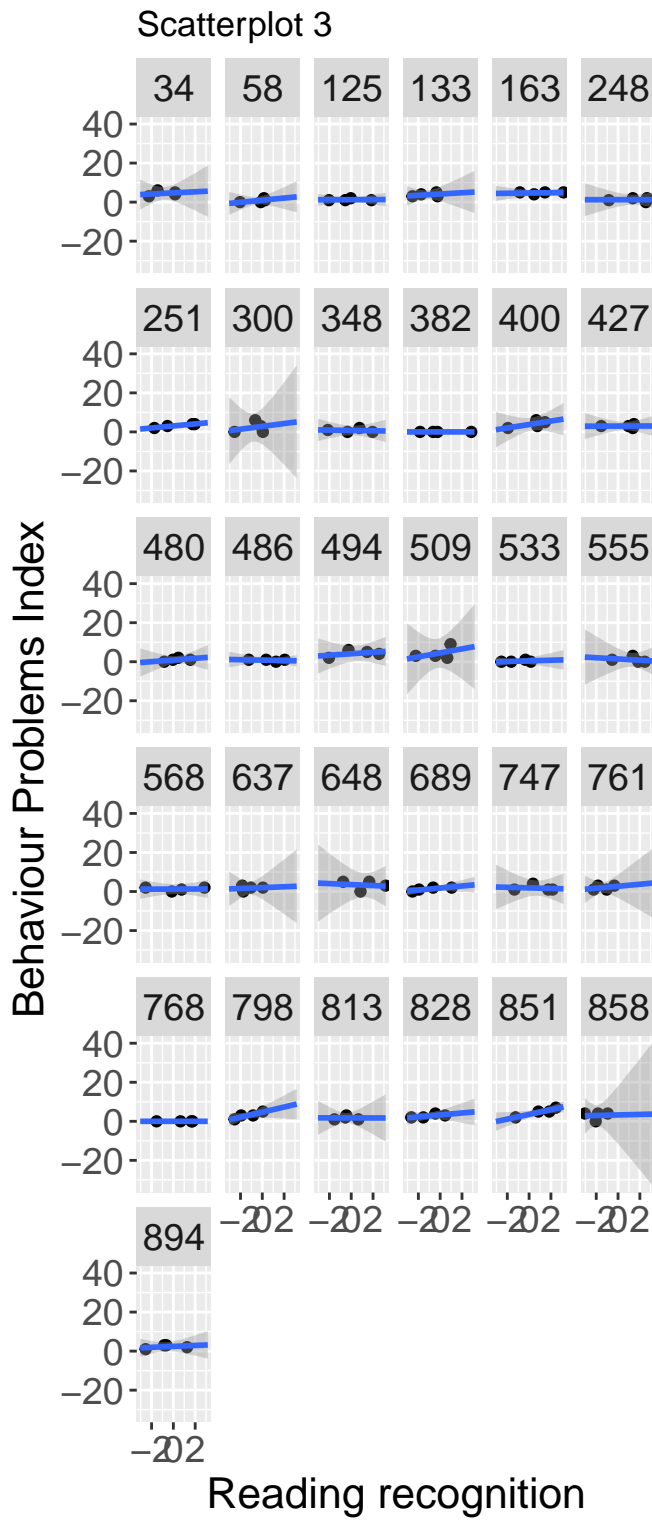
The plots for a subset of children show the relationships between the first level predictors and the outcome for each child separately (indicated by their id numbers). The relationship between *time* and *antisocial behavior* slightly varies between the children. The relationship between *reading recognition skills* and *antisocial behavior* is quite similar for different children in this subset and the slopes seem to be close to 0.

Scatterplot for cognitive stimulation and antisocial behavior
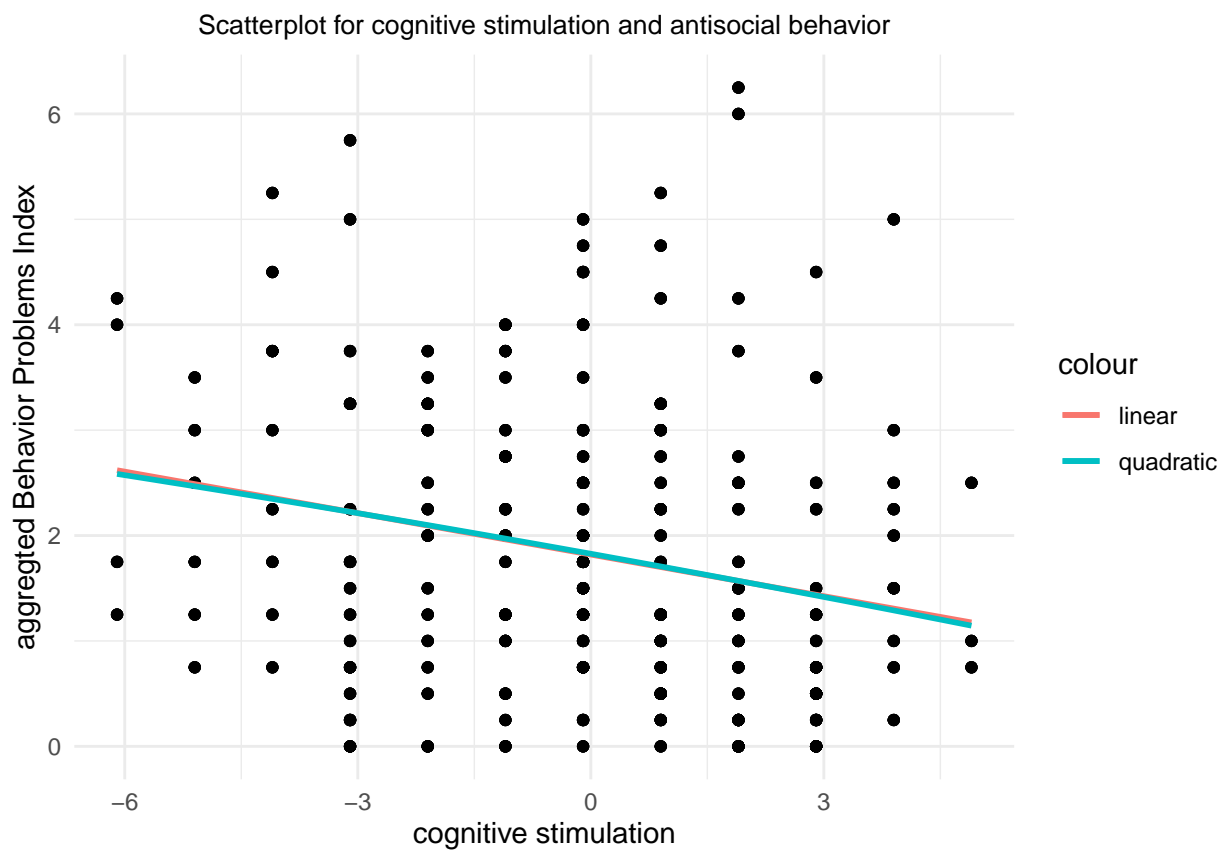
The *Scatterplot for cognitive stimulation and antisocial behavior* reveals that both linear and quadratic trend lines for the relationship between *cognitive stimulation* and *antisocial behavior* are again similar to each other; adding the quadratic trend to the model would not add additional information in predicting the outcome. We conclude that the relationship seems to be approximately linear.

Scatterplot for mother's age and antisocial behavior
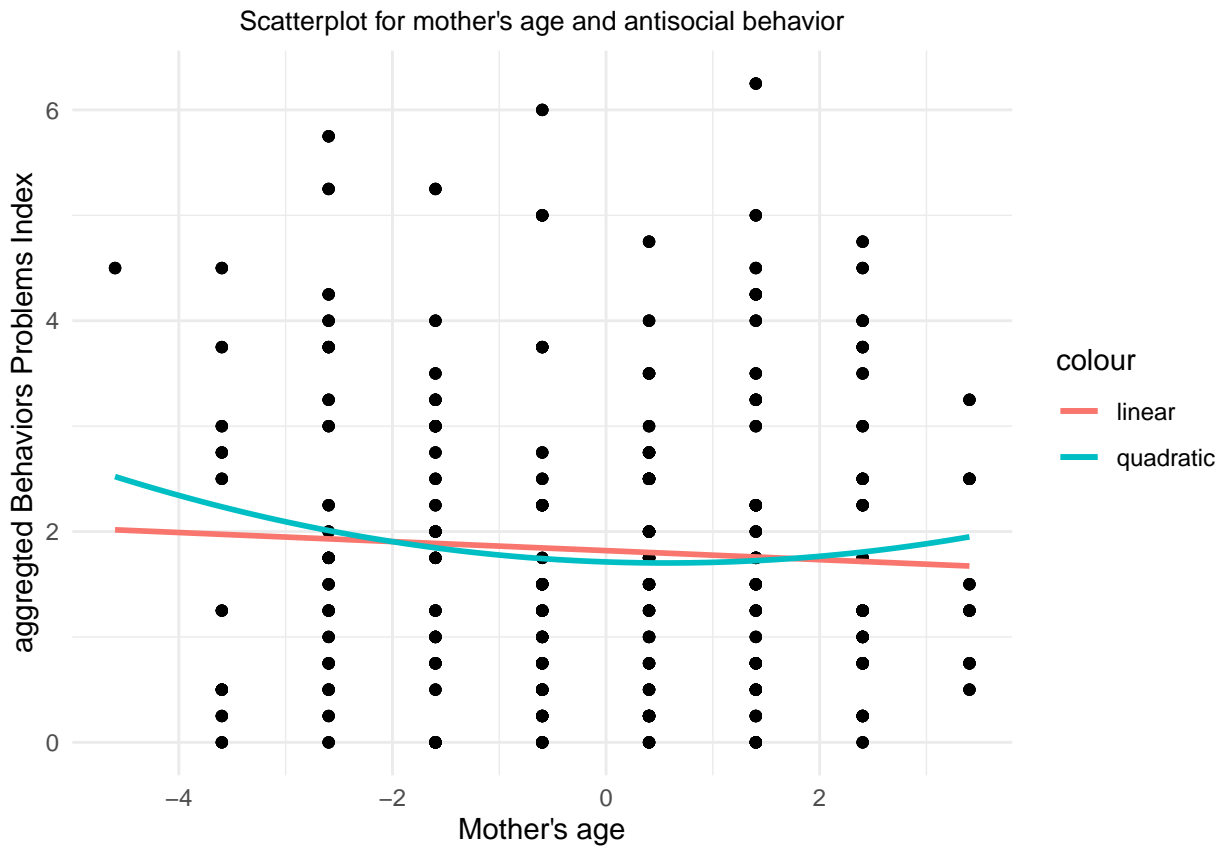
It can be seen from the *Scatterplot for mother's age and antisocial behavior* that the linear and quadratic trend lines for the relationship between *mother's age* and *antisocial behavior* differ and the quadratic trend line seems to fit the data better. The difference is not that big and including the quadratic term would probably not result in much higher explained variance in the outcome. We conclude that the linearity assumptions does not hold here, but the deviation does not seem to be too serious.

Check for outliers, report.

- Level 1: There are a few outliers at the middle two occasions for *time* and *antisocial behavior* and a few in the left middle part and upper middle part in the plot for *reading recognition* and *antisocial behavior*, but none of them seems too extreme. Given the size of the sample and the overall data, we do not expect them to be very influential.

- Level 2: There are a few outliers in the middle upper part of the plot for *cognitive stimulation* and *antisocial behavior*. Moreover, there are a few outliers in the left part (in the left extreme of the *read*'s scale) and in the upper middle part of the plot for *mother's age* and *antisocial behavior*. Again, we do not expect them to have much influence, given the size of the sample and the overall trends in the data.

2. Answer the question: should you perform a multilevel analysis?

What is the mixed model equation?

- Level-1 equation:

$$Anti_{ti} = \pi_{0i} + e_{ti}$$

- Level-2 equation:

$$\pi_{0i} = \beta_{00} + u_{0i}$$

- Mixed Model Equation :

$$Anti_{ti} = \beta_{00} + u_{0i} + e_{ti}$$

Provide and interpret the relevant results.

```
# intercept-only model
m0 <- lm(anti ~ 1, data=dat_long)
summary(m0)


# intercept-only model with the random intercept term
m1 <- lmer(anti ~ 1 + (1|id), REML=F, data=dat_long)
summary(m1)


# comparison of m0 and m1
anova(m1, m0)
```

Table 3

Single-level model and Intercept-only model

Model	M_0: single-level model	M_1: intercept only
Fixed part	Coefficient(SE)	Coefficient(SE)
Intercept (γ_{00})	1.82(.06)	1.82(.10)
Random part		
σ_e^2	3.32	1.74
σ_{u0}^2		1.58
Deviance	3569.5	3337.5
AIC	3573.5	3343.5
Deviance difference		231.97***

Note

***$p < .001$,**$p < .01$

Model 1 fits significantly better than model 0, $\chi^2(1) = 231.97$, $p < .001$. This points to the significant differences in the intercept (average *Behavior Problems Index*) between children ($\sigma_{u0}^2 > 0$).

What is the intraclass correlation?

The formula for ICC is:

$\rho = \frac{\sigma_{u0}^2}{(\sigma_e^2 + \sigma_{u0}^2)} = \frac{1.579}{(1.579 + 1.741)} = 0.48$.

```
ICC <- 1.579/(1.579+1.741)
cat(paste0("ICC = ", round(ICC, 3)))
```

```
## ICC = 0.476
```

The intraclass correlation is 0.48. We see that almost a half of the total unexplained variance represents the individual-level differences, which is quite high.

What is your conclusion regarding the overall question regarding the necessity of performing a multilevel analysis?

We have shown that the data has a multilevel structure:

- $\sigma_{u0}^2 > 0$, $ICC = 0.48$: a large part of variation in the outcome is on the child-level and accounting for this variation in the *antisocial behavior* across children by adding the random intercept term significantly improves the fit of the model;
- the observations are not independent: the correlation between Behavior Problems Index on two occasions for the same child is considerably large, 0.48.

Therefore, we should account for this dependence and variation by adopting a multilevel analysis.

3. Add the time-varying predictor(s).

We will proceed with adding time-varying predictors in two steps: - create the model 2 with the fixed predictor *time*. This model will be used in further analyses as a benchmark to calculate R^2 on level 1 and level 2; - create the model 3 by adding the fixed predictor *read*.

```
# add time as a fixed level 1 predictor
m2 <- lmer(anti ~ 1 + time + (1|id), REML=F, data=dat_long)
summary(m2)

# compare it with the random intercept-only model
anova(m2, m1)

# add read as a fixed level 1 predictor
m3 <- lmer(anti ~ 1 + time + read + (1|id), REML=F, data=dat_long)
summary(m3)

# compare the model with both level 1 predictors with the model with just time
anova(m3, m2)
```

Table 4

Model with time and with time and read

Model	M_2: model with time	M_3: model with read and time
Fixed part	Coefficient(SE)	Coefficient(SE)
Intercept (γ_{00})	1.55(.11)***	1.50(.15)***
Time	0.18(.04)***	0.21(.08)**
Reading recognition		-0.03(.06)
Random part		
σ_e^2	1.69	1.69
σ_{u0}^2	1.59	1.58
Deviance	3317.5	3317.2
AIC	3325.5	3327.2
Deviance difference[ab]	20.06***	0.29

Note

***$p < .001$,**$p < .01$

Provide and interpret the relevant results and provide your overall conclusion.

Model 2 with fixed time

The intercept, $b_0 = 1.55, t(400) = 13.872, p < .001$, can be interpreted as the expected Behavior Problems Index at the first occasion. The coefficient for *time* is significant, $b_{time} = 0.18, t(663) = 4.513, p < .001$. It means that with each one point increase in time, i.e. with each measurement occasion, the expected *antisocial behavior* score increases by 0.18.

Comparing the fixed *time* model with the random intercept-only model, we see that the former fits the data better, $\chi^2(1) = 20.06, p < .001$. AIC also indicates that model 2 ($AIC = 3325.5$) is better than model 1 ($AIC = 3343.5$).

Model 3 with fixed time and reading recognition

Both the intercept ($b_0 = 1.50, t(580) = 9.94, p < .001$) and the coefficient of *time* ($b_{time} = 0.21, t(882) = 2.73, p < .05$) are again significant, but their values are slightly higher. The intercept now represents the Behavior Problems Index that would be expected at the first measurement occasion when the *reading recognition* score at that occasion is average. The coefficient for *time* can be interpreted as follows: with each measurement occasion, the predicted *antisocial behavior* score goes up by .21 when *reading recognition* is held constant. We see that *reading recognition skills* is not significant ($b_{read} = -0.03, t(831) = -0.54, p > .05$).

Comparing the fixed *time* model with the model with both *time* and *read*, we see that adding *read* does not improve the fit of the model, $\chi^2(1) = 0.29, p > .05$. AIC values further confirm this, $AIC_{m2} = 3325.5, AIC_{m3} = 3327.2$.

Overall conclusion

- We observed a significant linear time trend, where Behavior Problems Index increases with each measurement occasion. Adding the *time* variable improves the fit of the model.

- *Reading recognition* does not provide additional information in predicting *antisocial behavior* and does not improve the fit of the model. It will not be included as a fixed predictor in subsequent models.

The better of these two models is the one with only *time*, model 2.

Its equation is:

$Anti_{ti} = \beta_{00} + \beta_{10}time_{ti} + u_{0i} + e_{ti}$

4. On which level or levels can you expect explained variance?

Generally, both level 1 and level 2 variance can be explained by the level 1 (time-varying) predictors. R^2 in the multilevel analysis of longitudinal data is calculated based on the reduction in unexplained variance components on each level compared to the model which includes only *time*. This is due to the underestimation of variance on level 2 and, consequently, overestimation of variance on level 1 in the random intercept-only model, which render it inappropriate as a benchmark model.

However, the only predictor that we added to the baseline model, that is *reading recognition skills*, is not significant, does not improve the model and will not be used in subsequent models as a fixed predictor. Therefore, we expect R^2 to be very low on both levels.

Calculate and interpret the explained variances.

The formula for the explained variance on level-1 is:

$R^2_{Level1} = \frac{\sigma^2_{e(model2)} - \sigma^2_{e(model3)}}{\sigma^2_{e(model2)}} = \frac{1.689 - 1.693}{1.689} = 0.00$

```
(1.689 - 1.693) / 1.689
```

```
## [1] -0.002368265
```

The formula for the explained variance on level-2 is:

$R^2_{Level2} = \frac{\sigma^2_{u0(model2)} - \sigma^2_{u0(model3)}}{\sigma^2_{u0(model2)}} = \frac{1.592 - 1.576}{1.592} = 0.01$

```
(1.592 - 1.576) / 1.592
```

```
## [1] 0.01005025
```

We see that explained variance on both level-1 ($R_1 = -0.002$) and level-2 ($R_2 = 0.01$) is very low, once again confirming that the model 3 does not explain additional variance on any level and *read* should not be included as a fixed predictor in the subsequent models.

We also observe that the explained variance for level 1 is (very close to 0 but) negative.

5. Add the time invariant predictor(s) to the model.

There are two variables at the child level, *mother's age* and *cognitive stimulation*.

```
# add momage and homecog as fixed level 2 predictors
m4a <- lmer(anti ~ 1 + time + momage + homecog + (1|id), REML=F, data=dat_long)
summary(m4a)

# compare the both level 2 predictors model with the best model without them
anova(m4a, m2)

# remove momage from the model and keep only homecog
m4b <- lmer(anti ~ 1 + time + homecog + (1|id), REML=F, data=dat_long)
summary(m4b)

# compare the only homecog model with the model with both level 2 predictors:
anova(m4b, m4a)

# compare the model with homecog and the best model with only level 1 predictors
anova(m4b, m2)
```

Table 5

Model with both level 2 predictors and with only one

Model	M_{4a}: homecog and momage	M_{4b}: homecog
Fixed part	Coefficient(SE)	Coefficient(SE)
Intercept	1.55(.11)***	1.55(.11)***
time	.18(.04)***	.18(.04)***
homecog	-0.13(.04)***	-0.13(.04)***
momage	.00(.05)	
Random part		
σ_e^2	1.69	1.69
σ_{u0}^2	1.49	1.49
Deviance	3305.8	3305.8
AIC	3317.8	3315.8
Deviance difference[a]	11.642**	11.641***

Note
***$p < .001$,**$p < .01$
[a] Here the deviance difference represents the difference compared to model 2

Provide and interpret the relevant results and provide your overall conclusion

Model 4a with momage and homecog

We notice that only the time-invariant variable *momage* is not a statistically significant predictor, $b_{momage} = -0.0009, t(221) = -0.02, p > .05$.

Comparing this model with the best model with only level 1 predictor, we see that the former model fits better, $\chi^2(2) = 11.642, p < .01$. However, since we now added 2 parameters, this significant reduction in the deviance of the model is probably due to the significance of the other level 2 predictor, *homecog*.

Therefore, we created a new model, model 4b, without *momage*, and will interpret the coefficients for this modified model.

Model 4b with homecog

The intercept is significant, $b_0 = 1.55, t(410) = 14.138, p < .001$. It represents the predicted Behavior Problems Index at the first occasion for a child who has an average *cognitive stimulation* score.

Both predictors, that is *time*, $b_{time} = 0.18, t(663) = 4.51, p < .001$, and *cognitive stimulation*, $b_{homecog} = -0.13, t(221) = -3.46, p < .001$, are significant.

We can interpret them as follows:

- with each one point increase in *time* (each measurement occasion), the expected *antisocial behavior* score goes up by 0.18, when *homecog* is kept constant;

- with each one unit increase in *cognitive stimulation*, the expected *anti* score at the first occasion decreases by 0.13.

For the sake of comparison, we can also check if the model is influenced by leaving *momage* out of it. As we see, the model with both *homecog* and *momage* does not fit significantly better than the model with only *homecog*, $\chi^2(1) = 0.00, p > .05$. Moreover, AIC values show that model 4b is better as well, $AIC_{m4a} = 3317.8, AIC_{m4b} = 3315.8$.

Comparing the model 4b, and the best model with only level 1 predictors, model 2, we see that the former fits better, $\chi^2(1) = 11.64, p < .001$. This model also has a lower AIC, $AIC_{m4b} = 3315.8$, than model 2, $AIC_{m2} = 3325.5$.

Overall conclusion

Cognitive stimulation is an important predictor of *antisocial behavior* on the child level and will be kept in the model. It explains a part of the differences between children' *antisocial behavior* at the first occasion. Conversely, *mother's age* does not provide much additional information in predicting these differences and will be eliminated from the model.

6. On which level or levels can you expect explained variance?

Since we have added only a level 2 predictor, which predicts the intercept (the outcome at the first occasion) for each child, we expect explained variance only on this level (that is, we expect it to explain only the variance of the intercept across children).

Calculate and interpret the explained variances.

The formula for the explained variances of level-2 is:

$R^2_{Level2} = \frac{\sigma^2_{u0(model2)} - \sigma^2_{u0(model4b)}}{\sigma^2_{u0(model2)}} = \frac{1.592 - 1.488}{1.592} = 0.07$

```
(1.592 - 1.488) / 1.592
```

```
## [1] 0.06532663
```

The proportion of explained variance on level-2 is 0.07. *Cognitive stimulation* can explain 7% of the variance in individuals' *antisocial behavior* at the first occasion.

7. For the time-varying predictor(s), check if the slope is fixed or random.

We will check if the slope is random for each of the time-varying predictors separately. Firstly, we should inspect the random slope for *time* and modify the model based on the results. Secondly, even though *reading recognition* was not a significant predictor of *antisocial behavior* when treated as fixed across children, it is possible that its slope varies across children and this variation might be non-negligible. Therefore, we will check if the slope of *read* is random as well.

What are the null- and alternative hypotheses?

- For *time*:

$$H_0 : \sigma^2_{u_1} = 0$$
$$H_1 : \sigma^2_{u_1} > 0$$

- For *reading recognition*:

$$H_0 : \sigma^2_{u_2} = 0$$
$$H_1 : \sigma^2_{u_2} > 0$$

Provide and interpret the relevant results.

```
# add the random slope term for the time variable
m5a <- lmer(anti ~ 1 + time + homecog + (1 + time|id), REML=F, data=dat_long)

summary(m5a)

# compare the time random-slope model with the best fixed-slope model
anova(m5a, m4b)

# add the random slope term for the read variable
m5b <- lmer(anti ~ 1 + time + homecog + read + (1 + time + read|id), REML=F, data=dat_long)

summary(m5b)

# compare the model with two random slopes and the model with only one
anova(m5b, m5a)

# compare the model with one random slope with the best model without it
anova(m5b, m4b)
```

Model 5a with random slope for time

The intercept, $b_0 = 1.55, t(219) = 16.26, p < .001$, can be interpreted as the expected outcome at the first occasion for a child with average *homecog* score.

The coefficient for *time* is significant, $b_{time} = 0.18, t(221) = 4.14, p < .001$. It shows that with each occasion, predicted Behavior Problems Index increases by 0.18 when *homecog* is kept constant. However, since we now have a varying slope for *time*, this coefficient is just the expected value across all children.

Lastly, the coefficient for *homecog* is significant as well, $b_{homecog} = -0.10, t(221) = -2.79, p < .01$. With each one point increase in *cognitive stimulation*, the expected *antisocial behavior* score at the first occasion goes down by 0.1.

Correlation between the intercepts and slopes for *time* is 0.41. It shows that children who have a higher Behavior Problems Index at the first occasion also have a higher rate of increase in this index over time.

Comparing this model with the model without the random slope, we see that the former fits the data significantly better, $\chi^2(2) = 26.56$, $p < .001$, and its AIC is lower, $AIC_{m5a} = 3293.3, AIC_{m4b} = 3315.8$. This shows that adding the random slope term and the covariance between the slope and the intercept improves the model.

Table 6

Models with random slopes

Model	M_{5a}: random slope for time	M_{5b}: random slope for read and time
Fixed part	Coefficient(SE)	Coefficient(SE)
Intercept	1.55(.1)***	1.54(.14)***
time	.18(.04)***	.19(.08)*
homecog	-.10(.04)**	-0.10(.04)**
read		-.00(.06)
Random part		
σ_e^2	1.52	1.52
σ_{u0}^2	.95	.92
σ_{u1}^2	.10	.08
σ_{u2}^2		.02
$\rho(u0, u1)$		.51
$\rho(u0, u2)$		.02
$\rho(u1, u2)$		-.02
Deviance	3279.3	3279.0
AIC	3293.3	3301.0
Deviance difference	26.56***	.26

Note
***$p < .001$,** $p < .01$

Model 5b with random slope for time and read

The intercept and coefficients for *time* and *homecog* are still significant, with slightly different values. *Read* is not significant, as expected, since its coefficient is now just the average value across all children, and it has already been shown as not predictive of *antisocial behavior* in model 3.

We see that adding, besides the coefficient for *read*, the random slope term, covariance between slope for *time* and slope for *read*, as well as the covariance between the intercept and slope for *read* to the model, does not reduce its deviance, $\chi^2(4) = 0.2609$, $p > .05$.

Therefore, model 5a is preferred and coefficients for model 5b will not be interpreted.

Provide an overall conclusion.

We can conclude that the relationship between *antisocial behavior* and *time* varies across children and we should account for that by allowing the random slope for *time*. *Read* does not have a varying slope and including it in the model, together with the other parameters, does not improve the fit of the model.

One thing to note here is that the hypotheses cannot be assessed directly with the test we have used. Since there are 2 additional parameters in m5a, σ_{u1}^2 and $cov(u1, uo)$, significant χ^2 test does not necessarily reflect significance of σ_{u1}^2. But even if only the covariance parameter is significant, the variance of the slope is implied.

Similarly, there are 4 additional parameters in m5b and non-significance of χ^2 test does not necessarily reflect non-significance of σ_{u2}^2. However, since the covariance parameters have to be included in the model with varying slopes, the result of this test does indicate that properly allowing the random slope for *read* does not bring significant additions to the model and there is not much information lost by not accounting for it.

8. If there is a random slope, set up a model that predicts the slope variation.

We will inspect if each of the time invariant predictors, *momage* and *homecog*, moderates the relationship between *time* and *anti*. Even though we have shown before that *mother's age* does not predict the differences in *antisocial behavior*

at the first occasion across children, it is possible that it predicts the rate of differences in the outcome between children over time (that is, of the slope of *time*).

Therefore, we will add the interactions between *cognitive stimulation* and *time* and between *mother's age* and *time* to the model and modify it based on the results.

```
# add the predictor of the slope for the time variable
m6a <- lmer(anti ~ 1 + time + homecog + time*homecog + time*momage + (1 + time|id), REML=F, data=dat_long)

summary(m6a) # momage*time is not significant

# compare the model with the two predictors of slope and the model without them
anova(m6a, m5a)

# keep only homecog*time
m6b <- lmer(anti ~ 1 + time + homecog + time*homecog + (1 + time|id), REML=F, data=dat_long)

summary(m6b)

# compare model with one predictor and the model with two predictors
anova(m6b, m6a)

# compare the model with homecog with the model without predictors for slope
anova(m6b, m5a)
```

Table 7

Model with predictors for the slope of time

Model	M_{6a}: with momage and homecog	M_{6b}: only homecog
Fixed part	Coefficient(SE)	Coefficient(SE)
Intercept	1.55(.1)***	1.55(.09)***
time	.18(.04)***	.18(.04)***
homecog	-.06(.04)	-.06(.04)
momage	-.02(.05)	
time x homecog	-.05(.02)**	-.05(.02)**
time x momage	.01(.02)	
Random part		
σ_e^2	1.53	1.53
σ_{u0}^2	.94	.94
σ_{u1}^2	.08	.08
$corr(u0, u1)$	.48	.47
Deviance	3272.1	3272.4
AIC	3292.1	3288.4
Deviance difference[a]	7.18	6.88**

Note
***$p < .001$,** $p < .01$
[a] The difference in deviance is calculated in comparison to model 5a.

Provide and interpret the relevant results and provide your overall conclusion.

Model 6a with time x momage and time x homecog

The intercept, $b_0 = 1.55, t(221) = 16.30, p < .001$, coefficient for *time*, $b_{time} = 0.18, t(221) = 4.21, p < .001$, and *time X homecog*, $b_{time \times homcog} = -0.05, t(221) = -2.69, p < .01$, are significant. As expected, *momage* is not significant, $b_{momage} = -0.02, t(221) = -0.36, p > .05$, but neither is the interaction between *time* and *momage*, $b_{time \times momage} = 0.01, t(221) = 0.52, p > .05$. Therefore, we will remove only this interaction from the model and interpret the coefficients

for the modified model. One thing to note is that *homecog* is now not significant either, $b_{homecog} = -0.06, t(221) = -1.49, p > .05$.

Comparing this model and the model without the predictors of slope for *time*, m5a, we see that the former does not fit significantly better, $\chi^2(3) = 7.18, p > .05$, which is due to non-significant coefficients of *momage* and *time X momage*. Its AIC is, however, slightly lower, $AIC_{m5a} = 3293.3, AIC_{m6a} = 3292.1$

Model 6b with time x homecog

The intercept, $b_0 = 1.55, t(221) = 16.30, p < .001$, coefficient for *time*, $b_{time} = 0.18, t(221) = 4.20, p < .001$, and *time X homecog*, $b_{time \times homecog} = -0.05, t(221) = -2.64, p < .01$, are again significant. We can interpret them as:

- the predicted Behavior Problems Index at the first occasion for a child with average *cognitive stimulation* score is 1.55;

- with each measurement occasion, the predicted *anti* score increases by 0.18 when the *cognitive stimulation* score is average;

- with each one point increase in *cognitive stimulation*, the coefficient for *time* decreases by 0.05. This means that the rate of change in *antisocial behavior* over time is lower (for positive slopes, it would get higher for negative slopes) for children who received more *cognitive stimulation* than for children with lower score on this variable.

Cognitive stimulation is not significant, $b_{homecog} = -0.06, t(221) = -1.49, p > .05$.

Correlation between the slopes of *time* and intercepts is 0.48, slightly lower than in the model without the predictors of the slope for *time*. Again, the interpretation remains the same in the sense that *antisocial behavior* increases more rapidly in children with higher *antisocial behavior* score at the first occasion than in the ones with the lower score at the beginning.

```
(0.09628-0.08397)/0.09628
```

```
## [1] 0.1278563
```

Cognitive stimulation explains 13% of variance in slopes of *time* (that is, rates of change in the outcome over time), $R^2 = 0.13$.

Comparing this model with the model with both interactions (m6a), we see that the latter does not fit the data better $\chi^2(2) = 0.3, p > .05$. This is further confirmed by AIC values, $AIC_{m6b} = 3288.4$, than of the model with both interactions, $AIC_{m6a} = 3292.1$.

Furthermore, we obtained a better-fitting model compared to the model without the slope predictors, m5a, $\chi^2(1) = 6.88, p < .05$.

Overall conclusion

We conclude that adding the cross-level interaction between *cognitive stimulation* and *time* improves the model. *Cognitive stimulation* can explain a part of the variation in the trends in *antisocial behavior* over time across children, but when we account for this, it does not predict the differences across children at the first occasion. Adding *mother's age* as a predictor of slope for *time* does not improve the model and this variable does not moderate the relationship between *time* and *antisocial behavior* nor does it explain the differences in *antisocial behavior* at the first occasion.

Therefore, model 6b is our final model.

9. Decide on a final model.

provide the separate level 1 and 2 model equations, as well as the mixed model equation.

- Equation of level 1:

$$Anti_{ti} = \pi_{0i} + \pi_{1i}time_{ti} + e_{ti}$$

- Equation of level 2:

$$\pi_{0i} = \beta_{00} + \beta_{01} homecog_i + u_{0i}$$
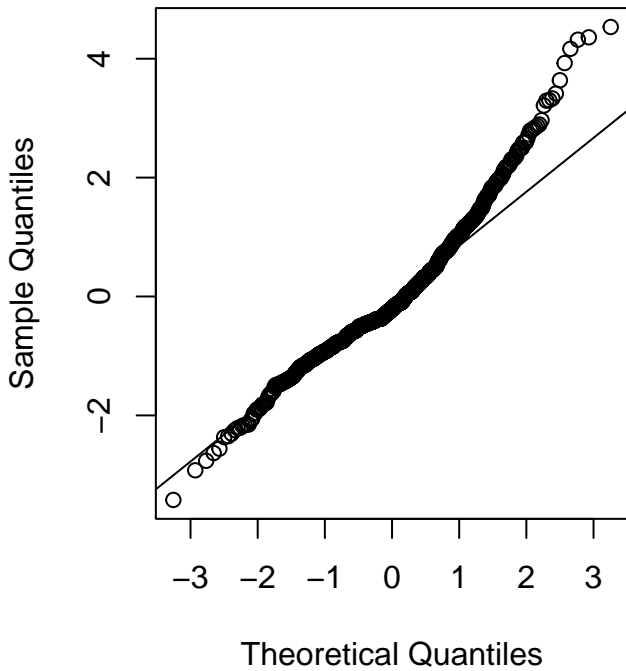
$$\pi_{1i} = \beta_{10} + \beta_{11} homecog_i + u_{1i}$$

Mixed Model Equation :

$$Anti_{ti} = \beta_{00} + \beta_{01} homecog_i + \beta_{10} time_{ti} + \beta_{11} homecog_i * time_{ti} + u_{0i} + u_{1i} time_{ti} + e_{ti}$$
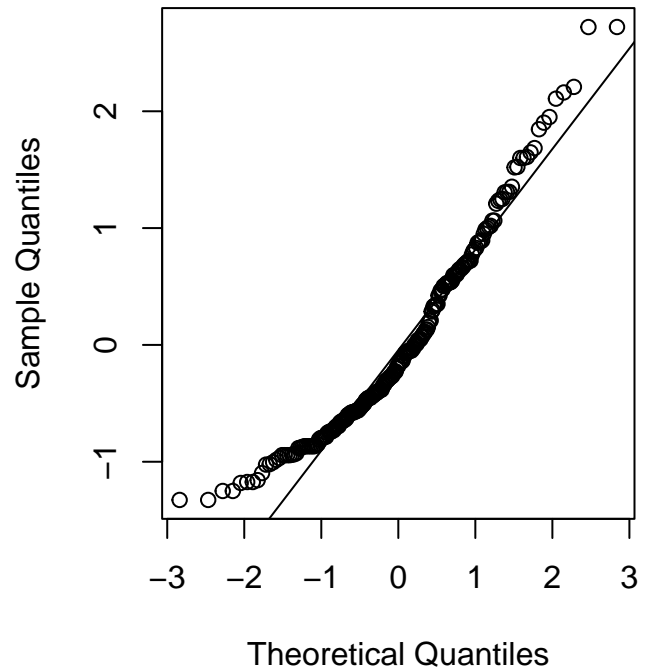
Check the normality assumption for both the level-1 and level-2 errors, report.

- First level residuals are approximately normally distributed with some outliers and minor deviations at the left bottom and upper right corner of the plot.
- The residuals for the intercept and slope of *time* also include some outliers on each of the extremes and the middle part of the plot does not follow the line as well as the residuals on the first level. But we can conclude that the deviations are not too extreme and they are approximately normally distributed.

Q−Q plot for the slope of time