SMiP 2024 MLM Example 3 *(complete syntax and output available for R electronically)*

This example shows variants of "explanatory" item response theory (IRT) models, which can be estimated as generalized multilevel models with a random person intercept and either fixed item effects (in which the model just has level-1 trials nested in level-2 persons) or a random item intercept (in which case 36 level-2 items are then crossed with 152 level-2 persons). These example data are from my dissertation: 5,323 trials of a change detection task assessing attentional search scored incorrect (correct=0) or correct (correct=1). The items are natural driving scenes whose changes varied by four features: continuous visual clutter (clutter), whether the change was relevant to driving (relevant), continuous brightness of the change (bright), and whether the change was made to a legible sign (sign). The subjects (labeled as "persons" below) are older adults who varied continuously in years of age (centered at 75) and on a z-scaled factor score measure of vision impairment. These analyses when conducted through multilevel software require a "stacked" (or "long") data format in which each item for each person is stored on a separate row. Syntax and output for 1PL or Rasch IRT models on wide-format data are also available in the online materials.

R Syntax for Importing and Preparing Data for Analysis (after loading packages *readxl*, *TeachingDemos*, *expss*, *psych*, *lme4*, *lmerTest*, and *performance*, as well as *mirt* for the IRT version):

```
# Define variables for working directory and data name -- CHANGE THESE
filesave = " C:\\Dropbox/Papers and Data/SMiP/SMiP_2024_MLM_Part2/"
filename = "Example3_Data.xlsx"
setwd(dir=filesave)

#  Import long-format example excel data file from sheet "long"
Example3 = read_excel(paste0(filesave,filename), sheet="long")
# Convert to data frame to use in analysis
Example3 = as.data.frame(Example3)

# Filter to only cases complete on all variables to be used below
Example3 = Example3[complete.cases(Example3[ ,
                    c("correct","clutter","relevant","bright","sign","age75","visimp")]),]

print("R Descriptive Statistics for Example Variables")
print(describe(x=Example3[ , c("correct","clutter","relevant","bright","sign","age75","visimp")],
               fast=TRUE), digits=3)

# Labels for original variables in comments only
#PersonID=  "PersonID: Person Identifier"
#PictureID= "PictureID: Picture Identifier"
#correct=   "correct: Correct Response (0=no, 1=yes)"
#clutter=   "clutter: Rated Clutter in Picture Z-Score"
#relevant=  "relevant: Change Relevant to Driving (0=no, 1=yes)"
#bright=    "bright: Rated Change Brightness Z-Score"
#sign=      "sign: Change to Legible Sign (0=no, 1=yes)"
#age75=     "age75: Age in Years (0=75)"
#visimp=    "visimp: Vision Factor Z-Score"
```

```
         vars    n   mean    sd median     min    max  range   skew kurtosis    se
correct     1 5323  0.783 0.412  1.000   0.000  1.000  1.000 -1.376   -0.108 0.006
clutter     2 5323  0.003 0.651 -0.118  -1.183  1.454  2.637  0.183   -0.709 0.009
relevant    3 5323  0.501 0.500  1.000   0.000  1.000  1.000 -0.004   -2.000 0.007
bright      4 5323 -0.023 0.429  0.043  -1.077  1.028  2.104 -0.041   -0.151 0.006
sign        5 5323  0.336 0.472  0.000   0.000  1.000  1.000  0.696   -1.516 0.006
age75       6 5323  0.091 4.565  0.129 -11.646 10.038 21.684  0.001   -0.308 0.063
visimp      7 5323 -0.024 0.907 -0.122  -2.525  2.518  5.043  0.228    0.568 0.012
```

Model 1: Single-Level Empty Means for Binary Correct Response
(*t* = trial, *p* = person, *i* = item)

Composite: $Log\left[\frac{prob(correct_{tpi}=1)}{prob(correct_{tpi}=0)}\right] = Logit\left(correct_{tpi}=1\right) = \gamma_{000}$

```
print("R Model 1: Single-Level Empty Means for Binary Correct Response")
Model1 = glm(data=Example3, family=binomial(link="logit"), formula=correct~1)
summary(Model1) # residual deviance = -2LL already

Coefficients:
            Estimate Std. Error z value  Pr(>|z|)
(Intercept) 1.285549   0.033273  38.636 < 2.2e-16  gamma000 in logits

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5563.39  on 5322  degrees of freedom
Residual deviance: 5553.39  on 5322  degrees of freedom → -2LL for model
AIC: 5564.39

print("Convert logits to probability via short inverse link")
Model1Prob=1/(1+exp(-1*coefficients(Model1))); Model1Prob

0.78339282  → gamma000 in probability
```

$$\text{Prob}(y=1) = \frac{\exp(1.286)}{1+\exp(1.286)} = .783$$

All two-level models from here use Laplace estimation (one adaptive quadrature point) for comparability.

Model 2: Random Persons Only, Empty Means

$$Logit\left(correct_{tpi}=1\right) = \gamma_{000} + U_{0p0}$$

```
print("R Model 2: Random Persons Only, Empty Means")
Model2 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~1+(1|PersonID))
print("Show -2LL with more precision, results, and ICC using 3.29=residual variance")
-2*logLik(Model2); summary(Model2); icc(Model2)

'log Lik.' 5481.6566 (df=2) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  5485.7   5498.8  -2740.8   5481.7     5321

Random effects:
 Groups   Name        Variance Std.Dev.
 PersonID (Intercept) 0.25002  0.50002        Var(U0p0)
Number of obs: 5323, groups:  PersonID, 152

Fixed effects:
            Estimate Std. Error z value  Pr(>|z|)
(Intercept) 1.354170   0.053927  25.111 < 2.2e-16  gamma000

# Intraclass Correlation Coefficient
    Adjusted ICC: 0.071
  Unadjusted ICC: 0.071
```

Model-scale ICC (in logits) for the correlation of trials from the same person:

$$\text{ICC} = \frac{0.250}{0.250 + 3.29} = .071$$

```
print("LRT for person random intercept"); anova(Model2, Model1)
       npar     AIC     BIC   logLik deviance   Chisq Df Pr(>Chisq)
Model1    1 5565.39 5571.97 -2781.69  5563.39
Model2    2 5485.66 5498.82 -2740.83  5481.66 81.7313  1 < 2.22e-16
```

For a mixture *p*-value with df = 0 and 1, cut this *p*-value in half

```
print("Convert logits to probability via shorter inverse link")
Model2Prob=1/(1+exp(-1*fixef(Model2))); Model2Prob
0.7948105
```

$$\text{Prob}(y=1) = \frac{\exp(1.354)}{1+\exp(1.354)} = .795$$

Model 3: Random Persons Only, Rasch Fixed Items (via a Categorical Item ID Predictor)

$$Logit(correct_{tpi} = 1) = \gamma_{00,1}(Pic1_i) + \gamma_{00,2}(Pic2_i) + \cdots + \gamma_{00,36}(Pic36_i) + U_{0p0}$$

```
print("R Model 3: Random Persons Only, Rasch Fixed Items")
print("Rasch version of 1PL: Person Trait Variance Estimated, Single Discrimination=1")
Model3 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~0+as.factor(PictureID)+(1|PersonID))
print("Show -2LL with more precision and results")
-2*logLik(Model3); summary(Model3)
print("Compute predicted mean per item given theta=0")
Model3Prob = 1/(1+exp(-1*(as.data.frame(fixef(Model3))))); Model3Prob
```

```
'log Lik.' 4786.0814 (df=37) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  4860.1   5103.5  -2393.0   4786.1     5286

Random effects:
 Groups   Name        Variance Std.Dev.
 PersonID (Intercept) 0.36839  0.60695  Var(U0p0)
```

Note that the level-2 person random intercept variance is **higher** after adding fixed item effects that explain level-1 residual variance). This is because the model has changed scale to reflect what the total variance must have been in order for a level-1 residual variance to remain $\pi^2/3 = 3.29$.

```
Fixed effects: → Item Easiness parameters
                       Estimate Std. Error z value  Pr(>|z|)    Probability (computed separately)
as.factor(PictureID)2    0.25066    0.17985  1.3937 0.1634083    0.56233874
as.factor(PictureID)6    3.53484    0.45781  7.7212 1.153e-14    0.97166297
as.factor(PictureID)7    1.11733    0.19888  5.6180 1.932e-08    0.75349340
as.factor(PictureID)10   2.27728    0.27282  8.3471 < 2.2e-16    0.90697818
as.factor(PictureID)11   0.97000    0.19333  5.0174 5.237e-07    0.72511901
as.factor(PictureID)13   0.14559    0.17749  0.8203 0.4120613    0.53633308
as.factor(PictureID)22   1.25016    0.20584  6.0735 1.252e-09    0.77732730
as.factor(PictureID)23   2.90294    0.34950  8.3059 < 2.2e-16    0.94799151
as.factor(PictureID)26   1.03855    0.19612  5.2956 1.186e-07    0.73857052
as.factor(PictureID)33   1.66216    0.22670  7.3319 2.268e-13    0.84052727
as.factor(PictureID)35   1.33291    0.20936  6.3666 1.933e-10    0.79132227
as.factor(PictureID)42   0.98644    0.19550  5.0456 4.521e-07    0.72838346
as.factor(PictureID)52   2.69255    0.32006  8.4127 < 2.2e-16    0.93658555
as.factor(PictureID)59   2.23255    0.27330  8.1689 3.113e-16    0.90313477
as.factor(PictureID)61   0.92676    0.19523  4.7471 2.063e-06    0.71641756
as.factor(PictureID)62  -0.36351    0.18188 -1.9986 0.0456479    0.41010896
as.factor(PictureID)66   2.50904    0.29758  8.4314 < 2.2e-16    0.92477304
as.factor(PictureID)97   1.34753    0.20871  6.4564 1.072e-10    0.79372610
as.factor(PictureID)117  3.18551    0.39185  8.1294 4.315e-16    0.96028523
as.factor(PictureID)123  2.42183    0.28851  8.3941 < 2.2e-16    0.91847705
as.factor(PictureID)128  1.19463    0.20192  5.9165 3.290e-09    0.76756790
as.factor(PictureID)135  3.35064    0.42078  7.9628 1.681e-15    0.96612587
as.factor(PictureID)136  1.28554    0.21107  6.0907 1.124e-09    0.78339141
as.factor(PictureID)137  1.46764    0.21572  6.8034 1.022e-11    0.81269870
as.factor(PictureID)140  0.40504    0.18267  2.2173 0.0266005    0.59989681
as.factor(PictureID)146  1.76550    0.23310  7.5739 3.623e-14    0.85389664
as.factor(PictureID)152  0.71602    0.18756  3.8175 0.0001348    0.67172892
as.factor(PictureID)155 -0.25789    0.18051 -1.4287 0.1531044    0.43588186
as.factor(PictureID)161  3.01986    0.36888  8.1867 2.686e-16    0.95346336
as.factor(PictureID)162  3.01476    0.36871  8.1766 2.920e-16    0.95323652
as.factor(PictureID)171  1.54787    0.22039  7.0232 2.169e-12    0.82460624
as.factor(PictureID)172  0.86338    0.19237  4.4882 7.183e-06    0.70336692
as.factor(PictureID)173  3.36228    0.42069  7.9923 1.324e-15    0.96650459
as.factor(PictureID)174  1.71728    0.22981  7.4727 7.856e-14    0.84777852
as.factor(PictureID)177  0.90055    0.19099  4.7152 2.415e-06    0.71106333
as.factor(PictureID)179  1.86825    0.24099  7.7523 9.023e-15    0.86625591
optimizer (Nelder_Mead) convergence code: 0 (OK)
Model failed to converge with max|grad| = 0.00710241 (tol = 0.002, component 1)
```

These intercepts give the expected logit of $y_{tpi} = 1$ for each picture (item) given person random intercept (otherwise known as theta) = 0.

This "fixed-items" model controls for differences in item easiness, but it does not allow any predictors of those differences—stay tuned for Model 5!

```
print("LRT for fixed item differences"); anova(Model3, Model2)
       npar     AIC     BIC   logLik deviance   Chisq Df Pr(>Chisq)
Model2    2 5485.66 5498.82 -2740.83  5481.66
Model3   37 4860.08 5103.53 -2393.04  4786.08 695.575 35 < 2.22e-16
```

Model 4: Random Persons, Empty Means Random Items

$$Logit(correct_{tpi} = 1) = \gamma_{000} + U_{0p0} + U_{00i}$$

```
print("R Model 4: Random Persons, Empty Means Random Items")
Model4 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~1+(1|PersonID)+(1|PictureID))
print("Show -2LL with more precision, results, and ICCs")
-2*logLik(Model4); summary(Model4); icc(Model4, by_group=TRUE)

'log Lik.' 4928.7898 (df=3) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  4934.8   4954.5  -2464.4   4928.8     5320

Random effects:
 Groups    Name        Variance Std.Dev.
 PersonID  (Intercept) 0.36335  0.60279  Var(U0p0)
 PictureID (Intercept) 0.98812  0.99404  Var(U00i)
Number of obs: 5323, groups:  PersonID, 152; PictureID, 36

Fixed effects:
            Estimate Std. Error z value  Pr(>|z|)
(Intercept)  1.60859    0.17817  9.0284 < 2.2e-16  gamma000

Group     |   ICC
-----------------
PersonID  | 0.078
PictureID | 0.213

print("LRT for item random intercept"); anova(Model4, Model2)
       npar     AIC     BIC   logLik deviance   Chisq Df Pr(>Chisq)
Model2    2 5485.66 5498.82 -2740.83  5481.66
Model4    3 4934.79 4954.53 -2464.39  4928.79 552.867  1 < 2.22e-16
```

For a mixture p-value with df = 0 and 1, cut this p-value in half

```
# Save each variance as an object to compute pseudo-R2
Model4SubIntVar  = as.data.frame(VarCorr(Model4))[1,4]
Model4ItemIntVar = as.data.frame(VarCorr(Model4))[2,4]
```

Model 5: Random Persons, Predicted Random Items (Predicted Item Easiness)

$$Logit(correct_{tpi} = 1) = \gamma_{000} + \gamma_{001}(clutter_i) + \gamma_{002}(relevant_i) + \gamma_{003}(bright_i) + \gamma_{004}(sign_i) + U_{0p0} + U_{00i}$$

```
print("R Model 5: Random Persons, LLTM-Predicted Random Items")
Model5 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~1+clutter+relevant+bright+sign+(1|PersonID)+(1|PictureID))
print("Show -2LL with more precision and results")
-2*logLik(Model5); summary(Model5)

'log Lik.' 4919.2565 (df=7) → -2LL for model
     AIC      BIC   logLik deviance df.resid
  4933.3   4979.3  -2459.6   4919.3     5316

Random effects:
 Groups    Name        Variance Std.Dev.
 PersonID  (Intercept) 0.36332  0.60276  Var(U0p0)
 PictureID (Intercept) 0.74393  0.86251  Var(U00i)

Fixed effects:
             Estimate Std. Error z value     Pr(>|z|)
(Intercept)  1.373626   0.264748  5.1884 0.0000002121  gamma000
clutter     -0.332720   0.246572 -1.3494      0.17721  gamma001
relevant     0.011631   0.434184  0.0268      0.97863  gamma002
bright       0.833050   0.508496  1.6383      0.10137  gamma003
sign         0.748701   0.342880  2.1836      0.02899  gamma004
```

```
# Save each variance as an object to compute pseudo-R2
Model5SubIntVar  = as.data.frame(VarCorr(Model5))[1,4]
Model5ItemIntVar = as.data.frame(VarCorr(Model5))[2,4]

# Compute pseudo-R2 for each variance
Model5SubIntR2  = (Model4SubIntVar-Model5SubIntVar)/Model4SubIntVar; Model5SubIntR2
[1] 0.000097249414

Model5ItemIntR2 = (Model4ItemIntVar-Model5ItemIntVar)/Model4ItemIntVar; Model5ItemIntR2
[1] 0.24712914
```

Do we need random item variance leftover? How does the exchangeable-item solution compare?

Model 6: Random Persons Only, Predicted Fixed Items (Predicted Item Easiness)

$$Logit\left(correct_{tpi}=1\right)=\gamma_{000}+\gamma_{001}(clutter_i)+\gamma_{002}(relevant_i)+\gamma_{003}(bright_i)+\gamma_{004}(sign_i)+U_{0p0}$$

```
print("R Model 6: Random Persons Only, LLTM-Predicted Fixed Items")
Model6 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~1+clutter+relevant+bright+sign+(1|PersonID))
print("Show -2LL with more precision and results")
-2*logLik(Model6); summary(Model6)

'log Lik.' 5320.9823 (df=6) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  5333.0   5372.5  -2660.5   5321.0     5317

Random effects:
 Groups   Name        Variance Std.Dev.
 PersonID (Intercept) 0.2716   0.52115  Var(U0p0)

Fixed effects:
             Estimate Std. Error z value  Pr(>|z|)
(Intercept)  1.096225   0.072716 15.0754 < 2.2e-16  gamma000
clutter     -0.267560   0.055835 -4.7920 1.651e-06  gamma001
relevant     0.206068   0.100041  2.0598   0.03941  gamma002
bright       0.508638   0.113652  4.4754 7.628e-06  gamma003
sign         0.665401   0.082787  8.0375 9.168e-16  gamma004
```

```
Fixed effects from Random Item Model 5:
             Estimate Std. Error      Pr(>|z|)
(Intercept)  1.373626   0.264748 0.0000002121
clutter     -0.332720   0.246572      0.17721
relevant     0.011631   0.434184      0.97863
bright       0.833050   0.508496      0.10137
sign         0.748701   0.342880      0.02899
```

```
print("LRT for random item variance leftover"); anova(Model5, Model6)
       npar     AIC     BIC  logLik deviance   Chisq Df Pr(>Chisq)
Model6    6 5332.98 5372.46 -2660.49  5320.98
Model5    7 4933.26 4979.32 -2459.63  4919.26 401.726  1 < 2.22e-16
```

Comparing Results: Fixed vs. Random Effects for Item Easiness Predictions

```
# Model 3 Rasch Fixed Items
Fixed_b = as.data.frame(fixef(Model3))

# Model 4 Random Items
Random_b = as.data.frame(ranef(Model4)$PictureID) + (fixef(Model4))

# Model 6 Predicted Fixed Items
Pred = as.data.frame(unique(Example3[,c("PictureID","clutter","relevant","bright","sign")]))
Fixed_Pred_b = as.data.frame(predict(Model6, newdata=Pred, re.form=NA, type="link"))

# Model 5 Predicted Random Items (random intercept not included)
Random_Pred_b = as.data.frame(predict(Model5, newdata=Pred, re.form=NA, type="link"))

# Combine results and rename
Compare_b = cbind(Fixed_b, Random_b, Fixed_Pred_b, Random_Pred_b)
colnames(Compare_b) = c("Fixed_b","Random_b","Fixed_Pred_b","Random_Pred_b")
print("Correlation of item easiness across models")
print(cor(Compare_b), digits=3)
```

```
               Fixed_b Random_b Fixed_Pred_b Random_Pred_b
Fixed_b          1.000    0.999        0.472         0.482
Random_b         0.999    1.000        0.475         0.483
Fixed_Pred_b     0.472    0.475        1.000         0.984
Random_Pred_b    0.482    0.483        0.984         1.000
```

$R^2 = .472^\wedge2 = .232$
$R^2 = .483^\wedge2 = .234$

The R^2 values for item easiness above are very close to the pseudo-R^2 = .247 for the proportion reduction in random item variance after including the 4 item predictors. The difference lies in the significance of the item predictor effects, whose SEs are based on the wrong error term in the random-persons-only predictive model.

Model 7: Predicted Random Persons, Predicted Random Items (Predicted Item Easiness)

$$Logit(correct_{tpi} = 1) = \gamma_{000} + \gamma_{001}(clutter_i) + \gamma_{002}(relevant_i) + \gamma_{003}(bright_i) + \gamma_{004}(sign_i) + \gamma_{010}(age_p - 75) + \gamma_{020}(visimp_p) + U_{0p0} + U_{00i}$$

```
print("R Model 7: Predicted Random Persons, Predicted Random Items")
Model7 = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
               correct~1+clutter+relevant+bright+sign+age75+visimp
                        +(1|PersonID)+(1|PictureID))
print("Show -2LL with more precision and results")
-2*logLik(Model7); summary(Model7)

'log Lik.' 4895.5295 (df=9) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  4913.5   4972.7  -2447.8   4895.5     5314

Random effects:
 Groups    Name        Variance Std.Dev.
 PersonID  (Intercept) 0.27856  0.52779  Var(U0p0)
 PictureID (Intercept) 0.74447  0.86283  Var(U00i)

Fixed effects:
             Estimate Std. Error z value      Pr(>|z|)
(Intercept)  1.375207   0.263895  5.2112 0.0000001876  gamma000
clutter     -0.332832   0.246709 -1.3491      0.17731  gamma001
relevant     0.010945   0.434567  0.0252      0.97991  gamma002
bright       0.833447   0.508942  1.6376      0.10150  gamma003
sign         0.750162   0.343147  2.1861      0.02881  gamma004
age75       -0.051826   0.012781 -4.0551 0.0000501161  gamma010
visimp      -0.131996   0.063596 -2.0755      0.03794  gamma020

# Save each variance as an object to compute pseudo-R2
Model7SubIntVar  = as.data.frame(VarCorr(Model7))[1,4]
Model7ItemIntVar = as.data.frame(VarCorr(Model7))[2,4]

# Compute pseudo-R2 for each variance
Model7SubIntR2  = (Model4SubIntVar- Model7SubIntVar) /Model4SubIntVar;  Model7SubIntR2
[1] 0.23335931

Model7ItemIntR2 = (Model4ItemIntVar-Model7ItemIntVar)/Model4ItemIntVar; Model7ItemIntR2
[1] 0.24657901
```

Model 8a: Adding a Random Slope of Item Predictor "Bright" over Persons *(Models 8bcd were NS)*

$$Logit(correct_{tpi} = 1) = \gamma_{000} + \gamma_{001}(clutter_i) + \gamma_{002}(relevant_i) + \gamma_{003}(bright_i) + \gamma_{004}(sign_i) + \gamma_{010}(age_p - 75) + \gamma_{020}(visimp_p) + U_{0p0} + U_{0p3}(bright_i) + U_{00i}$$

```
print("R Model 8a: Add Random Brightness Slope over Persons (NS)")
Model8a = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
                correct~1+clutter+relevant+bright+sign+age75+visimp+drivimp
                         +(1+bright|PersonID)+(1|PictureID))
```

```
print("Show -2LL with more precision and results")
-2*logLik(Model8a); summary(Model8a)

'log Lik.' 4893.018  (df=11) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  4915.0   4987.4  -2446.5   4893.0     5312

Random effects:
 Groups    Name        Variance Std.Dev. Corr
 PersonID  (Intercept) 0.27173  0.52128          Var(U0p0)
           bright      0.13871  0.37244  -0.321  Var(U0p3) Cor(U0p0,U0p3)
 PictureID (Intercept) 0.74750  0.86458          Var(U00i)

print("LRT for random brightness slope over persons"); anova(Model8a, Model7)
        npar     AIC     BIC   logLik deviance   Chisq Df Pr(>Chisq)
Model7     9 4913.53 4972.75 -2447.76  4895.53
Model8a   11 4915.02 4987.40 -2446.51  4893.02 2.51144  2    0.28487

print("LRT for random slope variance using a mixture-chi-square test")
Model8aDiff2LL = -2*(logLik(Model7)-logLik(Model8a))
Model8aDiffP1 = pchisq(Model8aDiff2LL, df=1, lower.tail=FALSE)
Model8aDiffP2 = pchisq(Model8aDiff2LL, df=2, lower.tail=FALSE)
Model8aDiffP12 = (.5*Model8aDiffP1) + (.5*Model8aDiffP2)
print("Test statistic and mixture p-values for df=1,2")
Model8aDiff2LL; Model8aDiffP12
'log Lik.' 2.511444 (df=9)  'log Lik.' 0.19894621 (df=9)
```

Model 9a: Adding a Random Slope of Person Predictor "Age75" over Items *(Models 7bcd were NS)*

$$Logit\left(correct_{tpi}=1\right)=\gamma_{000}+\gamma_{001}(clutter_i)+\gamma_{002}(relevant_i)+\gamma_{003}(bright_i)+\gamma_{004}(sign_i)$$
$$+\gamma_{010}\left(age_p-75\right)+\gamma_{020}\left(visimp_p\right)+U_{0p0}+U_{00i}+U_{01i}\left(age_p-75\right)$$

```
print("R Model 8a: Add Random Brightness Slope over Persons (NS)")
Model9a = glmer(data=Example3, family=binomial(link="logit"), nAGQ=1,
                correct~1+clutter+relevant+bright+sign+age75+visimp+drivimp
                         +(1|PersonID)+(1+age75|PictureID))
print("Show -2LL with more precision and results")
-2*logLik(Model9a); summary(Model9a)

'log Lik.' 4894.6949 (df=12) → -2LL for model

     AIC      BIC   logLik deviance df.resid
  4916.7   4989.1  -2447.3   4894.7     5312

Random effects:
Groups     Name          Variance   Std.Dev. Corr
 PersonID  (Intercept) 0.27974400 0.528908          Var(U0p0)
 PictureID (Intercept) 0.75298932 0.867750          Var(U00i)
           age75       0.00047998 0.021909 -0.248   Var(U01i) Cor(U00i,U01i)

print("LRT for random age slope over items"); anova(Model9a, Model7)
        npar     AIC     BIC   logLik deviance   Chisq Df Pr(>Chisq)
Model7     9 4913.53 4972.75 -2447.76  4895.53
Model9a   11 4916.69 4989.07 -2447.35  4894.69 0.83453  2    0.65885

print("LRT for random slope variance using a mixture-chi-square test")
Model9aDiff2LL = -2*(logLik(Model7)-logLik(Model9a))
Model9aDiffP1 = pchisq(Model9aDiff2LL, df=1, lower.tail=FALSE)
Model9aDiffP2 = pchisq(Model9aDiff2LL, df=2, lower.tail=FALSE)
Model9aDiffP12 = (.5*Model9aDiffP1) + (.5*Model9aDiffP2)
print("Test statistic and mixture p-values for df=1,2")
Model9aDiff2LL; Model9aDiffP12
'log Lik.' 0.83453069 (df=9)  'log Lik.' 0.50990589 (df=9)
```

Sample Results Section using R Output
[indicates notes about what to customize or also include; note that SE and p-values are not needed if you provide tables for the model solutions]

The extent to which item features could predict binary accuracy in a change detection task was examined in a series of multilevel models predicting 5,323 trial-level responses to 36 items from 152 persons. Binary accuracy was predicted using a logit link function and Bernoulli conditional outcome distribution. The models were estimated via full-information marginal maximum likelihood (MML) using the Laplace method (via the glmer function in R lme4). All fixed effects should be interpreted as unit-specific (i.e., as the fixed effect specifically for persons and items in which their corresponding random effect = 0). The significance of fixed effects was evaluated with Wald tests (i.e., the ratio of each estimate to its standard error using no denominator degrees of freedom) or likelihood ratio tests (i.e., the difference in model −2LL values). To evaluate new random effect variances and covariances, the likelihood ratio tests used a mixture of χ^2 distributions (with the two mixture degrees of freedom given in parentheses below) to determine the significance of the new random effect variances bounded at 0. Effect size was evaluated via pseduo-R^2 values for the proportion reduction in each variance component for level-2 random variance when appropriate [as well as odds ratios for individual slopes, not used here because I don't like them].

An empty means (no predictor) model with only a person random intercept variance had an intraclass correlation of ICC = .071 (using 3.29 as the logit-scale level-1 residual variance), indicating that 7.1% of the variance in accuracy was due to person mean differences, which was significant, $-2\Delta LL(0,1) = 81.73$, $p < .0001$. The extent of differences in item easiness was initially examined by treating items as fixed effects (i.e., in which item ID was a categorical predictor as a factor variable, otherwise known as a "Rasch" psychometric model). Significant differences in item easiness were observed, $-2\Delta LL(35) = 695.58$, $p < .0001$. To be able to quantify (and then predict) the extent of those item differences, we removed the categorical item ID predictor and instead added an item random intercept variance. The revised empty means model (with persons and items as crossed random effects at level 2) indicated that 7.9% of the variance in accuracy was due to person mean differences in ability and 20.6% was due to item mean differences in easiness; the latter was also significant, $-2\Delta LL(0,1) = 552.87$, $p < .0001$. We then attempted to predict those item easiness differences by using four item feaures: continuous visual clutter (z-score metric), whether the change was relevant to driving (binary), continuous brightness of the change (z-score metric), and whether the change was made to a legible sign (binary). Although the four item feaures accounted for 24.7% of the item random intercept variance, only one predictor was uniquely significant: Items with changes to legible signs had higher accuracy (logit estimate = 0.749, SE = 0.343, $p = .029$). Significant variance in item easiness remained, as indicated by a model comparison without the item random intercept variance, $-2\Delta LL(0,1) = 401.73$, $p < .0001$. Thus, items should be examined for further characteristics that translate into differences in easiness.

We then attempted to predict person ability differences by using three continuous person characteristics: age in years (centered at 75) and a z-scaled measure of vision impairment. The two person characteristics accounted for 23.3% of the person random intercept variance, and each predictor was uniquely significant: person mean accuracy was lower in persons who were older (logit estimate = −0.052, SE = 0.013, $p < .001$), or had greater vision impairment (logit estimate = −0.132, SE = 0.064, $p = .038$). [Significant variance in person ability remained, as indicated by a model comparison without the person random intercept variance…]. Thus, persons should be examined for further characteristics that translate into differences in ability.

We then examined the potential for person differences in the effects of the item features (i.e., a random slopes over level-2 persons). A person random slope variance for the effect of change brightness did not improve model fit, $-2\Delta LL(1,2) = 2.51$, $p = .199$, and nor did a person random slope variance for the other three item features. In parallel, we examined potential for item differences in the effects of the person features (i.e., a random slopes over level-2 items). An item random slope variance for the effect of person age did not improve model fit, $-2\Delta LL(1,2) = 0.835$, $p = .510$, and nor did an item random slope variance for the other person feature. Consequently, it appears that the effects of the item features were largely comparable across persons and that the effects of the person predictors were largely comparable across items.