# Guidelines for model fitting

*Laurel Brehm*

*August 2019*

*These are some best-practice guidelines for how to fit a mixed-effect model. There are reasons to do each and every step differently– but here's something that will help guide your analysis*

1. Formulate your question. What, specifically, do you want to compare?

   an ok question: *What is the effect of native language and word frequency on lexical decision reaction time?*

   a better question: *Is performing a lexical decision task in English equally easier than performing it in any other of the second languages we tested? Does it vary by word frequency– and does it vary by word frequency equally for each of the other languages?*

2. Think about how to operationalize the necessary variables:

   a. What's the necessary dependent measure? Is it continuous (linear model) or binomial (logistic model)? If it is neither: Is there a way to re-word your question to make the dependent measure continuous or binomial? If not: look into ordinal models, poisson models, etc.

   b. What are the predictors? How are they structured– continuous (e.g. word frequency, age, IQ. . . ), or categorical (e.g. gender, social class, native language. . . )?

   c. If one predictor is categorical: what comparisons do you want to be able to make? Look up a contrast coding scheme that makes those comparisons (e.g. in your notes, or on https://stats.idre.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/)

      *e.g. I want to look at the overall effect of frequency–so this means I need my other predictor to have zero at the average of all levels– not at some level of the data*

      *I also want to compare NativeLanguage is "Dutch" or "German" to "English"*

      so: *use simple effects coding to assess effect of "Dutch" and "German" compared to "English", preserving the intercept as the grand mean of all observations.*

   d. If there is more than one predictor– do you even want to allow the predictors to interact? (Are they likely to be independent, or is one effect likely to be contingent on another?)

   e. What are the random intercepts (aka, grouping variables)? (What do observations group over that will have no overall effect but might still account for variability? These are categorical variables: e.g., Participant, Item, School, City)

   f. Do any predictors make sense to vary as random slopes on top of the random intercepts (grouping variables)? Specifically– what repeats across my random intercept terms in my data set? These could be categorical or continuous variables. *e.g. (participant) native languages as a property of the items (item); frequencies as a property of the subjects?* Caution: interactions in random slopes require an awful lot of data! Think about whether you've really got enough observations before trying to include. If you're not sure– try fitting a model, and expect that there could be a convergence error.

3. Format your data in a way to test this question. Unlike in SPSS, each row needs to be: **one observation**

   *My data frame needs at least 5 columns to operationalize the above question:*

   *-one column for RT (listed out per observation)*

   *-one for native language (listed out for each subject, repeated across observations)*

*-one for frequency (associated with each word per observations)*

*-one for subject (per observation)*

*-one for item (per observation)*

Here's how this looks:

```
##         RT Frequency NativeLanguage Subject       Word
## 1 6.340359  4.859812        English      A1        owl
## 2 6.308098  4.605170        English      A1       mole
## 3 6.349139  4.997212        English      A1     cherry
## 4 6.186209  4.727388        English      A1       pear
## 5 6.025866  7.667626        English      A1        dog
## 6 6.180017  4.060443        English      A1 blackberry
```

4. Open R. Load libraries– lme4 at least, and anything else you want (such as effects, or tidyverse)

```
library(lme4) #for lmer / glmer
library(effects) #for effects extracting
library(lattice) # for qqmath plot
```

5. Load your data– if it's in the same folder as your code, you can say this:

```
data <- read.table('mydata.txt',header=T) ## for txt files
```

6. Check that your data looks right with a 'summary'– continuous variables are treated as numbers, everything else is not treated as numbers.

```
summary(lexdec)
```

```
##        RT          Frequency     NativeLanguage    Subject
##  Min.   :5.829   Min.   :1.792   English:948    A1     :  79
##  1st Qu.:6.215   1st Qu.:3.951   Other  :711    A2     :  79
##  Median :6.346   Median :4.754                  A3     :  79
##  Mean   :6.385   Mean   :4.751                  C      :  79
##  3rd Qu.:6.502   3rd Qu.:5.652                  D      :  79
##  Max.   :7.587   Max.   :7.772                  I      :  79
##                                                 (Other):1185
##         Word
##  almond   :  21
##  ant      :  21
##  apple    :  21
##  apricot  :  21
##  asparagus:  21
##  avocado  :  21
##  (Other)  :1533
```

To convert a number to a not-number (a factor), use the as.factor() function:

```
lexdec$Subject <- as.factor(lexdec$Subject)
```

To convert a not-number to a number, use a combination of as.numeric(as.character()):

```
lexdec$RT <- as.numeric(as.character(lexdec$RT))
```

If you have continuous predictors, it's good to center them.

```
lexdec$FreqC <-lexdec$Frequency - mean(lexdec$Frequency)
```

7. Translate your model-building thoughts into R. Take your time to think.

```
m1 <- lmer(RT ~ NativeLanguage*FreqC + (1 + FreqC | Subject)
           + (1 + NativeLanguage |Word), data=lexdec, REML=F)
```

8. Fit a model!

a. If the model gives an error or warning about 'convergence warnings' or 'singular fit', look at the random effects only. From here, you have 2 options:

   1. (advanced) Run rePCA() function to determine how many random effects are justified by your data. If there are e.g., 3 components that would model 95% of the variance in the random effects, pick the 3 terms that cover the most variance (have the largest number in the variance column of the random effects tier). Important: do not include interactions without also including main effects– even in the random effects tier.

   2. (more simple) Look at the random effects tier. Take out the single, highest-order term that accounts for the least variance. Re-fit model. Repeat as many times as needed. This is called model fitting by backwards selection.

b. If the model gives you a warning about 'Rescale variables' – look to see if your predictors are on a similar scale (e.g., if one ranges from 1-2000, but the other ranges from .001 to .0002. . . .rescale). Use scale() function. Re-fit model.

```
lexdec$Frequency <- scale(lexdec$Frequency)
```

c. If you get other types of errors, check whether your model is specified in the right way (are the correct things interacting? am I trying to account for the same thing twice?). Consider changing the optimizer, or running for more iterations; in either case. . . re-fit model.

[see https://rstudio-pubs-static.s3.amazonaws.com/33653_57fc7b8e5d484c909b615d8633c01d51.html for more details]

9. Once you have a converged model, take a look at the random effects again: are there any terms very close to zero, or correlations between random effects above |.9| ? These are evidence of overfitting.

If so, exclude term from the model (removing random slope interactions too if you are removing random slope main effects). Refit model as needed.

10. Look at your residuals:

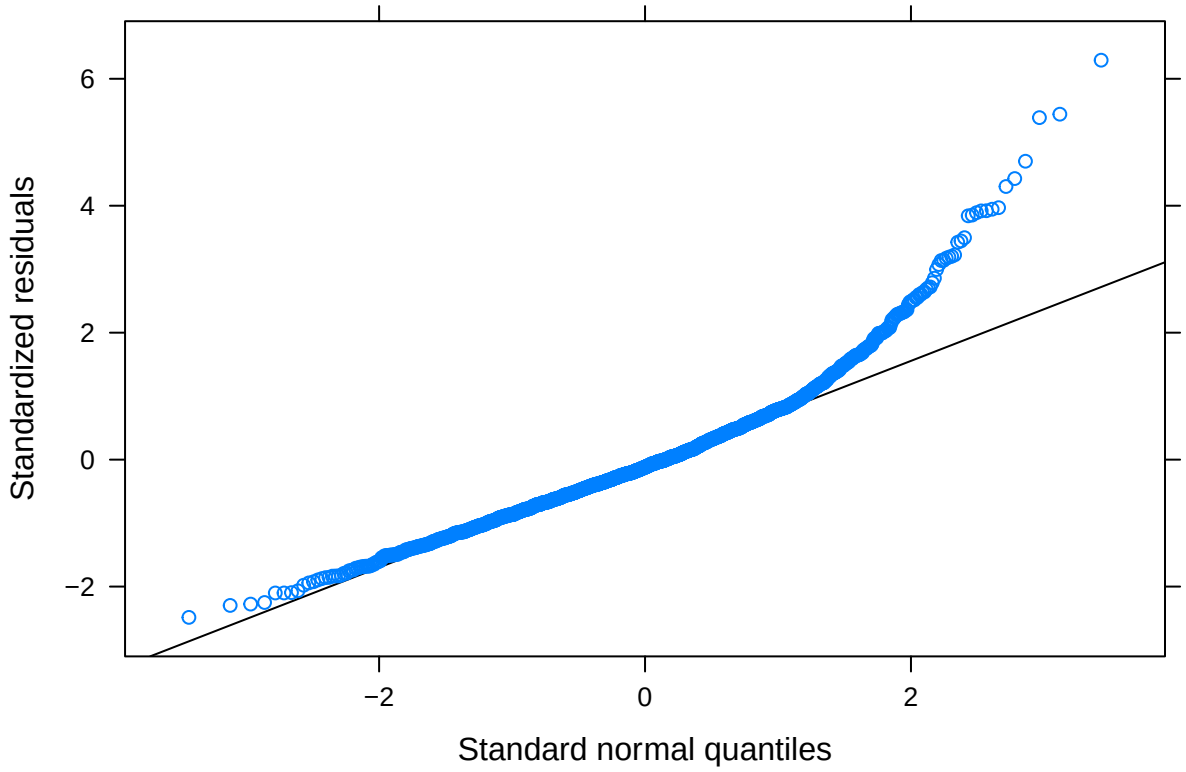
```
summary(m1)
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.4933 -0.6184 -0.1204  0.4764  6.2905
```

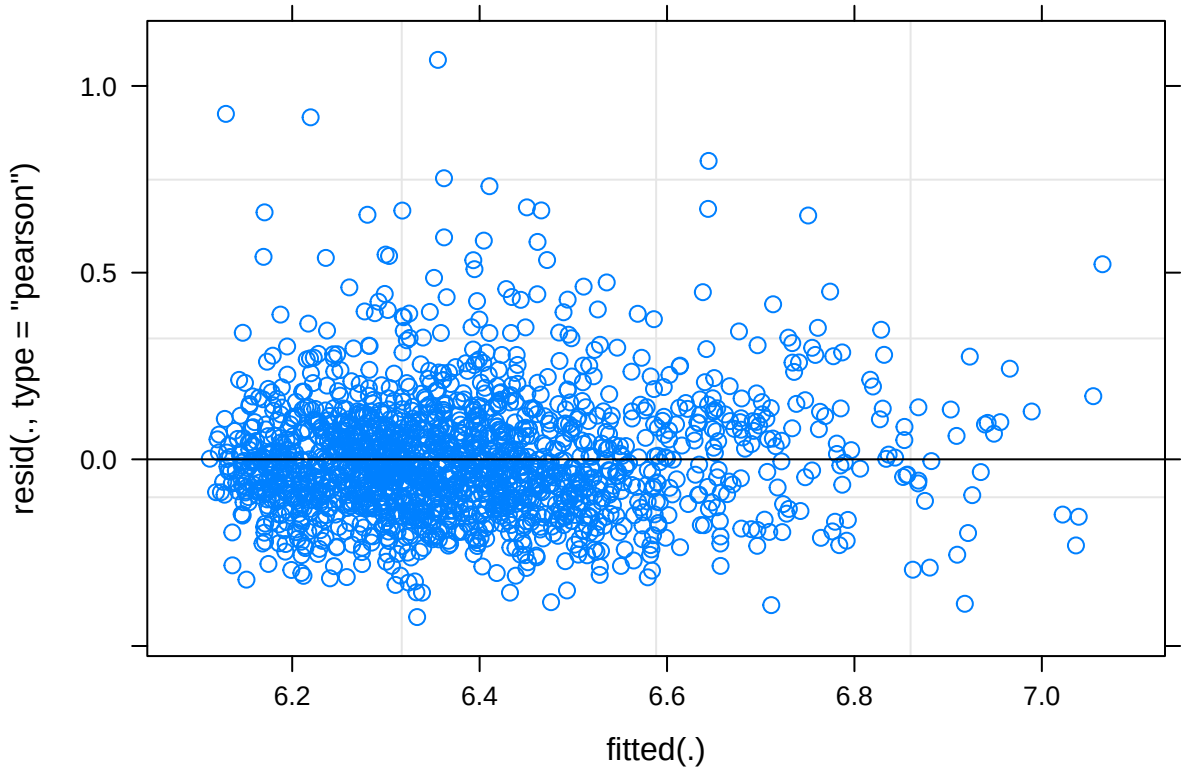
a. Numerically: are they approximately symmetrical about zero? If not: are there other sources of variability you should be accounting for? Add them to the model if so; if not, move on.

b. *For lmer models only:* Run a qq plot (should be as close to the line as possible; will probably have tails going up and down) and a fitted vs resid plot (should look like a fluffy cloud)

```
qqmath(m1)  #in the lattice package
```




```
plot(m1)
```




If the residuals are wonky, consider what you can do:

a. You could transform your DV– consider whether transforming will make your model interpretable. If so, transform, refit. If not, move on.

b. Are there other sources of variance hiding in your data that you could account for? If so, try to add parameters. If not, move on.

11. Now, look at the fixed effects. Interpret terms in light of your contrasts if you had them [remember that the way to interpret a model is, for each level of each variable: y = intercept + effect*Contrast per level]

Extract effects with the effects() package if you want to plot the fitted values on a sensible scale neatly.

```
Fixed effects:
                          Estimate Std. Error t value
(Intercept)                6.31831    0.03802 166.184
NativeLanguage             0.15582    0.05779   2.697
FreqC                     -0.03965    0.00892  -4.445
NativeLanguage:FreqC      -0.03503    0.01232  -2.843
```

```
plot(effect('NativeLanguage:FreqC',m1))
```

12. If you want to perform hypothesis testing, do one or more of the following:

a. run the function confint()

```
confint(m1)
```

b. calculate Type II p-values by model comparison (exclude effects, and interactions they participate in)

c. calculate Type III p-values by model comparison (exclude term only)

d. calculate p-values using Sattherwaite approximation or other.

13. Report what you did in your paper:

- Describe your software (R version ...) and your packages. Here's an easy way to get the citation for the package you used:

```
citation('lme4')
```

```
##
## To cite lme4 in publications use:
##
##   Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015).
##   Fitting Linear Mixed-Effects Models Using lme4. Journal of
##   Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
##
## A BibTeX entry for LaTeX users is
##
##   @Article{,
##     title = {Fitting Linear Mixed-Effects Models Using {lme4}},
##     author = {Douglas Bates and Martin M{\"a}chler and Ben Bolker and Steve Walker},
##     journal = {Journal of Statistical Software},
##     year = {2015},
##     volume = {67},
##     number = {1},
##     pages = {1--48},
##     doi = {10.18637/jss.v067.i01},
##   }
```

- Describe what went in to your model and how you selected your random effects (e.g. "We began with the maximal model justified by the structure of the data, random intercepts by P and Q, and random slopes for Z by P. We simplified as needed due to non-convergence, and to remove highly-correlated random effects (above |.9|) in order to avoid overfitting").
- Report the fixed and random effects from your model in a table. At minimum, include the fixed effect coefficients and SE terms, plus include the random effect variances or SDs. This would ideally go in the main body of the paper, but could also go in an appendix.
- Report whatever means and/or CIs you need to make your inferences clear in your paper. Supplement with figures and tables of means, variability, etc.