

# English Nominal Recursion Study, Part 1: General description of the data

1. General description of the data
2. Condition (ABCE) by Group (Adult-Child): Response is AllTargets
3. Trying to test the interaction of Adult-Child by item (Item is fixed)
4. Condition by Agegroup, DV is AllTargets: Limit the analysis to just kids
5. Trying to do better with Item nested within Condition, kids only.  
Fixed, then random.

```
> rm(list=ls()); options(scipen=999) # To avoid scientific notation
> # Install packages if necessary. Only need to do this once.
> # install.packages("lme4")
> # install.packages("car")
> # Load packages -- do this every time
> library(lme4) # For lmer function
Loading required package: Matrix
> library(car) # For F-tests with p-values, Wald chi-squared tests
>
> # Read data into a data frame
> rdata =
read.table("http://www.utstat.toronto.edu/~brunner/workshops/mixed/Recursion.data.txt",
header=T)
> head(rdata,30)
```

	Participant	AgeinMonths	Group	Agegroup	Condition	Item	StrictTarget	AllTargets	Years	Months	Days	KABC_Hand
1	1	48.5	Child	4yos	A	a1a	1	1	4	0	14	4
2	1	48.5	Child	4yos	A	a1b	0	0	4	0	14	4
3	1	48.5	Child	4yos	A	a2	0	0	4	0	14	4
4	1	48.5	Child	4yos	A	a3	0	0	4	0	14	4
5	1	48.5	Child	4yos	A	a4	0	0	4	0	14	4
6	1	48.5	Child	4yos	A	a5	NA	NA	4	0	14	4
7	1	48.5	Child	4yos	A	a6	0	0	4	0	14	4
8	1	48.5	Child	4yos	B	b1	1	1	4	0	14	4
9	1	48.5	Child	4yos	B	b2	0	0	4	0	14	4
10	1	48.5	Child	4yos	B	b3	1	1	4	0	14	4
11	1	48.5	Child	4yos	B	b4	0	0	4	0	14	4
12	1	48.5	Child	4yos	B	b5	0	0	4	0	14	4
13	1	48.5	Child	4yos	B	b6	0	0	4	0	14	4
14	1	48.5	Child	4yos	C	c1	0	0	4	0	14	4
15	1	48.5	Child	4yos	C	c2	0	0	4	0	14	4
16	1	48.5	Child	4yos	C	c3	0	0	4	0	14	4
17	1	48.5	Child	4yos	C	c4	0	0	4	0	14	4
18	1	48.5	Child	4yos	C	c5	0	0	4	0	14	4
19	1	48.5	Child	4yos	C	c6	0	0	4	0	14	4
20	1	48.5	Child	4yos	E	e1	0	0	4	0	14	4
21	1	48.5	Child	4yos	E	e2	0	0	4	0	14	4
22	1	48.5	Child	4yos	E	e3	NA	NA	4	0	14	4
23	1	48.5	Child	4yos	E	e4	0	0	4	0	14	4
24	1	48.5	Child	4yos	E	e5	0	0	4	0	14	4
25	1	48.5	Child	4yos	E	e6	0	0	4	0	14	4
26	2	63.9	Child	5yos	A	a1a	1	1	5	3	27	7
27	2	63.9	Child	5yos	A	a1b	1	1	5	3	27	7
28	2	63.9	Child	5yos	A	a2	1	1	5	3	27	7
29	2	63.9	Child	5yos	A	a3	0	0	5	3	27	7
30	2	63.9	Child	5yos	A	a4	0	0	5	3	27	7
	CELF_SS	CELF_WS	CELF_EV	CELF_Sent	PPVT	KBIT						
1	16	13	22	21	92	NA						
2	16	13	22	21	92	NA						
3	16	13	22	21	92	NA						
4	16	13	22	21	92	NA						
5	16	13	22	21	92	NA						
6	16	13	22	21	92	NA						
7	16	13	22	21	92	NA						
8	16	13	22	21	92	NA						

9	16	13	22	21	92	NA
10	16	13	22	21	92	NA
11	16	13	22	21	92	NA
12	16	13	22	21	92	NA
13	16	13	22	21	92	NA
14	16	13	22	21	92	NA
15	16	13	22	21	92	NA
16	16	13	22	21	92	NA
17	16	13	22	21	92	NA
18	16	13	22	21	92	NA
19	16	13	22	21	92	NA
20	16	13	22	21	92	NA
21	16	13	22	21	92	NA
22	16	13	22	21	92	NA
23	16	13	22	21	92	NA
24	16	13	22	21	92	NA
25	16	13	22	21	92	NA
26	19	22	29	34	96	NA
27	19	22	29	34	96	NA
28	19	22	29	34	96	NA
29	19	22	29	34	96	NA
30	19	22	29	34	96	NA

> summary(rdata)

Participant	AgeinMonths	Group	Agegroup	Condition	Item	StrictTarget
Min. : 1.00	Min. :48.50	Adult: 325	4yos :625	A:588	ala : 84	Min. :0.0000
1st Qu.: 23.75	1st Qu.:56.00	Child:1775	5yos :625	B:504	alb : 84	1st Qu.:0.0000
Median : 45.50	Median :63.00		6yos :525	C:504	a2 : 84	Median :0.0000
Mean :111.20	Mean :64.62		Adult:325	E:504	a3 : 84	Mean :0.1467
3rd Qu.:159.75	3rd Qu.:75.00				a4 : 84	3rd Qu.:0.0000
Max. :324.00	Max. :83.00				a5 : 84	Max. :1.0000
	NA's :325				(Other):1596	NA's :8
AllTargets	Years	Months	Days	KABC_Hand	CELF_SS	
Min. :0.0000	Min. :4.000	Min. : 0.000	Min. : 0.000	Min. : 3.000	Min. : 9.00	
1st Qu.:0.0000	1st Qu.:4.000	1st Qu.: 3.000	1st Qu.: 0.000	1st Qu.: 5.000	1st Qu.:15.00	
Median :0.0000	Median :5.000	Median : 4.000	Median : 7.000	Median : 6.000	Median :19.00	
Mean :0.2424	Mean :4.944	Mean : 4.972	Mean : 8.958	Mean : 6.901	Mean :17.63	
3rd Qu.:0.0000	3rd Qu.:6.000	3rd Qu.: 7.000	3rd Qu.:16.000	3rd Qu.: 8.000	3rd Qu.:20.00	
Max. :1.0000	Max. :6.000	Max. :11.000	Max. :28.000	Max. :16.000	Max. :22.00	
NA's :8	NA's :325	NA's :325	NA's :325	NA's :325	NA's :325	
CELF_WS	CELF_EV	CELF_Sent	PPVT	KBIT		
Min. : 9.00	Min. :10.00	Min. : 2.00	Min. : 56.0	Min. :15.00		
1st Qu.:17.00	1st Qu.:26.00	1st Qu.:20.00	1st Qu.: 96.0	1st Qu.:18.00		
Median :19.00	Median :30.00	Median :26.00	Median :107.0	Median :20.00		
Mean :18.55	Mean :29.07	Mean :24.93	Mean :107.8	Mean :20.65		
3rd Qu.:21.00	3rd Qu.:34.00	3rd Qu.:32.00	3rd Qu.:122.0	3rd Qu.:24.00		
Max. :24.00	Max. :40.00	Max. :40.00	Max. :146.0	Max. :29.00		
NA's :325	NA's :325	NA's :325	NA's :325	NA's :1600		

```
> attach(rdata)
> table(Item,Condition) # Item is nested within condition.
```

```
      Condition
Item  A  B  C  E
a1a  84  0  0  0
a1b  84  0  0  0
a2   84  0  0  0
a3   84  0  0  0
a4   84  0  0  0
a5   84  0  0  0
a6   84  0  0  0
b1   0  84  0  0
b2   0  84  0  0
b3   0  84  0  0
b4   0  84  0  0
b5   0  84  0  0
b6   0  84  0  0
c1   0  0  84  0
c2   0  0  84  0
c3   0  0  84  0
c4   0  0  84  0
c5   0  0  84  0
c6   0  0  84  0
e1   0  0  0  84
e2   0  0  0  84
e3   0  0  0  84
e4   0  0  0  84
e5   0  0  0  84
e6   0  0  0  84
```

```
> table(Participant,Condition)
```

```
      Condition
Participant A B C E
1          7 6 6 6
2          7 6 6 6
4          7 6 6 6
5          7 6 6 6
6          7 6 6 6
7          7 6 6 6
8          7 6 6 6
9          7 6 6 6
...
319       7 6 6 6
320       7 6 6 6
321       7 6 6 6
322       7 6 6 6
323       7 6 6 6
324       7 6 6 6
```

```
> table(StrictTarget, AllTargets) # AllTargets does include StrictTarget.
```

```
      AllTargets
StrictTarget 0 1
0 1585 200
1 0 307
```

```
> table(AllTargets,Item,Group)
, , Group = Adult
```

	Item																									
AllTargets	ala	alb	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6	c1	c2	c3	c4	c5	c6	e1	e2	e3	e4	e5	e6	
0	1	1	3	2	7	6	5	3	6	3	5	3	2	12	7	8	10	7	3	9	9	10	7	10	13	
1	12	11	10	11	6	7	8	10	7	10	8	10	11	1	6	5	3	6	10	4	4	2	6	3	0	

```
, , Group = Child
```

	Item																									
AllTargets	ala	alb	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6	c1	c2	c3	c4	c5	c6	e1	e2	e3	e4	e5	e6	
0	27	46	62	61	67	46	53	37	63	47	60	37	55	67	59	66	64	55	65	71	63	56	66	69	71	
1	44	25	9	9	4	23	16	34	8	24	11	34	16	4	12	5	7	16	6	0	8	14	5	2	0	

```
> # Note e6 has no yes responses,
> # e1 has no yes responses for children,
> table(StrictTarget,Item,Group)
, , Group = Adult
```

	Item																									
StrictTarget	ala	alb	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6	c1	c2	c3	c4	c5	c6	e1	e2	e3	e4	e5	e6	
0	4	1	4	2	8	9	7	8	10	8	8	8	7	13	10	9	12	12	5	11	9	11	11	12	13	
1	9	11	9	11	5	4	6	5	3	5	5	5	6	0	3	4	1	1	8	2	4	1	2	1	0	

```
, , Group = Child
```

	Item																									
StrictTarget	ala	alb	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6	c1	c2	c3	c4	c5	c6	e1	e2	e3	e4	e5	e6	
0	33	49	62	65	68	47	56	57	69	62	71	54	64	69	64	68	67	63	67	71	67	69	70	70	71	
1	38	22	9	5	3	22	13	14	2	9	0	17	7	2	7	3	4	8	4	0	4	1	1	1	0	

## English Nominal Recursion Study, Part 2: Condition (ABCE) by Group (Adult-Child): Response is AllTargets

```

> rm(list=ls()); options(scipen=999) # To avoid scientific notation
> # Install packages if necessary. Only need to do this once.
> # install.packages("lme4")
> # install.packages("car")
> # Load packages -- do this every time
> library(lme4) # For lmer function
Loading required package: Matrix
> library(car) # For F-tests with p-values, Wald chi-squared tests
>
> # Read data into a data frame
> rdata =
read.table("http://www.utstat.toronto.edu/~brunner/workshops/mixed/Recursion.data.txt",
header=T)
> head(rdata); attach(rdata)
  Participant AgeinMonths Group Agegroup Condition Item StrictTarget AllTargets Years Months Days KABC_Hand
1           1           48.5 Child    4yos         A  a1a             1           1         4         0    14         4
2           1           48.5 Child    4yos         A  alb             0           0         4         0    14         4
3           1           48.5 Child    4yos         A  a2             0           0         4         0    14         4
4           1           48.5 Child    4yos         A  a3             0           0         4         0    14         4
5           1           48.5 Child    4yos         A  a4             0           0         4         0    14         4
6           1           48.5 Child    4yos         A  a5             NA           NA         4         0    14         4
  CELF_SS CELF_WS CELF_EV CELF_Sent PPVT KBIT
1       16       13       22        21    92   NA
2       16       13       22        21    92   NA
3       16       13       22        21    92   NA
4       16       13       22        21    92   NA
5       16       13       22        21    92   NA
6       16       13       22        21    92   NA
> # Effect coding
> contrasts(Group) = contr.sum; contrasts(Condition) = contr.sum
>
> # Adult versus Child
> table1 = table(Group,AllTargets); table1
  AllTargets
Group      0      1
Adult    152    171
Child   1433    336
> prop.table(table1,1) # Row proportions
  AllTargets
Group      0      1
Adult 0.4705882 0.5294118
Child 0.8100622 0.1899378
> table2 = table(Group,StrictTarget); table2
  StrictTarget
Group      0      1
Adult    212    111
Child   1573    196
> prop.table(table2,1) # Row proportions
  StrictTarget
Group      0      1
Adult 0.6563467 0.3436533
Child 0.8892029 0.1107971

> # The difference between adults and children is huge. There is no need to
> # test it by itself.
>

```

```

> # Condition (ABCE) by Group (Adult-Child)
> # Response is AllTargets
> table3 = table(AllTargets,Condition,Group); table3
, , Group = Adult

      Condition
AllTargets  A   B   C   E
      0  25  22  47  58
      1  65  56  31  19

, , Group = Child

      Condition
AllTargets  A   B   C   E
      0 362 299 376 396
      1 130 127  50  29

> prop3A = prop.table(table3[, , 1], 2); prop3A # Adults
      Condition
AllTargets  A           B           C           E
      0 0.2777778 0.2820513 0.6025641 0.7532468
      1 0.7222222 0.7179487 0.3974359 0.2467532
> prop3C = prop.table(table3[, , 2], 2); prop3C # Children
      Condition
AllTargets  A           B           C           E
      0 0.73577236 0.70187793 0.88262911 0.93176471
      1 0.26422764 0.29812207 0.11737089 0.06823529
> # Percent3 will have percent AllTargets response
> Percent3 = 100*rbind(prop3A[2, ], prop3C[2, ])
> rownames(Percent3) = c("Adult", "Child")
> oddsA = Percent3[1, ]/(100-Percent3[1, ]); oddsC = Percent3[2, ]/(100-Percent3[2, ])
> OddsRatio = oddsA/oddsC
> Percent3 = round(rbind(Percent3, OddsRatio), 2); Percent3
      A   B   C   E
Adult  72.22 71.79 39.74 24.68
Child  26.42 29.81 11.74  6.82
OddsRatio  7.24  5.99  4.96  4.47
>
> # Fit and test
> model3 = glmer(AllTargets ~ Group*Condition + (1 | Participant), family=binomial)
Warning message:
In checkConv(attr("opt", "derivs"), opt$par, ctrl = control$checkConv, :
  Model failed to converge with max|grad| = 0.00197264 (tol = 0.001, component 1)
> # Failed to converge, but just barely. Test anyway.
> Anova(model3, type="III") # Each effect controlled for all others
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets
      Chisq Df          Pr(>Chisq)
(Intercept)  26.9935  1  0.000000204145 ***
Group        33.0605  1  0.000000008933 ***
Condition   133.7476  3 < 0.00000000000000022 ***
Group:Condition  1.2555  3  0.7397
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # P-value for interaction = 0.7397

```

```

>
> model3 = glmer(AllTargets ~ Group*Condition + (1 | Participant), family=binomial,
nAGQ=0)
> # nAGQ=0 requests a faster, even less exact approximation of maximum likelihood
> Anova(model3, type="III") # Each effect controlled for all others
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets
          Chisq Df          Pr(>Chisq)
(Intercept)  24.8942  1          0.00000060565 ***
Group        30.5422  1          0.00000003267 ***
Condition    130.1869  3 < 0.000000000000000022 ***
Group:Condition  1.0962  3          0.778
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # P-value for interaction = 0.778
>
> model3 = glmer(AllTargets ~ Group*Condition + (1 | Participant), family=binomial,
nAGQ=2)
> # It worked with MORE precision!
> Anova(model3, type="III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets
          Chisq Df          Pr(>Chisq)
(Intercept)  26.9101  1          0.000000213142 ***
Group        32.9325  1          0.000000009542 ***
Condition    132.5334  3 < 0.000000000000000022 ***
Group:Condition  1.2293  3          0.746
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # P-value for interaction = 0.746
> # Same conclusion: No evidence of interaction. Now we believe it.
> # Playing with the nAGQ parameter seems to be a good way to deal with
> # convergence problems. It's available when there's only a single random
> # effect, ouch. Otherwise glmer uses a LaPlace approximation, which is nAGQ=1.
>

```

# English Nominal Recursion Study, Part 3:

## Trying to test the interaction of Adult-Child by item (Item is fixed)

```
> rm(list=ls()); options(scipen=999) # To avoid scientific notation
> # Install packages if necessary. Only need to do this once.
> # install.packages("lme4")
> # install.packages("car")
> # Load packages -- do this every time
> library(lme4) # For lmer function
Loading required package: Matrix
> library(car) # For F-tests with p-values, Wald chi-squared tests
>
> # Read data into a data frame
> rdata =
read.table("http://www.utstat.toronto.edu/~brunner/workshops/mixed/Recursion.data.txt",
header=T)
> dim(rdata)
[1] 2100 18
> # Eliminate Items e1 and e6
> rdata = subset(rdata,Item != 'e1'); rdata = subset(rdata,Item != 'e6')
> dim(rdata)
[1] 1932 18
> head(rdata); attach(rdata)
  Participant AgeinMonths Group Agegroup Condition Item StrictTarget AllTargets
1           1           48.5 Child    4yos         A   a1a             1           1
2           1           48.5 Child    4yos         A   a1b             0           0
3           1           48.5 Child    4yos         A   a2              0           0
4           1           48.5 Child    4yos         A   a3              0           0
5           1           48.5 Child    4yos         A   a4              0           0
6           1           48.5 Child    4yos         A   a5              NA          NA
  Years Months Days KABC_Hand CELF_SS CELF_WS CELF_EV CELF_Sent PPVT KBIT
1     4     0    14         4      16     13     22      21     92  NA
2     4     0    14         4      16     13     22      21     92  NA
3     4     0    14         4      16     13     22      21     92  NA
4     4     0    14         4      16     13     22      21     92  NA
5     4     0    14         4      16     13     22      21     92  NA
6     4     0    14         4      16     13     22      21     92  NA
> table(Item)
Item
a1a a1b a2 a3 a4 a5 a6 b1 b2 b3 b4 b5 b6 c1 c2 c3 c4 c5 c6 e1
84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 0
e2 e3 e4 e5 e6
84 84 84 84 0
> Item = factor(Item); table(Item); length(table(Item))
Item
a1a a1b a2 a3 a4 a5 a6 b1 b2 b3 b4 b5 b6 c1 c2 c3 c4 c5 c6 e2
84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84
e3 e4 e5
84 84 84
[1] 23
> # Item has 23 levels now, okay.
> contrasts(Group) = contr.sum; contrasts(Condition) = contr.sum
> contrasts(Item) = contr.sum
```



```
>
> table4 = table(AllTargets,Item,Group); table4
, , Group = Adult

      Item
AllTargets a1a a1b a2 a3 a4 a5 a6 b1 b2 b3 b4 b5 b6 c1 c2 c3 c4 c5 c6 e2 e3 e4 e5
0      1   1  3  2  7  6  5  3  6  3  5  3  2 12  7  8 10  7  3  9 10  7 10
1     12  11 10 11  6  7  8 10  7 10  8 10 11  1  6  5  3  6 10  4  2  6  3

, , Group = Child

      Item
AllTargets a1a a1b a2 a3 a4 a5 a6 b1 b2 b3 b4 b5 b6 c1 c2 c3 c4 c5 c6 e2 e3 e4 e5
0     27  46 62 61 67 46 53 37 63 47 60 37 55 67 59 66 64 55 65 63 56 66 69
1     44  25  9  9  4 23 16 34  8 24 11 34 16  4 12  5  7 16  6  8 14  5  2
```

```
> model4 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial, nAGQ=1)
Warning messages:
1: In (function (fn, par, lower = rep.int(-Inf, n), upper = rep.int(Inf, :
failure to converge in 10000 evaluations
2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
Model failed to converge with max|grad| = 0.0254562 (tol = 0.001, component 1)
> # Didn't converge. Test anyway for comparison
> Anova(model4, type="III") # Each effect controlled for all others
Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
Response: AllTargets
      Chisq Df      Pr(>Chisq)
(Intercept) 15.196 1      0.00009692635 ***
Group        30.149 1      0.00000004001 ***
Item        142.811 22 < 0.000000000000000022 ***
Group:Item   40.040 22      0.0107 *
```

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
> # p-value for Group \* Item = 0.0107

```
>
> model4 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial, nAGQ=0)
> Anova(model4, type="III") # Each effect controlled for all others
Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
Response: AllTargets
      Chisq Df      Pr(>Chisq)
(Intercept) 13.695 1      0.0002151 ***
Group        27.544 1      0.0000001536 ***
Item        138.767 22 < 0.000000000000000022 ***
Group:Item   38.309 22      0.0169121 *
```

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
> # P-value for interaction = 0.0169121

```
>
> naive4 = glm(AllTargets ~ Group*Item, family=binomial) # Ignoring repeated
measures
> Anova(naive4, type="III") # Warning about fitted values near zero
Analysis of Deviance Table (Type III tests)
```

```
Response: AllTargets
      LR Chisq Df      Pr(>Chisq)
Group    33492  1 <0.000000000000000002 ***
Item      175 22 <0.000000000000000002 ***
Group:Item   39 22      0.0148 *
```

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
Warning messages:
1: glm.fit: algorithm did not converge
2: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```

> # P-value for interaction = 0.0148
>
> # Several values of nAGQ failed, and also these.

> # model4 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial, nAGQ=1,
> #               contrasts = list(Group="contr.treatment", Item="contr.treatment") )
> # model4 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial, nAGQ=1,
> #               contrasts = list(Group="contr.poly", Item="contr.poly") )
> # model4 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial, nAGQ=1,
> #               contrasts = list(Group="contr.helmert", Item="contr.helmert") )
>
> # There is some indication of an interaction between Group and Item, but I can't
get the model to actually fit. Rather than pursuing it, limit the analysis to just
kids.
>

#####
# So here's a collection of potential ways to deal with convergence problems.
# They did not accomplish anything for me with this data set, but it's nice to know
# how to do them.
#####

# The verbose option lets you see the progress of maximum likelihood.
glmer(AllTargets ~ Item + (1 | Participant), family=binomial, nAGQ=1, verbose=T)

# Choose a set of starting values for the search.
naive = glm(AllTargets ~ Item, family=binomial)
logregbetahats = naive$coef
summary(
glmer(AllTargets ~ Item + (1 | Participant), family=binomial, nAGQ=1, verbose=T,
start = list(fixef=logregbetahats, theta=1) )
) # End of summary
# Failed to converge after 10,000 iterations. This search converged when I let
# glmer choose the starting values.

# My brilliant idea of using a nAGQ=0 run as a source of starting values happens to
be the default.

# Try a different optimizer. The default is "bobyqa"
model44 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial,
               nAGQ=1, control = glmerControl(optimizer = "Nelder_Mead") )

# Specify maximum number of iterations as well as Nelder_Mead
model44 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial,
               nAGQ=1, control = glmerControl(optimizer = "Nelder_Mead",
               optCtrl = list(maxfun=100) )) # Worked!

# Try 100,000 iterations instead of 10,000.
model44 = glmer(AllTargets ~ Group*Item + (1 | Participant), family=binomial,
               nAGQ=1, control = glmerControl( optCtrl = list(maxfun=100000) ))

# What worked for me (sometimes) was trying an equivalent model.

```

## English Nominal Recursion Study, Part 4: Condition by Agegroup, DV is AllTargets: Limit the analysis to kids

```

> # 4. Condition by Agegroup, DV is AllTargets
> # Limit the analysis to just kids and no Item e6
>
> rm(list=ls()); options(scipen=999) # To avoid scientific notation
> # Install packages if necessary. Only need to do this once.
> # install.packages("lme4")
> # install.packages("car")
> # Load packages -- do this every time
> library(lme4) # For lmer function
Loading required package: Matrix
> library(car) # For F-tests with p-values, Wald chi-squared tests
>
> # Read data into a data frame
> rdata =
read.table("http://www.utstat.toronto.edu/~brunner/workshops/mixed/Recursion.data.txt",
header=T)
> # rdata = read.table("Recursion.data.txt", header=T) # Local copy of the data
> dim(rdata)
[1] 2100 18
> # Eliminate adults and Item e6
> rdata = subset(rdata, Item != 'e6'); rdata = subset(rdata, Group == 'Child')
> dim(rdata)
[1] 1704 18
> # Also eliminate 6 rows with NA for AllTargets and StrictTarget. This reduces
> # headaches later.
> rdata = subset(rdata, !is.na(AllTargets))
> dim(rdata)
[1] 1698 18
>
> head(rdata); attach(rdata)
Participant AgeinMonths Group Agegroup Condition Item StrictTarget AllTargets
1 1 48.5 Child 4yos A a1a 1 1
2 1 48.5 Child 4yos A a1b 0 0
3 1 48.5 Child 4yos A a2 0 0
4 1 48.5 Child 4yos A a3 0 0
5 1 48.5 Child 4yos A a4 0 0
7 1 48.5 Child 4yos A a6 0 0
Years Months Days KABC_Hand CELF_SS CELF_WS CELF_EV CELF_Sent PPVT KBIT
1 4 0 14 4 16 13 22 21 92 NA
2 4 0 14 4 16 13 22 21 92 NA
3 4 0 14 4 16 13 22 21 92 NA
4 4 0 14 4 16 13 22 21 92 NA
5 4 0 14 4 16 13 22 21 92 NA
7 4 0 14 4 16 13 22 21 92 NA
> Item = factor(Item); Agegroup = factor(Agegroup) # To eliminate empty levels
> # Effect coding
> contrasts(Agegroup) = contr.sum; contrasts(Condition) = contr.sum

```

```

>
> table5 = table(AllTargets,Condition,Agegroup); table5 # Looks like an interaction
, , Agegroup = 4yos

      Condition
AllTargets  A   B   C   E
0  141 122 140 120
1   31  28  10   4

, , Agegroup = 5yos

      Condition
AllTargets  A   B   C   E
0  139 112 139 121
1   34  38  11   4

, , Agegroup = 6yos

      Condition
AllTargets  A   B   C   E
0   82  65  97  84
1   65  61  29  21

> # Make table of percentages
> prop5.3 = prop.table(table5[, ,1],2)
> prop5.5 = prop.table(table5[, ,2],2)
> prop5.6 = prop.table(table5[, ,3],2)
> PercentRecursion = rbind(prop5.3[2,], prop5.5[2,], prop5.6[2,])
> PercentRecursion = round(100*PercentRecursion,2)
> rownames(PercentRecursion) = c("4yos","5yos","6yos"); PercentRecursion
      A      B      C      E
4yos 18.02 18.67  6.67  3.23
5yos 19.65 25.33  7.33  3.20
6yos 44.22 48.41 23.02 20.00
> round(addmargins(PercentRecursion, FUN=mean, quiet=T),2)
      A      B      C      E  mean
4yos 18.02 18.67  6.67  3.23 11.65
5yos 19.65 25.33  7.33  3.20 13.88
6yos 44.22 48.41 23.02 20.00 33.91
mean 27.30 30.80 12.34  8.81 19.81
>
> # Significance Tests
> model5 = glmer(AllTargets ~ Condition*Agegroup + (1 | Participant),
family=binomial, nAGQ=1)
> Anova(model5, type="III") # p-value for interaction = 0.721
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets

      Chisq Df          Pr(>Chisq)
(Intercept) 169.1986 1 < 0.000000000000000022 ***
Condition    87.5306 3 < 0.000000000000000022 ***
Agegroup     32.2982 2          0.00000009695 ***
Condition:Agegroup  3.6717 6          0.721
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> # Make a table of estimated population mean log odds
> X = model.matrix(model5) # The X matrix!
> summary5 = summary(model5)
> betahat = cbind(summary5$coef[,1]) # Estimated fixed effects as a column vector.
> estlogodds = X %*% betahat # Estimated pop mean log odds for each observation.
>
> aggmeans = aggregate(estlogodds, by = list(Agegroup,Condition), FUN = mean)
> aggmeans

```

```

  Group.1 Group.2      V1
1    4yos      A -1.78883979
2    5yos      A -1.66400533
3    6yos      A -0.26257195
4    4yos      B -1.74671600
5    5yos      B -1.30237792
6    6yos      B -0.06825281
7    4yos      C -3.02596960
8    5yos      C -2.91240497
9    6yos      C -1.37134313
10   4yos      E -3.82377098
11   5yos      E -3.82141335
12   6yos      E -1.57041266
> # aggregate(estlogodds, by = list(Agegroup,Condition), FUN = sd) # All zeros
> treatmeans = aggmeans[,3]
> dim(treatmeans) = c(3,4)
> rownames(treatmeans) = rownames(PercentRecursion)
> colnames(treatmeans) = colnames(PercentRecursion)
> round(addmargins(treatmeans,FUN=mean, quiet=T),2) # With marginal means
      A      B      C      E  mean
4yos -1.79 -1.75 -3.03 -3.82 -2.60
5yos -1.66 -1.30 -2.91 -3.82 -2.43
6yos -0.26 -0.07 -1.37 -1.57 -0.82
mean -1.24 -1.04 -2.44 -3.07 -1.95
> round(addmargins(PercentRecursion, FUN=mean, quiet=T),2) # For comparison
      A      B      C      E  mean
4yos 18.02 18.67  6.67  3.23 11.65
5yos 19.65 25.33  7.33  3.20 13.88
6yos 44.22 48.41 23.02 20.00 33.91
mean 27.30 30.80 12.34  8.81 19.81
>
> # Test pairwise differences between marginal means with a Bonferroni correction
> # First make a combination variable.
>
> # The combination variable AgeCond will have 12 values.
> n = length(AllTargets); n
[1] 1698
> AgeCond = character(n) # A character-valued variable of length n
> for(j in 1:n) AgeCond[j] = paste(Agegroup[j],Condition[j],sep='')
> freq = table(AgeCond); freq
AgeCond
4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
 172   150   150   124   173   150   150   125   147   126   126   105
>
> # Fit a no-intercept model on the combination variable.
> noint = glmer(AllTargets ~ 0+ AgeCond + (1 | Participant), family=binomial)
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
  Model failed to converge with max|grad| = 0.00346365 (tol = 0.001, component 1)

```

```
> nointsum = summary(noint); nointsum
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
Family: binomial ( logit )
Formula: AllTargets ~ 0 + AgeCond + (1 | Participant)
```

```
      AIC      BIC    logLik deviance df.resid
1439.5   1510.2   -706.7   1413.5     1685
```

```
Scaled residuals:
   Min       1Q   Median       3Q      Max
-2.1189 -0.4659 -0.2676 -0.1177  5.2976
```

```
Random effects:
 Groups          Name          Variance Std.Dev.
Participant (Intercept) 0.8784   0.9373
Number of obs: 1698, groups: Participant, 71
```

```
Fixed effects:
              Estimate Std. Error z value      Pr(>|z|)
AgeCond4yosA -1.78928    0.28882  -6.195 0.0000000005826662 ***
AgeCond4yosB -1.74756    0.29745  -5.875 0.0000000042263150 ***
AgeCond4yosC -3.02728    0.39232  -7.716 0.0000000000000120 ***
AgeCond4yosE -3.82506    0.54968  -6.959 0.00000000000034356 ***
AgeCond5yosA -1.66408    0.28286  -5.883 0.0000000040270322 ***
AgeCond5yosB -1.30203    0.28004  -4.649 0.0000033292673787 ***
AgeCond5yosC -2.91164    0.37958  -7.671 0.0000000000000171 ***
AgeCond5yosE -3.82161    0.54872  -6.965 0.00000000000032921 ***
AgeCond6yosA -0.26183    0.27119  -0.965          0.334
AgeCond6yosB -0.06788    0.27994  -0.242          0.808
AgeCond6yosC -1.37091    0.30457  -4.501 0.0000067564687411 ***
AgeCond6yosE -1.57043    0.32995  -4.760 0.0000019399829116 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> treatmeans # For comparison
              A          B          C          E
4yos -1.7888398 -1.74671600 -3.025970 -3.823771
5yos -1.6640053 -1.30237792 -2.912405 -3.821413
6yos -0.2625719 -0.06825281 -1.371343 -1.570413
> # In spite of the no convergence warning, the estimates are close to
> # the ones based on model5. Go with it.
>
> # Rows of Condmat are contrasts for pairwise comparisons of marginal
> # Condition means
> freq
AgeCond
4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
  172   150   150   124   173   150   150   125   147   126   126   105
> Condmat = rbind(
+ # 4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
+ c( 1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0, 0),
+ c( 1, 0, -1, 0, 1, 0, -1, 0, 1, 0, -1, 0),
+ c( 1, 0, 0, -1, 1, 0, 0, -1, 1, 0, 0, -1),
+ c( 0, 1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0),
+ c( 0, 1, 0, -1, 0, 1, 0, -1, 0, 1, 0, -1),
+ c( 0, 0, 1, -1, 0, 0, 1, -1, 0, 0, 1, -1))
> colnames(Condmat) = names(freq)
> rownames(Condmat) = c("AvsB", "AvsC", "AvsE", "BvsC", "BvsE", "CvsE")
```

```

> Condmat
      4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
AvsB      1      -1      0      0      1      -1      0      0      1      -1      0      0
AvsC      1      0      -1      0      1      0      -1      0      1      0      -1      0
AvsE      1      0      0      -1      1      0      0      -1      1      0      0      -1
BvsC      0      1      -1      0      0      1      -1      0      0      1      -1      0
BvsE      0      1      0      -1      0      1      0      -1      0      1      0      -1
CvsE      0      0      1      -1      0      0      1      -1      0      0      1      -1
> # Rows of Agemat are contrasts for pairwise comparisons of marginal
> # Agegroup means
> Agemat = rbind(
+ # 4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
+ c( 1,      1,      1,      1,      -1,      -1,      -1,      -1,      0,      0,      0,      0),
+ c( 1,      1,      1,      1,      0,      0,      0,      0,      -1,      -1,      -1,      -1),
+ c( 0,      0,      0,      0,      1,      1,      1,      1,      -1,      -1,      -1,      -1))
> colnames(Agemat) = names(freq)
> rownames(Agemat) = c("4vs5", "4vs6", "5vs6")
> Agemat
      4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
4vs5      1      1      1      1      -1      -1      -1      -1      0      0      0      0
4vs6      1      1      1      1      0      0      0      0      -1      -1      -1      -1
5vs6      0      0      0      0      1      1      1      1      -1      -1      -1      -1
>
> # Make pairwise comparison matrices. Chi-squared test statistics will be in the
> # upper triangle, and unadjusted p-values in the lower triangle.
> ConditionPairwise = diag(4)
> rownames(ConditionPairwise) = colnames(ConditionPairwise) = c("A","B","C","E")
>
> # Look at one test to see what's going on
> L = Condmat[1,] # Row 1
> testttest = linearHypothesis(noint,L) # Testing H0: L beta = 0
> testttest
Linear hypothesis test

Hypothesis:
AgeCond4yosA - AgeCond4yosB + AgeCond5yosA - AgeCond5yosB + AgeCond6yosA -
AgeCond6yosB = 0

Model 1: restricted model
Model 2: AllTargets ~ 0 + AgeCond + (1 | Participant)

  Df  Chisq Pr(>Chisq)
1
2  1 1.4925    0.2218
> c(testttest[2,2], testttest[2,3])
[1] 1.4924858 0.2218312

```

```

>
> # Fill the ConditionPairwise matrix
> rowno = 0
> for(i in 1:3)
+   {
+     for(j in (i+1):4)
+       {
+         rowno=rowno+1
+         L = Condmat[rowno,]
+         Ltest = linearHypothesis(noint,L) # Testing H0: L beta = 0
+         ConditionPairwise[i,j] = Ltest[2,2] # Test statistic
+         ConditionPairwise[j,i] = Ltest[2,3] # p-value
+       } # Next j
+     } # Next i
> AgePairwise = diag(3)
> rownames(AgePairwise) = colnames(AgePairwise) = c("4yos","5yos","6yos")
> rowno = 0
> for(i in 1:2)
+   {
+     for(j in (i+1):3)
+       {
+         rowno=rowno+1
+         L = Agemat[rowno,]
+         Ltest = linearHypothesis(noint,L) # Testing H0: L beta = 0
+         AgePairwise[i,j] = Ltest[2,2] # Test statistic
+         AgePairwise[j,i] = Ltest[2,3] # p-value
+       } # Next j
+     } # Next i
>
> # Here are the tables of means again.
> round(addmargins(treatmeans,FUN=mean, quiet=T),2) # Estimated population mean log
odds of recursion
      A      B      C      E  mean
4yos -1.79 -1.75 -3.03 -3.82 -2.60
5yos -1.66 -1.30 -2.91 -3.82 -2.43
6yos -0.26 -0.07 -1.37 -1.57 -0.82
mean -1.24 -1.04 -2.44 -3.07 -1.95
> round(addmargins(PercentRecursion, FUN=mean, quiet=T),2) # Percent recursion
      A      B      C      E  mean
4yos 18.02 18.67  6.67  3.23 11.65
5yos 19.65 25.33  7.33  3.20 13.88
6yos 44.22 48.41 23.02 20.00 33.91
mean 27.30 30.80 12.34  8.81 19.81
> # Pairwise comparisons between Condition means A, B, C, E
> ConditionPairwise
      A      B      C      E
A 1.0000000000000000 1.4924857721928317389 34.18923054 43.367609
B 0.22183117846160316 1.0000000000000000000 45.18013910 52.437336
C 0.00000000500049219 0.0000000000179718473 1.00000000 4.358665
E 0.000000000004536355 0.00000000000004441983 0.03682115 1.000000
> # Pairwise comparisons between Agegroup means 4,5,6 yrs old
> AgePairwise
      4yos      5yos      6yos
4yos 1.00000000000000 0.221685065362 25.54517
5yos 0.6377589717567 1.00000000000000 21.24834
6yos 0.0000004321467 0.000004034592 1.000000
> # To protect all 9 tests with a Bonferroni correction, compare p-values to
> 0.05/9
[1] 0.005555556

```



```

>
> # Try continuous age
>
> model6a = glmer( AllTargets ~ AgeinMonths + Condition + (1 | Participant),
+                 family=binomial, nAGQ=1,
+                 contrasts = list(Condition="contr.treatment") )
> summary(model6a) # Pretty nice
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
Family: binomial ( logit )
Formula: AllTargets ~ AgeinMonths + Condition + (1 | Participant)

      AIC      BIC   logLik deviance df.resid
 1429.2  1461.9   -708.6  1417.2    1692

Scaled residuals:
   Min       1Q   Median       3Q      Max
-2.2811 -0.4611 -0.2809 -0.1364  4.8863

Random effects:
 Groups      Name      Variance Std.Dev.
 Participant (Intercept) 0.8524  0.9233
Number of obs: 1698, groups: Participant, 71

Fixed effects:
              Estimate Std. Error z value      Pr(>|z|)
(Intercept) -6.19773     0.91381  -6.782 0.00000000001183 ***
AgeinMonths  0.07574     0.01367   5.542 0.00000002984295 ***
ConditionB   0.20885     0.16256   1.285      0.199
ConditionC  -1.18535     0.19702  -6.017 0.00000000178162 ***
ConditionE  -1.63912     0.23432  -6.995 0.00000000000265 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) AgnMnt CndtnB CndtnC
AgeinMonths -0.984
ConditionB   -0.101  0.015
ConditionC   -0.019 -0.048  0.391
ConditionE   -0.003 -0.053  0.328  0.296
>
> # One could include interaction terms (products), but the test would be for
> # departure from proportional odds; it's too strange.
>
> Anova(model6a, type="III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets
      Chisq Df      Pr(>Chisq)
(Intercept) 45.999  1 0.00000000001183 ***
AgeinMonths 30.718  1 0.00000002984295 ***
Condition   95.500  3 < 0.0000000000000022 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> 5.542^2
[1] 30.71376
> # Right, it's a Wald test: for df=1, Chi-square = Z-squared
>

```

```

> # Include the tests as covariates -- except KBIT (too many NAs)
> # This is very nice. You could have time-varying covariates,
> # Like concentration of ADHD drug in the kid's bloodstream if it were a
> # longitudinal study.
> model6b = glmer( AllTargets ~ AgeinMonths + KABC_Hand + CELF_SS + CELF_WS +
+                 CELF_EV + CELF_Sent + PPVT +
+                 Condition + (1 | Participant), family=binomial, nAGQ=1,
+                 contrasts = list(Condition="contr.treatment") )
Warning messages:
1: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
  Model failed to converge with max|grad| = 0.1301 (tol = 0.001, component 1)
2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
  Model is nearly unidentifiable: very large eigenvalue
- Rescale variables?; Model is nearly unidentifiable: large eigenvalue ratio
- Rescale variables?
>
> # Failure to converge, with suggestions about re-scaling ...
> # Max gradient does not look good.
>
> # Are the independent variables too highly correlated?
> cor(rdata[,c(2,12:17)])
      AgeinMonths KABC_Hand  CELF_SS  CELF_WS  CELF_EV CELF_Sent  PPVT
AgeinMonths  1.0000000  0.5412701  0.5341597  0.5775873  0.4491584  0.5511442  0.7372568
KABC_Hand    0.5412701  1.0000000  0.4944667  0.4625104  0.2813629  0.3725062  0.6203403
CELF_SS      0.5341597  0.4944667  1.0000000  0.7048561  0.5013098  0.6266722  0.7193117
CELF_WS      0.5775873  0.4625104  0.7048561  1.0000000  0.6186514  0.7858615  0.7133241
CELF_EV      0.4491584  0.2813629  0.5013098  0.6186514  1.0000000  0.4885895  0.6537793
CELF_Sent    0.5511442  0.3725062  0.6266722  0.7858615  0.4885895  1.0000000  0.6997446
PPVT         0.7372568  0.6203403  0.7193117  0.7133241  0.6537793  0.6997446  1.0000000
> # Looks quite nice
> # Try standardizing age and the test scores -- 7 of them
> zAgeinMonths = (AgeinMonths-mean(AgeinMonths))/sd(AgeinMonths)
> zKABC_Hand = (KABC_Hand-mean(KABC_Hand))/sd(KABC_Hand)
> zCELF_SS = (CELF_SS-mean(CELF_SS))/sd(CELF_SS)
> zCELF_WS = (CELF_WS-mean(CELF_WS))/sd(CELF_WS)
> zCELF_EV = (CELF_EV-mean(CELF_EV))/sd(CELF_EV)
> zCELF_Sent = (CELF_Sent-mean(CELF_Sent))/sd(CELF_Sent)
> zPPVT = (PPVT-mean(PPVT))/sd(PPVT)
>
> model6b = glmer( AllTargets ~ zAgeinMonths + zKABC_Hand + zCELF_SS + zCELF_WS +
+                 zCELF_EV + zCELF_Sent + zPPVT +
+                 Condition + (1 | Participant), family=binomial, nAGQ=1,
+                 contrasts = list(Condition="contr.treatment") )
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
  Model failed to converge with max|grad| = 0.0103166 (tol = 0.001, component 1)
> # The maximum gradient element is still above tolerance but it's much better.
>
> # I wonder about a no-intercept model
>
> model6c = glmer( AllTargets ~ 0 + zAgeinMonths + zKABC_Hand + zCELF_SS +
+                 zCELF_WS + zCELF_EV + zCELF_Sent + zPPVT +
+                 Condition + (1 | Participant), family=binomial, nAGQ=1)
> # Hoho!
> summary(model6c)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]

Family: binomial (logit)  
Formula: AllTargets ~ 0 + zAgeinMonths + zKABC\_Hand + zCELF\_SS + zCELF\_WS + zCELF\_EV + zCELF\_Sent + zPPVT + Condition + (1 | Participant)

AIC	BIC	logLik	deviance	df.resid
1395.4	1460.6	-685.7	1371.4	1686

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.0537	-0.4768	-0.2721	-0.0978	7.7211

Random effects:

Groups	Name	Variance	Std.Dev.
Participant	(Intercept)	0.2895	0.5381

Number of obs: 1698, groups: Participant, 71

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z )
zAgeinMonths	-0.004526	0.153270	-0.030	0.976442
zKABC_Hand	-0.188966	0.126655	-1.492	0.135705
zCELF_SS	0.254474	0.164264	1.549	0.121339
zCELF_WS	0.183961	0.208588	0.882	0.377812
zCELF_EV	-0.024733	0.157811	-0.157	0.875464
zCELF_Sent	0.690782	0.193351	3.573	0.000353 ***
zPPVT	0.365971	0.233511	1.567	0.117055
ConditionA	-1.378051	0.141966	-9.707	< 0.0000000000000002 ***
ConditionB	-1.170641	0.144448	-8.104	0.00000000000000531 ***
ConditionC	-2.555724	0.188686	-13.545	< 0.0000000000000002 ***
ConditionE	-3.004410	0.228419	-13.153	< 0.0000000000000002 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	zAgnMn	zKABC_	zCELF_SS	zCELF_W	zCELF_E	zCELF_Sn	zPPVT	CndtnA	CndtnB
zKABC_Hand	-0.086								
zCELF_SS	0.003	-0.075							
zCELF_WS	-0.100	-0.190	-0.275						
zCELF_EV	0.075	0.233	0.106	-0.380					
zCELF_Sent	-0.165	0.154	0.067	-0.456	0.061				
zPPVT	-0.462	-0.349	-0.336	0.054	-0.418	-0.186			
ConditionA	0.105	0.041	-0.096	-0.043	0.031	-0.181	-0.047		
ConditionB	0.109	0.041	-0.092	-0.038	0.034	-0.172	-0.047	0.346	
ConditionC	0.074	0.048	-0.082	-0.036	0.028	-0.163	-0.055	0.312	0.293
ConditionE	0.059	0.043	-0.069	-0.031	0.024	-0.139	-0.051	0.265	0.249

CndtnC

zKABC\_Hand  
zCELF\_SS  
zCELF\_WS  
zCELF\_EV  
zCELF\_Sent  
zPPVT  
ConditionA  
ConditionB  
ConditionC  
ConditionE 0.246

```
> Anova(model6c, type="III")
Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
Response: AllTargets
      Chisq Df      Pr(>Chisq)
zAgeinMonths  0.0009  1      0.9764417
zKABC_Hand    2.2260  1      0.1357053
zCELF_SS     2.4000  1      0.1213389
zCELF_WS     0.7778  1      0.3778120
zCELF_EV     0.0246  1      0.8754637
zCELF_Sent   12.7641  1      0.0003533 ***
zPPVT        2.4563  1      0.1170552
Condition    302.8184  4 < 0.00000000000000022 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> # Test for condition is undesirable.
```

```
>
> # Model 6b was not too bad.
> summary(model6b)
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
```

```
Family: binomial ( logit )
Formula: AllTargets ~ zAgeinMonths + zKABC_Hand + zCELF_SS + zCELF_WS +
zCELF_EV + zCELF_Sent + zPPVT + Condition + (1 | Participant)
```

```
      AIC      BIC  logLik deviance df.resid
1395.4  1460.6  -685.7  1371.4    1686
```

```
Scaled residuals:
      Min       1Q   Median       3Q      Max
-2.0533 -0.4770 -0.2722 -0.0977  7.7325
```

```
Random effects:
 Groups      Name      Variance Std.Dev.
Participant (Intercept) 0.2896  0.5382
Number of obs: 1698, groups: Participant, 71
```

```
Fixed effects:
      Estimate Std. Error z value      Pr(>|z|)
(Intercept) -1.378220  0.141984 -9.707 < 0.0000000000000002 ***
zAgeinMonths -0.005213  0.153291 -0.034  0.972870
zKABC_Hand   -0.188511  0.126663 -1.488  0.136676
zCELF_SS     0.253675  0.164275  1.544  0.122538
zCELF_WS     0.186115  0.208632  0.892  0.372355
zCELF_EV    -0.025301  0.157833 -0.160  0.872645
zCELF_Sent   0.690732  0.193376  3.572  0.000354 ***
zPPVT        0.365566  0.233541  1.565  0.117508
ConditionB   0.207371  0.163751  1.266  0.205377
ConditionC  -1.176929  0.197613 -5.956  0.00000000258932 ***
ConditionE  -1.626274  0.234765 -6.927  0.000000000000429 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Correlation of Fixed Effects:

	(Intr)	zAgnMn	zKABC_	zCELF_SS	zCELF_W	zCELF_E	zCELF_Sn	zPPVT	CndtnB
zAgeinMnths	0.106								
zKABC_Hand	0.041	-0.086							
zCELF_SS	-0.096	0.003	-0.075						
zCELF_WS	-0.043	-0.100	-0.190	-0.275					
zCELF_EV	0.031	0.076	0.233	0.106	-0.380				
zCELF_Sent	-0.181	-0.165	0.154	0.067	-0.455	0.060			
zPPVT	-0.047	-0.462	-0.349	-0.336	0.054	-0.418	-0.186		
ConditionB	-0.561	0.005	0.001	0.002	0.004	0.003	0.005	-0.001	
ConditionC	-0.421	-0.005	0.016	-0.009	-0.004	0.005	-0.026	-0.019	0.393
ConditionE	-0.346	-0.007	0.017	-0.009	-0.005	0.005	-0.025	-0.021	0.329

```

CndtnC
zAgeinMnths
zKABC_Hand
zCELF_SS
zCELF_WS
zCELF_EV
zCELF_Sent
zPPVT
ConditionB
ConditionC
ConditionE 0.297

```

convergence code: 0

Model failed to converge with max|grad| = 0.0103166 (tol = 0.001, component 1)

```
> Anova(model6b, type="III")
```

Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets

	Chisq	Df	Pr(>Chisq)
(Intercept)	94.2230	1	< 0.00000000000000022 ***
zAgeinMonths	0.0012	1	0.9728699
zKABC_Hand	2.2150	1	0.1366756
zCELF_SS	2.3846	1	0.1225382
zCELF_WS	0.7958	1	0.3723552
zCELF_EV	0.0257	1	0.8726451
zCELF_Sent	12.7589	1	0.0003543 ***
zPPVT	2.4502	1	0.1175078
Condition	93.6799	3	< 0.00000000000000022 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```

> # Look how close the z statistics are to Model 6c. It looks like CELF_Sent
> # is doing something even controlling for the other covariates. Age and
> # some of the other tests may be washing out each other's effects. Note that
> # this regression model is can't be correct. Age influences maturation, a
> # latent variable that is reflected in the test scores as well as use of
> # sophisticated grammar. Anyway, let's see if the other test scores are
> # significantly related to AllTargets when we control for age, CELF_Sent
> # and Condition. We could get a Wald test with linearHypothesis, but how
> # about a nice likelihood ratio test? Fit a restricted model without
> # zKABC_Hand + zCELF_SS + zCELF_WS + zCELF_EV + zPPVT
>

```

```

> restricted = glmer( AllTargets ~ 0 + zAgeinMonths + zCELF_Sent +
+ Condition + (1 | Participant), family=binomial, nAGQ=1)
> anova(restricted,model6c)

```

```

Data: NULL
Models:
restricted: AllTargets ~ 0 + zAgeinMonths + zCELF_Sent + Condition + (1 |
restricted: Participant)
model6c: AllTargets ~ 0 + zAgeinMonths + zKABC_Hand + zCELF_SS + zCELF_WS +
model6c: zCELF_EV + zCELF_Sent + zPPVT + Condition + (1 | Participant)
      Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
restricted  7 1396.2 1434.3 -691.1  1382.2      5  0.05559 .
model6c    12 1395.4 1460.7 -685.7  1371.4 10.796      5  0.05559 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(restricted)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
Family: binomial ( logit )
Formula: AllTargets ~ 0 + zAgeinMonths + zCELF_Sent + Condition + (1 |
Participant)

      AIC      BIC  logLik deviance df.resid
1396.2  1434.3  -691.1  1382.2     1691

Scaled residuals:
      Min       1Q   Median       3Q      Max
-2.0219 -0.4900 -0.2810 -0.1198  6.0955

Random effects:
 Groups      Name      Variance Std.Dev.
Participant (Intercept) 0.4021  0.6341
Number of obs: 1698, groups: Participant, 71

Fixed effects:
      Estimate Std. Error z value      Pr(>|z|)
zAgeinMonths  0.2236     0.1261  1.773      0.0762 .
zCELF_Sent    0.9788     0.1661  5.894 0.0000000037603410 ***
ConditionA   -1.3533     0.1458 -9.280 < 0.0000000000000002 ***
ConditionB   -1.1450     0.1482 -7.724 0.00000000000000113 ***
ConditionC   -2.5336     0.1914 -13.239 < 0.0000000000000002 ***
ConditionE   -2.9828     0.2305 -12.943 < 0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      zAgnMn zCELF_ CndtnA CndtnB CndtnC
zCELF_Sent -0.567
ConditionA  0.041 -0.249
ConditionB  0.050 -0.233  0.381
ConditionC  0.012 -0.236  0.338  0.320
ConditionE  0.003 -0.205  0.288  0.272  0.264
> # Humm, what do you think?
>
>

```

## English Nominal Recursion Study, Part 5: Item nested within Condition

```

> # 5. Trying to do better with Item nested within Condition
> # Limit the analysis to just kids and no Item e6.
> # Item e1 is also out for children.
> # Eliminate the 6 rows with NA for AllTargets and StrictTarget.
>
> rm(list=ls()); options(scipen=999) # To avoid scientific notation
> # Install packages if necessary. Only need to do this once.
> # install.packages("lme4")
> # install.packages("car")
> # Load packages -- do this every time
> library(lme4) # For lmer function
Loading required package: Matrix
> library(car) # For F-tests with p-values, Wald chi-squared tests
>
> # Read data into a data frame
> rdata =
read.table("http://www.utstat.toronto.edu/~brunner/workshops/mixed/Recursion.data.txt",
header=T)
> dim(rdata)
[1] 2100 18
> # Eliminate adults and Item e6
> rdata = subset(rdata,Item != 'e6'); rdata = subset(rdata,Item != 'e1')
> rdata = subset(rdata,Group == 'Child'); dim(rdata)
[1] 1633 18
> # Also eliminate 6 rows with NA for AllTargets and StrictTarget. This reduces headaches later.
> rdata = subset(rdata, !is.na(AllTargets))
> dim(rdata)
[1] 1627 18
>
> head(rdata); attach(rdata)
  Participant AgeinMonths Group Agegroup Condition Item StrictTarget AllTargets Years Months
1           1           48.5 Child    4yos        A   a1a             1           1           4           0
2           1           48.5 Child    4yos        A   a1b             0           0           4           0
3           1           48.5 Child    4yos        A   a2              0           0           4           0
4           1           48.5 Child    4yos        A   a3              0           0           4           0
5           1           48.5 Child    4yos        A   a4              0           0           4           0
7           1           48.5 Child    4yos        A   a6              0           0           4           0
  Days KABC_Hand CELF_SS CELF_WS CELF_EV CELF_Sent PPVT KBIT
1    14         4      16      13      22      21    92   NA
2    14         4      16      13      22      21    92   NA
3    14         4      16      13      22      21    92   NA
4    14         4      16      13      22      21    92   NA
5    14         4      16      13      22      21    92   NA
7    14         4      16      13      22      21    92   NA
> Item = factor(Item); Agegroup = factor(Agegroup) # To eliminate empty levels
>
> # Effect coding
> # contrasts(Agegroup) = contr.sum; contrasts(Condition) = contr.sum
> # contrasts(Item) = contr.sum
>
> # Quick look
> table(AllTargets,Condition)
      Condition
AllTargets  A   B   C   E
0    362 299 376 254
1    130 127  50  29

```

```

> itemtable = table(AllTargets,Item)
> round( prop.table( itemtable, 2 ), 3)
  Item
AllTargets  ala  alb  a2  a3  a4  a5  a6  b1  b2  b3  b4  b5  b6
0 0.380 0.648 0.873 0.871 0.944 0.667 0.768 0.521 0.887 0.662 0.845 0.521 0.775
1 0.620 0.352 0.127 0.129 0.056 0.333 0.232 0.479 0.113 0.338 0.155 0.479 0.225
  Item
AllTargets  c1  c2  c3  c4  c5  c6  e2  e3  e4  e5
0 0.944 0.831 0.930 0.901 0.775 0.915 0.887 0.800 0.930 0.972
1 0.056 0.169 0.070 0.099 0.225 0.085 0.113 0.200 0.070 0.028
>
> # model8b = glmer(AllTargets ~ Item + (1 | Participant), family=binomial, nAGQ=1) # Failed
> # model8b = glmer(AllTargets ~ Item + (1 | Participant), family=binomial, nAGQ=1,
> # contrasts = list(Item="contr.sum") ) # Failed
> model8b = glmer(AllTargets ~ 0 + Item + (1 | Participant), family=binomial, nAGQ=1) # No
intercept
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
Model failed to converge with max|grad| = 0.0020057 (tol = 0.001, component 1)
> # This version of 8b failed but not too badly: maxgrad = 0.0020057 tol = 0.001
> # Go with it.
>
> # Compose contrast matrices for Item nested within Condition. Note these are fixed effects.
>
> ItemWithinConditionA = rbind(
+ # ala alb a2 a3 a4 a5 a6 b1 b2 b3 b4 b5 b6 c1 c2 c3 c4 c5 c6 e2 e3 e4 e5
+ c( 1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 1, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 1, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 1, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 1, 0, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0))
> ItemWithinConditionB = rbind(
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0))
> ItemWithinConditionC = rbind(
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, -1, 0, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, -1, 0, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, -1, 0, 0, 0, 0))
> ItemWithinConditionE = rbind(
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, -1, 0, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, -1, 0),
+ c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, -1))
> ItemWithinCondition = rbind(ItemWithinConditionA,
+ ItemWithinConditionB,
+ ItemWithinConditionC,
+ ItemWithinConditionE)
>

```



```
> # Test Item within Condition overall
> linearHypothesis(model8b,ItemWithinCondition)
Linear hypothesis test
```

```
Hypothesis:
Itemala - Itemalb = 0
Itemala - Itema2 = 0
Itemala - Itema3 = 0
Itemala - Itema4 = 0
Itemala - Itema5 = 0
Itemala - Itema6 = 0
Itemb1 - Itemb2 = 0
Itemb1 - Itemb3 = 0
Itemb1 - Itemb4 = 0
Itemb1 - Itemb5 = 0
Itemb1 - Itemb6 = 0
Itemc1 - Itemc2 = 0
Itemc1 - Itemc3 = 0
Itemc1 - Itemc4 = 0
Itemc1 - Itemc5 = 0
Itemc1 - Itemc6 = 0
Iteme2 - Iteme3 = 0
Iteme2 - Iteme4 = 0
Iteme2 - Iteme5 = 0
```

```
Model 1: restricted model
Model 2: AllTargets ~ 0 + Item + (1 | Participant)
```

```
   Df  Chisq      Pr(>Chisq)
1    1      NA              NA
2  19 153.88 < 0.00000000000000022 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> # Now within each condition
> linearHypothesis(model8b,ItemWithinConditionA)
Linear hypothesis test
```

```
Hypothesis:
Itemala - Itemalb = 0
Itemala - Itema2 = 0
Itemala - Itema3 = 0
Itemala - Itema4 = 0
Itemala - Itema5 = 0
Itemala - Itema6 = 0
```

```
Model 1: restricted model
Model 2: AllTargets ~ 0 + Item + (1 | Participant)
```

```
   Df  Chisq      Pr(>Chisq)
1    1      NA              NA
2    6  84.235 0.0000000000000004755 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> linearHypothesis(model8b, ItemWithinConditionB)
Linear hypothesis test
```

```
Hypothesis:
Itemb1 - Itemb2 = 0
Itemb1 - Itemb3 = 0
Itemb1 - Itemb4 = 0
Itemb1 - Itemb5 = 0
Itemb1 - Itemb6 = 0
```

```
Model 1: restricted model
Model 2: AllTargets ~ 0 + Item + (1 | Participant)
```

```
   Df Chisq    Pr(>Chisq)
1    5 52.72 0.000000000384 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> linearHypothesis(model8b, ItemWithinConditionC)
Linear hypothesis test
```

```
Hypothesis:
Itemc1 - Itemc2 = 0
Itemc1 - Itemc3 = 0
Itemc1 - Itemc4 = 0
Itemc1 - Itemc5 = 0
Itemc1 - Itemc6 = 0
```

```
Model 1: restricted model
Model 2: AllTargets ~ 0 + Item + (1 | Participant)
```

```
   Df  Chisq Pr(>Chisq)
1    5 17.076 0.004358 **
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> linearHypothesis(model8b, ItemWithinConditionE)
Linear hypothesis test
```

```
Hypothesis:
Iteme2 - Iteme3 = 0
Iteme2 - Iteme4 = 0
Iteme2 - Iteme5 = 0
```

```
Model 1: restricted model
Model 2: AllTargets ~ 0 + Item + (1 | Participant)
```

```
   Df  Chisq Pr(>Chisq)
1    3 12.767 0.005169 **
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

> # Pairwise comparisons of items within conditions are easy and probably a good idea.
>
> # There is an easier but less straightforward way to test for Item within Condition.
> # Fit a no-intercept model (model8a) for just Condition. Then 8b is a more refined version
> # of 8a and the difference is Item nested within Condition.
> # Compare this likelihood ratio test to the Wald test with Chi-squared = 153.88, df = 19.
> model8a = glmer(AllTargets ~ 0 + Condition + (1 | Participant), family=binomial, nAGQ=1)
> anova(model8a,model8b)
Data: NULL
Models:
model8a: AllTargets ~ 0 + Condition + (1 | Participant)
model8b: AllTargets ~ 0 + Item + (1 | Participant)
      Df    AIC    BIC  logLik deviance Chisq Chi Df          Pr(>Chisq)
model8a  5 1438.9 1465.9 -714.46   1428.9
model8b 24 1283.9 1413.4 -617.96   1235.9   193    19 < 0.000000000000000022 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Chisq = 193, df = 19, p = 0.000000000000000022
> # Note the df is 19 = 6 + 5 + 5 + 3, exactly right for Item within Condition.
>
> #####
> # Now try Item as a random effect. Items are naturally nested within Condition by
> # the way they are named. If Item = 1, 2, etc. meant different things within each
> # condition, this would have to be indicated in the glmer syntax.
>
> model8c = glmer(AllTargets ~ Condition + (1 | Participant) + (1 | Item),
+               family=binomial, nAGQ=1)
> summary(model8c) # Note variance estimates.
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial (logit)
Formula: AllTargets ~ Condition + (1 | Participant) + (1 | Item)

      AIC      BIC  logLik deviance df.resid
1316.7  1349.1  -652.4  1304.7    1621

Scaled residuals:
  Min       1Q   Median       3Q      Max
-3.0788 -0.4232 -0.2354 -0.0896  9.6590

Random effects:
 Groups      Name      Variance Std.Dev.
Participant (Intercept)  2.200    1.4832
Item        (Intercept)  0.928    0.9634
Number of obs: 1627, groups: Participant, 71; Item, 23

Fixed effects:
      Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.6308     0.4291  -3.800 0.000145 ***
ConditionB   0.3002     0.5666   0.530 0.596288
ConditionC  -1.2063     0.5782  -2.086 0.036973 *
ConditionE  -1.4517     0.6599  -2.200 0.027806 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) CndtnB CndtnC
ConditionB -0.617
ConditionC -0.597  0.456
ConditionE -0.522  0.399  0.395

```

```

> # Let's add Agegroup and the interaction. Use effect coding
> contrasts(Agegroup) = contr.sum; contrasts(Condition) = contr.sum
>
> model8d = glmer(AllTargets ~ Agegroup*Condition + (1 | Participant) + (1 | Item),
+               family=binomial)
> summary(model8d)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: AllTargets ~ Agegroup * Condition + (1 | Participant) + (1 | Item)

      AIC      BIC   logLik deviance df.resid
 1302.8  1378.3  -637.4  1274.8    1613

Scaled residuals:
      Min       1Q   Median       3Q      Max
-2.9291 -0.4168 -0.2287 -0.0886  6.8836

Random effects:
 Groups      Name      Variance Std.Dev.
Participant (Intercept) 1.3271  1.1520
Item        (Intercept) 0.9278  0.9632
Number of obs: 1627, groups: Participant, 71; Item, 23

Fixed effects:
              Estimate Std. Error z value      Pr(>|z|)
(Intercept)   -2.17144    0.27222  -7.977 0.0000000000000015 ***
Agegroup1     -0.75868    0.24154  -3.141  0.00168 **
Agegroup2     -0.56451    0.23941  -2.358  0.01838 *
Condition1     0.63492    0.35544   1.786  0.07405 .
Condition2     0.93663    0.37062   2.527  0.01150 *
Condition3    -0.55403    0.38070  -1.455  0.14559
Agegroup1:Condition1 0.06970    0.18700   0.373  0.70937
Agegroup2:Condition1 0.02820    0.18417   0.153  0.87829
Agegroup1:Condition2 -0.07506    0.18857  -0.398  0.69057
Agegroup2:Condition2 0.24722    0.18340   1.348  0.17766
Agegroup1:Condition3 0.10638    0.22623   0.470  0.63817
Agegroup2:Condition3 0.03258    0.22223   0.147  0.88346
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) Aggrp1 Aggrp2 Cndtn1 Cndtn2 Cndtn3 Ag1:C1 Ag2:C1 Ag1:C2 Ag2:C2 Ag1:C3
Agegroup1  0.046
Agegroup2  0.029 -0.519
Condition1 -0.149 -0.023 -0.019
Condition2 -0.094 -0.021 -0.026 -0.220
Condition3 -0.035  0.000 -0.003 -0.247 -0.275
Aggrp1:Cnd1 -0.034 -0.231  0.168  0.052  0.013 -0.005
Aggrp2:Cnd1 -0.029  0.169 -0.231  0.037  0.022  0.002 -0.597
Aggrp1:Cnd2 -0.028 -0.221  0.170  0.013  0.050 -0.006  0.036 -0.079
Aggrp2:Cnd2 -0.046  0.169 -0.243  0.025  0.025  0.008 -0.080  0.057 -0.589
Aggrp1:Cnd3 -0.005 -0.015  0.012 -0.004 -0.004  0.070 -0.170  0.083 -0.175  0.081
Aggrp2:Cnd3 -0.007  0.013 -0.019  0.002  0.007  0.050  0.083 -0.162  0.080 -0.159 -0.635
> Anova(model8d, type="III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: AllTargets
      Chisq Df      Pr(>Chisq)
(Intercept)  63.6298  1 0.000000000000001501 ***
Agegroup     31.6554  2 0.000000133693299449 ***
Condition    12.1541  3      0.006873 **
Agegroup:Condition  3.9003  6      0.690166
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

```

> # To see what's happening, earlier code may be copied almost exactly.
>
> # Make a table of estimated population mean log odds
> X = model.matrix(model8d) # The X matrix
> summary8 = summary(model8d)
> betahat = cbind(summary8$coef[,1]) # Estimated fixed effects as a column vector.
> estlogodds = X %*% betahat # Estimated pop mean log odds for each observation.
>
> # Just look at marginal estimated population mean log odds
> aggregate(estlogodds, by = list(Agegroup), FUN = mean)
  Group.1      V1
1     4yos -2.8030692
2     5yos -2.5936661
3     6yos -0.7719827
> aggregate(estlogodds, by = list(Condition), FUN = mean)
  Group.1      V1
1         A -1.599870
2         B -1.299659
3         C -2.792190
4         E -3.283842
>
> # Test pairwise differences between marginal means with a Bonferroni correction
> # First make a combination variable.
>
> # The combination variable AgeCond will have 12 values.
> n = length(AllTargets); n
[1] 1627
> AgeCond = character(n) # A character-valued variable of length n
> for(j in 1:n) AgeCond[j] = paste(Agegroup[j],Condition[j],sep='')
> freq = table(AgeCond); freq
AgeCond
4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
  172   150   150   99   173   150   150   100   147   126   126   84
>
> # Fit a no-intercept model on the combination variable.
> noint = glmer(AllTargets ~ 0 + AgeCond + (1 | Participant), family=binomial)
Warning message:
In checkConv(attr("opt", "derivs"), opt$par, ctrl = control$checkConv, :
  Model failed to converge with max|grad| = 0.00251889 (tol = 0.001, component 1)
> nointsum = summary(noint); nointsum
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: AllTargets ~ 0 + AgeCond + (1 | Participant)

      AIC      BIC   logLik deviance df.resid
1424.4  1494.5  -699.2  1398.4    1614

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.1596 -0.4796 -0.2731 -0.1325  4.8617

Random effects:
 Groups      Name      Variance Std.Dev.
Participant (Intercept) 0.8895  0.9431
Number of obs: 1627, groups: Participant, 71

Fixed effects:
      Estimate Std. Error z value Pr(>|z|)
AgeCond4yosA -1.79096    0.28970  -6.182 0.0000000006323229 ***
AgeCond4yosB -1.74933    0.29831  -5.864 0.0000000045157394 ***
AgeCond4yosC -3.02881    0.39289  -7.709 0.0000000000000127 ***
AgeCond4yosE -3.58580    0.55346  -6.479 0.0000000000923854 ***
AgeCond5yosA -1.66621    0.28376  -5.872 0.0000000043084812 ***
AgeCond5yosB -1.30435    0.28097  -4.642 0.0000034442719780 ***
AgeCond5yosC -2.91566    0.38038  -7.665 0.0000000000000179 ***

```

```

AgeCond5yosE -3.58566      0.55218    -6.494  0.0000000000837954 ***
AgeCond6yosA -0.26245      0.27232    -0.964      0.335165
AgeCond6yosB -0.06752      0.28104    -0.240      0.810132
AgeCond6yosC -1.37449      0.30578    -4.495  0.0000069576380643 ***
AgeCond6yosE -1.25218      0.33821    -3.702      0.000214 ***

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Correlation of Fixed Effects:

```

      AgCn4A AgCn4B AgCn4C AgCn4E AgCn5A AgCn5B AgCn5C AgCn5E AgCn6A AgCn6B AgCn6C
AgeCond4ysB 0.473
AgeCond4ysC 0.375  0.364
AgeCond4ysE 0.269  0.261  0.213
AgeCond5ysA 0.025  0.024  0.025  0.019
AgeCond5ysB 0.022  0.021  0.022  0.016  0.496
AgeCond5ysC 0.026  0.025  0.025  0.019  0.387  0.383
AgeCond5ysE 0.019  0.018  0.019  0.014  0.269  0.266  0.215
AgeCond6ysA 0.002  0.002  0.002  0.001  0.002  0.001  0.002  0.001
AgeCond6ysB 0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.556
AgeCond6ysC 0.009  0.009  0.009  0.007  0.008  0.007  0.008  0.006  0.512  0.494
AgeCond6ysE 0.007  0.007  0.007  0.005  0.007  0.006  0.007  0.005  0.463  0.447  0.422

```

convergence code: 0

Model failed to converge with max|grad| = 0.00251889 (tol = 0.001, component 1)

```
> # In spite of the no convergence warning, go with it.
```

```
>
> # Rows of Condmat are contrasts for pairwise comparisons of marginal
> # Condition means
> freq
```

```
AgeCond
4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
  172   150   150   99   173   150   150   100   147   126   126   84
```

```
> Condmat = rbind(
+ # 4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
+ c( 1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0, 0),
+ c( 1, 0, -1, 0, 1, 0, -1, 0, 1, 0, -1, 0),
+ c( 1, 0, 0, -1, 1, 0, 0, -1, 1, 0, 0, -1),
+ c( 0, 1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0),
+ c( 0, 1, 0, -1, 0, 1, 0, -1, 0, 1, 0, -1),
+ c( 0, 0, 1, -1, 0, 0, 1, -1, 0, 0, 1, -1))
```

```
> colnames(Condmat) = names(freq)
```

```
> rownames(Condmat) = c("AvsB", "AvsC", "AvsE", "BvsC", "BvsE", "CvsE")
```

```
> Condmat
```

```

      4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
AvsB      1    -1     0     0     1    -1     0     0     1    -1     0     0
AvsC      1     0    -1     0     1     0    -1     0     1     0    -1     0
AvsE      1     0     0    -1     1     0     0    -1     1     0     0    -1
BvsC      0     1    -1     0     0     1    -1     0     0     1    -1     0
BvsE      0     1     0    -1     0     1     0    -1     0     1     0    -1
CvsE      0     0     1    -1     0     0     1    -1     0     0     1    -1

```

```
> # Rows of Agemat are contrasts for pairwise comparisons of marginal
```

```
> # Agegroup means
```

```
> Agemat = rbind(
+ # 4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
+ c( 1, 1, 1, 1, -1, -1, -1, -1, 0, 0, 0, 0),
+ c( 1, 1, 1, 1, 0, 0, 0, 0, -1, -1, -1, -1),
+ c( 0, 0, 0, 0, 1, 1, 1, 1, -1, -1, -1, -1))
```

```
> colnames(Agemat) = names(freq)
```

```
> rownames(Agemat) = c("4vs5", "4vs6", "5vs6")
```

```

> Agemat
  4yosA 4yosB 4yosC 4yosE 5yosA 5yosB 5yosC 5yosE 6yosA 6yosB 6yosC 6yosE
4vs5      1      1      1      1      -1     -1     -1     -1      0      0      0      0
4vs6      1      1      1      1      0      0      0      0     -1     -1     -1     -1
5vs6      0      0      0      0      1      1      1      1     -1     -1     -1     -1
>
> # Make pairwise comparison matrices. Chi-squared test statistics will be in the
> # upper triangle, and unadjusted p-values in the lower triangle.
> ConditionPairwise = diag(4)
> rownames(ConditionPairwise) = colnames(ConditionPairwise) = c("A","B","C","E")
>
> # Fill the ConditionPairwise matrix
> rowno = 0
> for(i in 1:3)
+   {
+     for(j in (i+1):4)
+       {
+         rowno=rowno+1
+         L = Condmat[rowno,]
+         Lttest = linearHypothesis(noint,L) # Testing H0: L beta = 0
+         ConditionPairwise[i,j] = Lttest[2,2] # Test statistic
+         ConditionPairwise[j,i] = Lttest[2,3] # p-value
+       } # Next j
+     } # Next i
> AgePairwise = diag(3)
> rownames(AgePairwise) = colnames(AgePairwise) = c("4yos","5yos","6yos")
> rowno = 0
> for(i in 1:2)
+   {
+     for(j in (i+1):3)
+       {
+         rowno=rowno+1
+         L = Agemat[rowno,]
+         Lttest = linearHypothesis(noint,L) # Testing H0: L beta = 0
+         AgePairwise[i,j] = Lttest[2,2] # Test statistic
+         AgePairwise[j,i] = Lttest[2,3] # p-value
+       } # Next j
+     } # Next i
>
> # Here are the marginal means again.
> aggregate(estlogodds, by = list(Agegroup), FUN = mean)
Group.1      V1
1      4yos -2.8030692
2      5yos -2.5936661
3      6yos -0.7719827
> aggregate(estlogodds, by = list(Condition), FUN = mean)
Group.1      V1
1      A -1.599870
2      B -1.299659
3      C -2.792190
4      E -3.283842
>
> # Pairwise comparisons between Condition means A, B, C, E
> ConditionPairwise
      A      B      C      E
A 1.0000000000000000 1.49427865008775673 34.2397307 31.159399
B 0.221553788003150 1.00000000000000000 45.2413706 38.967502
C 0.0000000004872398 0.00000000001741864 1.0000000 1.439196
E 0.0000000023768574 0.00000000043091968 0.2302695 1.000000

```

```
> # Pairwise comparisons between Agegroup means 4,5,6 yrs old
> AgePairwise
      4yos      5yos      6yos
4yos 1.000000000000 0.215093680964 25.79351
5yos 0.6428039891223 1.000000000000 21.54594
6yos 0.0000003799642 0.000003454529 1.00000
> # To protect all 9 tests with a Bonferroni correction, compare p-values to
> 0.05/9
[1] 0.005555556
>
> # Same conclusions: 6year olds use more recursion than 4 and 5, and AB elicits more
> # recursion than CE. Could start adding covariates ...
>
```

This handout was prepared by Jerry Brunner, Department of Statistical Sciences, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The OpenOffice.org document is available from the course website:

<http://www.utstat.toronto.edu/~brunner/oldclass/appliedf17>