Reanalysing Hsu & Bishop

DVM Bishop

15/05/2018

Background

The usual way of looking for implicit learning in serial-reaction-time (SRT) task has been to compare performance when there is a repeated pattern in the stimuli (learning phase) with performance when the pattern is broken (random phase). Typically this has been done by comparing RT for the last block in a pattern phase with the first block in the random phase: if the pattern was learned, we should see ‘rebound’ when it is broken. This approach was adopted in a paper by Hsu and Bishop (2014), which used a ‘Green Monster’ task developed by Tomblin et al (2007). The child was presented with four squares in a line and had to press a key to correspond to the square in which a green monster appeared. After an initial random phase of 100 trials, there were 200 trials where the repeated sequence 2 4 1 3 4 2 1 4 3 1 was used, and then a further 100 random trials. (Note: this sequence was wrongly described in the published paper: this is the correct sequence).

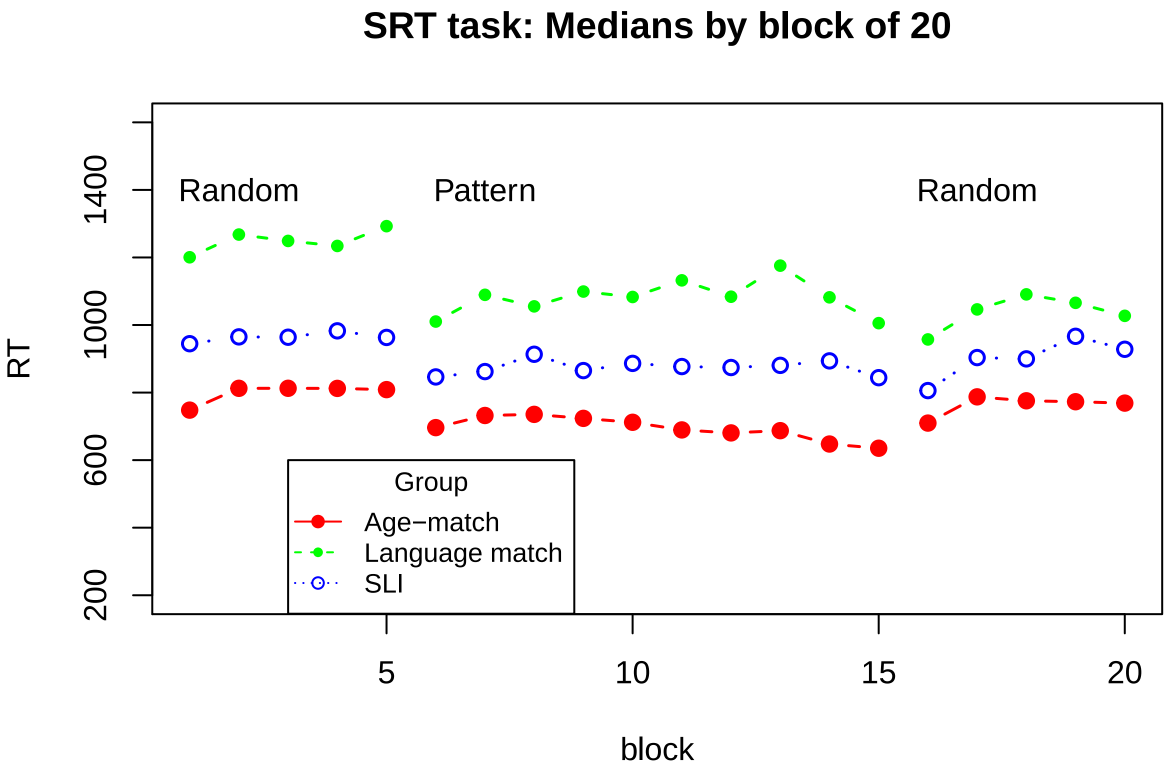
In a later paper, Kuppuraj et al (2018) used a task designed to overcome some issues with previous research. In particular, a square array of four pictures was used and the task was to select the named picture. Using this method, it was possible to interleave different types of sequential dependency to compare learning for short sequences that varied in the statistical structure of dependencies. We also adopted a new analytic approach that used all of the data, regression discontinuity. This uses a test for the difference in slope between a pattern phase and a random phase. Although regression discontinuity is usually applied to group data, the method of comparing two slopes can be used for data from a single case, and it gives a t-value for each participant and condition that can be evaluated to determine whether learning occurred (i.e. whether the difference in the two slopes is statistically reliable).

In this paper, the main goal is apply the regression discontinuity approach to the original Hsu and Bishop dataset, to compare whether its sensitivity relative to the original method of measuring implicit learning.

The definitive version of the dataset for Hsu and Bishop (2014), alldata_GM.xlsx, is available on Open Science Framework here: https://osf.io/396ai/wiki/home/. A csv version with just the worksheet showing responses and RTs is saved here.

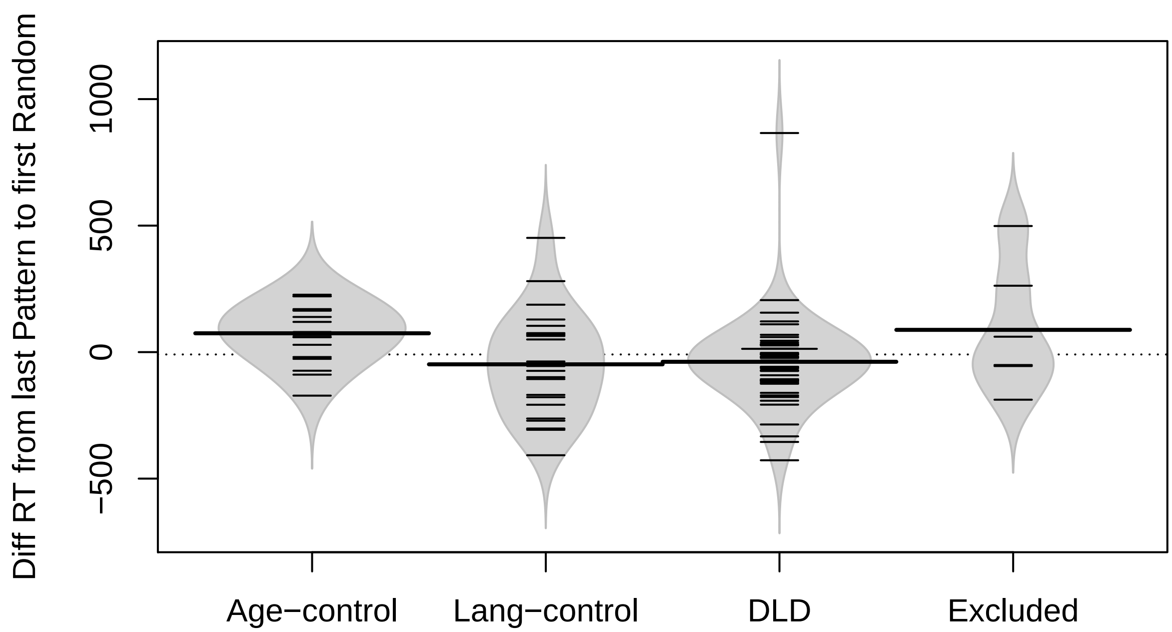
Data summary for three groups

This plot shows the data processed as in the original Hsu and Bishop paper: i.e. medians for correct responses extracted for blocks of 20 trials.



Original subtraction method

We compute the difference between the last datapoint for the pattern blocks and the first datapoint for the random blocks, and compare this for the 3 groups. NB. We use the acronym DLD for 'developmental language disorder', which has superseded SLI (for 'specific language impairment') -see Bishop et al (2017).

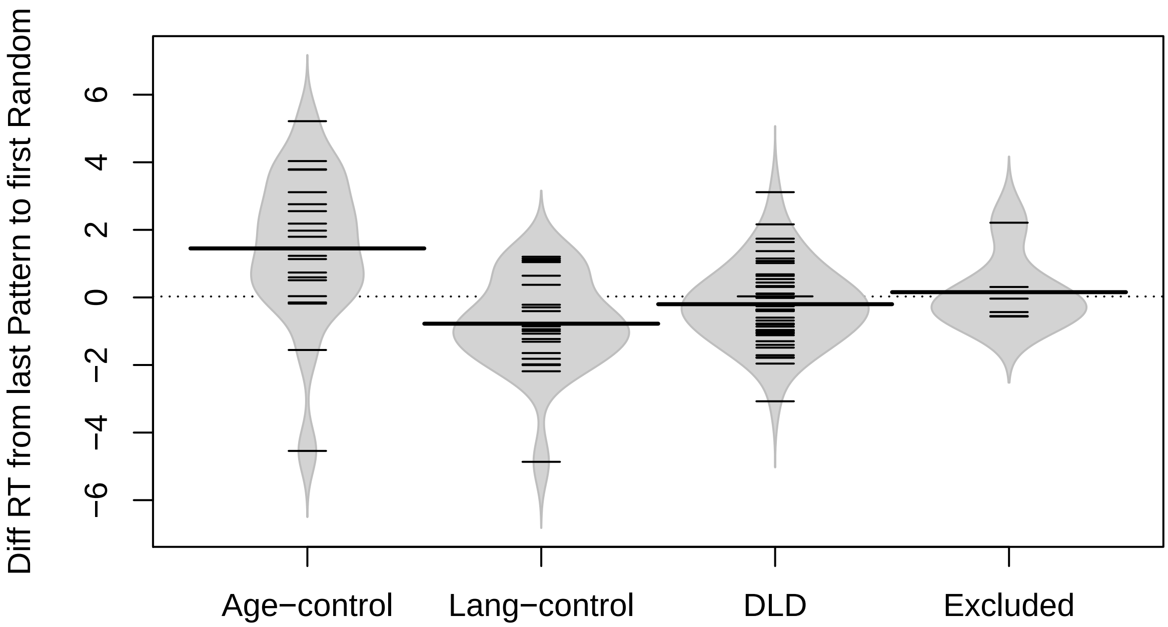


Regression discontinuity approach

We compare the original approach with the method of Kuppuraj and Bishop. The package rddtools (Stigler & Quast, 2015) is used to compare the slopes of two portions of data. We do this for the pattern and second random sequence. The t-value for this comparison can be used as an index of learning. This table shows the N children from each group who learned (p < .05).

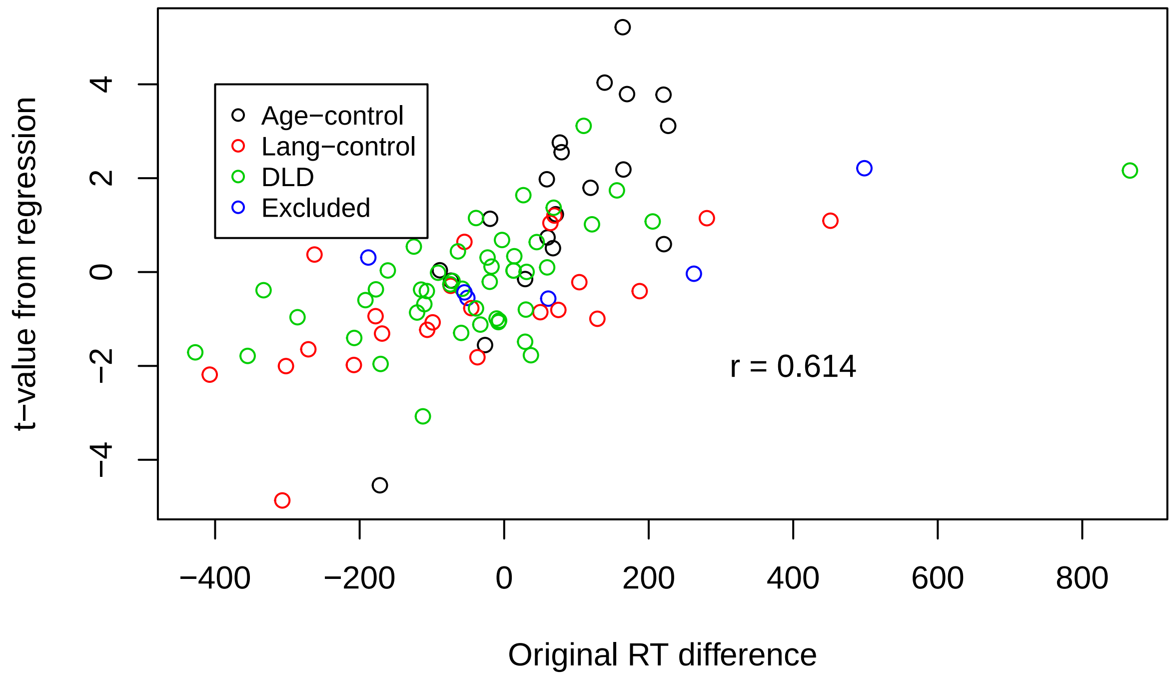
```
##
##                 Not learned Learned
##   Age-control            10      10
##   Lang-control           18       5
##   DLD                    43       4
##   Excluded                5       1
```

Now we look at the distribution of scores for t-values. This appears to give clearer differentiation between groups than the original analysis.



Compare discontinuity regression results with subtraction results

The relationship between the two measures is shown here.



Anova with diffc and with reg_t for 3 groups

```
## [1] "Original analysis on difference score"

##              Df  Sum Sq Mean Sq F value Pr(>F)
## group         2  210744  105372   3.344 0.0399 *
## Residuals    87 2741374   31510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## [1] "New analysis on t-value for regression discontinuity"

##              Df Sum Sq Mean Sq F value   Pr(>F)
## group         2  57.86  28.931    12.7 1.44e-05 ***
## Residuals    87 198.13   2.277
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Compute effect sizes (d) for different methods

Pairwise comparisons with age-control for (i) DLD/age-con and (ii) lang-con/age-con

```
##          group  N Mean.diff       sd
## 1  Age-control 20  74.45000 109.6852
## 2 Lang-control 23 -48.30435 206.5936
## 3          DLD 47 -38.18085 184.9681

##          group  N     Mean.t       sd
## 1  Age-control 20  1.4508557 2.210791
## 2 Lang-control 23 -0.7777486 1.391864
## 3          DLD 47 -0.2001901 1.166966
```

The regression method yields larger effect sizes, as would be expected given the ANOVA results. With the original method, for the DLD vs age-matched controls, Cohen's d is 0.68, and for age-matched controls vs language-matched controls it is 0.73.

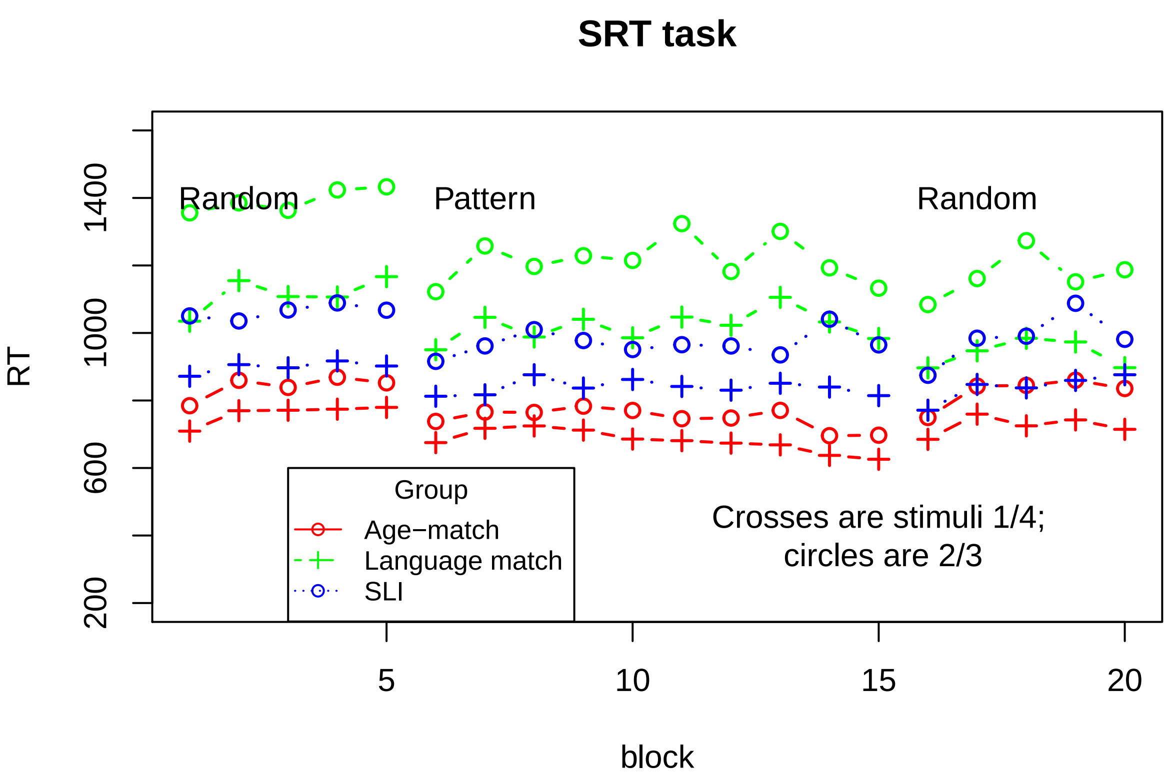
With the regression discontinuity method, for the DLD vs age-matched controls, Cohen's d is 1.07, and for age-matched controls vs language-matched controls it is 1.23.

Conclusion

The regression discontinuity approach appears more sensitive to detecting the rebound effect in a serial RT task than the traditional approach of subtracting the RT from the final block of the pattern phase from the first block of the subsequent random phase.

Additional analysis of item-specific effects

The sequence used in the Green Monster task was 10 items long, raising concerns that the individual responses were not of equal frequency. In the sequence 2 4 1 3 4 2 1 4 3 1, numbers 2 and 3 occur twice, whereas 1 and 4 occur thrice. We were interested in the possibility that this difference in frequency of occurrence might influence the RTs. In fact, when the responses were subdivided into two subgroups, corresponding to stimuli 2 and 3 vs stimuli 1 and 4, it was apparent from observation that responses to the latter stimuli were indeed more rapid. However, this difference was clearly present in the first random block, when all stimuli were equally frequent. This suggests that the effect is not due to stimulus frequency, but rather to perceptual crowding: the four stimuli were arranged in a linear row, with 1 and 4 at the ends. They would therefore be more discriminable than stimuli 2 and 3, which had another box on either side.

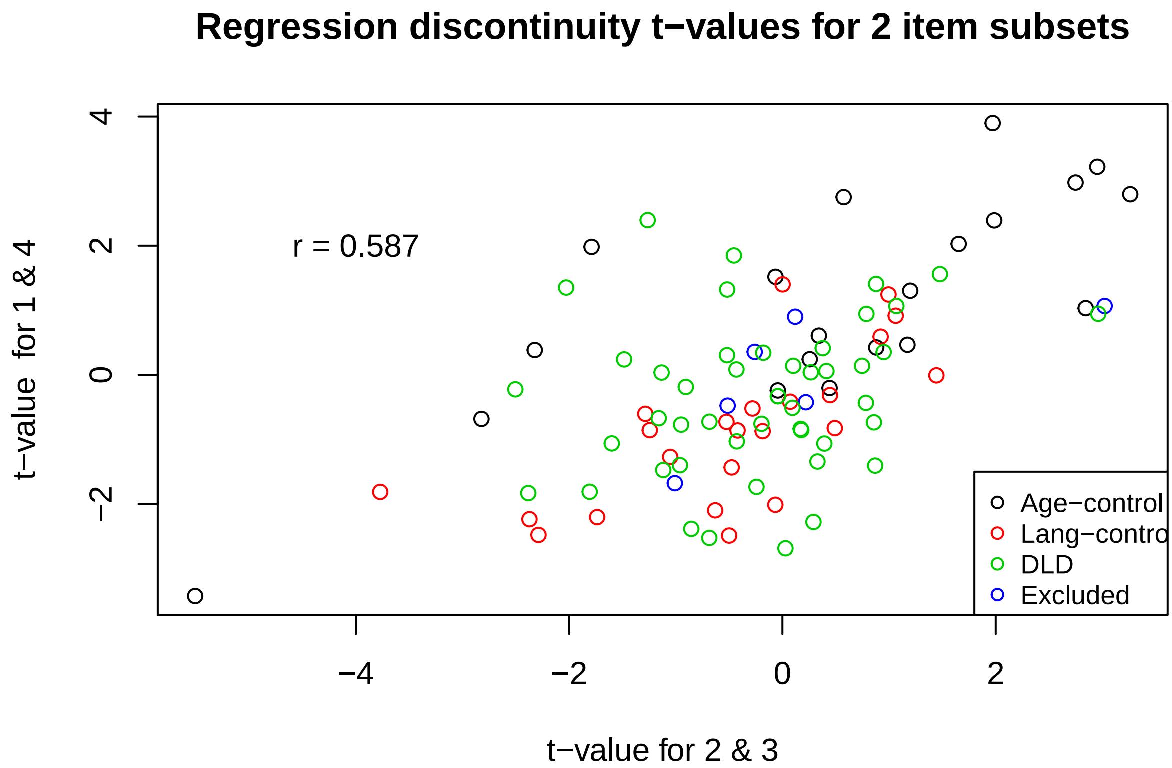


This provides an opportunity to look at the regression discontinuity t-value separately for each subset. This will be based on half as much data so will be noisier, but could give an indication of split half reliability of the method.

The first table shows agreement in terms of categorical classification in relation to $p < .05$: this is not very good.

```
##
##           nolearn learn
##   nolearn      83     5
##   learn         3     5
```

The correlation between the t-values for the two item sets is moderate, r = 0.587.



References

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE Consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. Journal of Child Psychology and Psychiatry, 58(10), 1068-1080. doi:10.1111/jcpp.12721

Hsu, H. J., & Bishop, D. V. M. (2014). Sequence- specific procedural learning deficits in children with specific language impairment. Developmental Science, 17(3), 352-365. doi:10.1111/desc.12125.

Kuppuraj, S., Duta, M., Thompson, P. A., & Bishop, D. V. M. (2018). Online incidental statistical learning of audiovisual word sequences in adults - a registered report Royal Society Open Science, 5(2), 171678. doi:10.1098/rsos.171678

Stigler, M & Quast, B. (2015). rddtools: Toolbox for Regression Discontinuity Design ('RDD'). R package version 0.4.0. https://CRAN.R-project.org/package=rddtools

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. Language Learning and Development,, 3, 269-293.

Session information

```r
sessionInfo()
```

```
## R version 3.3.3 (2017-03-06)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: macOS  10.13.4
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] bindrcpp_0.2    rddtools_0.4.0  np_0.60-7       AER_1.2-5
##  [5] survival_2.40-1 sandwich_2.3-4  lmtest_0.9-35   zoo_1.7-14
##  [9] car_2.1-4       beanplot_1.2    stargazer_5.2.1 dplyr_0.7.2
## [13] purrr_0.2.2.2   readr_1.1.1     tidyr_0.6.3     tibble_1.3.3
## [17] ggplot2_2.2.1   tidyverse_1.1.1 plyr_1.8.4
##
## loaded via a namespace (and not attached):
##  [1] httr_1.2.1         jsonlite_1.2       splines_3.3.3
##  [4] modelr_0.1.0       Formula_1.2-1      assertthat_0.1
##  [7] yaml_2.1.14        backports_1.0.5    lattice_0.20-34
## [10] quantreg_5.29      glue_1.1.0         quadprog_1.5-5
## [13] digest_0.6.12      rvest_0.3.2        minqa_1.2.4
## [16] colorspace_1.3-2   htmltools_0.3.6    Matrix_1.2-8
## [19] psych_1.7.3.21     pkgconfig_2.0.1    rdd_0.57
## [22] broom_0.4.2        SparseM_1.74       haven_1.0.0
## [25] scales_0.4.1       locpol_0.6-0       lme4_1.1-13
## [28] cubature_1.3-11    MatrixModels_0.4-1 mgcv_1.8-17
## [31] nnet_7.3-12        lazyeval_0.2.0     pbkrtest_0.4-6
## [34] mnormt_1.5-5       magrittr_1.5       readxl_0.1.1
## [37] evaluate_0.10.1    nlme_3.1-131       MASS_7.3-45
## [40] forcats_0.2.0      xml2_1.1.1         foreign_0.8-67
## [43] tools_3.3.3        hms_0.3            stringr_1.2.0
## [46] munsell_0.4.3      rlang_0.1.1        grid_3.3.3
## [49] nloptr_1.0.4       rmarkdown_1.3      boot_1.3-18
## [52] gtable_0.2.0       reshape2_1.4.2     R6_2.2.0
## [55] lubridate_1.6.0    knitr_1.16         bindr_0.1
## [58] rprojroot_1.2      KernSmooth_2.23-15 stringi_1.1.5
## [61] parallel_3.3.3     Rcpp_0.12.12
```