

# GKN - Contactmoment 4

## Multilevel Analyse (Deel 2)

Sven De Maeyer & Bea Mertens

25/11/2021

1 / 41

## Random Intercepts

*Herhaling...*

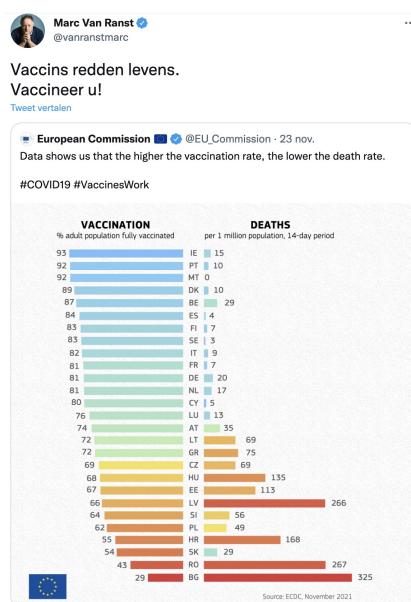
2 / 41

# Multilevel structuur...



3 / 41

## Waarom multilevel?



4 / 41

# Nulmodel

```
Vlaanderen_1_2_3$Leesvaardigheid <- Vlaanderen_1_2_3$ASRREA01  
Vlaanderen_1_2_3$LeesvaardigheidZ <- scale(Vlaanderen_1_2_3$ASRREA01)
```

## Model zonder enige voorspellers. OOK (*Onconditioneel model*)

### Rol van School en Klas bij leesscores in PIRLS?

```
PIRLS_Model0 <- lmer(Leesvaardigheid ~ 1 + (1|IDSCHOOL) + (1|IDCLASS),  
                      data = Vlaanderen_1_2_3,  
                      REML = F)
```

5 / 41

# Nulmodel: resultaten

```
summary(PIRLS_Model0)

Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method [lmerModLmerTest]
Formula: Leesvaardigheid ~ 1 + (1 | IDSCHOOL) + (1 | IDCLASS)
Data: Vlaanderen_1_2_3

AIC      BIC      logLik deviance df.resid
56785.5 56811.7 -28388.7 56777.5     5194

Scaled residuals:
    Min      1Q  Median      3Q      Max 
-3.4014 -0.6589  0.0139  0.7004  3.7464

Random effects:
 Groups   Name        Variance Std.Dev. 
IDCLASS (Intercept) 166.1    12.89  
IDSCHOOL (Intercept) 490.4    22.14  
Residual            3016.6   54.92  
Number of obs: 5198, groups: IDCLASS, 277; IDSCHOOL, 148

Fixed effects:
            Estimate Std. Error    df t value Pr(>|t|)    
(Intercept) 523.846    2.175 143.213  240.8 <2e-16 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Hoe groot is de impact van klassen/scholen?

6 / 41

# Nulmodel: ICC

Effectgrootte, uitgedrukt in %

```
# ICC klasniveau  
166.1/(166.1 + 490.4 + 3016.6)
```

```
[1] 0.04522066
```

```
# ICC schoolniveau  
490.4/(166.1 + 490.4 + 3016.6)
```

```
[1] 0.1335112
```

7 / 41

## Random intercept model

**Model met voorspellers. OF Conditioneel model**

Rol van **School** en **Klas** bij leesscores in PIRLS,

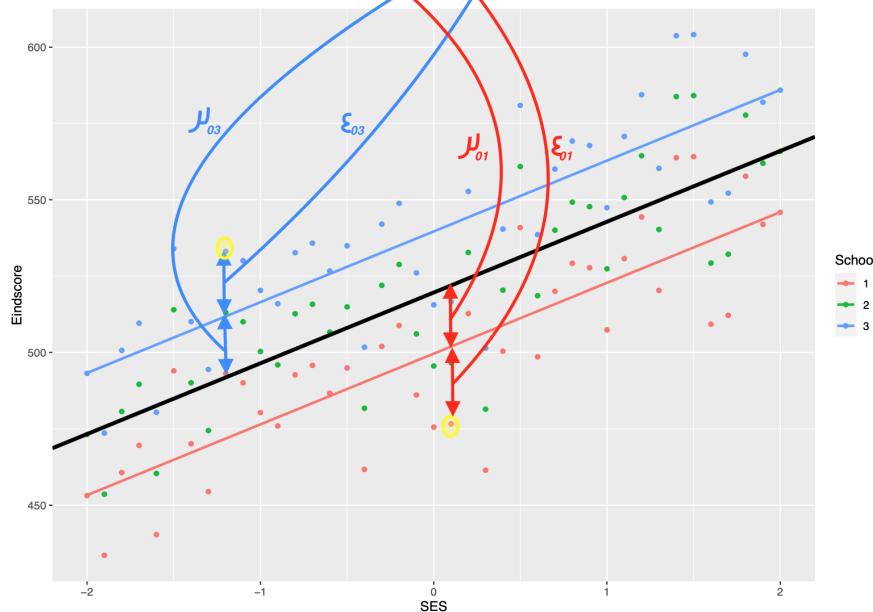
na controle voor **Ouders\_GraagLezenZ** en **Gender**?

```
PIRLS_ModelRI <- lmer(  
  Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ +  
    (1|IDSCHOOL) + (1|IDCLASS),  
  data = Vlaanderen_1_2_3, REML = F)
```

8 / 41

# Random intercept model: formule

$$Eindscore_{ij} = \beta_0 + \beta_1 * SES_{ij} + (\mu_{0j} + \epsilon_{ij})$$



9 / 41

# Random intercept model: resultaten

```
summary(PIRLS_ModelRI)

Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method [lmerModLmerTest]
Formula: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 | IDSCHOOL) +
(1 | IDCLASS)
Data: Vlaanderen_1_2_3

      AIC      BIC      logLik deviance df.resid
50414.6 50453.3 -25201.3 50402.6     4629

Scaled residuals:
    Min      1Q  Median      3Q      Max 
-3.5382 -0.6457  0.0261  0.6831  3.6867 

Random effects:
 Groups   Name        Variance Std.Dev.
IDCLASS (Intercept) 151.2    12.30
IDSCHOOL (Intercept) 420.8    20.51
Residual            2871.3   53.58
Number of obs: 4635, groups: IDCLASS, 275; IDSCHOOL, 147

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)    
(Intercept) 520.8727   2.2258 186.6044 234.011 < 2e-16 ***
GenderGirls 10.1501    1.6083 4492.7054   6.311 3.04e-10 ***
Ouders_GraagLezenZ 11.4031    0.8162 4546.6389  13.972 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
          (Intr) GndrGr 
GenderGirls -0.372
```

10 / 41

# Random intercept model: resultaten

```
# ICC klasniveau  
151.2/(151.2 + 420.8 + 2871.3)
```

```
[1] 0.04391136
```

```
# ICC schoolniveau  
420.8/(151.2 + 420.8 + 2871.3)
```

```
[1] 0.1222083
```

11 / 41

## Random intercept model: resultaten (via tab\_model)

```
library(sjPlot)  
tab_model(PIRLS_ModelRI,  
          show.se = TRUE,  
          show.ci = FALSE,  
          show.icc = FALSE,  
          show.aic = FALSE,  
          show.r2 = FALSE)
```

Leesvaardigheid			
Predictors	Estimates	std. Error	p
(Intercept)	520.87	2.23	<0.001
Gender: Girls	10.15	1.61	<0.001
Ouders Graag Lezen Z	11.40	0.82	<0.001

Random Effects			
$\sigma^2$	2871.34		
$\tau_{00}$ IDCLASS	151.19		
$\tau_{00}$ IDSCHOOL	420.81		
N IDSCHOOL	147		
N IDCLASS	275		
Observations	4635		

12 / 41

# Random intercept model: significantie random effecten?

Via de functie `confint` uit pakket `lme4` kan je betrouwbaarheidsintervallen voor de random effecten opvragen

```
# Eerst via VarCorr( ) enkel de varianties uit de output opnieuw opvragen...
VarCorr(PIRLS_ModelRI)
```

```
Groups      Name      Std.Dev.
IDCLASS    (Intercept) 12.296
IDSCHOOL   (Intercept) 20.514
Residual           53.585
```

```
# Vervolgens via confint( ) de CI's berekenen
ModelRI_CI <- confint(PIRLS_ModelRI, level = .95)
ModelRI_CI
```

	2.5 %	97.5 %
.sig01	9.045146	15.72210
.sig02	16.949180	24.39561
.sigma	52.479054	54.73002
(Intercept)	516.484187	525.25623
GenderGirls	6.997138	13.30317
Ouders_GraagLezenZ	9.800949	13.00573

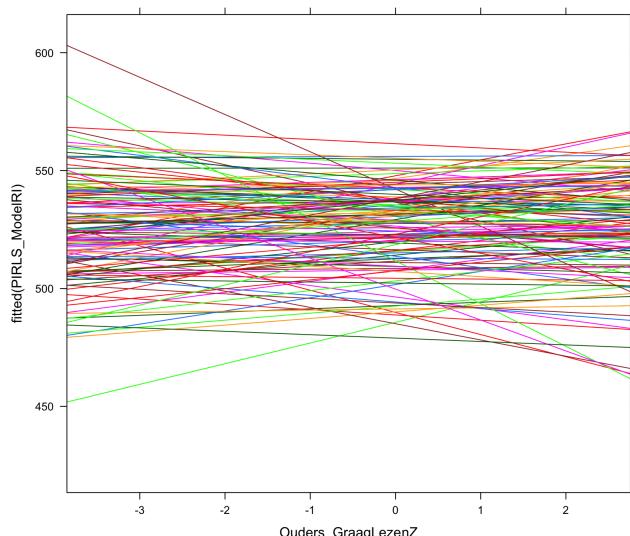
13 / 41

## Random intercept model: visualisatie (1)

In OLP staat functie `xyplot( )` uit het pakket `lattice...`

Geeft hier een raar resultaat. Hoe komt dat?

```
library(lattice)
xyplot(fitted(PIRLS_ModelRI) ~ Ouders_GraagLezenZ,
       group=IDSCHOOL,
       data=Vlaanderen_1_2_3,
       type='r')
```



14 / 41

## Random intercept model: visualisatie (2)

`xypplot()` uit `lattice` wordt gebruikt om "fitted values" (dus voorspellingen op basis van heel ons model) te plotten

In ons model zit:

- **Ouders\_GraagLezenZ** en **IDSCHOOL** (wat we willen plotten)
- maar ook **Gender** en **IDCLASS**

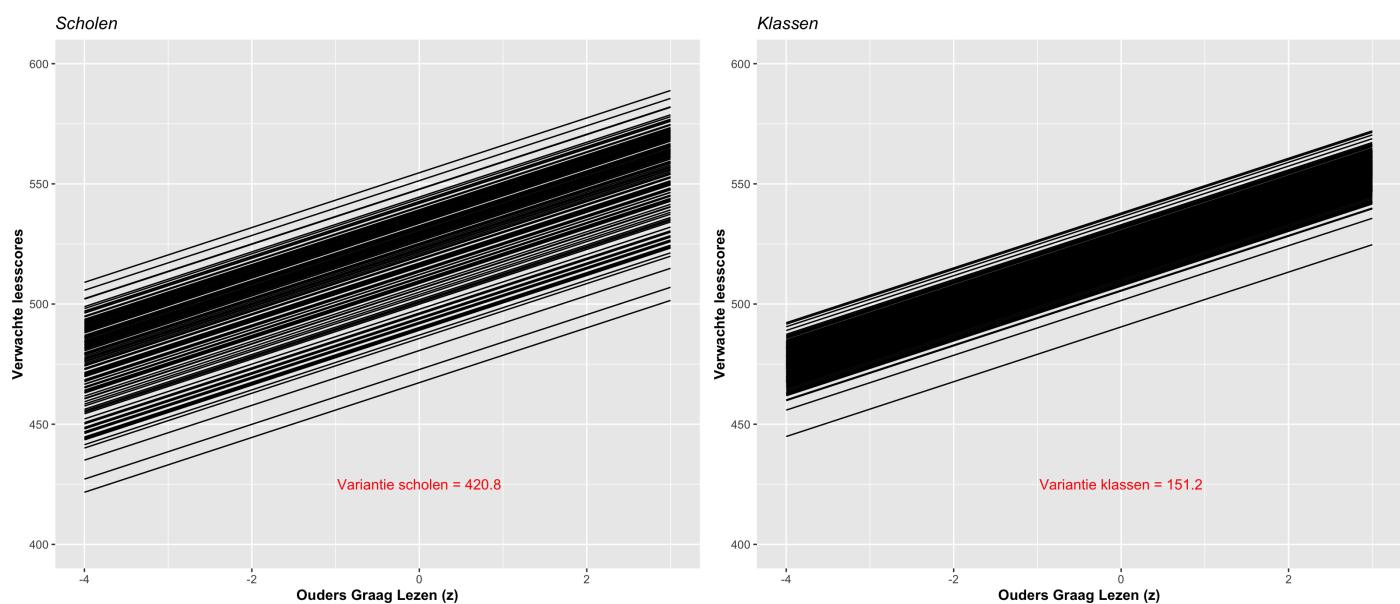
Vandaar dat de lijntjes niet mooi parallel lopen...

Dus werkt alleen bij model met 1 voorspeller en enkel 2 levels (bv. kinderen genest in scholen)

15 / 41

## Random intercept model: visualisatie (3)

We kunnen het wel plotten... maar met de nodige complexiteit



16 / 41

# Random intercept model: visualisatie (4)

## Vb. effect van scholen tonen

```
# Pakket ggeffects nodig om voorspelde scores te berekenen
library(ggeffects)

# Random effecten (school en klas) moeten gebaseerd zijn op categorische variabele.
# Dus hier eerst nieuwe gecategoriseerde variant maken
Vlaanderen_1_2_3$School <- as.factor(Vlaanderen_1_2_3$IDSCHOOL)
Vlaanderen_1_2_3$Class <- as.factor(Vlaanderen_1_2_3$IDCLASS)

# Model opnieuw schatten met gecategoriseerde variant
PIRLS_ModelRI <- lmer(Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1|School) + (1|Class),
                        data = Vlaanderen_1_2_3, REML = F)

# Voorspelde scores op niveau van school vragen met als voorspeller Ouders_GraagLezenZ
# en wegschrijven in een nieuw object met naam Pred
Pred <- ggpredict(PIRLS_Model2, terms = c("Ouders_GraagLezenZ", "School"), type = "re")
```

17 / 41

# Random intercept model: visualisatie (4)

## Vb. effect van scholen tonen (vervolg)

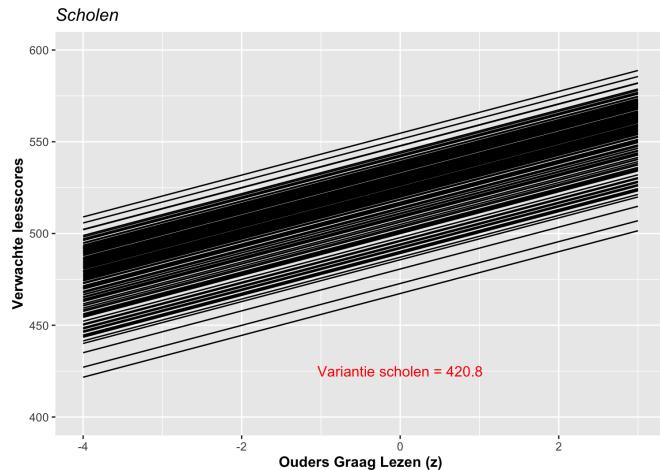
```
# Pakket ggplot2 nodig om de grafiek te maken
library(ggplot2)

# Basis leggen voor de grafiek (gebaseerd op object Pred dat we hierboven maakten);
# aes() om de x-as, y-as en groepvariabele te maken;
# geom_line() om aan te geven dat we lijntjes willen;
# scale_y_continuous() om de range van y-as te bepalen;
# theme() om de titels te lay-outen
P1 <- ggplot(Pred, aes(x=predicted, group=group)) +
  geom_line() +
  scale_y_continuous(limits = c(400, 600)) +
  theme(axis.title = element_text(face="bold"), plot.title = element_text(face="italic"))

# P1 'printen' met titel 'Scholen' en aangepaste x- en y-as labels
P1 + ggtitle("Scholen") + labs(y="Verwachte leesscores", x = "Ouders Graag Lezen (z)")
```

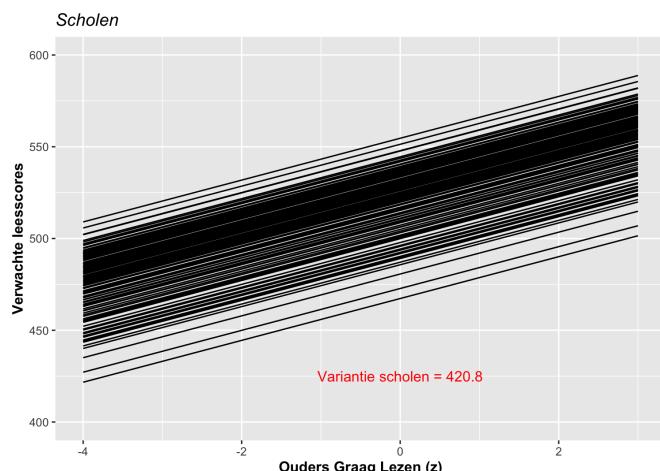
18 / 41

# Wat leren we conceptueel van dit model?



19 / 41

# Wat leren we conceptueel van dit model?



Zijn sommige scholen beter dan andere in het wegwerken van verschillen op basis van achtergrond (Ouders\_GraagLezenZ in dit geval)?

= vraag naar **differentiële effectiviteit**

20 / 41

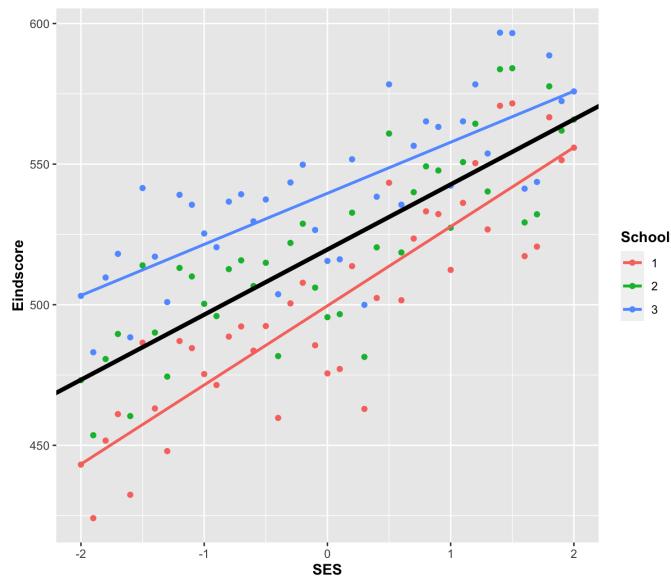
# Random slopes

*Spaghetti time ...*

21 / 41

## Basisidee (1)

$$Eindscore_{ij} = \beta_0 + \beta_1 * SES_{ij} + (\mu_{0j} + \mu_{1j} * SES_{ij} + \epsilon_{ij})$$



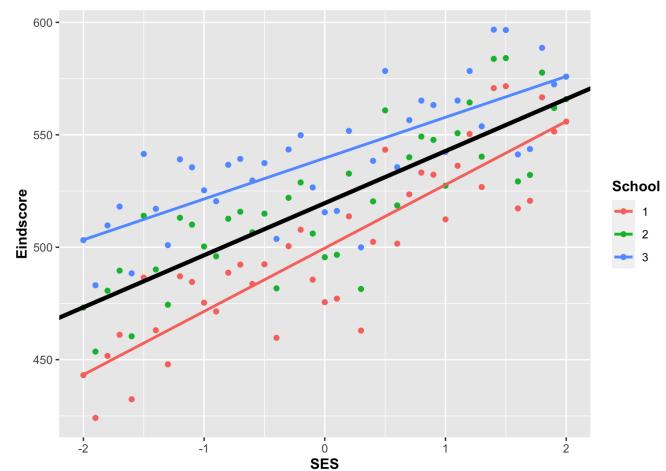
22 / 41

## Basisidee (2)

3 scholen, dus 3 verschillende intercepten en slopes:

$$Eindscore_{ij} = \beta_0 + \beta_1 * SES_{ij} + (\mu_{0j} + \mu_{1j} * SES_{ij} + \epsilon_{ij})$$

	Intercept	$\mu_{0j}$	Slope	$\mu_{1j}$
School 1	500.00	-19.62	30	5
School 2	519.62	0.00	25	0
School 3	539.24	19.62	20	-5



23 / 41

## PIRLS voorbeeld 1

Effect van **Ouders\_GraagLezenZ** in elke school even sterk?

```
PIRLS_ModelRS_lezen <- lmer(Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ +
  (1 + Ouders_GraagLezenZ|IDSCHOOL) +
  (1|IDCLASS),
  data = Vlaanderen_1_2_3, REML = F)
```

24 / 41

# PIRLS voorbeeld 1: resultaten

```
summary(PIRLS_ModelRS_lezen)

Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method [lmerModLmerTest]
Formula:
Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Ouders_GraagLezenZ |
  IDSCCHOOL) + (1 | IDCLASS)
Data: Vlaanderen_1_2_3

AIC      BIC      logLik deviance df.resid
50416.6 50468.2 -25200.3 50400.6     4627

Scaled residuals:
    Min      1Q  Median      3Q      Max 
-3.5351 -0.6483  0.0292  0.6893  3.6977 

Random effects:
Groups   Name        Variance Std.Dev. Corr
IDCLASS (Intercept) 148.018 12.17
IDSCCHOOL (Intercept) 422.989 20.57
          Ouders_GraagLezenZ 1.989  1.41   -1.00
Residual             2870.164 53.57
Number of obs: 4635, groups: IDCLASS, 275; IDSCCHOOL, 147

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)    
(Intercept) 520.9351   2.2253 186.4369 234.092 < 2e-16 ***
GenderGirls 10.1115    1.6083 4492.9865   6.287 3.55e-10 ***
Ouders_GraagLezenZ 11.5123   0.8247 1472.4862  13.959 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
```

25 / 41

# PIRLS voorbeeld 1: resultaten

Leesvaardigheid			
Predictors	Estimates	std. Error	p
(Intercept)	520.94	2.23	<0.001
Gender: Girls	10.11	1.61	<0.001
Ouders Graag Lezen Z	11.51	0.82	<0.001
Random Effects			
$\sigma^2$	2870.16		
$\tau_{00}$ IDCLASS	148.02		
$\tau_{00}$ IDSCCHOOL	422.99		
$\tau_{11}$ IDSCCHOOL.Ouders_GraagLezenZ	1.99		
$\tau_{01}$ IDSCCHOOL	-1.00		
N IDSCCHOOL	147		
N IDCLASS	275		
Observations	4635		

26 / 41

# PIRLS voorbeeld 1: modelvergelijking

Welk model is best?

```
anova(PIRLS_ModelRI, PIRLS_ModelRS_lezen, test="Chi")  
  
Data: Vlaanderen_1_2_3  
Models:  
PIRLS_ModelRI: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 | IDSCHOOL) + (1 | IDCLASS)  
PIRLS_ModelRS_lezen: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Ouders_GraagLezenZ | IDSCHOOL) + (1 | IDCLASS)  
npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)  
PIRLS_ModelRI      6 50415 50453 -25201    50403  
PIRLS_ModelRS_lezen 8 50417 50468 -25200    50401 1.986  2     0.3705
```

27 / 41

# PIRLS voorbeeld 2

Onderzoeksvraag: Is het verschil tussen jongens en meisjes in elke school even sterk?

```
PIRLS_ModelRS_gender <- lmer(Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ +  
                           (1 + Gender|IDSCHOOL) + (1|IDCLASS),  
                           data = Vlaanderen_1_2_3, REML = F)
```

28 / 41

## PIRLS voorbeeld 2: modelvergelijking (1)

Welk model is best, met of zonder random slopes voor gender?

```
anova(PIRLS_ModelRI, PIRLS_ModelRS_gender, test="Chi")  
  
Data: Vlaanderen_1_2_3  
Models:  
PIRLS_ModelRI: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 | IDSCHOOL) + (1 | IDCLASS)  
PIRLS_ModelRS_gender: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Gender | IDSCHOOL) + (1 | IDCLASS)  
npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)  
PIRLS_ModelRI      6 50415 50453 -25201     50403  
PIRLS_ModelRS_gender 8 50409 50461 -25197     50393 9.3953  2   0.009117 **  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

29 / 41

## PIRLS voorbeeld 2: modelvergelijking (2)

Welk model is best, met random slopes voor gender of Ouders\_GraagLezenZ?

```
anova(PIRLS_ModelRS_lezen, PIRLS_ModelRS_gender, test="Chi")  
  
Data: Vlaanderen_1_2_3  
Models:  
PIRLS_ModelRS_lezen: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Ouders_GraagLezenZ | IDSCHOOL) + (1 | IDCLASS)  
PIRLS_ModelRS_gender: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Gender | IDSCHOOL) + (1 | IDCLASS)  
npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)  
PIRLS_ModelRS_lezen  8 50417 50468 -25200     50401  
PIRLS_ModelRS_gender 8 50409 50461 -25197     50393 7.4094  0
```

30 / 41

## PIRLS voorbeeld 2: resultaten

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method [lmerModLmerTest]
Formula: Leesvaardigheid ~ 1 + Gender + Ouders_GraagLezenZ + (1 + Gender |
   IDSCCHOOL) + (1 | IDCLASS)
Data: Vlaanderen_1_2_3

AIC      BIC      logLik deviance df.resid
50409.2 50460.8 -25196.6 50393.2     4627

Scaled residuals:
    Min      1Q  Median      3Q     Max 
-3.5110 -0.6452  0.0260  0.6758  3.6793 

Random effects:
Groups      Name        Variance Std.Dev. Corr
IDCLASS (Intercept) 152.9    12.37
IDSCCHOOL (Intercept) 526.1    22.94
             GenderGirls 120.1    10.96  -0.54
Residual           2841.3   53.30

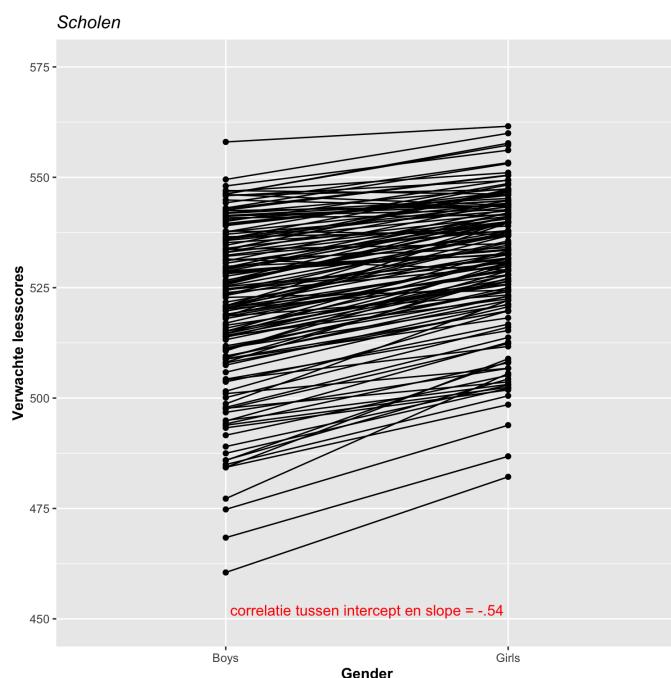
Number of obs: 4635, groups: IDCLASS, 275; IDSCCHOOL, 147

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)    
(Intercept) 520.715    2.398 139.471 217.126 < 2e-16 ***
GenderGirls  10.403    1.883 117.373  5.524 2.03e-07 ***
Ouders_GraagLezenZ 11.476    0.815 4539.061 14.082 < 2e-16 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
  (Intr) GndrGr
GenderGirls -0.522
Odrs_GrgLzz  0.010 -0.013
```

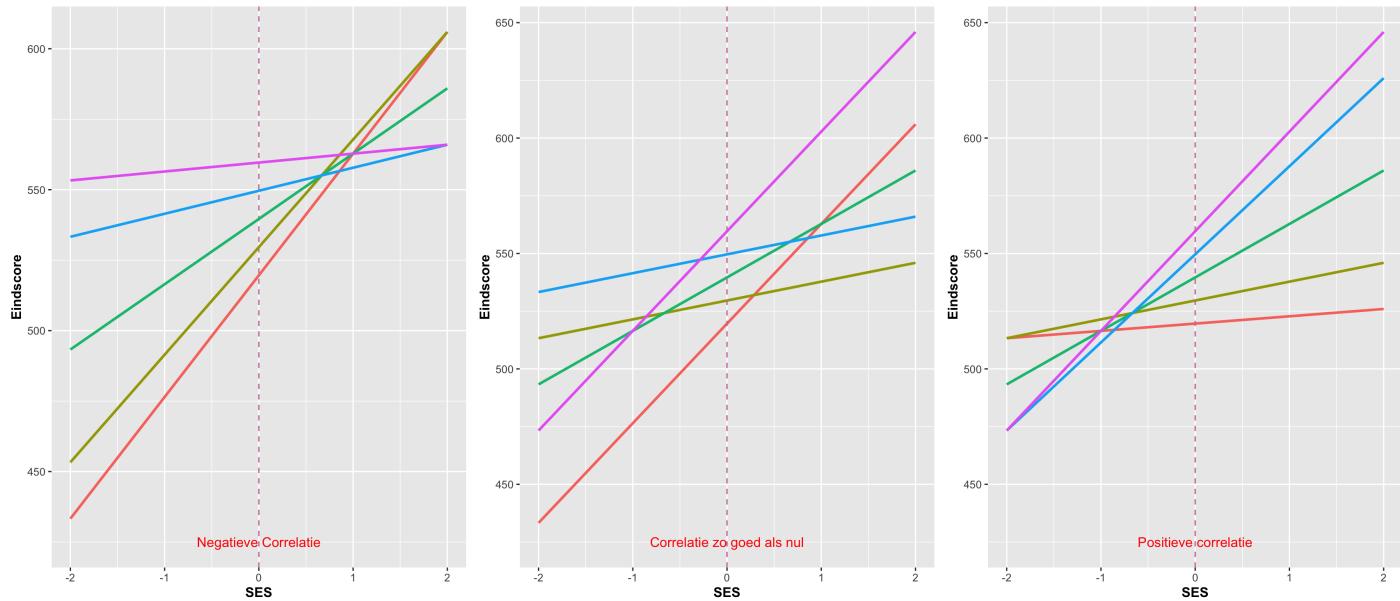
31 / 41

## PIRLS voorbeeld 2: visualisatie



32 / 41

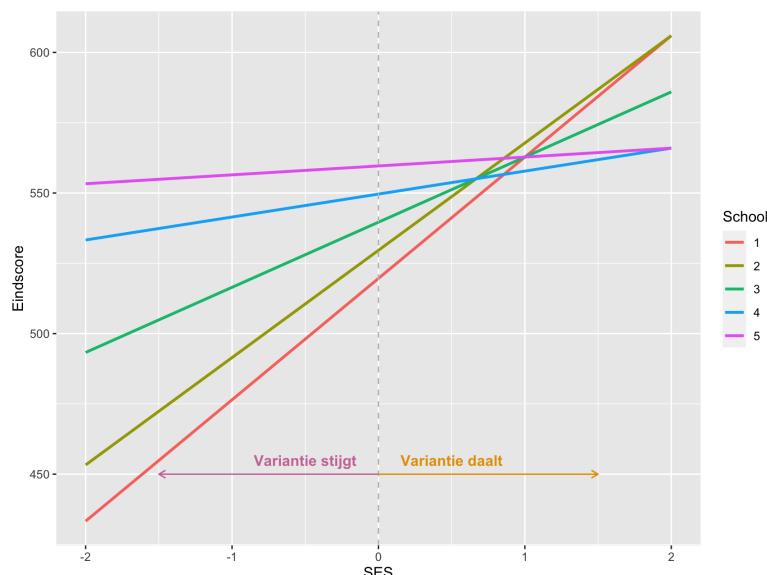
# Betekenis correlatie intercept & slope



33 / 41

## Let op met interpretatie intercept varianties

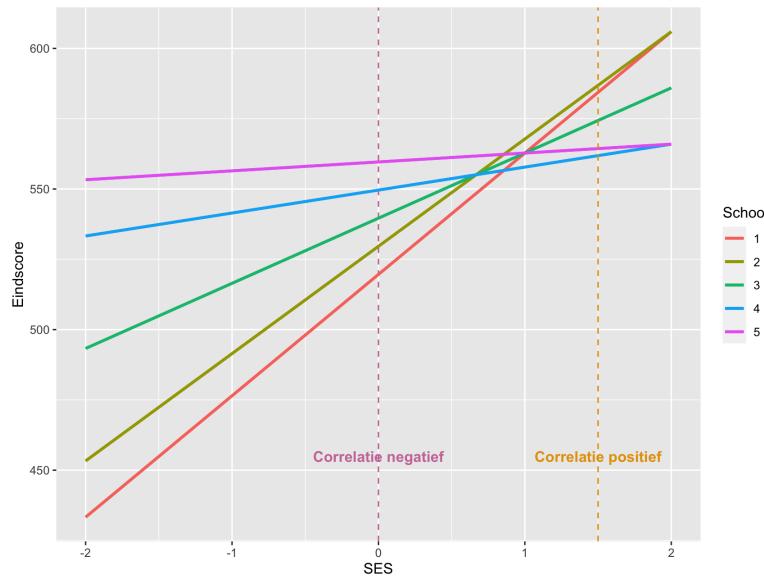
Intercept variantie afhankelijk schaal SES! (cfr. waar is de nul...?)



34 / 41

# Let op met interpretatie correlatie tss'en intercept en slope

Covariantie (dus ook correlatie) tussen Intercept en Slope afhankelijk schaal SES!



35 / 41

ICC?!

Formule bij random intercept model:

$$ICC = \frac{\sigma_{\mu_{0j}}^2}{\sigma_{\mu_{0j}}^2 + \sigma_{\epsilon_{0ij}}^2} = \frac{var_{scholen}}{var_{scholen} + var_{ln}}$$

Maar wat is de totale variantie in een random slope model? 🤔

Hangt af van...

36 / 41

## Enkele praktische richtlijnen (1)

- Denk goed na over de schaal van de variabelen
  - onafhankelijke variabelen evt centreren?
  - eventueel z-scores maken van de onafhankelijke variabelen?
-  Beperkte variantie in een random intercepts model op een bep. niveau is **GEEN** reden om geen random slopes te schatten indien dit inhoudelijk toch relevant zou zijn
- Tracht spaarzaam te zijn! Dus geen overload aan random slopes in een model...

37 / 41

## Enkele praktische richtlijnen (2)

- Vertaal onderzoeksvragen in (reeks) van modellen
- Hanteer modelvergelijkingen
- Rapporteer ALLEEN parameterschattingen van het best passend model
- Maak gebruik van tabellen en visualisaties

38 / 41

# Rapportage: toepassing

*Is het effect van leesplezier op leesvaardigheid afhankelijk van de school waarin een leerling les volgt, na controle voor geslacht en betrokkenheid?*

```
# Alle kwantitatieve variabelen standaardiseren
Vlaanderen_1_2_3$LeesplezierZ <- scale(Vlaanderen_1_2_3$ASBGERL)
Vlaanderen_1_2_3$BetrokkenheidZ <- scale(Vlaanderen_1_2_3$ASBHPLR)

# Model schatten
PIRLS_Toepassing <- lmer(LeesvaardigheidZ ~ 1 + Gender + LeesplezierZ + BetrokkenheidZ +
                           (1 + LeesplezierZ|IDSCHOOL) + (1|IDCLASS),
                           data = Vlaanderen_1_2_3, REML = F)

# Nulmodel met afhankelijke ook als z-score
PIRLS_M0z <- lmer(LeesvaardigheidZ ~ 1 +
                     (1 |IDSCHOOL) + (1|IDCLASS),
                     data = Vlaanderen_1_2_3, REML = F)
```

39 / 41

## Relevante informatie? Wat in tabel? Visualisaties?

```
tab_model(
  PIRLS_M0z,
  PIRLS_Toepassing,
  show.ci = F,
  show.aic = T,
  show.icc = F,
  show.se = T,
  show.r2 = F,
  show.ngroups = F)
```

Predictors	Leesvaardigheid Z		Leesvaardigheid Z	
	Estimates	std. Error	Estimates	std. Error
(Intercept)	-0.01	0.04	-0.05	0.04
Gender: Girls			0.16 ***	0.03
Leesplezier Z			0.04 *	0.02
Betrokkenheid Z			0.19 ***	0.01
Random Effects				
$\sigma^2$	0.83		0.78	
$\tau_{00}$	0.05	IDCLASS	0.04	IDCLASS
	0.13	IDSCHOOL	0.12	IDSCHOOL
$\tau_{11}$			0.00	IDSCHOOL.LeesplezierZ
$Q_{01}$			-0.30	IDSCHOOL
Observations	5198		4619	
AIC	14149.725		12337.919	

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

40 / 41

# Time to pRactice!

**Oefeningen en respons terug te vinden op BB**