Week4 MLM

Zoë Hawks

9/28/2017

Contents

Problem 1: 1
Problem 2: 2
Problem 3: 2
Problem 4A: 6
Problem 4B: 7
Problem 5: 8
Problem 6: 9

Problem 1:

Run a series of models using a time-invariant nominal covariate. a) where the covariate only predicts the intercept b) predicts both intercept and slope c) is rescaled eg centering. For all models, how does your model change from model to model. What is your final model? *(From last week's homework, so code is suppressed.)*

Part A: Fixed effect estimates

- `Intercept` = 40.57; mean of Controls at mean age (age.centered = 0)
- `Age.centered` = 2.05; increase in words correct/year controlling for Group
- `Group` = -1.4; difference in words correct between Control and PKU controlling for mean age
- `Pseudo-R2` = .38 (marginal; fixed) and .69 (conditional; fixed + random)

Part B: Fixed effect estimates

- `Intercept` = 40.93; mean of Controls at mean age (age.centered = 0)
- `Age.centered` = 2.33; increase in words correct/year for Controls
- `Group` = -2.25; difference in words correct between Control and PKU age mean age (age.centered = 0)
- `Age.centered:GROUP` = -.77; difference in slope between Control and PKU
- `Pseudo-R2` = .38 (marginal; fixed) and .70 (conditional; fixed + random)

Part C: Fixed effect estimates

- `Intercept` = 38.69; mean of PKU at mean age (age.centered = 0)
- `Age.centered` = 1.56; increase in words correct/year for PKU
- `Group` = 2.25; difference in words correct between PKU and Controls at mean age (age.centered = 0)
- `Age.centered:GROUP` = .77; difference in slope between PKU and Control
- `Pseudo-R2` = .38 (marginal; fixed) and .70 (conditional; fixed + random)

Likelihood ratio test suggests that simpler model, where covariate only predicts intercept, is preferred.

Problem 2:

Run a series of models using a time-invariant continuous covariate. *(From last week's homework, so output is suppressed.)*

Part A: Fixed effect estimates

- `Intercept` = 37.46; mean across all participants at means levels of age (i.e., age.centered = 0) and when baselinePho = 0
- `Age.centered` = 1.91; slope (increase in words correct/year) controlling for baselinePho
- `baselinePho` = .13; slope (increase in words correct/unit of baselinePho) controlling for Age.centered
- `Pseudo-R2` = .39 (marginal; fixed) and .68 (conditional; fixed + random)

Part B: Fixed effect estimates

- `Intercept` = 37.46; mean across all participants at mean age (age.centered = 0) and when baselinePho = 0
- `Age.centered` = 1.80; slope (increase in words correct/year) when baselinePho = 0
- `baselinePho` = .13; slope (increase in words correct/unit of baselinePho) at mean age (age.centered = 0)
- `Age.centered:baselinePho` = .006; extent to which relationship between age.centered and Sem_Total_Correct changes at different levels of baselinePho (change is minimal)
- `Pseudo-R2` = .39 (marginal; fixed) and .68 (conditional; fixed + random)

Part C: Fixed effect estimates

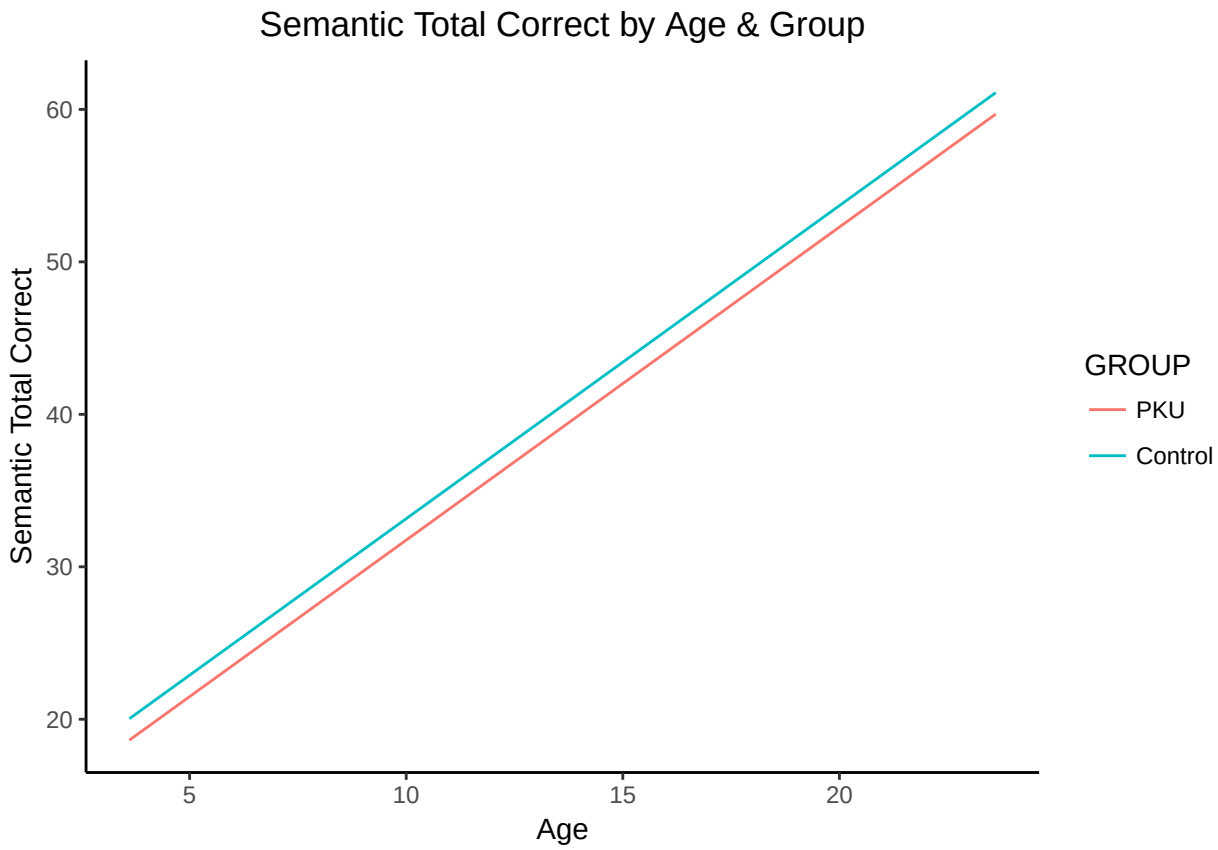
- `Intercept` = 39.94; mean across all participants at means levels of age and baselinePho (i.e., age.centered and baselinePho.centered = 0)
- `Age.centered` = 1.90; slope (increase in words correct/year) at mean levels of baselinePho (baselinePho.centered = 0)
- `baselinePho` = .13; slope (increase in words correct/unit of baselinePho) at mean age (age.centered = 0)
- `Age.centered:baselinePho` = .006; extent to which relationship between age.centered and Sem_Total_Correct changes at different levels of baselinePho (change is minimal; this does not change as a function of centering)
- `Pseudo-R2` = .39 (marginal; fixed) and .68 (conditional; fixed + random)

Likelihood ratio test suggests that simpler model, where covariate only predicts intercept, is preferred.

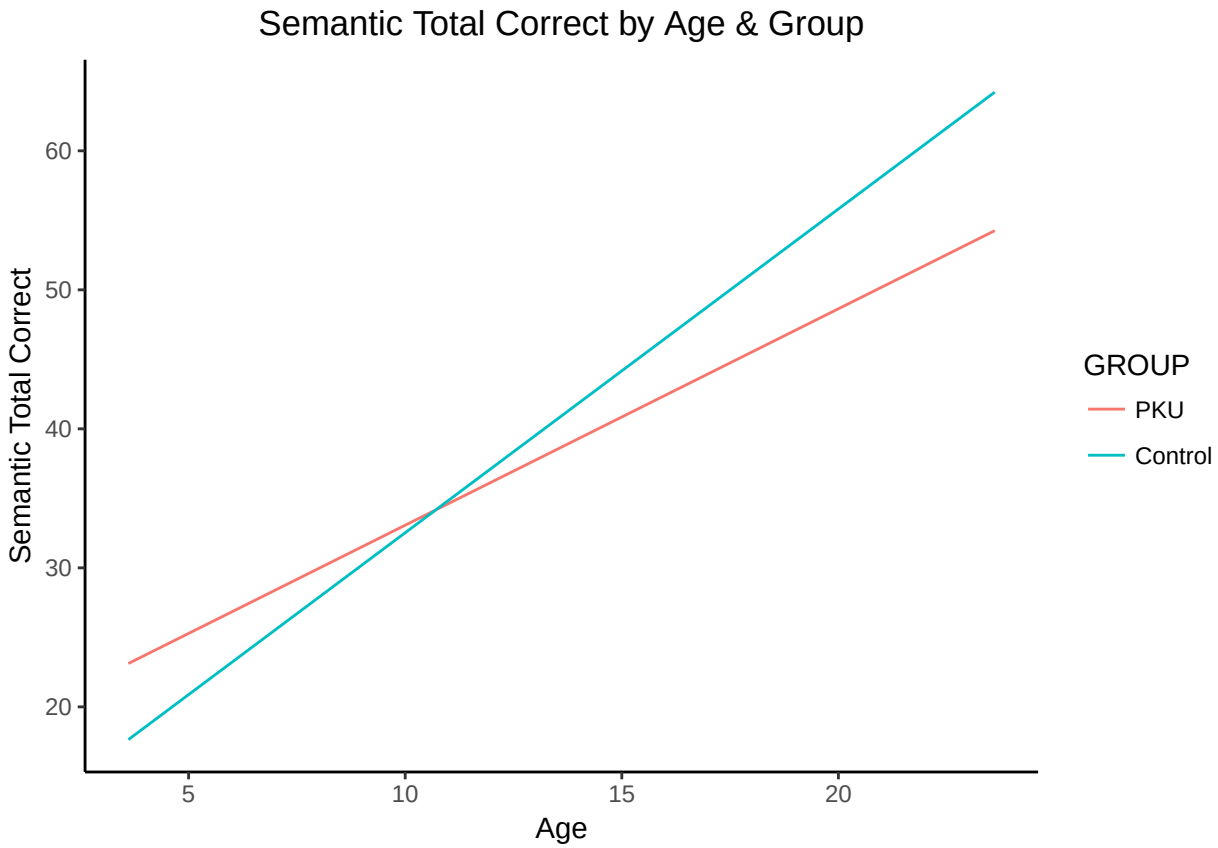
Problem 3:

Graph both of your final models for the continuous and nominal models above.

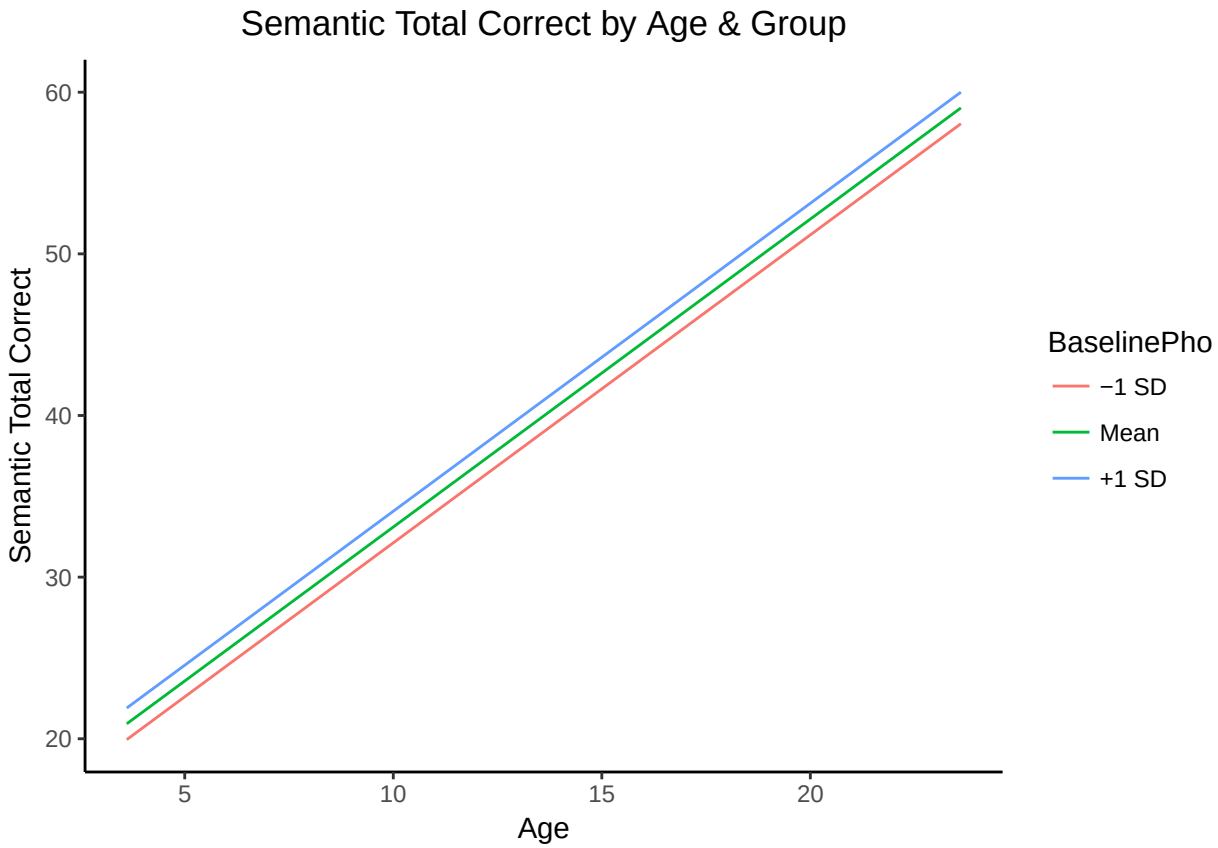
```r
#Graphing nominal, where covariate only predicts the intercept
wide_to_long_merged$GROUP <- relevel(wide_to_long_merged$GROUP, ref = "PKU")
fixed.frame <- data.frame(expand.grid(age.centered = seq(-10, 10, 4),
                                      GROUP = c("PKU", "Control"))) %>%
  mutate(pred = predict(lin.nom1, newdata = ., re.form = NA))
fixed.frame <- fixed.frame %>% mutate(age.new = age.centered + 13.61)

library(ggplot2)
a <- ggplot(aes(x = age.new, y = pred, color = GROUP), data = fixed.frame) +
  geom_line() +
  labs(x = "Age", y = "Semantic Total Correct", title = "Semantic Total Correct by Age & Group") +
  theme_classic() +
  theme(plot.title = element_text(hjust = .5))
a
```



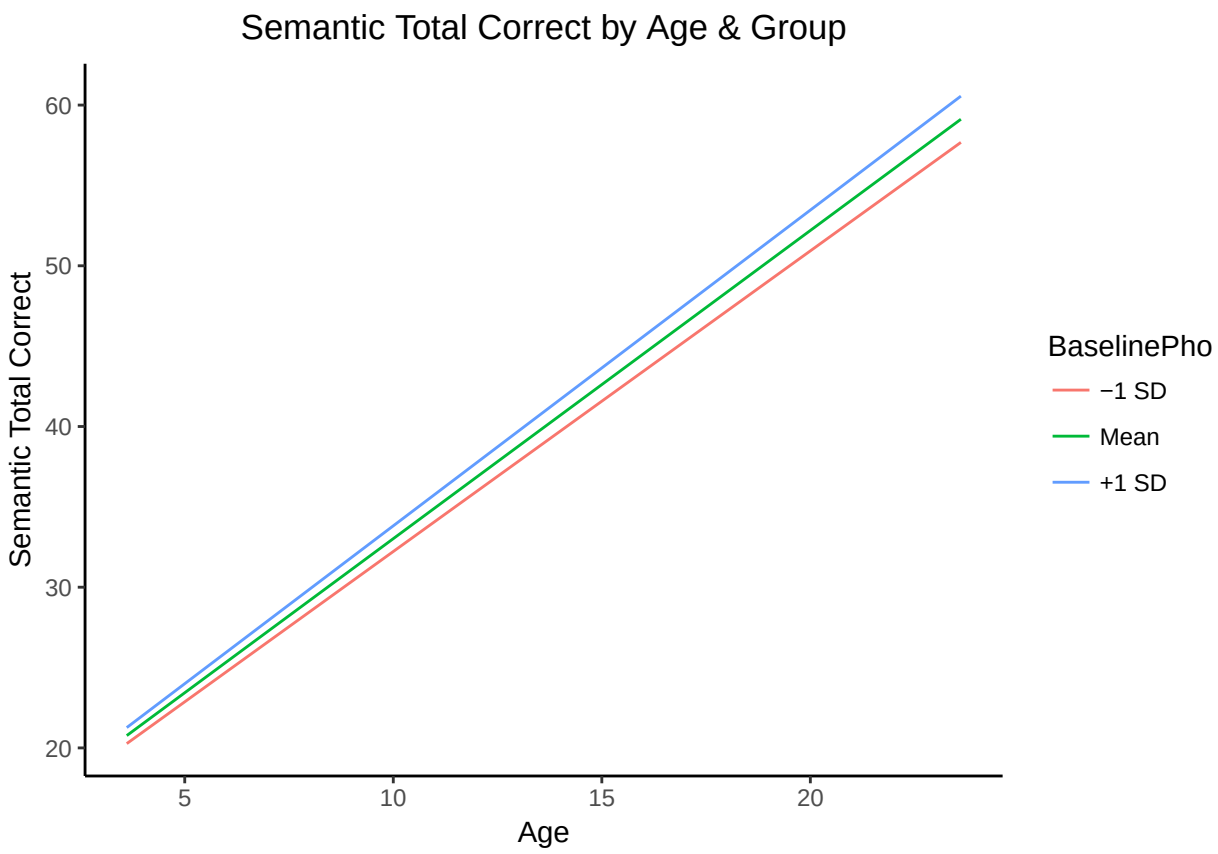

```r
#Graphing nominal, where covariate predicts both the slope and the intercept
fixed.frame2 <- data.frame(expand.grid(age.centered = seq(-10, 10, 4),
                                      GROUP = c("PKU", "Control"))) %>%
  mutate(pred = predict(lin.nom2, newdata = ., re.form = NA))
fixed.frame2 <- fixed.frame2 %>% mutate(age.new = age.centered + 13.61)

b <- ggplot(aes(x = age.new, y = pred, color = GROUP), data = fixed.frame2) +
  geom_line() +
  labs(x = "Age", y = "Semantic Total Correct", title = "Semantic Total Correct by Age & Group") +
  theme_classic() +
  theme(plot.title = element_text(hjust = .5))
b
```




```
#Graphing continuous, where covariate only predicts the intercept
fixed.frame3 <- wide_to_long_merged %>%
  summarise(mean = mean(baselinePho, na.rm = T),
            sd = sd(baselinePho, na.rm = T))

fixed.frame3 <-
  data.frame(
    expand.grid(
      age.centered = seq(-10, 10, 4),
      baselinePho = c(fixed.frame3$mean - fixed.frame3$sd,
                      fixed.frame3$mean,
                      fixed.frame3$mean + fixed.frame3$sd))) %>%
  mutate(pred = predict(lin.cont1, newdata = ., re.form = NA))
fixed.frame3 <- fixed.frame3 %>% mutate(age.new = age.centered + 13.61,
                                        BaselinePho = factor(baselinePho,
                                                             levels = unique(baselinePho),
                                                             labels = c("-1 SD", "Mean", "+1 SD")))

c <- ggplot(aes(x = age.new, y = pred, color = BaselinePho), data = fixed.frame3) +
  geom_line() +
  labs(x = "Age", y = "Semantic Total Correct", title = "Semantic Total Correct by Age & Group") +
  theme_classic() +
  theme(plot.title = element_text(hjust = .5))
c
```




```
#Graphing continuous, where covariate predicts both the slope and the intercept
fixed.frame4 <- wide_to_long_merged %>%
  summarise(mean = mean(baselinePho, na.rm = T),
            sd = sd(baselinePho, na.rm = T))

fixed.frame4 <-
  data.frame(
    expand.grid(
      age.centered = seq(-10, 10, 4),
      baselinePho = c(fixed.frame4$mean - fixed.frame4$sd,
                      fixed.frame4$mean,
                      fixed.frame4$mean + fixed.frame4$sd))) %>%
  mutate(pred = predict(lin.cont2, newdata = ., re.form = NA))
fixed.frame4 <- fixed.frame4 %>% mutate(age.new = age.centered + 13.61,
                                        BaselinePho = factor(baselinePho,
                                                             levels = unique(baselinePho),
                                                             labels = c("-1 SD", "Mean", "+1 SD")))

d <- ggplot(aes(x = age.new, y = pred, color = BaselinePho), data = fixed.frame4) +
  geom_line() +
  labs(x = "Age", y = "Semantic Total Correct", title = "Semantic Total Correct by Age & Group") +
  theme_classic() +
  theme(plot.title = element_text(hjust = .5))
d
```



Problem 4A:

Calculate confidence intervals around your models

```
##                          .rownames      X2.5..    X97.5..
## 1               sd_(Intercept)|ID2  3.81170172  7.1532688
## 2 cor_age.centered.(Intercept)|ID2 -0.99999999  1.0000000
## 3              sd_age.centered|ID2  0.07387441  1.7007404
## 4                            sigma  5.24382976  6.7492485
## 5                      (Intercept) 38.75310103 43.0554870
## 6                    age.centered  1.68410123  2.9832330
## 7                       GROUPPKU -5.86648814  1.2318661
## 8           age.centered:GROUPPKU -1.88150613  0.3071001
```

```
##                          .rownames      X2.5..     X97.5..
## 1               sd_(Intercept)|ID2  3.69045660  7.08475365
## 2 cor_age.centered.(Intercept)|ID2 -0.44826654  1.00000000
## 3              sd_age.centered|ID2  0.08235269  1.53979563
## 4                            sigma  5.35809118  6.89182786
## 5                      (Intercept) 32.44401946 42.56905355
## 6                    age.centered  0.38271472  3.23090244
## 7                     baselinePho -0.12362644  0.36905831
## 8        age.centered:baselinePho -0.06361907  0.07301523
```

Table 1: Group Model

	Semantic TC	
Term	Estimate	CI
Fixed		
(Intercept)	40.93	(38.88, 43.01)
Age	2.33	(1.71, 2.95)
Group	-2.25	(-5.82, 1.48)
Interaction	-0.77	(-1.85, 0.34)
Random		
τ_{00}	30.65	(12.49, 50.91)
τ_{11}	0.89	(0.01, 2.47)
τ_{10}	0.34	(1.00, 1.00)
$\hat{\sigma^2}$	36.43	(27.42, 46.82)
Model		
ICC	0.46	NA
R^2_m	0.38	NA
R^2_c	0.70	NA

Problem 4B:

Create tables for both of your continous and nominal models above.

```
#Nominal table
library(reghelper)
library(plyr)
library(dplyr)
library(papaja)

tab <- table_fun(lin.nom2)
#age.new <- as.character(round(as.numeric(tab$estimate[2]) +
#                              mean(wide_to_long_merged$age, na.rm = T), 2))
tab2 <- tab %>% mutate(#estimate = mapvalues(estimate, from = tab$estimate[2], to = age.new),
  term = mapvalues(term, from = tab$term[2:4], to = c("Age", "Group",
                                                      "Interaction")))
papaja::apa_table(tab2 %>% select(-type), caption = "Group Model", na_string = "",
                  stub_indents = list(Fixed = c(1:4), Random = c(5:8), Model = c(9:11)),
                  col_spanners = list(`Semantic TC` = c(2,3)),
                  col.names = c("Term", "Estimate", "CI"),
                  align = c("l", "c", "c"))
```

```
#Continuous table
tab3 <- table_fun(lin.cont2)
#age.new <- as.character(round(as.numeric(tab3$estimate[2]) +
#                              mean(wide_to_long_merged$age, na.rm = T), 2))
tab4 <- tab3 %>%  mutate(#estimate = mapvalues(estimate, from = tab3$estimate[2], to = age.new),
  term = mapvalues(term, from = tab3$term[2:4], to = c("Age",
                                                       "BaselinePho", "Interaction")))
papaja::apa_table(tab4 %>% select(-type), caption = "Continuous Table", na_string = "",
                  stub_indents = list(Fixed = c(1:4), Random = c(5:8), Model = c(9:11)),
                  col_spanners = list(`Semantic TC` = c(2,3)),
                  col.names = c("Term", "Estimate", "CI"),
                  align = c("l", "c", "c"))
```

Table 2: Continuous Table

	Semantic TC	
Term	Estimate	CI
Fixed		
(Intercept)	37.46	(31.81, 42.93)
Age	1.80	(0.27, 3.23)
BaselinePho	0.13	(-0.12, 0.40)
Interaction	0.01	(-0.06, 0.08)
Random		
τ_{00}	30.20	(13.30, 50.73)
τ_{11}	0.67	(0.00, 2.20)
τ_{10}	0.44	(1.00, 1.00)
$\hat{\sigma^2}$	38.10	(28.81, 47.90)
Model		
ICC	0.45	NA
R^2_m	0.39	NA
R^2_c	0.68	NA

Problem 5:

Include both types of covariates in a single model. How does your interpretation of parameters change? *(From last week's homework, so output is suppressed.)*

Fixed effect estimates:

- `Intercept` = 40.94; mean of Controls at mean levels of age and baselinePho (i.e., age.centered and baselinePho.centered = 0)
- `Age.centered` = 2.25; slope of Controls (increase in words correct/year) at mean levels of baselinePho (i.e., baselinePho.centered = 0)
- `baselinePho.centered` = .10; slope of Controls (increase in words correct/unit of baselinePho) at mean levels of age (i.e., age.centered = 0)
- `GROUP` = -2.73; difference in words correct between Control and PKU at mean levels of age and baselinePho (i.e., when age.centered and baselinePho.centered = 0)
- `age.centered:baselinePho.centered` = -.02; extent to which relationship between age.centered and Sem_TotalCorrect changes at different levels of baselinePho.centered, when group = Controls
- `age.centered:GROUP` = -.63; extent to which relationship between age.centered and Sem_TotalCorrect changes at different levels of GROUP, when baselinePho.centered = 0
- `baselinePho.centered:GROUP` = -.05; extent to which relationship between baselinePho.centered and Sem_TotalCorrect changes at different levels of GROUP, when age.centered = 0
- `age.centered:baselinePho.centered:GROUP` = .09; extent to which the interaction between age.centered and baselinePho.centered changes at different levels of GROUP
- `Pseudo-R2` = .39 (marginal; fixed) and .70 (conditional; fixed + random)

Problem 6:

If you have one available, introduce a time-varying covariate.

Fixed effect estimates:

- `Intercept` = 40.19; mean of Controls at mean levels of age and Pho_TC (i.e., age.centered and Pho_TC.centered = 0)
- `Age.centered` = 1.64; slope of Controls (increase in words correct/year) at mean levels of Pho_TC (i.e., Pho_TC.centered = 0)
- `Pho_TotalCorrect.centered` = .37; slope of Controls (increase in words correct/unit of Pho_TC) at mean levels of age (i.e., age.centered = 0)
- `GROUP` = -1.08; difference in words correct between Control and PKU at mean levels of age and Pho_TC (i.e., when age.centered and Pho_TC.centered = 0)
- `age.centered:Pho_TotalCorrect.centered` = -.03; extent to which relationship between age.centered and Sem_TotalCorrect changes at different levels of Pho_TC.centered, when group = Controls
- `age.centered:GROUP` = -.34; extent to which relationship between age.centered and Sem_TotalCorrect changes at different levels of GROUP, when Pho_TC.centered = 0
- `Pho_TotalCorrect.centered:GROUP` = -.13; extent to which relationship between Pho_TC.centered and Sem_TotalCorrect changes at different levels of GROUP, when age.centered = 0
- `age.centered:Pho_TotalCorrect.centered:GROUP` = .02; extent to which the interaction between age.centered and Pho_TC.centered changes at different levels of GROUP
- `Pseudo-R2` = .31 (marginal; fixed) and .70 (conditional; fixed + random)

```
wide_to_long_merged$GROUP <- relevel(wide_to_long_merged$GROUP, ref = "Control")
wide_to_long_merged$Pho_TotalCorrect <- as.numeric(wide_to_long_merged$Pho_TotalCorrect)
wide_to_long_merged$Pho_TotalCorrect.centered <- wide_to_long_merged$Pho_TotalCorrect -
  mean(wide_to_long_merged$Pho_TotalCorrect, na.rm = T)

lin.comb.cont <- lmer(Sem_TotalCorrect ~ age.centered*Pho_TotalCorrect.centered*GROUP +
                    (age.centered | ID2), data = wide_to_long_merged)
summary(lin.comb.cont)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Sem_TotalCorrect ~ age.centered * Pho_TotalCorrect.centered *
##     GROUP + (age.centered | ID2)
##    Data: wide_to_long_merged
##
## REML criterion at convergence: 1372.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.12944 -0.62402 -0.00917  0.46883  2.69170
##
## Random effects:
##  Groups   Name         Variance Std.Dev. Corr
##  ID2      (Intercept)  31.4544  5.6084
##           age.centered  0.3161  0.5622   0.38
```

```
##  Residual              34.9040  5.9080
## Number of obs: 200, groups:  ID2, 67
##
## Fixed effects:
##                                                 Estimate Std. Error
## (Intercept)                                     40.19235    1.13881
## age.centered                                     1.63629    0.32621
## Pho_TotalCorrect.centered                        0.37119    0.09666
## GROUPPKU                                        -1.07938    1.88156
## age.centered:Pho_TotalCorrect.centered          -0.03130    0.02451
## age.centered:GROUPPKU                           -0.33708    0.57300
## Pho_TotalCorrect.centered:GROUPPKU              -0.12852    0.19327
## age.centered:Pho_TotalCorrect.centered:GROUPPKU  0.02072    0.05490
##                                                 t value
## (Intercept)                                       35.29
## age.centered                                       5.02
## Pho_TotalCorrect.centered                          3.84
## GROUPPKU                                          -0.57
## age.centered:Pho_TotalCorrect.centered            -1.28
## age.centered:GROUPPKU                             -0.59
## Pho_TotalCorrect.centered:GROUPPKU                -0.66
## age.centered:Pho_TotalCorrect.centered:GROUPPKU    0.38
##
## Correlation of Fixed Effects:
##             (Intr) ag.cnt Ph_TC. GROUPP ag.:P_TC. a.:GRO P_TC.:
## age.centerd  0.394
## Ph_TtlCrrc. -0.236 -0.508
## GROUPPKU    -0.605 -0.239  0.143
## ag.cn:P_TC. -0.345 -0.147  0.053  0.209
## a.:GROUPPKU -0.224 -0.569  0.289  0.150  0.084
## P_TC.:GROUP  0.118  0.254 -0.500  0.091 -0.026    -0.361
## a.:P_TC.:GR  0.154  0.066 -0.023 -0.235 -0.446     0.290  0.032
```

```
anova(lin.comb.cont)
```

```
## Analysis of Variance Table
##                                            Df  Sum Sq Mean Sq F value
## age.centered                                1 2678.63 2678.63 76.7428
## Pho_TotalCorrect.centered                   1  662.49  662.49 18.9803
## GROUP                                       1    0.17    0.17  0.0050
## age.centered:Pho_TotalCorrect.centered      1   34.84   34.84  0.9982
## age.centered:GROUP                          1   40.39   40.39  1.1571
## Pho_TotalCorrect.centered:GROUP             1   16.02   16.02  0.4591
## age.centered:Pho_TotalCorrect.centered:GROUP  1    4.97    4.97  0.1425
```

```
r.squaredGLMM(lin.comb.cont)
```

```
##       R2m       R2c
## 0.4116268 0.7007126
```