

# Recomendaciones y herramientas para el uso de modelos lineales con efectos mixtos

*Federico Alvarez*

*26 de junio de 2019*

## 1 Introducción

Los modelos lineales con efectos mixtos se usan de manera bastante extendida en varios campos científicos. En el caso de las ciencias cognitivas y, fundamentalmente, en el caso de la psicolingüística, su uso se ha ido volviendo más y más extendido, dado que permiten dar cuenta de factores que afectan consistentemente a los estudios del campo, como es el hecho de que en varios diseños experimentales se toman múltiples muestras por sujeto o por ítem experimental, lo cual viola el supuesto de independencia necesario para la aplicación de métodos que no tomen en cuenta efectos aleatorios.

Tomando esto en cuenta, en esta reunión quiero señalar cómo abordar dos cuestiones que dificultan la implementación de estos modelos: cómo determinar la estructura de efectos aleatorios y cómo obtener valores de  $p$ .

## 2 Consideraciones sobre la estructura de efectos aleatorios

### 2.1 Descripción

En Barr (2013) se conduce una serie de simulaciones dirigidas a determinar qué procedimientos son más apropiados en términos de conservatividad y de sensibilidad a la hora de trabajar con modelos de efectos mixtos. Las diferentes condiciones simuladas tienen que ver con estructuras de efectos aleatorios de complejidad decreciente, y también consideraron diferentes procedimientos de selección de modelos.

En el estudio consideran siete modelos: 1) Sin efectos aleatorios, 2) con intercepto aleatorio por sujeto, 3) con intercepto y pendiente aleatoria por sujeto, 4) con intercepto y pendiente aleatorias por sujeto e intercepto aleatorio por ítem, 5) con interceptos aleatorios por sujeto y por ítem, 6) con intercepto y pendiente aleatorios sin correlación por sujeto y con intercepto aleatorio por ítem, y 7) con pendiente aleatoria sin correlaciones por sujeto y con intercepto aleatorio por ítem.

Si considero una variable dependiente  $Y$ , una variable independientes  $X$  y variables de agrupamiento sujeto e ítem, las fórmulas de los modelos se escribirían así:

1. Sin efectos aleatorios (no es un modelo mixto)

$$Y \sim X$$

2. Con intercepto aleatorio por sujeto

$$Y \sim X + (1|\text{sujeto})$$

3. Con intercepto y pendiente aleatorios por sujeto

$$Y \sim X + (1 + X|\text{sujeto})$$

4. Con intercepto y pendiente aleatorios por sujeto e intercepto aleatorio por ítem

$$Y \sim X + (1 + X|\text{sujeto}) + (1|\text{ítem})$$

5. Con interceptos aleatorios por sujeto y por ítem

$$Y \sim X + (1|\text{sujeto}) + (1|\text{ítem})$$

6. Con pendiente e intercepto aleatorios no correlacionados por sujeto e intercepto aleatorio por ítem

$$Y \sim X + (1|\text{sujeto}) + (0 + X|\text{sujeto}) + (1|\text{ítem})$$

7. Con pendiente aleatoria sin correlaciones por sujeto y con intercepto aleatorio por ítem.

$$Y \sim X + (0 + X|\text{sujeto}) + (1|\text{ítem})$$

Las simulaciones se hicieron sobre datasets de 24 sujetos con 12 y 24 observaciones por sujeto. La variable dependiente era continua y era una muestra aleatoria correspondiente a la ausencia de efecto o a un efecto de .8, en tanto que la variable dependiente era un factor de dos niveles.

Los resultados que obtienen apuntan a que los modelos máximos (4) y los modelos cuasi-máximos (6 y 7) son los menos anticonservadores y con mejor sensibilidad, en tanto que los peores son los modelos que sólo incluyen interceptos aleatorios (5). Proponen casos intermedios que abarcan modelos seleccionados por procedimientos hacia adelante o hacia atrás (también toman en consideración anovas con efectos aleatorios por sujeto o por ítem y su combinación). Los autores consideran que para usar modelos con efectos mixtos es importante que la estructura de efectos mixtos se adecúe al diseño del estudio y que, al momento de publicación del artículo, los investigadores se han preocupado más por no sobreparametrizar los modelos en función de los datos que por no subparametrizarlos en función del diseño.

## 2.2 ¿Cómo afecta esto al modo en el que usamos los modelos mixtos?

Parte de la discusión hace foco en cómo identificar la estructura máxima de efectos aleatorios. Según los autores, la necesidad de usar pendientes o interceptos aleatorios depende fuertemente de la unidad de muestreo (sujetos o ítems). Cuando un factor aleatorio se observa una única vez por unidad (por ejemplo, cuando cada sujeto ve un ítem experimental una única vez), basta con considerar un intercepto aleatorio para ese factor. En cambio, si un factor aleatorio se repite múltiples veces en la unidad de muestreo (o sea, cuando hay medidas repetidas de cada sujeto para cada condición), es necesario considerar pendientes aleatorias (hay salvedades para casos muy desbalanceados).

Otra cuestión de relevancia abordada en el artículo tiene que ver con los fallos de convergencia en el modelo. Los autores dicen que las chances de que un modelo converja van a depender del tamaño de los efectos aleatorios y de si la cantidad de observaciones es suficiente para estimar los efectos aleatorios. En general, mientras mayor sea la muestra, más chances habrá de que el modelo converja. Sin embargo, no siempre es posible tomar un experimento a un número alto de sujetos, o emplear un alto número de ítems por experimento. Además, frecuentemente es más difícil que un modelo converja cuando la variable dependiente es categórica que si es numérica (continua). Es importante no caer en el uso de modelos que sólo tengan interceptos aleatorios. Cuando un modelo no converge, el primer paso siempre debe ser la inspección de los datos, a ver si hay algún error que dificulta el análisis. También recomiendan la eliminación de outliers y el centrado y recodificación de los predictores. Otro consejo es aumentar la cantidad de iteraciones en el procedimiento de estimación. Finalmente, también recomiendan considerar si un imbalance en los niveles de los factores aleatorios puede ser lo que cause el problema de convergencia: en un caso en el que hay uno o varios ítems o sujetos de los cuales se tiene una cantidad reducida de observaciones, puede ser buena idea eliminarlos y conservar la estructura máxima. Si aún así se vuelve necesario simplificar el modelo, lo que se recomienda es eliminar las correlaciones aleatorias, o incluso los interceptos aleatorios, pero no las pendientes aleatorias, dado que los modelos sin correlaciones (ni interceptos) se desempeñan de manera razonablemente similar a los modelos máximos para distintos tamaños de muestras y de efectos. Un caso de mayor complejidad es el que se da cuando hay una cantidad elevada de efectos fijos. Según los autores aún falta investigación sobre las estrategias a seguir en esta situación, pero una heurística recomendada es que para cualquier efecto fijo de interés, es necesario considerar los efectos aleatorios. Si un análisis con varios efectos fijos no converge, se pueden realizar múltiples análisis que incluyan todos los efectos fijos, pero los efectos aleatorios completos de uno sólo. Si todo esto falla, se puede proceder a partir de un modelo mínimo al cual se le van adicionando efectos aleatorios (en el paper discuten un algoritmo para realizar una selección hacia adelante óptima, no temino de cazarle la onda). Una observación final sobre el tema es que los modelos máximos van a tener mayores chances de converger cuando los efectos aleatorios son mayores, con lo cual una metodología posible ante la falta de convergencia es la parametrización de un modelo con efectos aleatorios

de ítems y otro con efectos aleatorios de los sujetos, de manera análoga a el empleo de anova por ítems y por sujetos.

Para concluir, consideran también la relevancia de cómo se reportan los modelos de efectos mixtos. Dadas todas las observaciones previas sobre las diferencias posibles en la estructura de efectos aleatorios, es muy importante explicitar la estructura elegida, ya sea a través de la matriz de varianza-covarianza o a través de la aclaración de que se empleó una estructura máxima, indicando para qué variables se consideraron las pendientes aleatorias y, si no se trata de una estructura máxima, explicar y justificar la simplificación elegida. En el caso de haber empleado un procedimiento de selección de modelos hacia adelante, es importante reportar los criterios empleados para el testeo y la inclusión de los efectos aleatorios, así como los cambios en los supuestos sobre la generación de los datos resultantes de exclusión de determinados efectos más allá de “la incorporación de la pendiente aleatoria para la variable X no mejoró el desempeño del modelo”.

### 3 lmerTest y valores de $p$

Un obstáculo común para utilizar modelos mixtos es la dificultad para interpretarlos en términos de significancia estadística. Puntualmente en el caso de los modelos parametrizados utilizando el paquete lme4 de R (Bates 2010) (no quisiera generalizar a otras implementaciones) la salida del modelo no incluye valores de  $p$  para los efectos fijos, con lo cual se dificulta saber en qué casos esas variables tienen un efecto significativo en la variable dependiente. Frente a este problema existe una variedad de propuestas que intentan dar ayuda a la inferencia sobre las variables involucradas, como puede ser la comparación entre modelos de acuerdo con alguna métrica de información como AIC, o empleando un test de significancia sobre la verosimilitud del modelo a través de la función anova().

El output por defecto de un modelo lineal de efectos mixtos luce así:

```
modelo <- lmer(puntaje ~ fct_lect * Concordancia +
              (1+fct_lect*Concordancia|sujeto) +
              (1|oracion),
              datos_experimento)

## boundary (singular) fit: see ?isSingular
summary(modelo)

## Linear mixed model fit by REML ['lmerMod']
## Formula:
## puntaje ~ fct_lect * Concordancia + (1 + fct_lect * Concordancia |
##   sujeto) + (1 | oracion)
##   Data: datos_experimento
##
## REML criterion at convergence: 14979
##
## Scaled residuals:
##   Min       1Q   Median       3Q      Max
## -3.4322 -0.5389  0.0700  0.5401  3.5154
##
## Random effects:
##   Groups   Name                Variance Std.Dev. Corr
##   sujeto   (Intercept)          0.819852  0.90546
##           fct_lectD            0.007871  0.08872  -1.00
##           ConcordanciaT        0.172977  0.41590  -0.65  0.65
##           ConcordanciaU        1.086386  1.04230  -0.40  0.40  0.16
##           fct_lectD:ConcordanciaT 0.362198  0.60183   0.20 -0.20 -0.16
##           fct_lectD:ConcordanciaU 0.117024  0.34209   0.16 -0.16 -0.49
##   oracion (Intercept)          0.078266  0.27976
```

```

## Residual                      1.099053 1.04836
##
##
##
##
## 0.18
## -0.46 0.54
##
##
## Number of obs: 4680, groups: sujeto, 260; oracion, 18
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)      3.9141    0.1751  22.357
## fct_lectD         0.3051    0.2346   1.301
## ConcordanciaT     0.1769    0.2359   0.750
## ConcordanciaU    -1.5744    0.2433  -6.472
## fct_lectD:ConcordanciaT -0.9090    0.3337  -2.724
## fct_lectD:ConcordanciaU -0.5923    0.3323  -1.782
##
## Correlation of Fixed Effects:
##              (Intr) fct_1D CncrdT CncrdU f_D:CT
## fct_lectD    -0.677
## ConcordancT -0.688  0.499
## ConcordancU -0.680  0.484  0.484
## fct_lctD:CT  0.478 -0.703 -0.700 -0.333
## fct_lctD:CU  0.476 -0.706 -0.354 -0.688  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular

```

```
anova(modelo)
```

```

## Analysis of Variance Table
##
##              Df  Sum Sq Mean Sq F value
## fct_lect      1   1.939   1.939  1.7646
## Concordancia  2 139.641  69.821 63.5280
## fct_lect:Concordancia 2   8.413   4.206  3.8273

```

Si queremos verificar el efecto de una variable a través de la comparación con un modelo que no la incluye tenemos que reparametrizar el modelo para que use máxima verosimilitud (ML) en vez de su versión restringida (REML), y esto luce así:

```

modelo_ML <- lmer(
  puntaje ~ fct_lect * Concordancia +
    (1+fct_lect*Concordancia|sujeto) +
    (1|oracion),
  datos_experimento, REML = FALSE
)

```

```

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.040192
## (tol = 0.002, component 1)

```

```

modelo_ML_noInter <- lmer(
  puntaje ~ fct_lect + Concordancia +

```

```

      (1+fct_lect*Concordancia|sujeto) +
      (1|oracion),
datos_experimento, REML = FALSE
)

```

```

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.071807
## (tol = 0.002, component 1)

```

```

modelo_ML_noMainConc <- lmer(
  puntaje ~ fct_lect + fct_lect:Concordancia +
    (1+fct_lect*Concordancia|sujeto) +
    (1|oracion),
  datos_experimento, REML = FALSE
)

```

```

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0592972
## (tol = 0.002, component 1)

```

```

comp1 <- anova(modelo_ML, modelo_ML_noInter)
comp1

```

```

## Data: datos_experimento
## Models:
## modelo_ML_noInter: puntaje ~ fct_lect + Concordancia + (1 + fct_lect * Concordancia |
## modelo_ML_noInter:      sujeto) + (1 | oracion)
## modelo_ML: puntaje ~ fct_lect * Concordancia + (1 + fct_lect * Concordancia |
## modelo_ML:      sujeto) + (1 | oracion)
##
##           Df  AIC   BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## modelo_ML_noInter 27 15028 15202 -7486.8    14974
## modelo_ML          29 15023 15210 -7482.5    14965 8.6091     2    0.01351
##
## modelo_ML_noInter
## modelo_ML          *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

comp2 <- anova(modelo_ML, modelo_ML_noMainConc)
comp2

```

```

## Data: datos_experimento
## Models:
## modelo_ML: puntaje ~ fct_lect * Concordancia + (1 + fct_lect * Concordancia |
## modelo_ML:      sujeto) + (1 | oracion)
## modelo_ML_noMainConc: puntaje ~ fct_lect + fct_lect:Concordancia + (1 + fct_lect *
## modelo_ML_noMainConc:      Concordancia | sujeto) + (1 | oracion)
##
##           Df  AIC   BIC logLik deviance Chisq Chi Df
## modelo_ML          29 15023 15210 -7482.5    14965
## modelo_ML_noMainConc 29 15023 15210 -7482.5    14965     0     0
##
##           Pr(>Chisq)
## modelo_ML
## modelo_ML_noMainConc          1

```

Sin embargo, esta metodología puede ser anticonservadora, es decir, es posible que tenga una tasa de errores de tipo I superior al estándar de  $\alpha = .05$ . En Luke (2016) se realiza una serie de experimentos sobre datos simulados con el fin de comparar diferentes maneras de obtener valores de  $p$  para los efectos fijos. Las

metodologías comparadas son:

- El método “*t-as-z*”, según el cual el valor de *t* correspondiente a cada variable se interpreta como perteneciente a una distribución normal, dado que la distribución de *t* con grados de libertad infinitos se aproxima a *z*. De acuerdo con este método, un predictor es significativo cuando su valor de *t* es igual o superior a 1.96.
- La prueba de tasa de verosimilitud, que es el resultado de la comparación de modelos mediante la función `anova()` que vimos más arriba.
- Los métodos de *Satterthwaite* y *Kenward-Roger* para la aproximación de los grados de libertad para una prueba *F*: en los modelos mixtos no es claro cuáles son los grados de libertad relevantes para el modelo nulo según el factor de agrupamiento (efecto mixto). Tanto *Satterthwaite* como *Kenward-Roger* son maneras de estimar los grados de libertad de la distribución nula.
- *Bootstrapping* paramétrico, que implica el muestreo con reemplazo sobre la propia muestra.

Estos métodos se combinan a su vez con la parametrización de los modelos mixtos según ML o REML. El resultado de estos experimentos es que los métodos de *Satterthwaite* y *Kenward-Roger* para modelos parametrizados empleando verosimilitud máxima restringida son los menos anticonservativos y que, por lo menos con muestras en un rango acotado, no poseen una sensibilidad reducida, o sea, no poseen una tasa de error de tipo II mayor a la de otros métodos.

Tanto el método de *Satterthwaite* como el de *Kenward-Roger* se hallan implementados en el paquete `lmerTest` (Kuznetsova, Brockhoff, and Christensen 2017). Este paquete extiende los métodos `summary()` y `anova()` de `lme4` agregando dos columnas a los efectos fijos, una correspondiente al valor de *F* y otra al valor de *p*. Si llamamos a la librería y reparametrizamos el modelo, obtenemos el siguiente output:

OUTPUT

```
library(lmerTest)
```

```
##  
## Attaching package: 'lmerTest'  
  
## The following object is masked from 'package:lme4':  
##  
##   lmer  
  
## The following object is masked from 'package:stats':  
##  
##   step
```

```
modelo_test <- lmer(  
  puntaje ~ fct_lect * Concordancia +  
    (1+fct_lect*Concordancia|sujeto) +  
    (1|oracion),  
  datos_experimento)
```

```
## boundary (singular) fit: see ?isSingular  
  
## Warning: Model failed to converge with 2 negative eigenvalues: -2.2e-01  
## -1.8e+00
```

```
summary(modelo_test)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula:  
## puntaje ~ fct_lect * Concordancia + (1 + fct_lect * Concordancia |  
##   sujeto) + (1 | oracion)
```

```

## Data: datos_experimento
##
## REML criterion at convergence: 14979
##
## Scaled residuals:
##   Min       1Q   Median       3Q      Max
## -3.4322 -0.5389  0.0700  0.5401  3.5154
##
## Random effects:
##   Groups   Name                Variance Std.Dev. Corr
##   sujeto  (Intercept)           0.819852  0.90546
##           fct_lectD             0.007871  0.08872  -1.00
##           ConcordanciaT         0.172977  0.41590  -0.65  0.65
##           ConcordanciaU         1.086386  1.04230  -0.40  0.40  0.16
##           fct_lectD:ConcordanciaT 0.362198  0.60183   0.20 -0.20 -0.16
##           fct_lectD:ConcordanciaU 0.117024  0.34209   0.16 -0.16 -0.49
##   oracion (Intercept)           0.078266  0.27976
## Residual                        1.099053  1.04836
##
##
##
##
##   0.18
##  -0.46  0.54
##
##
## Number of obs: 4680, groups: sujeto, 260; oracion, 18
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)      3.9141    0.1751  14.9285  22.357 6.86e-13 ***
## fct_lectD         0.3051    0.2346  12.0431   1.301  0.2177
## ConcordanciaT     0.1769    0.2359  12.3211   0.750  0.4674
## ConcordanciaU    -1.5744    0.2433  13.9183  -6.472 1.51e-05 ***
## fct_lectD:ConcordanciaT -0.9090    0.3337  12.3358  -2.724  0.0181 *
## fct_lectD:ConcordanciaU -0.5923    0.3323  12.1282  -1.782  0.0997 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) fct_1D CncrdT CncrdU f_D:CT
## fct_lectD    -0.677
## ConcordancT  -0.688  0.499
## ConcordancU  -0.680  0.484  0.484
## fct_lctD:CT  0.478 -0.703 -0.700 -0.333
## fct_lctD:CU  0.476 -0.706 -0.354 -0.688  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular

```

```
anova(modelo_test)
```

```

## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## fct_lect      2.251   2.251     1  12.412  2.0483  0.17707

```

```
## Concordancia          140.034  70.017      2 14.380 63.7065 7.143e-08 ***
## fct_lect:Concordancia   8.413   4.206      2 12.236  3.8273  0.05118 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
modelo_test_ML <- lmer(
  puntaje ~ fct_lect * Concordancia +
    (1+fct_lect*Concordancia|sujeto) +
    (1|oracion),
  datos_experimento, REML = FALSE)
```

```
## Warning in checkConv(attr("opt", "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.040192
## (tol = 0.002, component 1)
```

```
summary(modelo_test_ML)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
## Satterthwaite's method [lmerModLmerTest]
## Formula:
## puntaje ~ fct_lect * Concordancia + (1 + fct_lect * Concordancia |
## sujeto) + (1 | oracion)
## Data: datos_experimento
##
##      AIC      BIC   logLik deviance df.resid
## 15023.0 15210.1 -7482.5 14965.0     4651
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.4364 -0.5398  0.0656  0.5444  3.5107
##
## Random effects:
## Groups   Name                Variance Std.Dev. Corr
## sujeto  (Intercept)          0.82526  0.9084
##          fct_lectD           0.03956  0.1989  -0.53
##          ConcordanciaT       0.17288  0.4158  -0.66  0.42
##          ConcordanciaU       1.02840  1.0141  -0.38 -0.09  0.09
##          fct_lectD:ConcordanciaT 0.29470  0.5429   0.21  0.03 -0.11
##          fct_lectD:ConcordanciaU 0.06015  0.2453   0.18  0.31 -0.54
## oracion (Intercept)          0.05306  0.2303
## Residual                    1.09518  1.0465
##
##
##
##      0.35
##     -0.36  0.34
##
##
## Number of obs: 4680, groups: sujeto, 260; oracion, 18
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)      3.9141    0.1492 24.0877 26.232 < 2e-16 ***
```



```
## fct_lectD          0.3051      0.1958 17.9227   1.558 0.13660
## ConcordanciaT     0.1769      0.1971 18.3994   0.898 0.38096
## ConcordanciaU    -1.5744      0.2053 21.6114  -7.670 1.33e-07 ***
## fct_lectD:ConcordanciaT -0.9090      0.2784 18.3067  -3.265 0.00423 **
## fct_lectD:ConcordanciaU -0.5923      0.2768 17.8914  -2.140 0.04637 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) fct_lD CncrdT CncrdU f_D:CT
## fct_lectD  -0.666
## ConcordancT -0.682  0.498
## ConcordancU -0.668  0.473  0.476
## fct_lctD:CT  0.469 -0.700 -0.698 -0.321
## fct_lctD:CU  0.466 -0.704 -0.354 -0.678  0.498
## convergence code: 0
## Model failed to converge with max|grad| = 0.040192 (tol = 0.002, component 1)
```

```
anova(modelo_test_ML)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## fct_lect          3.184   3.184     1 18.881  2.9075  0.10457
## Concordancia    189.307  94.654     2 22.754 86.4271 2.345e-11 ***
## fct_lect:Concordancia 12.062   6.031     2 18.149  5.5070  0.01352 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Como podemos ver, esto nos permite obtener valores de p para cada efecto de manera muy sencilla, y es importante tener en cuenta si se está empleando ML o REML porque los resultados cambian. En Agmon, Loewenstein, and Grodzinsky (2019) podemos ver cómo se reportan estos resultados.

[...] we fitted a linear mixed effects model using R's lmer function (Bates et al. 2015), with the logarithmic transformation of RT as the dependent variable. Polarity, Type, Standard and Truth-value were used as fixed effects, as well as all interactions. Random intercepts and slopes of Polarity, Type, Truth-value, Polarity  $\times$  Type and Polarity  $\times$  Truth-value were included, adjusted by subjects. P-values were obtained using R's lmerTest package (Kuznetsova, Brockhoff & Christensen 2017). A significant main effect of Polarity was found ( $t = 16.2$ ,  $p < 0.0001$ ) as well as a significant Polarity  $\times$  Type interaction ( $t = 4.5$ ,  $p < 0.0001$ ). This interaction stems from a stronger Polarity effect for quantifiers, in each of the Standard levels, as can be visualized in Figure 2.

## 4 Graficar nuestros modelos

La librería sjPlot (Lüdtke 2019) nos permite graficar nuestros modelos directamente sin necesidad de extraer los resultados previamente y podemos seleccionar diferentes elementos para visualizar.

Por ejemplo, aquí tenemos los efectos principales de nuestro modelo:

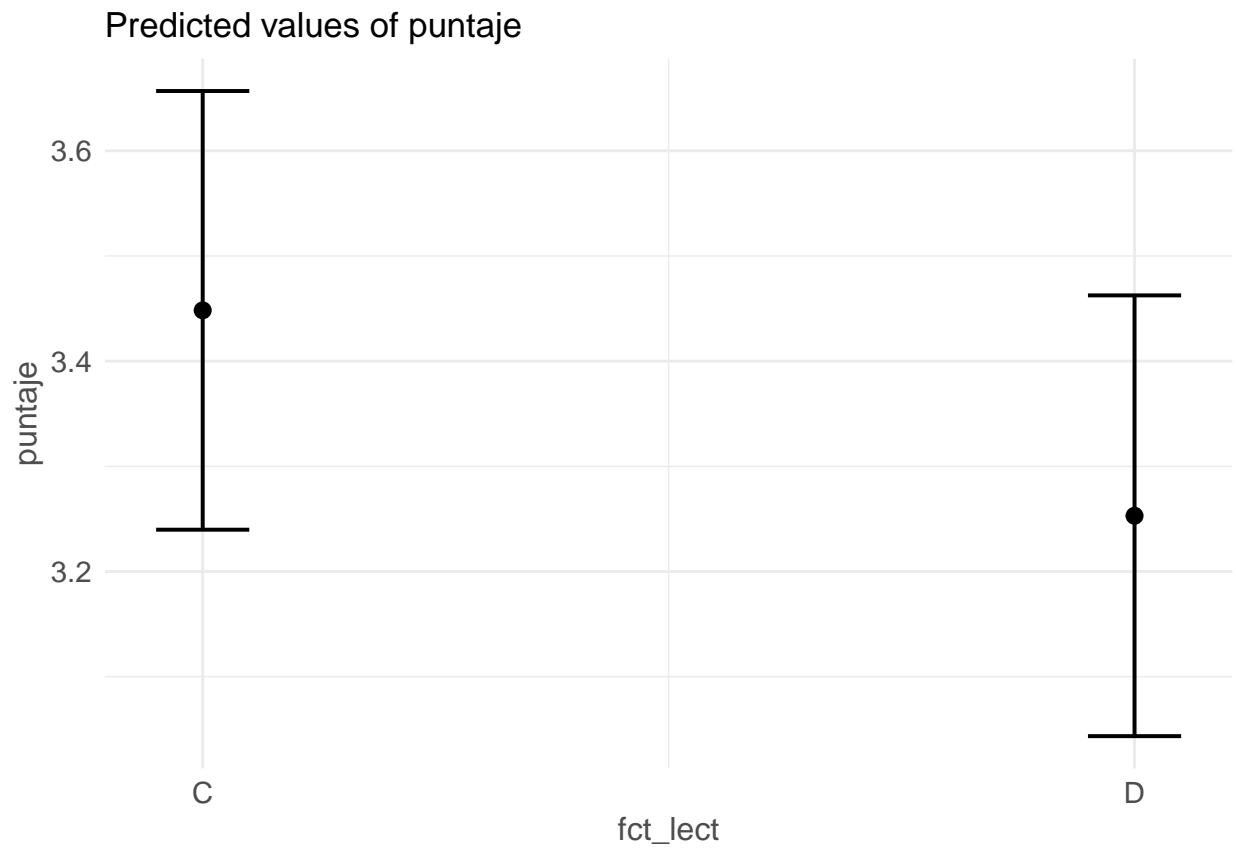
```
library(sjPlot)
```

```
## Learn more about sjPlot with 'browseVignettes("sjPlot")'.
```

```
set_theme(base = theme_minimal())
```

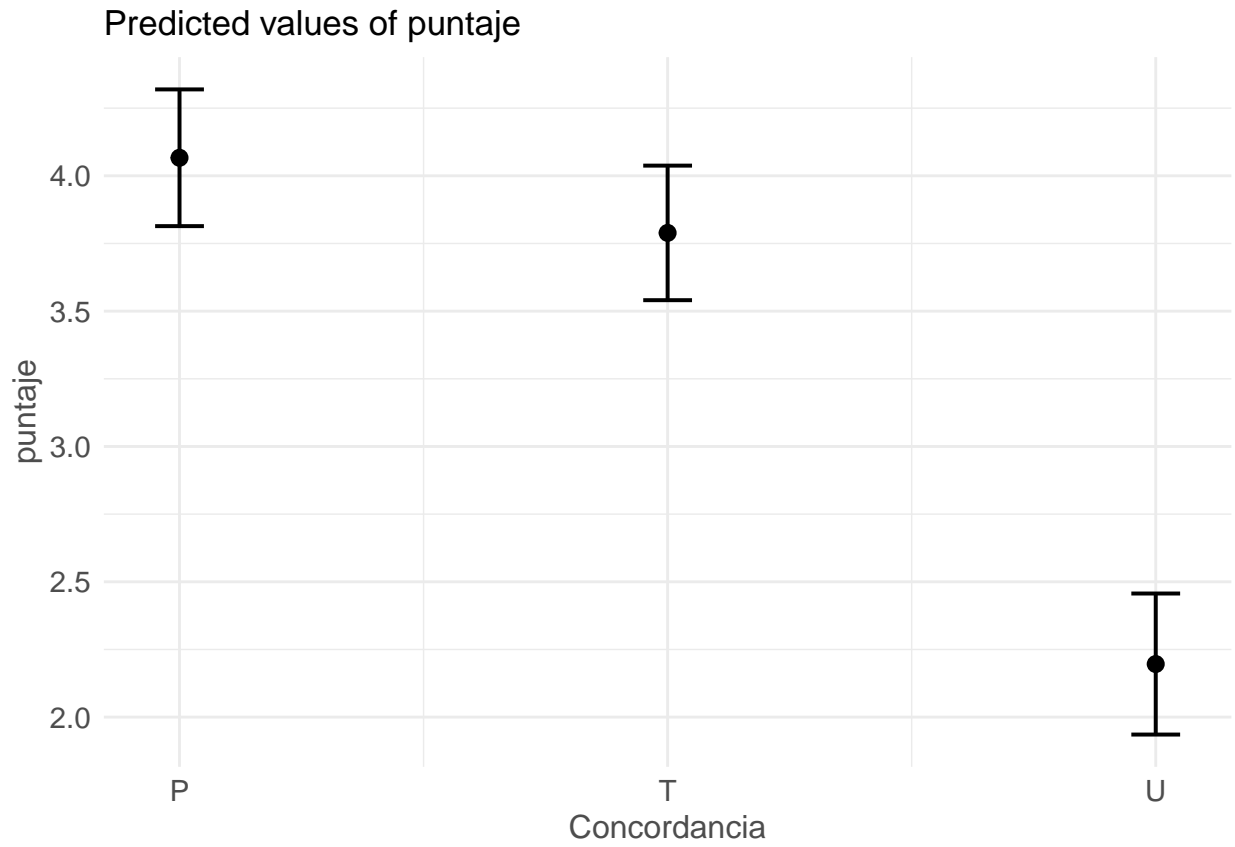
```
plot_model(modelo_test, type = 'eff')
```

```
## $fct_lect
```



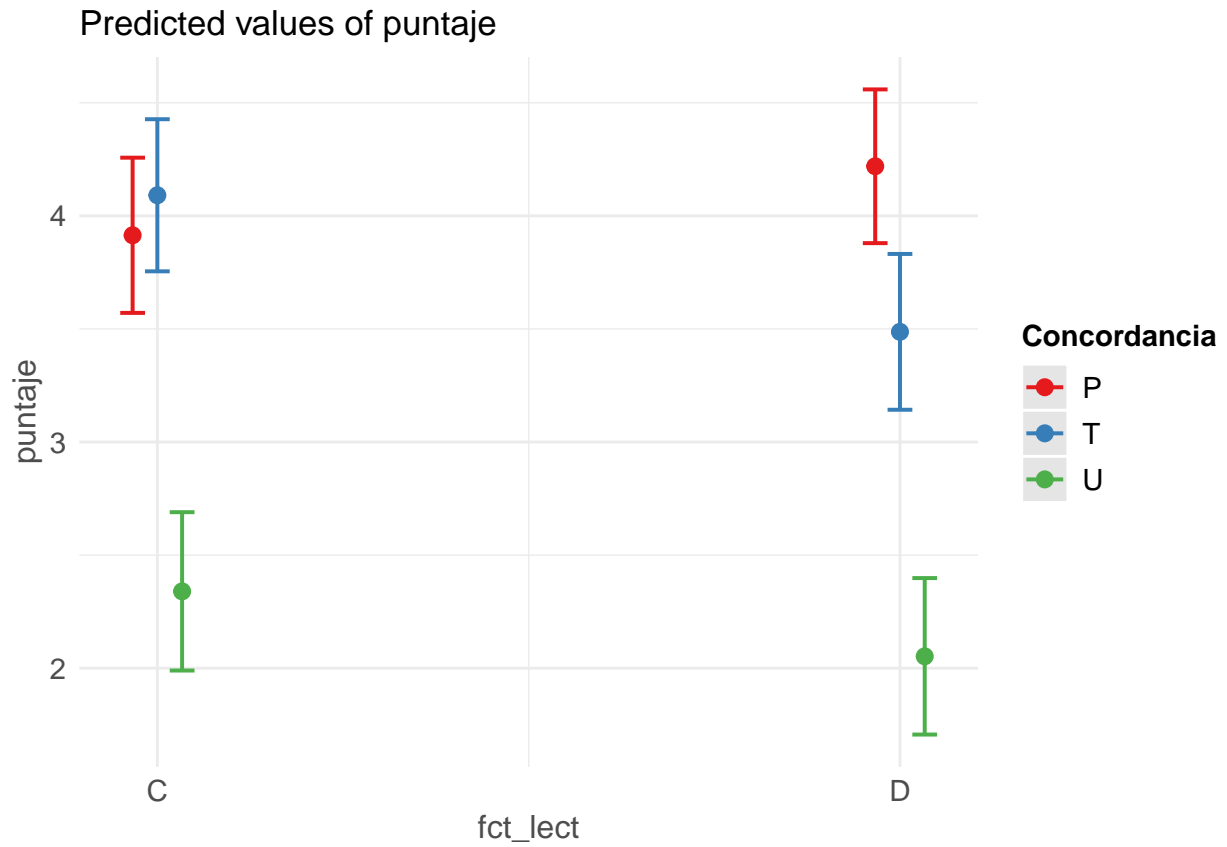
```
##
```

```
## $Concordancia
```



Y aquí las interacciones (por alguna razón genera dos plots iguales, así que me quedé con uno solo):

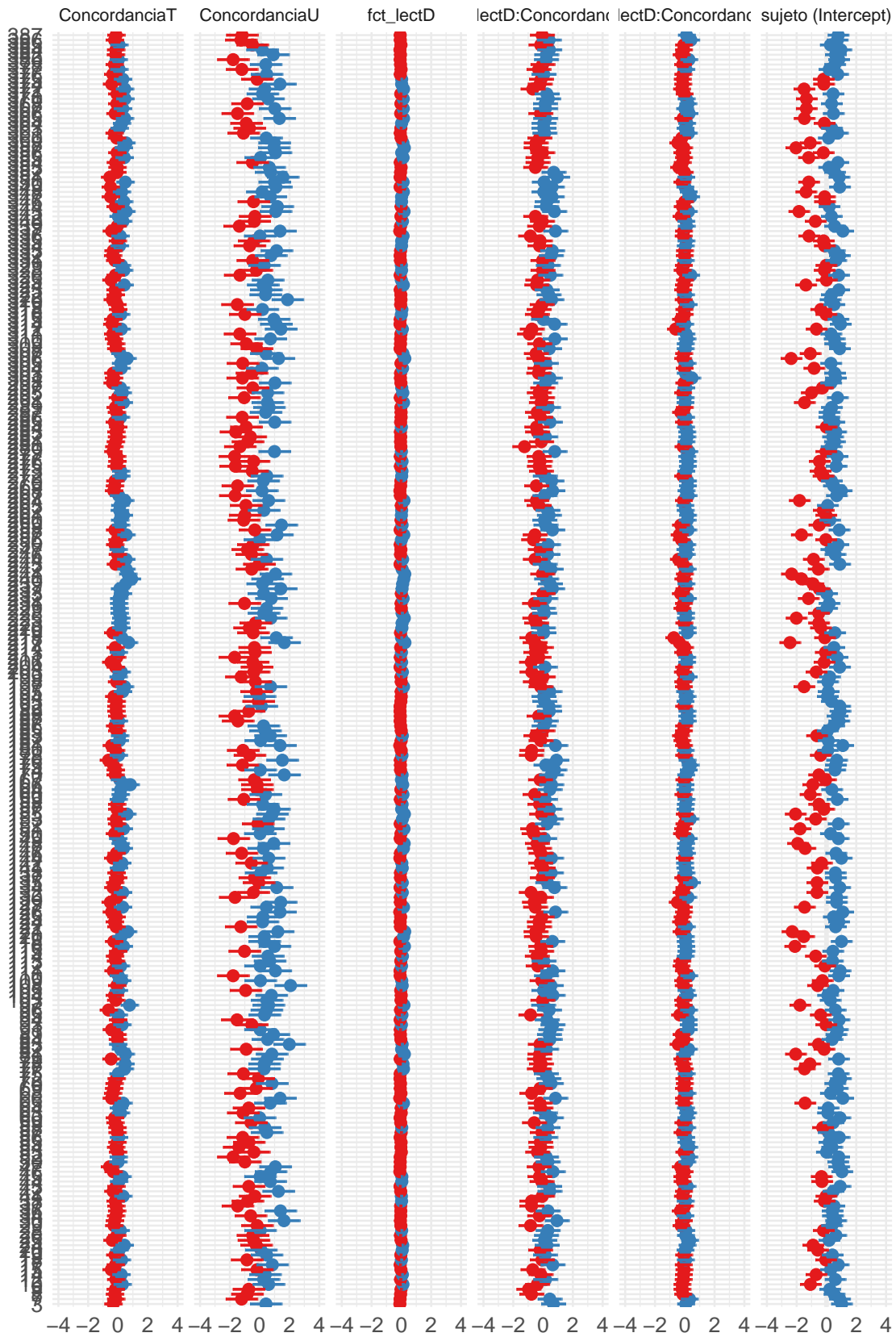
```
plot_model(modelo_test, type = 'int')[[1]]
```



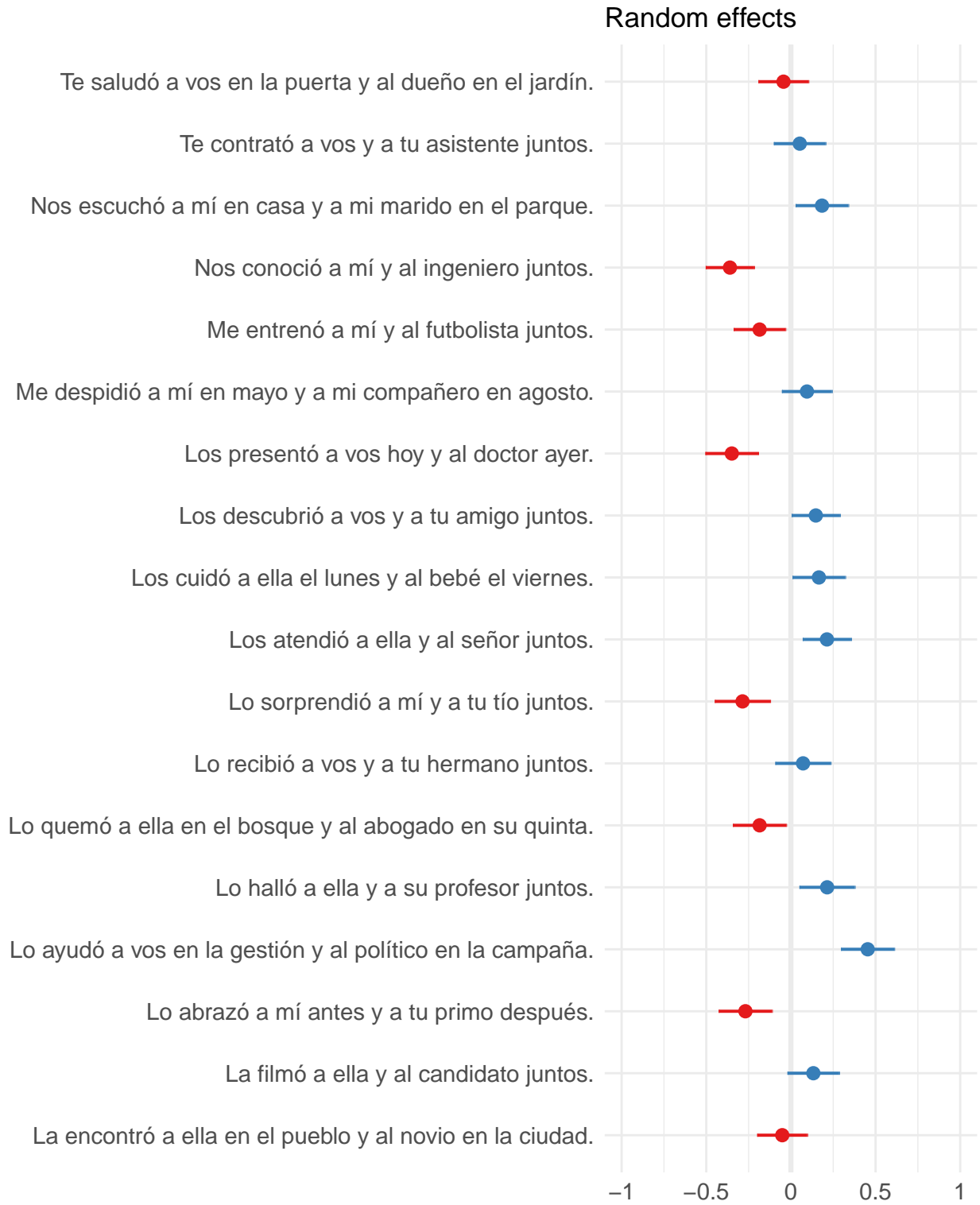
Incluso es posible graficar los efectos aleatorios (el gráfico 1 muestra efectos aleatorios por sujeto y el 2 por ítem)

```
aleatorios <- plot_model(modelo_test, type = 're')
aleatorios[[1]]
```

# Random effects



aleatorios[[2]]



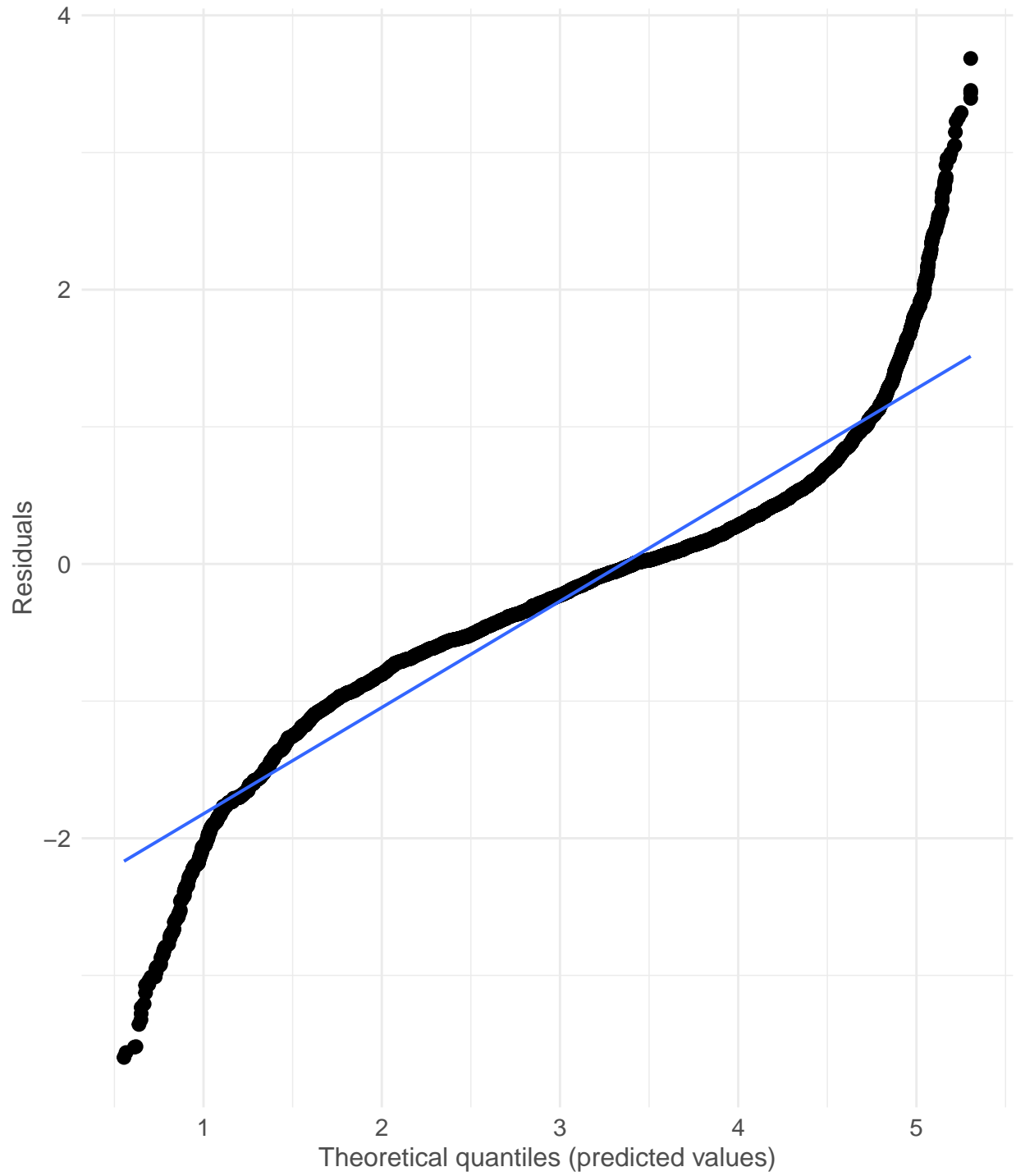
También se pueden visualizar diagnósticos del modelo, como homocedasticidad y normalidad de la varianza residual.

```
plot_model(modelo_test, type = 'diag')
```

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

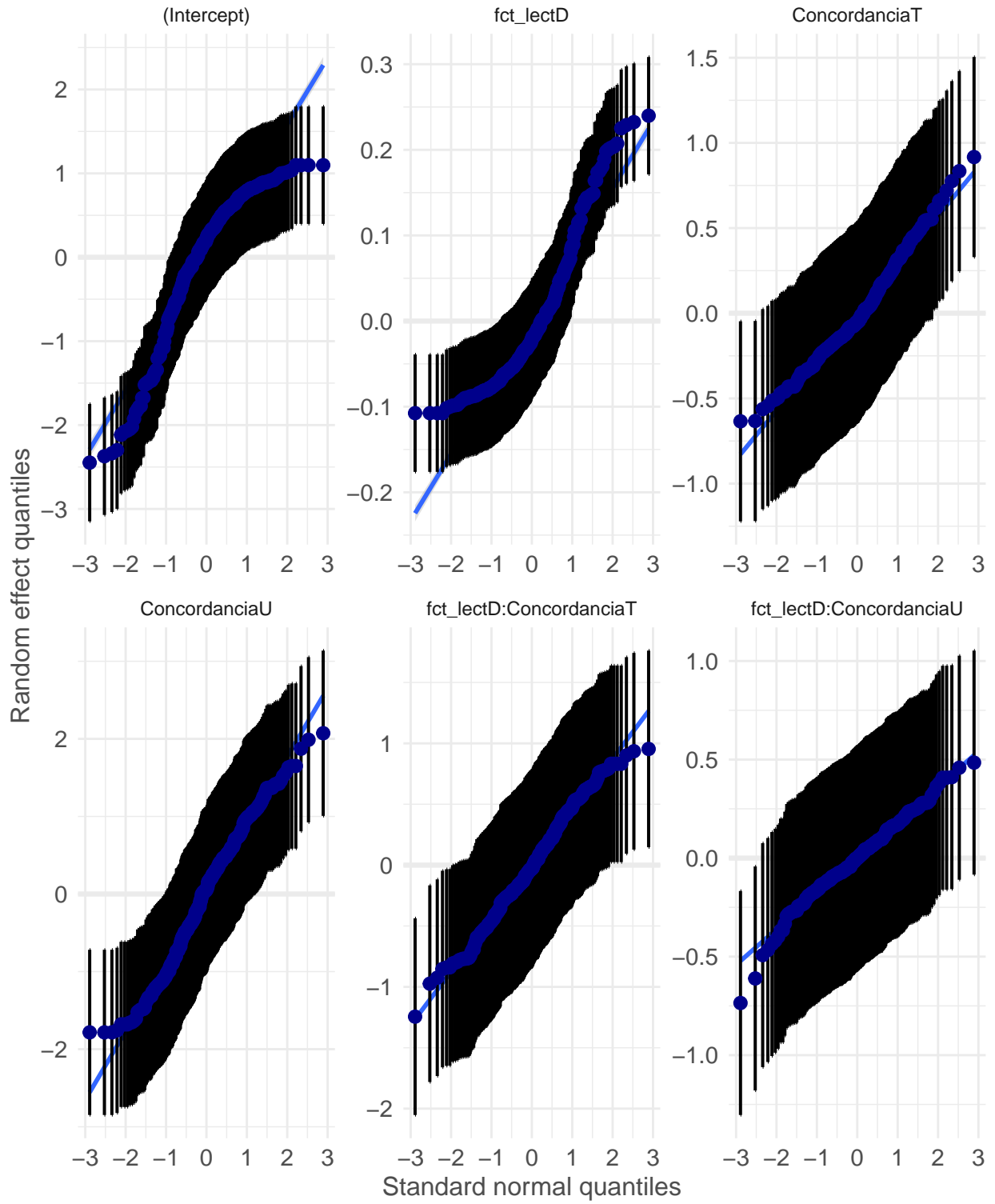
Non-normality of residuals and outliers

Dots should be plotted along the line



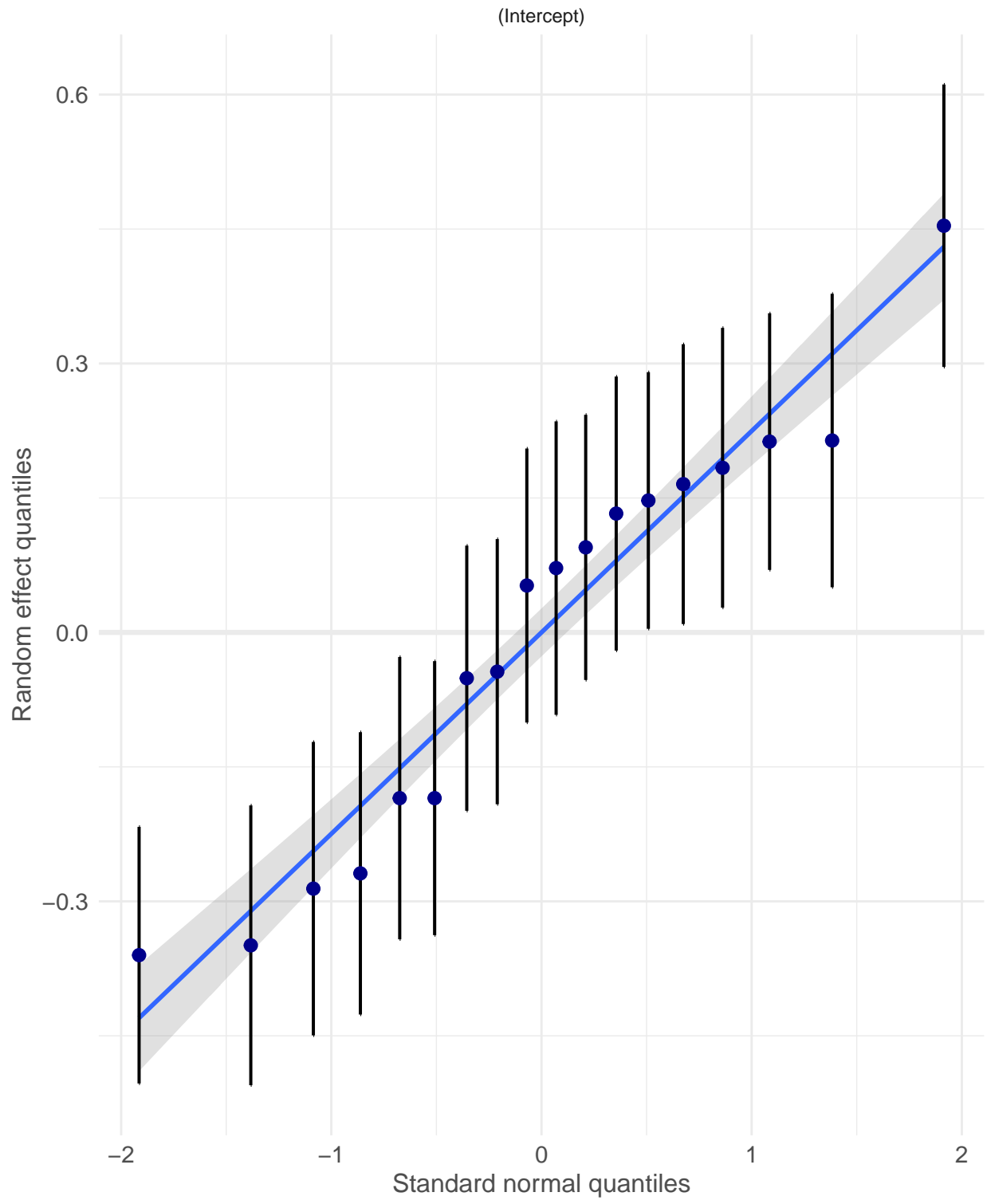
```
##
```

```
## [[2]]
## [[2]]$sujeto
```



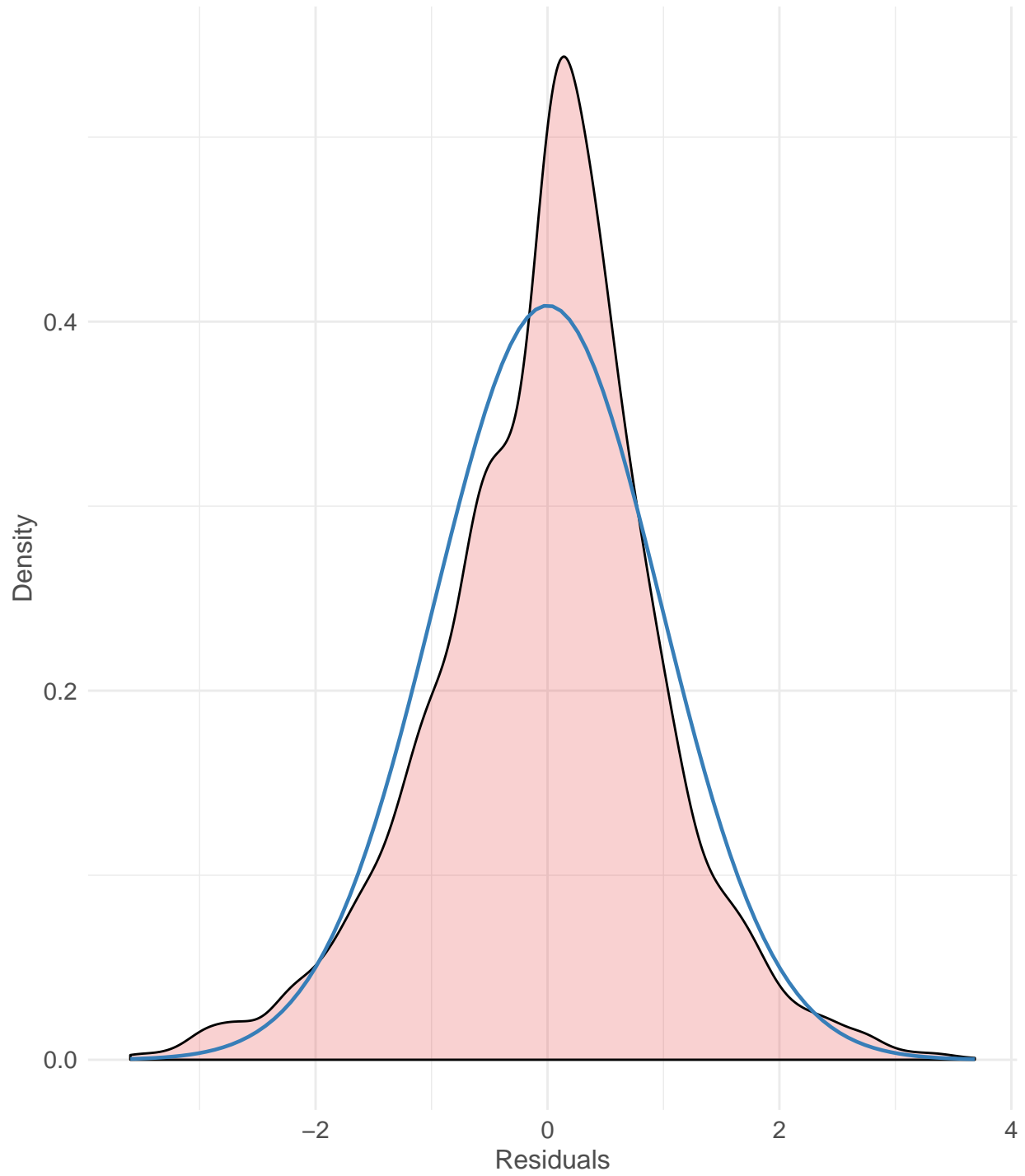
```
##
## [[2]]$oracion
```





```
##  
##  
## [[3]]
```

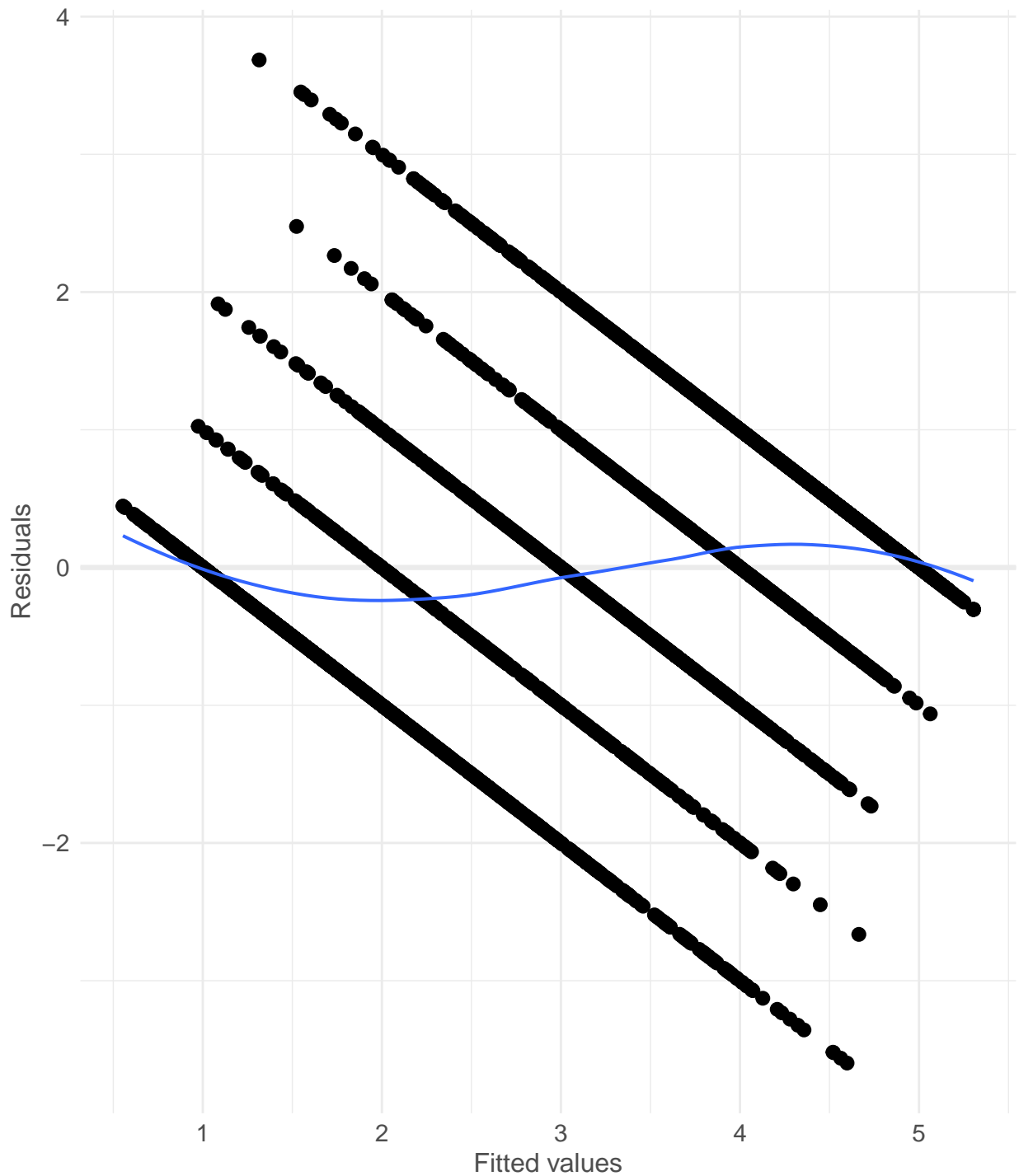
Non-normality of residuals  
Distribution should look like normal curve



```
##  
## [[4]]
```

## Homoscedasticity (constant variance of residuals)

Amount and distance of points scattered above/below line is equal or randomly spread



## 5 ¿Qué queda afuera?

Un tema importante que no pude cubrir acá es que, del mismo modo que la regresión lineal sin efectos mixtos, los modelos mixtos pueden generalizarse. De ese modo, es posible parametrizar regresiones logísticas

o *poisson*, o de otros tipos con efectos mixtos. En `lme4` puede usarse la función `glmer()`, que emplea una sintaxis similar a la de `lmer()` con un parámetro adicional para indicar qué tipo de modelo generalizado desea usarse.

Hay muchos tipos diferentes de modelos generalizados con efectos mixtos, y para casos en los que la variable independiente posee alguna clase de autocorrelación temporal existen los modelos generalizados aditivos, que permiten agregar términos de suavizado sobre medidas en una secuencia temporal y son útiles para fonética, lingüística histórica, y para medidas de procesamiento online. El paquete `mgecv` (Wood 2011, 2004, 2017, 2003; Wood, Pya, and Säfken 2016) permite parametrizar modelos de este tipo con términos de suavizado para los efectos aleatorios.

## Bibliografía

- Agmon, G., Y. Loewenstein, and Y. Grodzinsky. 2019. “Measuring the Cognitive Cost of Downward Monotonicity by Controlling for Negative Polarity.” *Glossa: A Journal of General Linguistics*. <https://doi.org/10.5334/gjgl.770>.
- Barr, R.; Scheepers, D. J.; Levy. 2013. “Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal.” *Journal of Memory and Language* 68 (3). <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bates, D. M. 2010. *lme4: Mixed-Effects Modeling with R*.
- Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. “lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software* 82 (13): 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Lüdtke, D. 2019. *SjPlot: Data Visualization for Statistics in Social Science*. <https://doi.org/10.5281/zenodo.1308157>.
- Luke, S. G. 2016. “Evaluating Significance in Linear Mixed-Effects Models in R.” *Behavior Research Methods*, September. <https://doi.org/10.3758/s13428-016-0809-y>.
- Wood, S. N. 2003. “Thin-Plate Regression Splines.” *Journal of the Royal Statistical Society (B)* 65 (1): 95–114.
- . 2004. “Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models.” *Journal of the American Statistical Association* 99 (467): 673–86.
- . 2011. “Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models.” *Journal of the Royal Statistical Society (B)* 73 (1): 3–36.
- . 2017. *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman; Hall/CRC.
- Wood, S. N., N. Pya, and B. Säfken. 2016. “Smoothing Parameter and Model Selection for General Smooth Models (with Discussion).” *Journal of the American Statistical Association* 111: 1548–75.