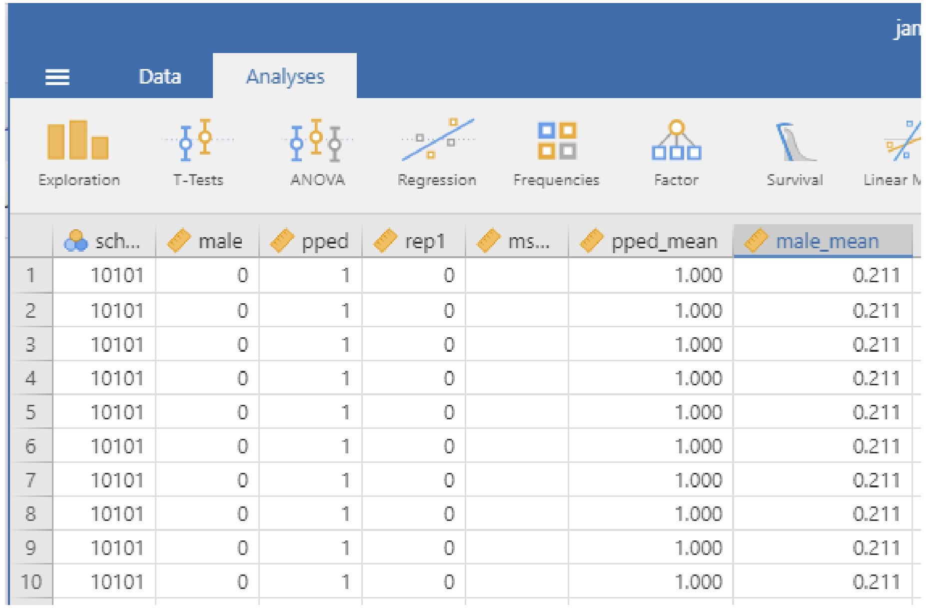
Multilevel binary logistic regression using jamovi

Mike Crowson, Ph.D.
March 11, 2020
(updated)

This demonstration is based on an example and data associated with Hox (2010; Chapter 6). Download SPSS data file (which can be read into jamovi) here: https://drive.google.com/open?id=1F7DifHMpKr2vlR7mx7V_B_kU_nSAg75o.



	sch...	male	pped	rep1	ms...	pped_mean	male_mean
1	10101	0	1	0		1.000	0.211
2	10101	0	1	0		1.000	0.211
3	10101	0	1	0		1.000	0.211
4	10101	0	1	0		1.000	0.211
5	10101	0	1	0		1.000	0.211
6	10101	0	1	0		1.000	0.211
7	10101	0	1	0		1.000	0.211
8	10101	0	1	0		1.000	0.211
9	10101	0	1	0		1.000	0.211
10	10101	0	1	0		1.000	0.211

We will be predicting the likelihood of a child repeating a grade during their primary education.

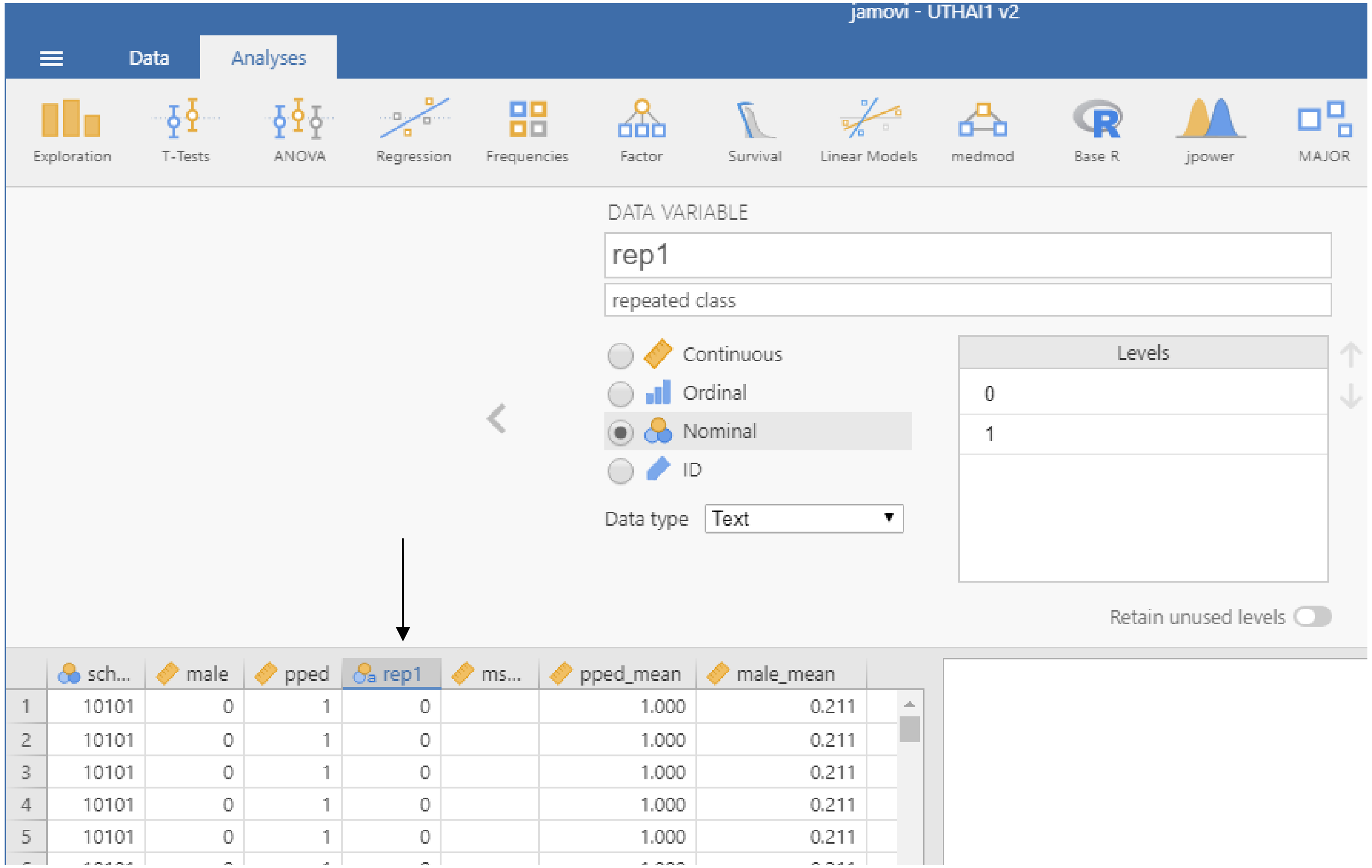
Level 1 (student-level) variables:

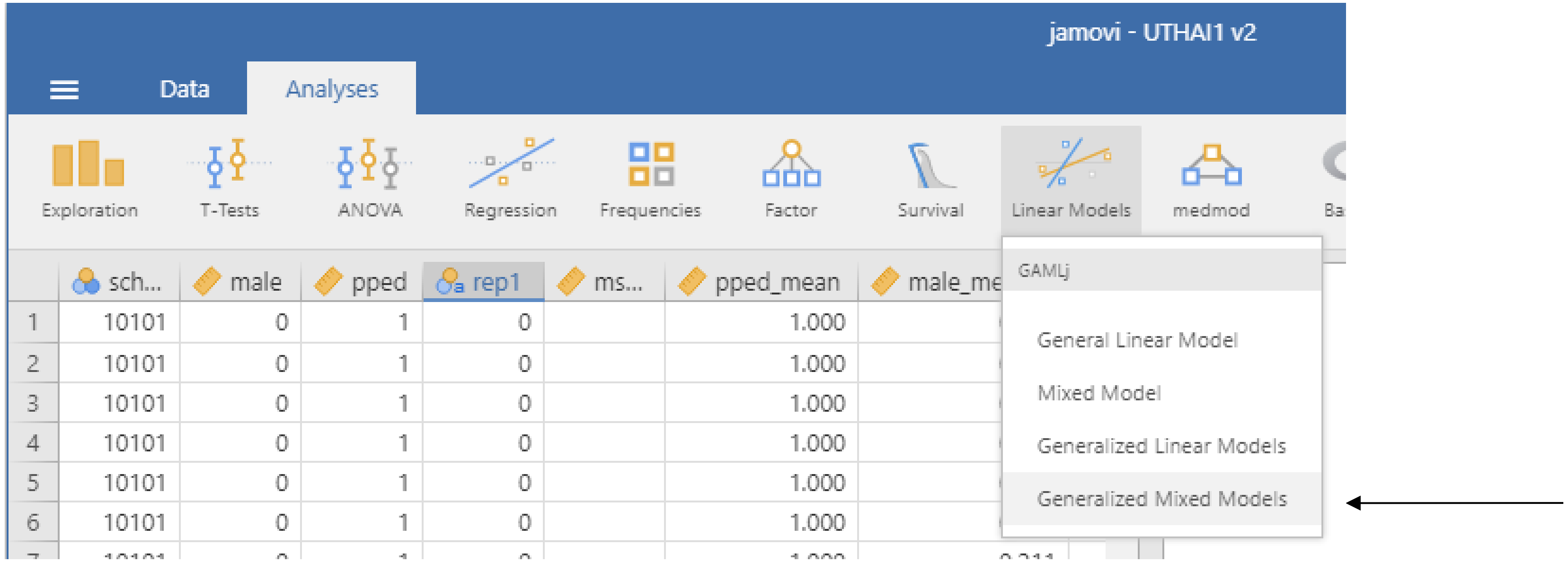
'rep1' = binary level 1 outcome (coded 1=child repeated a grade during primary education, 0=child did not repeat)
'male' = binary gender variable (coded 1=male, 0=female)
'pped' = binary variable indicating whether a child had preschool education (coded 1=yes, 0=no).

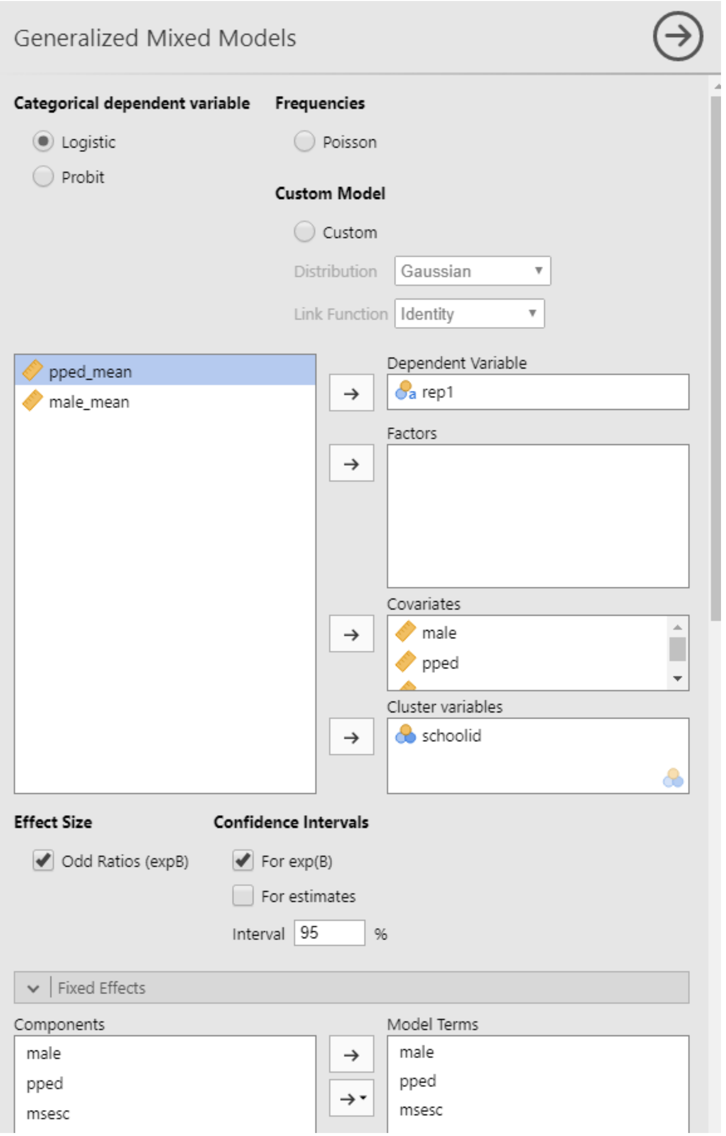
Level 2 (school-level) variables:
'msesc' = average SES associated with a school
'pped_mean' = proportion of students in a school who had preschool education (not in original dataset, but added for this demo)
'male_mean' = proportion of males at a school (not in original dataset, but added for this demo)

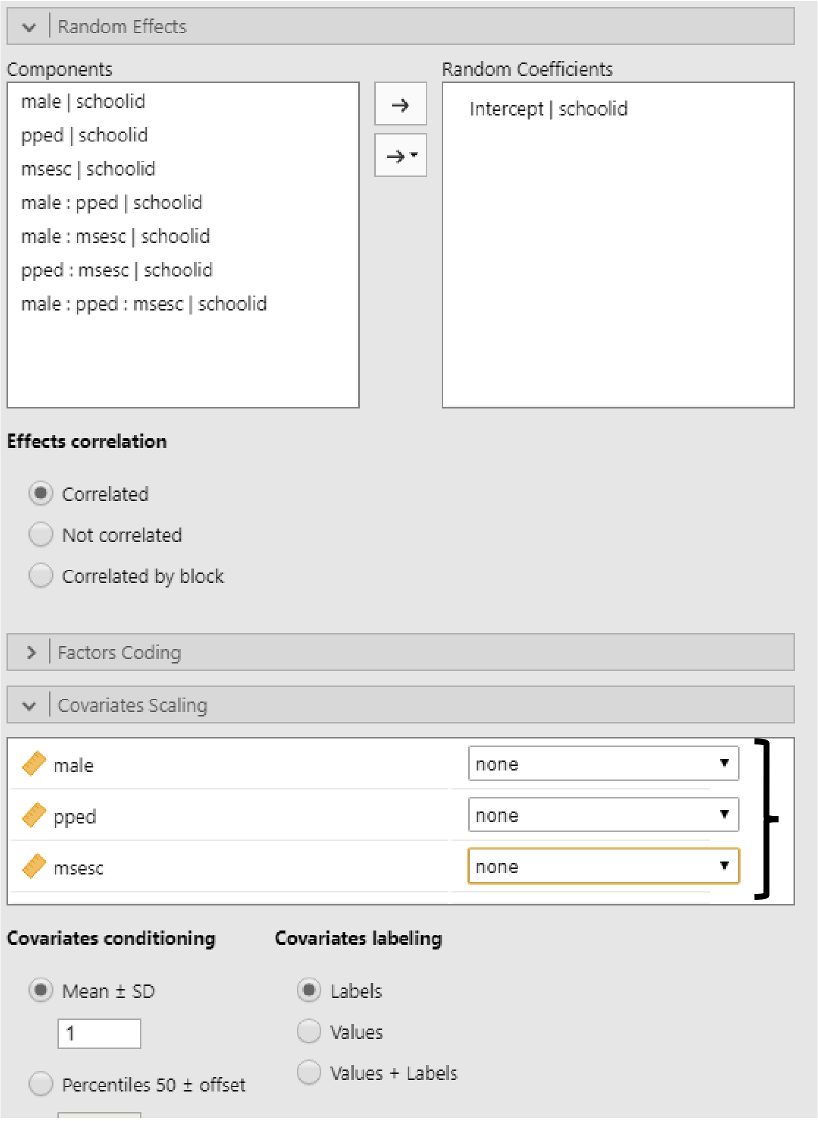
Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd edition). Routledge: New York.

Before trying to run the analysis, make sure that the Level 1 outcome is recognized as categorical. Here, I am changing it from the default continuous (after the data were read in from SPSS) to nominal.




jamovi - UTHAI1 v2
Data
Analyses
Exploration
T-Tests
ANOVA
Regression
Frequencies
Factor
Survival
Linear Models
medmod
sch...
male
pped
rep1
ms...
pped_mean
male_me
GAMLj
General Linear Model
Mixed Model
Generalized Linear Models
Generalized Mixed Models
1 10101 0 1 0 1.000
2 10101 0 1 0 1.000
3 10101 0 1 0 1.000
4 10101 0 1 0 1.000
5 10101 0 1 0 1.000
6 10101 0 1 0 1.000





Hox (2010) model with gender, pre-school education, and school mean SES as predictors of grade repeat.

Here, we have not centered the Level 1 or Level 2 predictors (to remain consistent with Hox's presentation).

Generalized Mixed Model

Model Info

Info	Value	Comment
Model Type	Logistic	Model for binary y
Call	glm	rep1 ~ 1 + male + pped + msesc + (1 | schoolid)
Link function	Logit	Log of the odd of y=1 over y=0
Direction	P(y=1)/P(y=0)	P(rep1 = 1) / P(rep1 = 0)
Distribution	Binomial	Dichotomous event distribution of y
LogLikel.	−2723.5310	Less is better
R-squared	0.0391	Marginal
R-squared	0.3547	Conditional
AIC	5457.0600	Less is better
BIC	5491.6860	Less is better
Deviance	4732.0637	Conditional
Residual DF	7511.0000	

[3]

The R-squared marginal represents the percentage decrease in the error of approximation (see https://gamlj.github.io/glmmixed_example1.html & https://gamlj.github.io/gzlm_specs.html) based on the fixed predictors in the model. The R-squared conditional indicates the percentage decrease in the error of approximation based on all fixed and random effects.

We interpret the Marginal R-squared as indicating that the inclusion of the fixed effects increases the ability to predict retention by 3.9% (see similar example of interpretation here: https://gamlj.github.io/gzlm_example2.html).

Model Results

Fixed Effect Omnibus tests

	X^2	df	p
male	52.07	1.00	< .001
pped	41.21	1.00	< .001
msesc	1.96	1.00	0.162

Fixed Effects Parameter Estimates

				95% Exp(B) Confidence Interval			
Names	Estimate	SE	exp(B)	Lower	Upper	z	p
(Intercept)	−2.233	0.1025	0.107	0.0877	0.131	−21.78	< .001
male	0.533	0.0738	1.703	1.4738	1.968	7.22	< .001
pped	−0.624	0.0973	0.536	0.4426	0.648	−6.42	< .001
msesc	−0.294	0.2104	0.745	0.4933	1.125	−1.40	0.162

Random Components

Groups	Name	SD	Variance
schoolid	(Intercept)	1.27	1.61
Residuals		1.00	1.00

Note. Number of Obs: 7516 , groups: schoolid , 356

In this model, there is no Level 1 residual variance since the "underlying probability distribution associated with Y_{ij} [the binary outcome] is not normally distributed" (Heck et al., 2015, p. 136). That is why the 1.0 appears here.

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.233	0.1025	0.107	0.0877	0.131	−21.78	< .001
male	0.533	0.0738	1.703	1.4738	1.968	7.22	< .001
pped	−0.624	0.0973	0.536	0.4426	0.648	−6.42	< .001
msesc	−0.294	0.2104	0.745	0.4933	1.125	−1.40	0.162

There are slight difference between table and computations below due to rounding.

Since ‘male’ is coded 0=female, 1=male & ‘pped’ is coded 0=did not attend preschool, 1=attended preschool, then the intercept in the model is the predicted log odds (logit) of retention for female students who did not attend preschool and who attend a school with an average SES of 0. The odds (i.e., probability of retention/probability of non-retention) for these individuals is .107 (see exp(B) column). [Note, the exp(B) column for the intercept contains *odds* pertaining to target group membership for persons with a value of 0 on all predictors. Subsequently, they are ratios of odds – hence odds ratios – representing multiplicative change in odds per unit increase on a predictor.] The predicted probability of retention for these individuals is: . The probability of non-retention is 1-.097 = .903. The odds allow us to say that the persons having a value of 0 on all predictors were .107 times as likely to be retained as to not be retained. Or, if we take the inverse of the odds, we can say persons having a value of 0 on all predictors were 1/.107 = 9.3 times more likely to not be retained than to be retained.

Note 1: An Excel file containing a worksheet (see sheet 1) for this slide and the next couple of slides can be downloaded here: https://drive.google.com/open?id=16lfYRe9f2BUX7EstKqFECghCfe8trTjj

Note 2: See Osborne (2015) for details on computations and verbiage when referring to logistic regression results.

[Osborne, J.W. (2015). *Best practices in logistic regression*. Sage: Los Angeles.]

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.233	0.1025	0.107	0.0877	0.131	−21.78	< .001
male	0.533	0.0738	1.703	1.4738	1.968	7.22	< .001
pped	−0.624	0.0973	0.536	0.4426	0.648	−6.42	< .001
msesc	−0.294	0.2104	0.745	0.4933	1.125	−1.40	0.162

There are slight difference between table and computations below due to rounding.

'Male' is a positive and significant predictor (b=.533, p<.001) of the likelihood of retention, indicating that males were more likely to be retained than females (reflected in the intercept estimate of -2.233). Formally speaking, the regression slope is interpreted as the predicted *difference* in logits (in this case, .533) between males and females. This difference remains the same so long as values on the remaining predictors are held constant between groups.

The odds associated with retention for males scoring 0 on the remaining predictors can be computed as: . The probability of retention for these males, therefore, will be: The probability of non-retention for these males is 1-.154 = .846.

The odds ratio you see in the exp(B) column for this predictor is a ratio of the odds for males (odds=.183) to odds for female (odds=.107) = .183/.107 = 1.70. This will be the same ratio of odds assuming the remaining predictors are held constant at some value between the two groups.

The odds of .182 indicates that males scoring 0 on the remaining predictors were .182 times as likely to be retained than to not be retained. If we compute its inverse, we can say that these males were 1/.182 = 5.495 times more likely to not be retained than to be retained.

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.233	0.1025	0.107	0.0877	0.131	−21.78	< .001
male	0.533	0.0738	1.703	1.4738	1.968	7.22	< .001
pped	−0.624	0.0973	0.536	0.4426	0.648	−6.42	< .001
msesc	−0.294	0.2104	0.745	0.4933	1.125	−1.40	0.162

There are slight difference between table and computations below due to rounding.

'pped' is a negative and significant predictor ($b=-.624$, $p<.001$) of the likelihood of retention, indicating that students who had received preschool education were less likely to be retained than those who had not received it. Formally speaking, the regression slope is interpreted as the predicted *difference* in logits (in this case, -.624) between those students who had received preschool (coded 1) and those who had not (coded 0; and who are reflected in the intercept estimate of -2.233). This difference remains the same so long as values on the remaining predictors are held constant between groups.

The odds associated with retention for students who had received preschool education and had a value of 0 on the remaining predictors is:. The probability of retention for these students, therefore, is: The probability of non-retention for these students is 1-.054 = .946.

The odds ratio in the exp(B) column is computed as the odds for pped=1 divided by the odds when pped=0, or simply .057/.107 = .533. This will be the same ratio of odds assuming the remaining predictors are held constant at some value between the two groups.

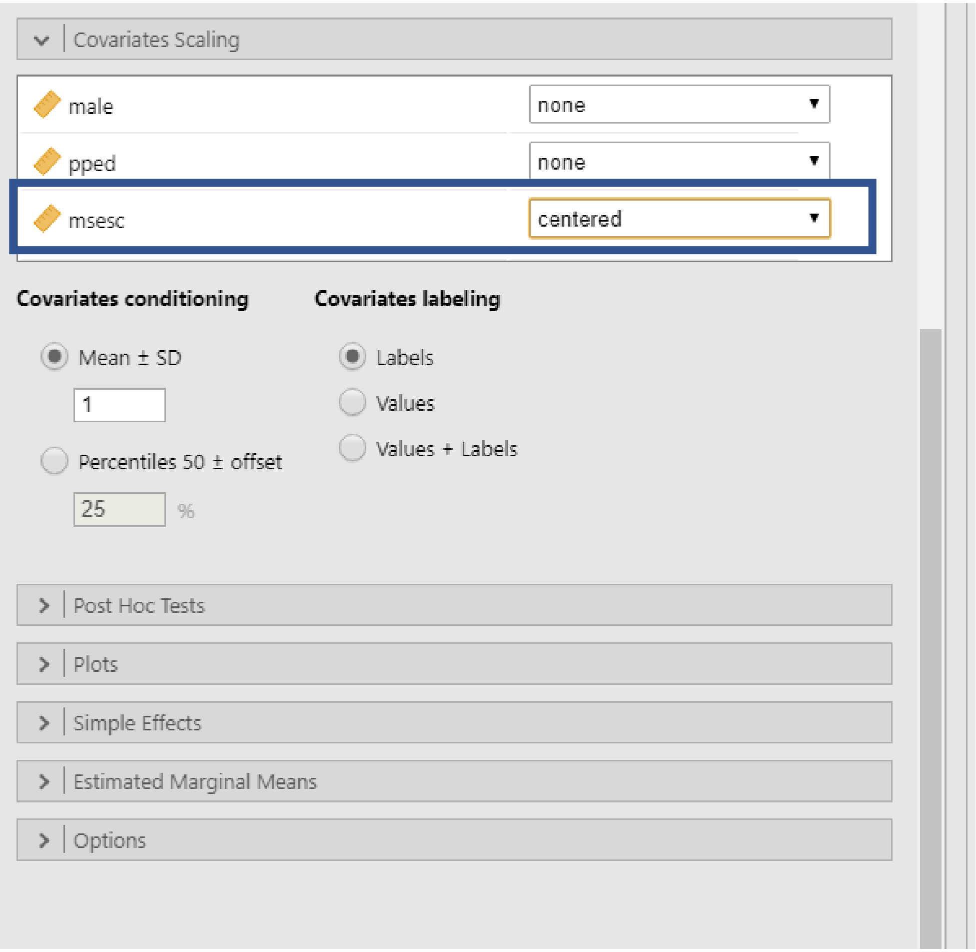
The odds of retention (i.e., .057) for students with pped=1 and scoring 0 on the remaining predictors indicates that students who had preschool education were .057 times as likely to be retained than to not be retained. If we compute the inverse of the odds, we can say that students who had preschool education were 1/.057 = 17.545 times more likely to not be retained than to be retained.

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval		z	p
				Lower	Upper		
(Intercept)	−2.233	0.1025	0.107	0.0877	0.131	−21.78	< .001
male	0.533	0.0738	1.703	1.4738	1.968	7.22	< .001
pped	−0.624	0.0973	0.536	0.4426	0.648	−6.42	< .001
msesc	−0.294	0.2104	0.745	0.4933	1.125	−1.40	0.162

There are slight difference between table and computations below due to rounding.

The regression slope for ‘msesc’ is negative, but non-significant (b=-.294, p=.162). Had it been significant, we would have been inferring that the likelihood for retention would be lower in schools with higher average SES.



[3]

Model Results

Fixed Effect Omnibus tests

	X^2	df	p
male	49.51	1.00	< .001
pped	39.30	1.00	< .001
msesc	1.92	1.00	0.165

Fixed Effects Parameter Estimates

				95% Exp(B) Confidence Interval			
Names	Estimate	SE	exp(B)	Lower	Upper	z	p
(Intercept)	−2.242	0.1059	0.106	0.0863	0.131	−21.17	< .001
male	0.534	0.0759	1.706	1.4700	1.979	7.04	< .001
pped	−0.627	0.0999	0.534	0.4394	0.650	−6.27	< .001
msesc	−0.297	0.2141	0.743	0.4884	1.130	−1.39	0.165

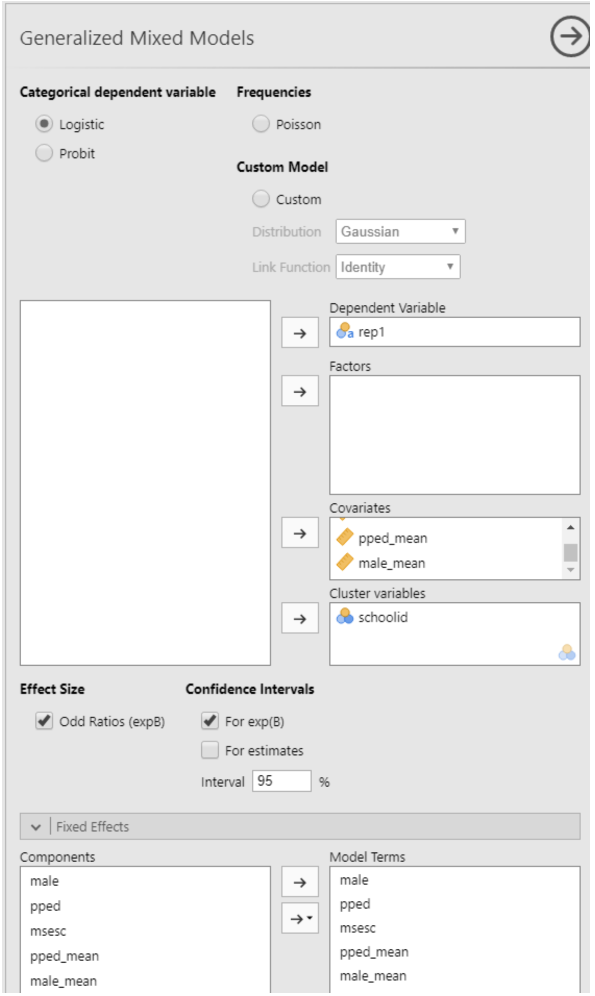
Random Components

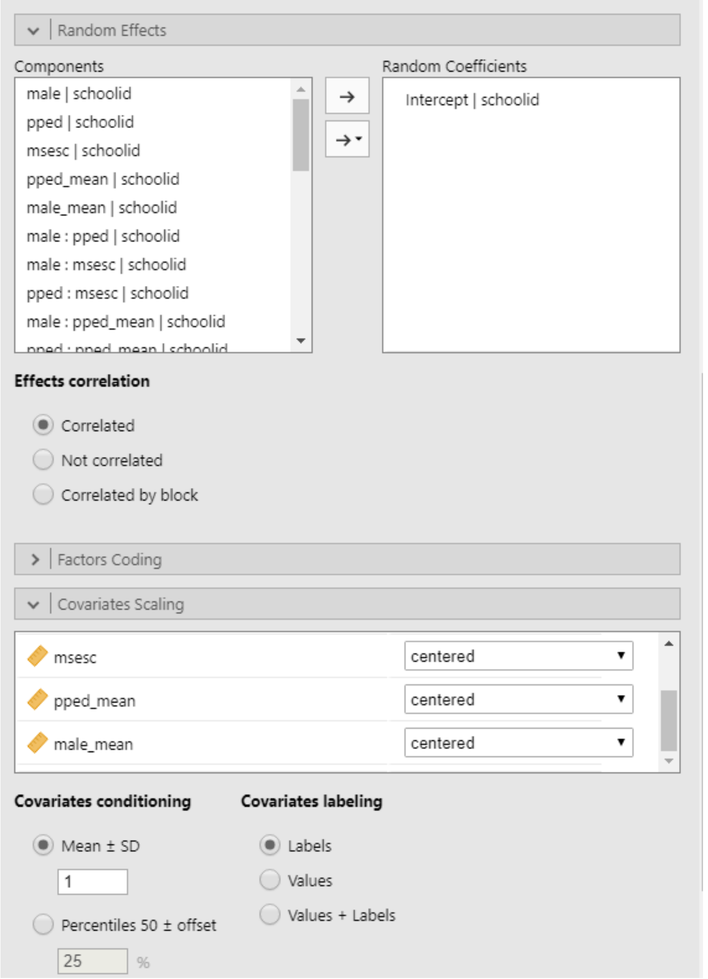
Groups	Name	SD	Variance
schoolid	(Intercept)	1.27	1.62
Residuals		1.00	1.00

Note. Number of Obs: 7516 , groups: schoolid , 356

If we mean center the Level 2 predictor, then the intercept is the predicted logit associated with retention for females who received no preschool and who were attending a school whose SES was at the mean of all the schools.

A limitation of the previous model is that there was no control for between-school differences on the predictors at Level 1. As such, the slopes for the within-group effects of the Level 1 predictors may be biased due to conflation of within-group variation and between-group variation. To account for this possibility, we will add in school-level means (aggregates) for the 'male' and 'pped' variables at Level 2.





We will leave all Level 2 (school-level) predictors centered in the model.

Generalized Mixed Model

Model Info

Info	Value	Comment
Model Type	Logistic	Model for binary y
Call	glm	rep1 ~ 1 + male + pped + msesc + pped_mean + male_mean + (1 | schoolid)
Link function	Logit	Log of the odd of y=1 over y=0
Direction	P(y=1)/P(y=0)	P(rep1 = 1) / P(rep1 = 0)
Distribution	Binomial	Dichotomous event distribution of y
LogLikel.	−2723.1776	Less is better
R-squared	0.0378	Marginal
R-squared	0.3531	Conditional
AIC	5460.3600	Less is better
BIC	5508.8288	Less is better
Deviance	4732.1638	Conditional
Residual DF	7509.0000	

[3]

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2356	0.1064	0.107	0.0868	0.132	−21.013	< .001
male	0.5404	0.0744	1.717	1.4838	1.986	7.265	< .001
pped	−0.6329	0.1079	0.531	0.4298	0.656	−5.863	< .001
msesc	−0.3161	0.2169	0.729	0.4766	1.115	−1.458	0.145
pped_mean	0.0459	0.2415	1.047	0.6522	1.681	0.190	0.849
male_mean	−0.4781	0.5699	0.620	0.2029	1.894	−0.839	0.401

Random Components

Groups	Name	SD	Variance
schoolid	(Intercept)	1.27	1.60
Residuals		1.00	1.00

Note. Number of Obs: 7516 , groups: schoolid , 356

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2356	0.1064	0.107	0.0868	0.132	−21.013	< .001
male	0.5404	0.0744	1.717	1.4838	1.986	7.265	< .001
pped	−0.6329	0.1079	0.531	0.4298	0.656	−5.863	< .001
msesc	−0.3161	0.2169	0.729	0.4766	1.115	−1.458	0.145
pped_mean	0.0459	0.2415	1.047	0.6522	1.681	0.190	0.849
male_mean	−0.4781	0.5699	0.620	0.2029	1.894	−0.839	0.401

We see that males were more likely to experience retention during primary school than females (b=.54, p<.001). The odds ratio (in the exp(B) column) indicates that the odds of retention for males are 1.717 times that of the odds for females.

The logit for males having a value of 0 on the remaining predictors can be computed as -2.2356+.5404 = -1.6952. The odds for these males is: .1835. The probability of retention among the males is: . The probability of non-retention is 1-.155 = .845.

Note: An Excel file containing a worksheet for this slide can be downloaded here: https://drive.google.com/open?id=16lfYRe9f2BUX7EstKqFECghCfe8trTjj . From here on out, I will not be computing all coefficients, etc. as shown above. The remaining calculations are represented on Sheet 2 from the Excel file.

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2356	0.1064	0.107	0.0868	0.132	−21.013	< .001
male	0.5404	0.0744	1.717	1.4838	1.986	7.265	< .001
pped	−0.6329	0.1079	0.531	0.4298	0.656	−5.863	< .001
msesc	−0.3161	0.2169	0.729	0.4766	1.115	−1.458	0.145
pped_mean	0.0459	0.2415	1.047	0.6522	1.681	0.190	0.849
male_mean	−0.4781	0.5699	0.620	0.2029	1.894	−0.839	0.401

We see that students who received preschool education were less likely to experience retention during primary school than those without preschool education (b=-.6329, p<.001). The odds ratio (in the exp(B) column) indicates that the odds of retention for students receiving preschool education are .531 times that of students not receiving preschool education.

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2356	0.1064	0.107	0.0868	0.132	−21.013	< .001
male	0.5404	0.0744	1.717	1.4838	1.986	7.265	< .001
pped	−0.6329	0.1079	0.531	0.4298	0.656	−5.863	< .001
msesc	−0.3161	0.2169	0.729	0.4766	1.115	−1.458	0.145
pped_mean	0.0459	0.2415	1.047	0.6522	1.681	0.190	0.849
male_mean	−0.4781	0.5699	0.620	0.2029	1.894	−0.839	0.401

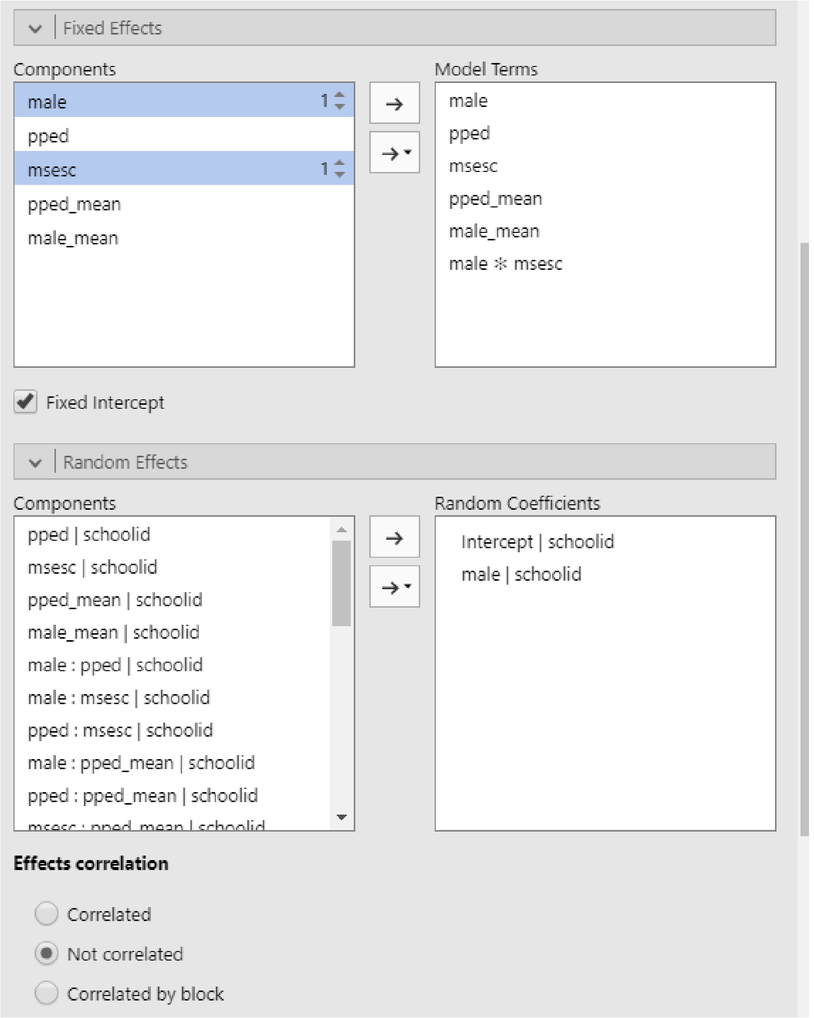
Random Components

Groups	Name	SD	Variance
schoolid	(Intercept)	1.27	1.60
Residuals		1.00	1.00

Note. Number of Obs: 7516 , groups: schoolid , 356

Notably, none of the Level 2 predictors were significant in the model.

Now, we will test one more model. We will allow the slope for ‘male’ to randomly vary across schools. Moreover, we will add a cross-level interaction: msesc*male. In other words, we are testing whether the effect of gender on the likelihood of retention varies across schools that themselves vary on average SES.



For this demo, we will assume no correlation between random intercepts and slopes across schools (Level 2 units). Moreover, we are leaving the previous centering of the Level 2 predictors in place.

Generalized Mixed Model

Model Info

Info	Value	Comment
Model Type	Logistic	Model for binary y
Call	glm	rep1 ~ 1 + male + pped + msesc + pped_mean + male_mean + male:msesc + (1 | schoolid) + (0 + male | schoolid)
Link function	Logit	Log of the odd of y=1 over y=0
Direction	P(y=1)/P(y=0)	P(rep1 = 1) / P(rep1 = 0)
Distribution	Binomial	Dichotomous event distribution of y
LogLikel.	−2720.8627	Less is better
R-squared	0.0367	Marginal
R-squared	0.3590	Conditional
AIC	5459.7300	Less is better
BIC	5522.0485	Less is better
Deviance	4670.2720	Conditional
Residual DF	7507.0000	

[3]

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2335	0.1107	0.107	0.0863	0.133	−20.182	< .001
male	0.5069	0.0890	1.660	1.3943	1.977	5.693	< .001
pped	−0.6386	0.1120	0.528	0.4240	0.658	−5.703	< .001
msesc	−0.4968	0.2471	0.608	0.3749	0.987	−2.011	0.044
pped_mean	0.0672	0.2478	1.070	0.6581	1.738	0.271	0.786
male_mean	−0.4725	0.5848	0.623	0.1982	1.961	−0.808	0.419
male ✻ msesc	0.2970	0.2117	1.346	0.8888	2.038	1.403	0.161

Random Components

Groups	Name	SD	Variance
schoolid	(Intercept)	1.249	1.559
	male	0.433	0.188
Residuals		1.000	1.000

Note. Number of Obs: 7516 , groups: schoolid , 356

Variance of intercepts and slopes are estimated

Fixed Effects Parameter Estimates

Names	Estimate	SE	exp(B)	95% Exp(B) Confidence Interval Lower	Upper	z	p
(Intercept)	−2.2335	0.1107	0.107	0.0863	0.133	−20.182	< .001
male	0.5069	0.0890	1.660	1.3943	1.977	5.693	< .001
pped	−0.6386	0.1120	0.528	0.4240	0.658	−5.703	< .001
msesc	−0.4968	0.2471	0.608	0.3749	0.987	−2.011	0.044
pped_mean	0.0672	0.2478	1.070	0.6581	1.738	0.271	0.786
male_mean	−0.4725	0.5848	0.623	0.1982	1.961	−0.808	0.419
male ✻ msesc	0.2970	0.2117	1.346	0.8888	2.038	1.403	0.161

The cross-level interaction was not statistically significant.

Ordinarily if it was, then you would seek to probe that interaction. Although we didn’t find a significant interaction, In jamovi, you can do this using the Plots option and Simple Effects option.