Multinomial

Yu Du

2020/12/18

```
#datasets.
eng3<-read.csv('eng3.csv')
spa3<-read.csv('spa3.csv')
```

```
#eng3
```

```
#spa3
```

```
eng3$accuracy_spanish<-spa3$accuracy
```

```
mult<-eng3
```

```
colnames(mult)[4]<-'accuracy_english'
```

```
#Change the order of columns.
mult <- mult[, c(1, 2, 3,4,10,5,6,7,8,9)]
```

```
#One NA in the dataset, removed
mult<- mult[-115,]
mult$response <- NA
```

```
#Create levels of responses for the multinomial model
for (i in 1:1187) {
  if (mult$accuracy_english[i]=="1"){
    if(mult$accuracy_spanish[i]=="0"){
     mult$response[i] <- "right in English"
    } else {
       mult$response[i] <- "both right"
    }
  } else {
    if(mult$accuracy_spanish[i]=="0"){
      mult$response[i] <- "both wrong"
    } else{
      mult$response[i] <- "right in Spanish"
    }
  }
}
```

```
#mult
```

```
#Fit the first multinomial model
#b1 <- brm (response ~ cognate+ diff+(1|subject) , data=mult,
                       # family="categorical")
```

```
#Take long time to run.
b1<-read_rds("b1.rds")
```

```
print(summary(b1))
```

```
##  Family: categorical
##   Links: mubothwrong = logit; murightinEnglish = logit; murightinSpanish = logit
## Formula: response ~ cognate + diff + (1 | subject)
##    Data: mult (Number of observations: 1187)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Group-Level Effects:
## ~subject (Number of levels: 27)
##                                 Estimate Est.Error l-95% CI u-95% CI Rhat
## sd(mubothwrong_Intercept)           3.04      0.52     2.19     4.21 1.00
## sd(murightinEnglish_Intercept)      2.06      0.41     1.42     2.95 1.00
## sd(murightinSpanish_Intercept)      2.79      0.59     1.87     4.12 1.00
##                                 Bulk_ESS Tail_ESS
## sd(mubothwrong_Intercept)           1268     1883
## sd(murightinEnglish_Intercept)      1450     2378
## sd(murightinSpanish_Intercept)      1672     2302
##
## Population-Level Effects:
##                             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## mubothwrong_Intercept         -2.89      0.67    -4.23    -1.60 1.00     1114
## murightinEnglish_Intercept    -1.41      0.53    -2.50    -0.39 1.00     1189
## murightinSpanish_Intercept    -2.95      0.73    -4.47    -1.61 1.00     1352
## mubothwrong_cognate           -1.79      0.25    -2.27    -1.31 1.00     3514
## mubothwrong_diff               1.52      0.10     1.33     1.73 1.00     3121
## murightinEnglish_cognate      -1.33      0.23    -1.79    -0.87 1.00     3711
## murightinEnglish_diff          0.73      0.09     0.56     0.92 1.00     3098
## murightinSpanish_cognate      -0.93      0.27    -1.47    -0.37 1.00     3516
## murightinSpanish_diff          0.76      0.11     0.55     0.98 1.00     3610
##                             Tail_ESS
## mubothwrong_Intercept           2320
## murightinEnglish_Intercept      1984
## murightinSpanish_Intercept      2044
## mubothwrong_cognate             3267
## mubothwrong_diff                2988
## murightinEnglish_cognate        3287
## murightinEnglish_diff           3509
## murightinSpanish_cognate        3170
## murightinSpanish_diff           2947
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
#saveRDS(b1,"b1.rds")
```

```
#Fit the second multinomial model with L2AoA predictor and Language predictor
#b3 <-brm (response ~ cognate+ diff+(1|subject)+L2AoA+Language, data=mult,family="categorical")
#print(summary(b3))
```

```
#saveRDS(b3,"b2.rds")
```

```r
loo1<-loo(b1)
#loo2<-loo(b3)
```

```r
#loo_compare(loo1,loo2)
```

```r
#Get some warnings when compare loo of b1 and b3, so refit b3.
#b4 <-brm (response ~ cognate+ diff+(1|subject)+L2AoA+Language, data=mult,family="categorical",save_all
b4<-read_rds("b4.rds")
print(summary(b4))
```

```
##  Family: categorical
##   Links: mubothwrong = logit; murightinEnglish = logit; murightinSpanish = logit
## Formula: response ~ cognate + diff + (1 | subject) + L2AoA + Language
##    Data: mult (Number of observations: 1187)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Group-Level Effects:
## ~subject (Number of levels: 27)
##                                 Estimate Est.Error l-95% CI u-95% CI Rhat
## sd(mubothwrong_Intercept)           3.13      0.58     2.20     4.48 1.00
## sd(murightinEnglish_Intercept)      1.96      0.41     1.30     2.89 1.00
## sd(murightinSpanish_Intercept)      2.28      0.53     1.45     3.51 1.00
##                                 Bulk_ESS Tail_ESS
## sd(mubothwrong_Intercept)           1293     1709
## sd(murightinEnglish_Intercept)      1658     2620
## sd(murightinSpanish_Intercept)      1916     2520
##
## Population-Level Effects:
##                             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## mubothwrong_Intercept         -4.41      1.26    -6.88    -1.85 1.00     1110
## murightinEnglish_Intercept    -0.52      0.94    -2.36     1.31 1.00     1608
## murightinSpanish_Intercept    -5.48      1.16    -7.98    -3.37 1.00     2089
## mubothwrong_cognate           -1.79      0.26    -2.30    -1.30 1.00     3043
## mubothwrong_diff               1.53      0.11     1.33     1.74 1.00     2877
## mubothwrong_L2AoA              0.10      0.09    -0.07     0.27 1.00     1457
## mubothwrong_Language           2.06      1.60    -1.04     5.14 1.00     1537
## murightinEnglish_cognate      -1.33      0.24    -1.83    -0.87 1.00     3500
## murightinEnglish_diff          0.73      0.10     0.55     0.93 1.00     3008
## murightinEnglish_L2AoA        -0.12      0.07    -0.26     0.02 1.00     2015
## murightinEnglish_Language      1.18      1.13    -0.99     3.40 1.00     1543
## murightinSpanish_cognate      -0.93      0.29    -1.49    -0.37 1.00     3823
## murightinSpanish_diff          0.76      0.11     0.55     1.00 1.00     3398
## murightinSpanish_L2AoA         0.23      0.07     0.10     0.39 1.00     2169
## murightinSpanish_Language     -0.26      1.54    -3.41     2.70 1.00     2316
##                             Tail_ESS
## mubothwrong_Intercept           1869
## murightinEnglish_Intercept      2615
## murightinSpanish_Intercept      2313
## mubothwrong_cognate             3279
## mubothwrong_diff                3113
## mubothwrong_L2AoA               2042
## mubothwrong_Language            2181
## murightinEnglish_cognate        3263
```

```
## murightinEnglish_diff          3046
## murightinEnglish_L2AoA         2371
## murightinEnglish_Language      2366
## murightinSpanish_cognate       3087
## murightinSpanish_diff          3399
## murightinSpanish_L2AoA         2355
## murightinSpanish_Language      2489
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
loo3<-loo(b4,moment_match = TRUE)
```

```
## Warning: Some Pareto k diagnostic values are slightly high. See help('pareto-k-diagnostic') for deta
```

```
#saveRDS(b4,"b4.rds")
```

```
loo(b1,b4)
```

```
## Warning: Found 1 observations with a pareto_k > 0.7 in model 'b4'. It is
## recommended to set 'moment_match = TRUE' in order to perform moment matching for
## problematic observations.
```

```
## Output of model 'b1':
##
## Computed from 4000 by 1187 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -893.7 26.0
## p_loo        66.9  3.6
## looic      1787.4 51.9
## ------
## Monte Carlo SE of elpd_loo is 0.2.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     1178  99.2%   1406
##  (0.5, 0.7]   (ok)          9   0.8%   380
##    (0.7, 1]   (bad)         0   0.0%   <NA>
##    (1, Inf)   (very bad)    0   0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'b4':
##
## Computed from 4000 by 1187 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -892.7 26.3
## p_loo        67.0  3.8
## looic      1785.4 52.5
## ------
## Monte Carlo SE of elpd_loo is NA.
##
```

```
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     1180  99.4%   1255
##  (0.5, 0.7]   (ok)          6   0.5%   243
##    (0.7, 1]   (bad)         1   0.1%   337
##    (1, Inf)   (very bad)    0   0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##    elpd_diff se_diff
## b4  0.0       0.0
## b1 -1.0       1.5
```

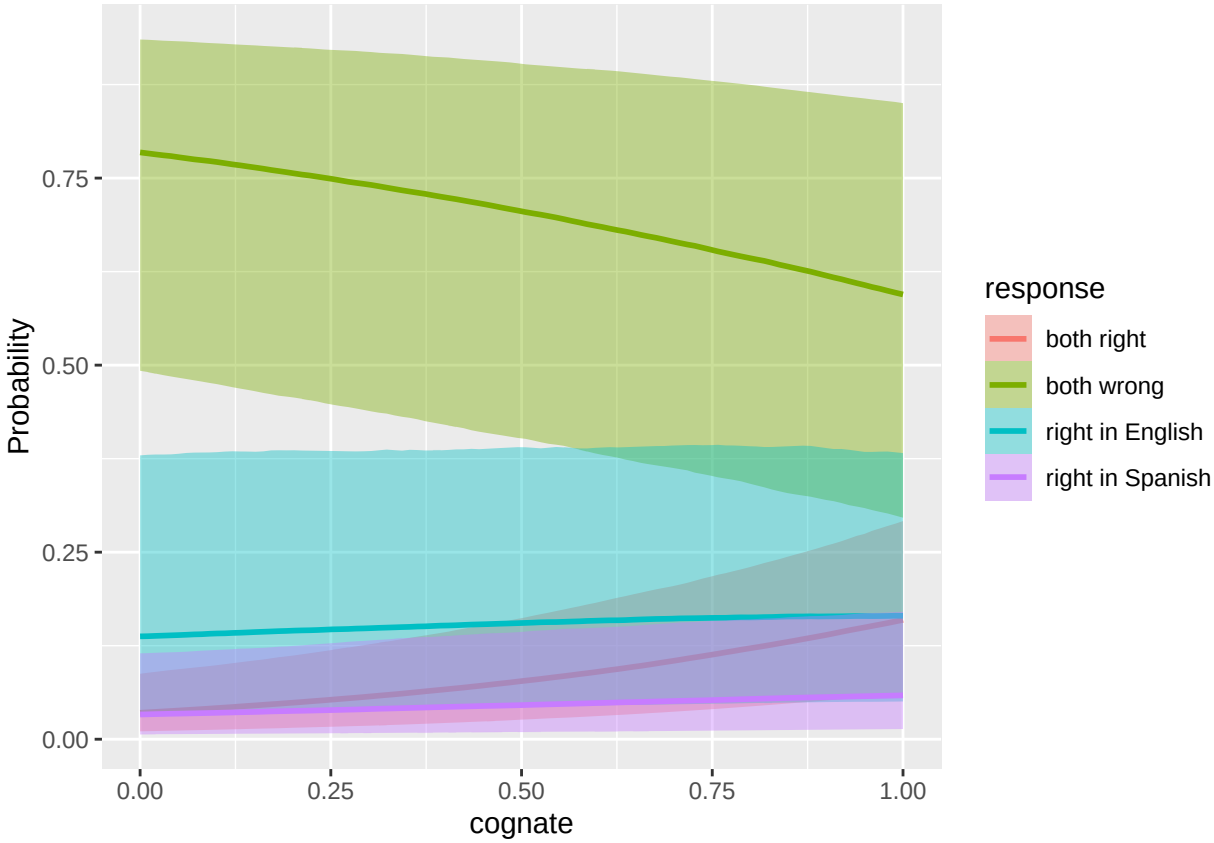
```
#loo compare b1,b4
loo_compare(loo1,loo3)
```

```
##    elpd_diff se_diff
## b4  0.0       0.0
## b1 -1.1       1.5
```

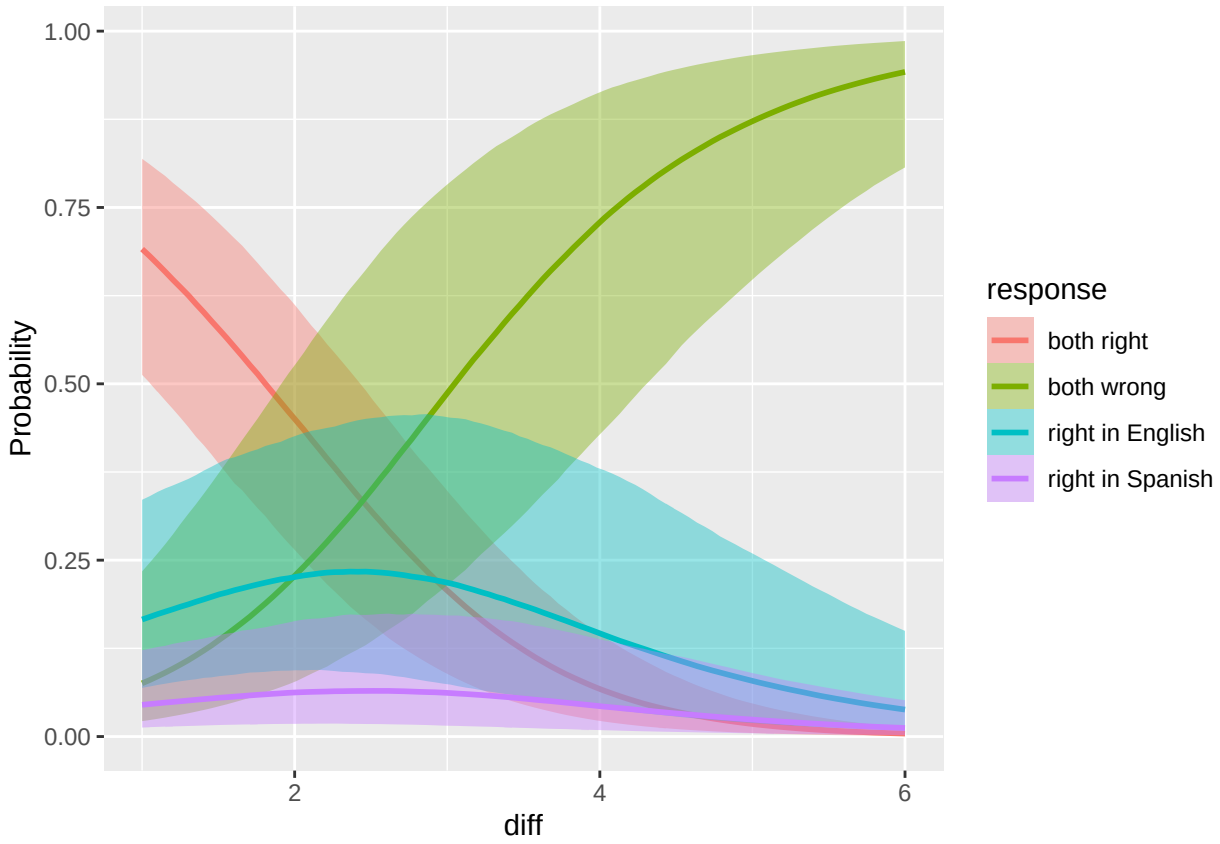
b4 is better from loo_compare.

```
#plot(conditional_effects(b1, effects = "cognate",categorical=TRUE))
```

```
# relationship between response and cognate.
plot(conditional_effects(b4, effects = "cognate",categorical=TRUE))
```

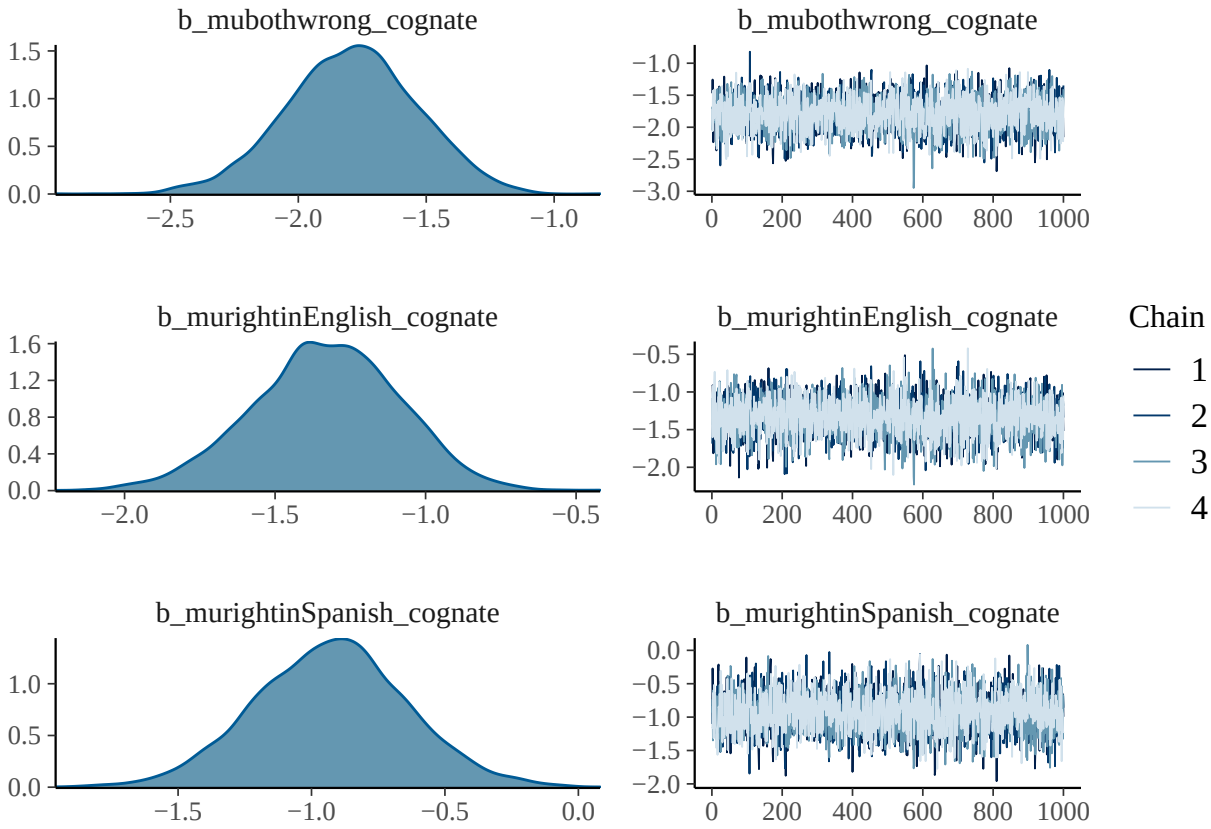



```
# relationship between response and difficulty levels.
plot(conditional_effects(b4, effects = "diff",categorical=TRUE))
```



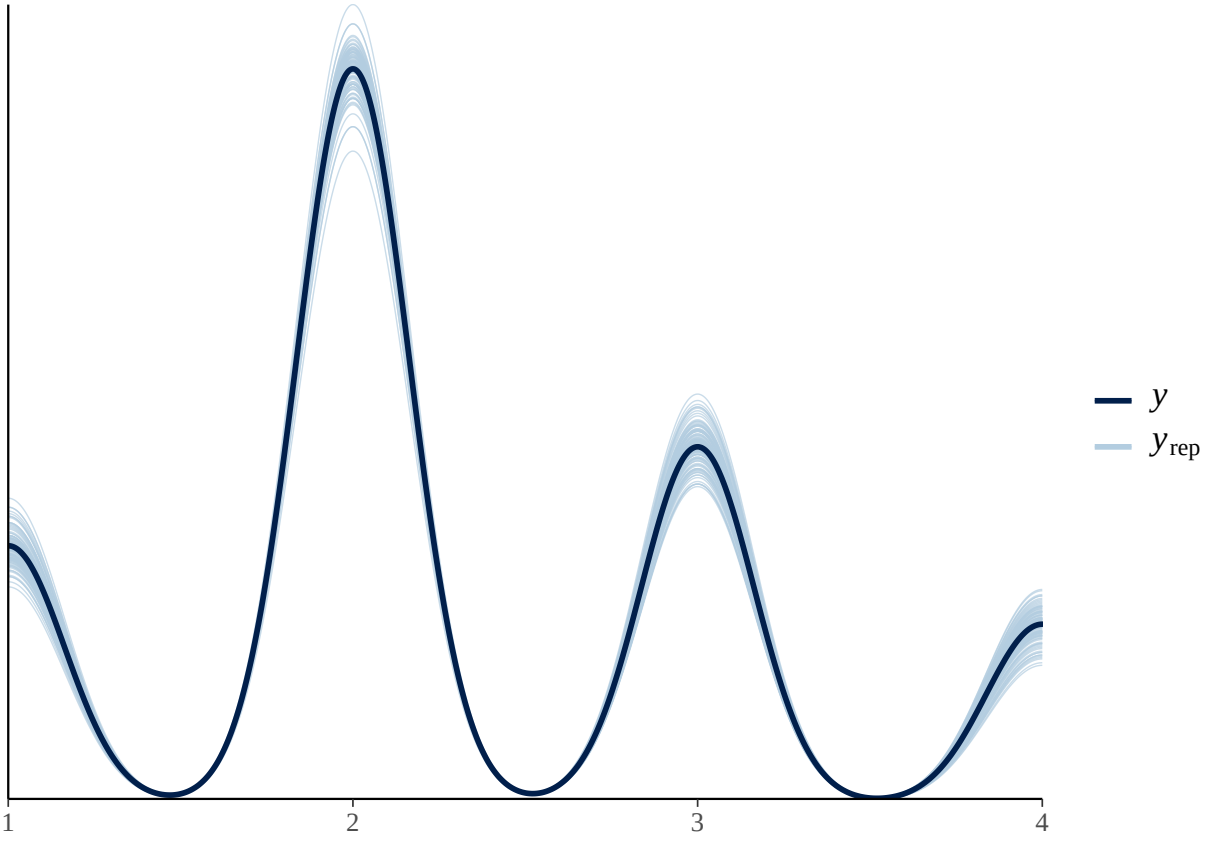

```
#plot(b1, pars = c("cognate"))
```

```
#Trace and density plots of "cognate"
plot(b4, pars = c("cognate"))
```




```
#pp_check(b1,nsamples=100)
```

```
pp_check(b4,nsamples=100)
```




```
fixef(b4)
```

```
##                                Estimate  Est.Error        Q2.5       Q97.5
## mubothwrong_Intercept      -4.4057066 1.26494397 -6.87731305 -1.85434061
## murightinEnglish_Intercept -0.5233598 0.94065620 -2.36130244  1.31080616
## murightinSpanish_Intercept -5.4810273 1.16163121 -7.97902528 -3.37201970
## mubothwrong_cognate        -1.7920185 0.25704147 -2.29918702 -1.30466531
## mubothwrong_diff            1.5311991 0.10586846  1.33140711  1.74287607
## mubothwrong_L2AoA           0.1037413 0.08661631 -0.06618078  0.27387528
## mubothwrong_Language        2.0643988 1.59633641 -1.04331867  5.13946655
## murightinEnglish_cognate   -1.3308796 0.24475438 -1.82828778 -0.87337473
## murightinEnglish_diff       0.7329256 0.09625024  0.55330282  0.93023910
## murightinEnglish_L2AoA     -0.1152004 0.07059270 -0.25848701  0.01728053
## murightinEnglish_Language   1.1782309 1.13124772 -0.98722499  3.39874652
## murightinSpanish_cognate   -0.9279763 0.28535249 -1.48694351 -0.36648914
## murightinSpanish_diff       0.7637423 0.11184371  0.54626180  0.99678768
## murightinSpanish_L2AoA      0.2336463 0.07353648  0.09903893  0.39097991
## murightinSpanish_Language  -0.2640576 1.54328422 -3.40955053  2.70440961
```