Appendices

There is one series of appendices for each software, introduced in the following order: Stata (pp. 2-4), R (pp. 5-8), Mplus (pp. 9-14), and SPSS (pp. 15-20). Each series of appendices gives the *general* commands to be used in each stage of the procedure (for the Justin Bieber study-*specific* commands, please see the datasets and syntax files). Across all appendices, the following terminology is used:

- "**id_cluster**" to the name of the level-2 identifier,
- "**outcome**" corresponds to the name of the binary outcome variable (Y_{ij})
- "**lvl1_iv**" corresponds to a level-1 predictor variable (x_{1ij}, x_{2ij}, …, or x_{Nij})
- "**lvl2_ iv**" corresponds to a level-2 predictor variable ($\mathbf{X1}_j$, $\mathbf{X2}_j$, …, or $\mathbf{XK}_j$)

Note: If you notice a mistake in the Stata or SPSS-related Appendices and/or syntax files, please report it to Nicolas Sommet (nicolas.sommet@unil.ch). If you notice a mistake in the Stata or SPSS-related Appendices and/or syntax files, please report it to Davide Morselli (davide.morselli@unil.ch).

Appendixes for Stata Users

Sub-Appendix A: Centering Variables (PRELIMINARY PHASE)

Below are the commands to grand-mean center a (level-1) variable (i.e., subtracting the *general* mean of the predictor variable). The new variable is "**lvl1_iv**_gmc".

```
egen grand_mean_lvl1_iv = mean(lvl1_iv)
gen lvl1_iv_gmc = lvl1_iv - grand_mean_lvl1_iv
```

Below are the commands to cluster-mean center a (level-1) variable (i.e., subtracting the *cluster-specific* mean of the predictor variable). The new variable is "**lvl1_iv**_cmc".

```
sort id_cluster
by id_cluster: egen cluster_mean_lvl1_iv = mean(lvl1_iv)
gen lvl1_iv_cmc = lvl1_iv - cluster_mean_lvl1_iv
```

Sub-Appendix B: Building the Empty Model (STEP #1)

Below are the commands to run an empty model, that is, a model containing no predictors, and calculate the intraclass correlation coefficient (ICC; the degree of homogeneity of the outcome within clusters).

```
xtmelogit outcome || id_cluster:, var
estat icc
```

Sub-Appendix C: Building the Intermediate Models (STEP #2)

If you focus on the between-observation effect of the (level-1) variable, you can use the grand-mean centered variable ("**lvl1_iv**_gmc"). If you focus on the within-cluster effect, use the cluster-mean centered variable ("**lvl1_iv**_cmc"). We arbitrarily choose to use the cluster-mean centered variable ("**lvl1_iv**_cmc") and we suppose that the level-2 variable is (grand-)mean centered ("**lvl2_iv**").

Below are the commands to run the constrained intermediate model (CIM; the model contains *all* level-1 variables, *all* level-2 variables well as *all intra*-level interactions) and store its deviance (the smaller the deviance, the better the fit).

```
xtmelogit outcome lvl1_iv_cmc lvl2_iv || id_cluster:, var
display "FYI: The deviance of the CIM is:" -2*e(ll)
estimate store CIM
```

Below are the commands to run the augmented intermediate model (AIM; the model similar to the constrained intermediate model with the exception that it estimates the random slope variance of the relevant lower-level variable(s)) and store its deviance.

```
xtmelogit outcome lvl1_iv_cmc lvl2_iv || id_cluster: lvl1_iv, var
display "FYI: The deviance of the AIM is:" -2*e(ll)
estimate store AIM
```

Below is the command to determine whether considering the variation of the level-1 variable from one cluster to another improves the model. The software performs a likelihood-ratio test LR χ^2 (1) comparing the deviance of the CIM with the deviance of the AIM.

```
lrtest CIM AIM
```

Notes: If you have one or more additional level-1 variable(s) and/or level-2 variable(s), just list them with the others (in the fixed part) along with the *intra*-level interaction; if you want to estimate the random slope variance of an additional level-1 variable, the AIM will contain the two random slope variance components: You will first compare the AIM with a CIM1 including the random slope variance component of the first level-1 variable, and *then* with a CIM2 including the random slope variance component of the second level-1 variable.

Sub-Appendix D: Building the Final Model (STEP #3)

Below is the command to run the final model (adding the cross-level interaction(s)).

```
xtmelogit outcome c.lvl1_iv_cmc##c.lvl2_iv || id_cluster: lvl1_iv_cmc, or var
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = - 1SD below the mean).

```
egen sd_lvl1_iv = sd(lvl1_iv_cmc)
gen lvl1_iv_m1SD = lvl1_iv_cmc + sd_ lvl1_iv
xtmelogit outcome c.lvl1_iv_m1SD##c.lvl2_iv || id_cluster: lvl1_iv_cmc, or var
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = + 1SD above the mean).

```
egen sd_lvl1_iv = sd(lvl1_iv_cmc)
gen lvl1_iv_p1SD = lvl1_iv_cmc - sd_ lvl1_iv
```

`xtmelogit` **outcome** `c.`**lvl1_iv**`_p1SD##c.`**lvl2_iv** `||` **id_cluster**`:` **lvl1_iv**`_cmc, or var`

Note: we have decided here to consider the random slope variance component.

Sub-Appendix E: Reshaping your data (for repeated measure and/or longitudinal data)

This sub-appendix *only* apply to those having repeated measure design and/or longitudinal data in which observations are nested in participants. Below is the command to reshape the dataset so as to have one line per observation (rather than per participant).

`reshape long` **outcome lvl1_iv**`, i(`**id_cluster**`) j(id_observation)`

Notes: In the original dataset, there should be *n* columns for the outcome variable, namely "outcome1," "outcome2," "…," "outcome*n*," as well as for the level-1 variable, namely "lvl1-iv1," "lvl1-iv2," "…," "lvl1-iv*n*"; "**id_cluster**" is here the participant identifier and "id_observation" will be a new variable identifying the level-1 observations.

Sub-Appendix F: Running a Two-Level Cross-Classified Multilevel Logistic Model

This sub-appendix is an *example* of a final multilevel logistic model for a three-level cross-classified data structure. Below is the command of such a model, in which observations are nested in non-hierarchical clusters: "**id_cluster_2a**" and "**id_cluster_2b**."

`xtmelogit` **outcome lvl1_iv lvl2_iv** `|| _all: R.`**id_cluster_2a** `||` **id_cluster_2b:**`, or var`

Notes: the interaction term is not considered; we have decided here not to consider any random slope variance component.

Sub-Appendix G: Running a Multilevel Poisson Model

This sub-appendix is an *example* of a final multilevel poison model. Below is the command of such a model, in which "**outcome_count**" is a count variable (typically corresponding to a number of occurrences)

`xtmepoisson` **outcome_count lvl1_iv lvl2_iv** `||` **id_cluster**`:, irr var`

Notes: the interaction term is not considered; we have decided here not to consider any random slope variance component.

Appendixes for R Users

Sub-Appendix A: Centering Variables (PRELIMINARY PHASE)

Below are the commands to grand-mean center a (level-1) variable (i.e., subtracting the *general* mean of the predictor variable). The new variable is "**lvl1_iv**_gmc".

```
grand_mean_lvl1_iv = mean(your_data$lvl1_iv, na.rm=T)
your_data$lvl1_iv_gmc = your_data$lvl1_iv - grand_mean_lvl1_iv
```

Below are the commands to cluster-mean center a (level-1) variable (i.e., subtracting the *cluster-specific* mean of the predictor variable). The new variable is "**lvl1_iv**_cmc".

```
cluster_mean_lvl1_iv <- data.frame(tapply(your_data$lvl1_iv, your_data$id_cluster, mean))
cluster_mean_lvl1_iv$id_cluster <- rownames(cluster_mean_lvl1_iv)
names(cluster_mean_lvl1_iv)<- c("cluster_mean_lvl1_iv","id_cluster")
your_data <- merge(your_data, cluster_mean_lvl1_iv, by="id_cluster")
your_data$lvl1_iv_cmc <- your_data$lvl1_iv – your_data$cluster_mean_lvl1_iv
```

Sub-Appendix B: Building the Empty Model (STEP #1)

Below are the commands to run an empty model, that is, a model containing no predictors, and calculate the intraclass correlation coefficient (ICC; the degree of homogeneity of the outcome within clusters).

```
M0 <- glmer(outcome ~ ( 1 | id_cluster), data=your_data, family = "binomial")
summary(M0)
icc <- M0@theta[1]^2/ (M0@theta[1]^2 + (3.14159^2/3))
icc
```

Sub-Appendix C: Building the Intermediate Models (STEP #2)

If you focus on the between-observation effect of the (level-1) variable, you can use the grand-mean centered variable ("**lvl1_iv**_gmc"). If you focus on the within-cluster effect, use the cluster-mean centered variable ("**lvl1_iv**_cmc"). We arbitrarily choose to use the cluster-mean centered variable ("**lvl1_iv**_cmc") and we suppose that the level-2 variable is (grand-)mean centered ("**lvl2_iv**").

Below are the commands to run the constrained intermediate model (CIM; the model contains *all* level-1 variables, *all* level-2 variables well as *all intra*-level interactions).

```
CIM <- glmer(outcome ~ lvl1_iv_cmc + lvl2_iv + (1 | id_cluster), data = your_data, family = "binomial")
```

```
summary(CIM)
```

Below are the commands to run the augmented intermediate model (AIM; the model similar to the constrained intermediate model with the exception that it estimates the random slope variance of the relevant lower-level variable(s)).

```
AIM <- glmer(outcome ~ lvl1_iv_cmc + lvl2_iv + (1 + lvl1_iv_cmc || id_cluster), data = your_data, family = "binomial")
summary(AIM)
```

Below is the command to determine whether considering the variation of the level-1 variable from one cluster to another improves the model. The software performs a likelihood-ratio test LR χ^2 (1) comparing the deviance of the CIM with the deviance of the AIM.

```
anova(CIM, AIM)
```

Notes: Sometimes the glmer function prompts a warning message of non-convergence. To understand if it is a problem, run the analysis with different optimizers (e.g., by adding `control=glmerControl(optimizer="bobyqa")`, to the model) and compare the results. In most cases, the differences are extremely small (at the fourth decimal place) and do not compromise the interpretation of the results (see "lme4 convergence warnings: Troubleshooting," 2017); if you have one or more additional level-1 variable(s) and/or level-2 variable(s), just list them with the others (in the fixed part) along with the *intra*-level interaction; if you want to estimate the random slope variance of an additional level-1 variable, the AIM will contain the two random slope variance components: You will first compare the AIM with a CIM1 including the random slope variance component of the first level-1 variable, and *then* with a CIM2 including the random slope variance component of the second level-1 variable.

Sub-Appendix D: Building the Final Model (STEP #3)

Below are the commands to run the final model (adding the cross-level interaction(s)).

```
FM <- glmer(outcome ~ lvl1_iv_cmc + lvl2_iv + (1 + lvl1_iv_cmc || id_cluster) + lvl1_iv_cmc:lvl2_iv, data = your_data, family = "binomial")
summary(FM)
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = - 1SD below the mean).

```
sd_lvl1_iv_cmc <- sd(your_data$lvl1_iv_cmc)
your_data$lvl1_iv_m1SD <- your_data$lvl1_iv_cmc + sd_lvl1_iv_cmc
```

```
FM_m1Syour_data <- glmer(outcome ~ lvl1_iv_m1SD + lvl2_iv + (1 +
lvl1_iv_cmc || id_cluster) + lvl1_iv_m1SD:lvl2_iv, data = your_data, family =
"binomial")
summary(FM_m1SD)
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = + 1SD above the mean).

```
sd_lvl1_iv_cmc <- sd(your_data$lvl1_iv_cmc)
your_data$lvl1_iv_p1SD <- your_data$lvl1_iv_cmc - sd_lvl1_iv_cmc
FM_p1Syour_data <- glmer(outcome ~ lvl1_iv_p1SD + lvl2_iv + (1 +
lvl1_iv_cmc || id_cluster) + lvl1_iv_p1SD:lvl2_iv, data = your_data, family =
"binomial")
summary(FM_p1SD)
```

Note: we have decided here to consider the random slope variance component.

Sub-Appendix E: Reshaping your data (for repeated measure and/or longitudinal data)

This sub-appendix *only* apply to those having repeated measure design and/or longitudinal data in which observations are nested in participants. Below is the command to reshape the dataset so as to have one line per observation (rather than per participant).

```
reshape(data = your_data, direction = "long", idvar =
"id_observation", timevar = "id_cluster", varying = c("outcome1",
"outcome2" … "outcomen" ,"lvl1-iv1", "lvl1-iv2" … "lvl1-ivn"), sep =
"")
```

Notes: In the original dataset, there should be *n* columns for the outcome variable, namely "outcome1," "outcome2," "…," "outcome*n*," as well as for the level-1 variable, namely "lvl1-iv1," "lvl1-iv2," "…," "lvl1-iv*n*"; "**id_cluster**" is here the participant identifier and "id_observation" will be a new variable identifying the level-1 observations.

Sub-Appendix F: Running a Two-Level Cross-Classified Multilevel Logistic Model

This sub-appendix is an *example* of a (random intercept) multilevel logistic model for a two-level cross-classified data structure. Below is the command of such a model, in which observations are nested in non-hierarchical clusters: "**id_cluster_2a**" and "**id_cluster_2b**."

```
CCM <- glmer(outcome ~ lvl1_iv + lvl2_iv + (1 | id_cluster_2a) +
(1 | id_cluster_2b), data = your_data, family = "binomial")
summary(CCM)
```

Notes: the interaction term is not considered; we have decided here not to consider any random slope variance component.

Sub-Appendix G: Running a Multilevel Poisson Model

This sub-appendix is an *example* of a final multilevel poison model. Below is the command of such a model, in which "**outcome_count**" is a count variable (typically corresponding to a number of occurrences)

```
PM <- glmer( outcome_count ~ lvl1_iv + lvl2_iv + (1 |
id_cluster_2a), data = your_data, family = "poisson")
summary(PM)
```

Notes: the interaction term is not considered; we have decided here not to consider any random slope variance component.

Appendixes for Mplus Users

Sub-Appendix A: Centering Variables (PRELIMINARY PHASE)

Below are the commands to grand-mean center a (level-1) variable (i.e., subtracting the *general* mean of the predictor variable).

```
DEFINE:
CENTER lvl1_iv (GRANDMEAN);
```

Below are the commands to cluster-mean center a (level-1) variable (i.e., subtracting the *cluster-specific* mean of the predictor variable).

```
DEFINE:
CENTER lvl1_iv (GROUPMEAN);
```

Sub-Appendix B: Building the Empty Model (STEP #1)

Below are the commands to run an empty model, that is, a model containing no predictors, and calculate the intraclass correlation coefficient (ICC; the degree of homogeneity of the outcome within clusters).

```
VARIABLE:
USEVARIABLES = outcome;
CLUSTER = id_cluster;
CATEGORICAL = outcome;
ANALYSIS:
TYPE = twolevel random;
ESTIMATOR = MLR;
MODEL:
%WITHIN%
%BETWEEN%
outcome (p1);
MODEL CONSTRAINTS:
new ICC;
ICC = p1 / (p1 + (3.14159^2/3));
```

Sub-Appendix C: Building the Intermediate Models (STEP #2)

If you focus on the between-observation effect of the (level-1) variable, you can use the grand-mean centered variable (see Sub-Appendix A). If you focus on the within-cluster effect, use the cluster-mean centered variable (see Sub-Appendix A). We arbitrarily choose to use the cluster-mean centered variable and we suppose that the level-2 variable is (grand-)mean centered ("**lvl2_iv**").

Below are the commands to run the constrained intermediate model (CIM; the model contains *all* level-1 variables, *all* level-2 variables well as *all intra*-level interactions) and store its deviance (the smaller the deviance, the better the fit).

```
VARIABLE:
USEVARIABLES = outcome lvl1_iv lvl2_iv;
BETWEEN = lvl2_iv;
WITHIN = lvl1_iv;
CLUSTER = id_cluster;
CATEGORICAL = outcome;
ANALYSIS:
TYPE = twolevel random;
ESTIMATOR = MLR;
MODEL:
%WITHIN%
outcome ON lvl1_iv;
%BETWEEN%
outcome ON lvl2_iv;
```

To run the augmented intermediate model (AIM; the model similar to the constrained intermediate model with the exception that it estimates the random slope variance of the relevant lower-level variable(s)) substitute to the MODEL section of the previous syntax the following:

```
MODEL:
%WITHIN%
S1 | outcome ON lvl1_iv;
%BETWEEN%
outcome ON lvl2_iv;
```

Notes: To determine whether considering the variation of the level-1 variable from one cluster to another improves the model, you have compare the deviance of the CIM with the deviance of the AIM using a likelihood-ratio test LR χ^2 (1). The deviance is computed as the -2*loglikelihood of each model in the sections `MODEL FIT INFORMATION (H0 Value)`; the LR χ^2 (1) has to be calculated by hand using the formula given in Eq. 9; If you have one or more additional level-1 variable(s) and/or level-2 variable(s), just list them with the others (in the fixed part) along with the *intra*-level interaction; if you want to estimate the random slope variance of an additional level-1 variable, the AIM will contain the two random slope variance components: You will first compare the AIM with a CIM1 including the random slope variance component of the first level-1 variable, and *then* with a CIM2 including the random slope variance component of the second level-1 variable.

Sub-Appendix D: Building the Final Model (STEP #3)

Below is the command to run the final model (adding the cross-level interaction(s)) and odds ratios.

```
MODEL:
%WITHIN%
S1 |outcome ON lvl1_iv;
%BETWEEN%
[s1] (l1);
outcome ON lvl2_iv (l2);
S1 ON lvl2_iv (l3);
MODEL CONSTRAINTS:
new OR_lvl1_iv OR_lvl2_iv OR3_interaction;
OR_lvl1_iv = exp(l1);
OR_lvl2_iv = exp(l2);
OR3_interaction = exp(l3);
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = - 1SD below the mean).

First, add the following lines to the respective sections:

```
DEFINE:
lvl1_iv_m1SD = lvl1_iv + sqrt(variance of the outcome which can be found in the output of the PRELIMINARY PHASE)

VARIABLE:
USEVARIABLES = lvl1_iv_m1SD;
```

N.b., the line `WITHIN =` **outcome** has to be replaced by `WITHIN =` **lvl1_iv**`_m1SD`.

Finally, the following lines specify the model for - 1 SD outcome.

```
MODEL:
%WITHIN%
s1 | outcome ON lvl1_iv_m1SD;
```

```
%BETWEEN%
outcome ON lvl2_iv (l4);
s1 ON lvl2_iv;
MODEL CONSTRAINTS:
new OR_m1SD;
OR_m1SD = exp(l4);
```

Below are the commands to calculate the first simple slope run, namely the effect of “**lvl2_iv**” when “**lvl1_iv**” = + 1SD above the mean).

First add the following lines to the respective sections:

```
DEFINE:
lvl1_iv_p1SD = lvl1_iv - sqrt(variance of the outcome which can be found in the output of the PRELIMINARY PHASE)
```

```
VARIABLE:
USEVARIABLES = lvl1_iv_p1SD;
```

N.b., the line `WITHIN = outcome` has to be replaced by `WITHIN = lvl1_iv_p1SD`.

Finally, the following lines specify the model for - 1 SD outcome.

```
MODEL:
%WITHIN%
s1 | outcome ON lvl1_iv_p1SD;
%BETWEEN%
outcome ON lvl2_iv (l5);
s1 ON lvl2_iv;
MODEL CONSTRAINTS:
new OR_p1SD;
OR_p1SD = exp(l5);
```

Note: we consider here the random slope variance component of the cluster-mean centered variable +/- 1 SD because Mplus does not allow the simultaneous use of different scales of the same variable. The change are generally minor.

Sub-Appendix E: Reshaping your data (for repeated measure and/or longitudinal data)

The current version of Mplus cannot perform data reshaping.

Sub-Appendix F: Running a Two-Level Cross-Classified Multilevel Logistic Model

This sub-appendix is an *example* of a (random intercept) multilevel logistic model for a two-level cross-classified data structure. Below is the command of such a model, in which observations are nested in non-hierarchical clusters: "**id_cluster_2a**" and "**id_cluster_2b**."

```
VARIABLE:
USEVARIABLES = outcome lvl1_iv lvl2_iv ;
CLUSTER = id_cluster_2a id_cluster_2b;
WITHIN = lvl1_iv;
BETWEEN = (id_cluster_2a) lvl2_iv;
ANALYSIS: TYPE = CROSSCLASSIFIED RANDOM;
MODEL:
%WITHIN%
outcome ON lvl1_iv;
%BETWEEN id_cluster_2a%
outcome ON lvl2_iv;
```

Notes: the interaction term is not considered; we have decided here not to consider any random slope variance component.

Sub-Appendix G: Running a Multilevel Poisson Model

To run a multilevel poison model in Mplus is sufficient to specify the outcome variable as count instead of categorical in the VARIABLE section of the syntax. Below is the command of such a model, in which "**outcome_count**" is a count variable (typically corresponding to a number of occurrences)

```
VARIABLE:
USEVARIABLES = outcome_count lvl1_iv lvl2_iv ;
BETWEEN = lvl2_iv;
WITHIN = lvl1_iv;
CLUSTER = id_cluster;
COUNT = outcome_count;
ANALYSIS:
TYPE = twolevel random;
ESTIMATOR = MLR;
MODEL:
%WITHIN%
outcome_count ON lvl1_iv;
```

```
%BETWEEN%
outcome_count ON lvl2_iv;
```

Appendixes for SPSS Users

Sub-Appendix A: Centering Variables (PRELIMINARY PHASE)

Below are the commands to grand-mean center a (level-1) variable (i.e., subtracting the *general* mean of the predictor variable). The new variable is "**lvl1_iv**_gmc".

```
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=
  /grand_mean_lvl1_iv =MEAN(lvl1_iv).
COMPUTE lvl1_iv_gmc= lvl1_iv - grand_mean_lvl1_iv.
EXECUTE.
```

Below are the commands to cluster-mean center a (level-1) variable (i.e., subtracting the *cluster-specific* mean of the predictor variable). The new variable is "**lvl1_iv**_cmc".

```
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=id_cluster
  /cluster_mean_lvl1_iv =MEAN(lvl1_iv).
COMPUTE lvl1_iv_cmc= lvl1_iv - cluster_mean_lvl1_iv.
EXECUTE.
```

Sub-Appendix B: Building the Empty Model (STEP #1)

Below are the commands to run an empty model, that is, a model containing no predictors, and calculate the intraclass correlation coefficient (ICC; the degree of homogeneity of the outcome within clusters).

```
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=id_cluster
  /FIELDS TARGET= outcome TRIALS=NONE OFFSET=NONE
  /TARGET_OPTIONS DISTRIBUTION=MULTINOMIAL LINK=LOGIT
  /FIXED USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
    CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
    SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Notes: By default, SPSS does not use "1" as the target category; TARGET_CATEGORY_ORDER=ASCENDING allows you to predict the odds that the outcome variable equals "1" instead of "0" (rather than the reverse); SPSS does not provide you with the ICC; the random intercept variance (i.e., the level-2 residual) var(u_{0j}) is [var(intercept)]² (in the seventh windows of the Model Viewer "Covariance parameters," Block 1) and the ICC has to be calculated by hand using the formula given in Eq. 6.

Sub-Appendix C: Building the Intermediate Models (STEP #2)

If you focus on the between-observation effect of the (level-1) variable, you can use the grand-mean centered variable ("**lvl1_iv**_gmc"). If you focus on the within-cluster effect, use the cluster-mean centered variable ("**lvl1_iv**_cmc"). We arbitrarily choose to use the cluster-mean centered variable ("**lvl1_iv**_cmc") and we suppose that the level-2 variable is (grand-)mean centered ("**lvl2_iv**").

Below is the command to run the constrained intermediate model (CIM; the model contains *all* level-1 variables, *all* level-2 variables well as *all intra*-level interactions).

```
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=id_cluster  /FIELDS TARGET= outcome TRIALS=NONE
OFFSET=NONE
  /TARGET_OPTIONS REFERENCE=1 DISTRIBUTION=MULTINOMIAL LINK=LOGIT
  /FIXED EFFECTS=lvl1_iv_cmc  lvl2_iv
USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
    CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
    SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Below is the command to run the run the augmented intermediate model (AIM; the model similar to the constrained intermediate model with the exception that it estimates the random slope variance of the relevant lower-level variable(s)).

```
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=id_cluster  /FIELDS TARGET= outcome TRIALS=NONE
OFFSET=NONE
  /TARGET_OPTIONS REFERENCE=1 DISTRIBUTION=MULTINOMIAL LINK=LOGIT
  /FIXED EFFECTS=lvl1_iv_cmc  lvl2_iv
```

```
USE_INTERCEPT=TRUE
 /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
 /RANDOM EFFECTS=lvl1_iv_cmc USE_INTERCEPT=FALSE
COVARIANCE_TYPE=VARIANCE_COMPONENTS
 /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
  CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
  SINGULAR=0.000000000001
 /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Notes: If you have one or more additional level-1 variable(s) and/or level-2 variable(s), just list them with the others (in the fixed part) along with the *intra*-level interaction; when a multilevel *logistic* regression model includes the estimation of a random slope variance component, SPSS will often—if not always—fail to estimate model parameters; thus, you will not be able to achieve STEP #2 and the three-step procedure *cannot* be completed; In SPSS, "-2 log pseudo likelihood" (in the first window of the Model Viewer) should be taken as the model deviance statistics; the LR χ^2 (1) has to be calculated by hand using the formula given in Eq. 9.

Sub-Appendix D: Building the Final Model (STEP #3)

Below are the commands to compute the cross-level interaction and to run the final model (adding the cross-level interaction(s)).

```
COMPUTE lvl1_iv_cmcXlvl2_iv = lvl1_iv_cmc * lvl2_iv.

GENLINMIXED
 /DATA_STRUCTURE SUBJECTS=id_cluster  /FIELDS TARGET= outcome TRIALS=NONE
OFFSET=NONE
 /TARGET_OPTIONS REFERENCE=1 DISTRIBUTION=MULTINOMIAL LINK=LOGIT
 /FIXED EFFECTS=lvl1_iv_cmc  lvl2_iv lvl1_iv_cmcXlvl2_iv
USE_INTERCEPT=TRUE
 /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
 /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
  CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
  SINGULAR=0.000000000001
 /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = - 1SD below the mean).

```
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=
  /sd_lvl1_iv =SD(lvl1_iv_cmc).
COMPUTE lvl1_iv_m1SD= lvl1_iv_cmc + sd_lvl1_iv.
COMPUTE lvl1_iv_m1SDX lvl2_iv = lvl1_iv_m1SD*lvl2_iv.
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=id_cluster /FIELDS TARGET= outcome TRIALS=NONE
OFFSET=NONE
  /TARGET_OPTIONS REFERENCE=1 DISTRIBUTION=MULTINOMIAL LINK=LOGIT
  /FIXED EFFECTS= lvl1_iv_m1SD lvl2_iv lvl1_iv_m1SDX lvl2_iv
USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
    CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
    SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Below are the commands to calculate the first simple slope run, namely the effect of "**lvl2_iv**" when "**lvl1_iv**" = + 1SD above the mean).

```
COMPUTE lvl1_iv_p1SD= lvl1_iv_cmc - sd_lvl1_iv.
COMPUTE lvl1_iv_p1SDX lvl2_iv = lvl1_iv_p1SD*lvl2_iv.
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=id_cluster /FIELDS TARGET= outcome TRIALS=NONE
OFFSET=NONE
  /TARGET_OPTIONS REFERENCE=1 DISTRIBUTION=MULTINOMIAL LINK=LOGIT
  /FIXED EFFECTS= lvl1_iv_p1SD lvl2_iv lvl1_iv_p1SDX lvl2_iv
USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE SUBJECTS=id_cluster
COVARIANCE_TYPE=VARIANCE_COMPONENTS
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=ASCENDING MAX_ITERATIONS=100
```

```
        CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL
PCONVERGE=0.000001(ABSOLUTE) SCORING=0
        SINGULAR=0.000000000001
       /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Note: we have decided here not to consider random slope variance.

Sub-Appendix E: Reshaping your data (for repeated measure and/or longitudinal data)

This sub-appendix *only* apply to those having repeated measure design and/or longitudinal data in which observations are nested in participants. Below is the command to reshape the dataset so as to have one line per observation (rather than per participant).

```
VARSTOCASES
/MAKE outcome FROM outcome1 outcome2 outcomen
/MAKE lvl1_iv FROM lvl1_iv1 lvl1_iv2 lvl1_ivn
/INDEX = id_observation.
```

Notes: In the original dataset, there should be *n* columns for the outcome variable, namely "outcome1," "outcome2," "…," "outcome*n*," as well as for the level-1 variable, namely "lvl1-iv1," "lvl1-iv2," "…," "lvl1-iv*n*"; "id_observation" will be a new variable identifying the level-1 observations.