Online Appendices A-G

Online Appendices: Table of Contents

Glossary of Notation

Operators:

$E_j\left[.\right]$ is the expectation over randomly sampled clusters.

$E_{i|j}\left[.\right]$ is the expectation over randomly sampled cases within cluster *j*.

$A\left[.\right]$ is either the unweighted or weighted average (defined in appendices).

Variables:

Here, we use the *focal* descriptor to denote the coding variable that is the focus of the derivation. We use *non-focal* to describe all other coding variables that may be present in the model but are not responsible for isolating group *f*.

d_{fij} is a focal dummy code and $\bar{d}_{f.j}$ is its cluster mean.

d_{hij} is a generic non-focal dummy code and $\bar{d}_{h.j}$ is its cluster mean.

c_{fij} is a focal contrast code and $\bar{c}_{f.j}$ is its cluster mean.

c_{hij} is a generic non-focal contrast code and $\bar{c}_{h.j}$ is its cluster mean.

e_{fij} is a focal effect code (unweighted or weighted) and $\bar{e}_{f.j}$ is its cluster mean.

e_{hij} is a generic non-focal effect code (unweighted or weighted) and $\bar{e}_{h.j}$ is its cluster mean.

Error terms:

u_{0j} is the level-2 error term, where $u_{0j} \overset{i.i.d.}{\sim} N(0, \tau_{00})$

r_{ij} is the level-1 error term, where $r_{ij} \overset{i.i.d.}{\sim} N(0, \sigma_e^2)$

Online Appendix A: Dummy Codes

$$y_{ij} = \beta_{0j} + \sum_{h=1}^{k-1} \beta_{hj} \left(d_{hij} - \bar{d}_{h.j} \right) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} + u_{0j}$$

$$\beta_{hj} = \gamma_{h0}$$

$$y_{ij} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \left(d_{hij} - \bar{d}_{h.j} \right) + u_{0j} + r_{ij}$$

With dummy codes our focus is on group mean differences for each group $h = 1 \ldots (k - 1)$ relative to reference group k, so we need expected values of group means for group k and a generic focal group f when ICC of all dummy codes = 0 and when ICC of all dummy codes = 1.

First, when ICC of all dummy codes = 1, start with the focal group (group f):

$$\begin{aligned} E\left(y_{ij}\right)\Big|_{g=f,ICC=1} &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\left(d_{hij} - \bar{d}_{h.j} \right)}_{0} \right) \right) + \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j} + r_{ij} \right) \right)}_{0} \\ &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} \right) \right) \\ &= \gamma_{00} + \sum_{\substack{h=1 \\ h \neq f}}^{k-1} \gamma_{0h} \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{h.j} \right) \right)}_{0} + \gamma_{0f} \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{f.j} \right) \right)}_{1} \\ &= \gamma_{00} + \gamma_{0f} \end{aligned}$$

Then move on to the reference group (group k):

$$\begin{aligned} E\left(y_{ij}\right)\Big|_{g=k,ICC=1} &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\left(d_{hij} - \bar{d}_{h.j} \right)}_{0} \right) \right) + \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j} + r_{ij} \right) \right)}_{0} \\ &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{d}_{h.j} \right) \right) \\ &= \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{h.j} \right) \right)}_{0} \\ &= \gamma_{00} \end{aligned}$$

Therefore:
$\gamma_{00} + \gamma_{0f}$ is the model-implied mean of generic focal group f.
γ_{00} is the model-implied mean of reference group k.

γ_{0f} is the model-implied mean difference between groups *k* and *f*.
This derivation did not require committing to balanced clusters or equal group sizes.

Second, when ICC of all dummy codes = 0, start with the focal group (group *f*):

$$\begin{aligned}
E\left(y_{ij}\right)\Big|_{g=f,ICC=0} &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\underbrace{\gamma_{0h}}_{0}\bar{d}_{h.j}+\sum_{h=1}^{k-1}\gamma_{h0}\left(d_{hij}-\bar{d}_{h.j}\right)\right)\right)+\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j}+r_{ij}\right)\right)}_{0} \\
&= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\left(d_{hij}-\bar{d}_{h.j}\right)\right)\right) \\
&= \gamma_{00}+\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(d_{hij}-\bar{d}_{h.j}\right)\right)+\gamma_{f0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(d_{fij}-\bar{d}_{f.j}\right)\right) \\
&= \gamma_{00}+\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(d_{hij}\right)\right)}_{0}-\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{h.j}\right)\right)}_{p_h}+\gamma_{f0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(d_{fij}\right)\right)}_{1}-\gamma_{f0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{f.j}\right)\right)}_{p_f} \\
&= \gamma_{00}-\sum_{h=1}^{k-1}\gamma_{h0}p_h+\gamma_{f0}
\end{aligned}$$

Then move on to the reference group (group *k*):

$$\begin{aligned}
E\left(y_{ij}\right)\Big|_{g=k,ICC=0} &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\underbrace{\gamma_{0h}}_{0}\bar{d}_{h.j}+\sum_{h=1}^{k-1}\gamma_{h0}\left(d_{hij}-\bar{d}_{h.j}\right)\right)\right)+\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j}+r_{ij}\right)\right)}_{0} \\
&= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\left(d_{hij}-\bar{d}_{h.j}\right)\right)\right) \\
&= \gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(d_{hij}-\bar{d}_{h.j}\right)\right) \\
&= \gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(d_{hij}\right)\right)}_{0}-\sum_{h=1}^{k-1}\gamma_{h0}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{d}_{h.j}\right)\right)}_{p_h} \\
&= \gamma_{00}-\sum_{h=1}^{k-1}\gamma_{h0}p_h
\end{aligned}$$

Therefore:

$\gamma_{00}-\sum_{h=1}^{k-1}\gamma_{h0}p_h+\gamma_{f0}$ is the model-implied expected within-group mean of generic focal group *f*.

$\gamma_{00}-\sum_{h=1}^{k-1}\gamma_{h0}p_h$ is the model-implied expected within-group mean of reference group *k*.

γ_{f0} is the model-implied within-group mean difference between groups *k* and *f*.
This derivation did not require committing to balanced clusters or equal group sizes.

Online Appendix B: Contrast Codes

$$y_{ij} = \beta_{0j} + \sum_{h=1}^{k-1} \beta_{hj} \left(c_{hij} - \overline{c}_{h.j} \right) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \overline{c}_{h.j} + u_{0j}$$

$$\beta_{hj} = \gamma_{h0}$$

$$y_{ij} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \overline{c}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \left(c_{hij} - \overline{c}_{h.j} \right) + u_{0j} + r_{ij}$$

With contrast codes our focus is on differences between the unweighted mean of the focal group(s) and the unweighted mean of the base group(s). We need to derive these quantities when ICC of all contrast codes = 0 and when ICC of all contrast codes = 1.

In the following, let f_0 denote a generic set of groups not involved in a comparison. Let f_1 and f_2 denote the focal group(s) and reference group(s), respectively.

There are three types of contrast codes to consider. First is the contrast code of interest, which compares the group(s) in f_1 to the group(s) in f_2; this code is denoted c_f. Second are the other contrast codes that take on non-zero values for the group(s) in f_1 and/or f_2; these are denoted as c_h where $c_h \neq 0$. Third are the other contrast codes that take on zero values for the group(s) in f_1 and/or f_2; these are denoted as c_h where $c_h = 0$, and can be disregarded.

First, when ICC of all contrast codes = 1, start with all the groups involved in the contrast.

$$E\left(y_{ij}\right)\Big|_{\substack{g \in f_1, f_2 \\ ICC=1}} = \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \overline{c}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\left(c_{hij} - \overline{c}_{h.j} \right)}_{0} \right) \right) + \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j} + r_{ij} \right) \right)}_{0}$$

$$= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \gamma_{0f} \overline{c}_{f.j} + \sum_{\substack{h=1 \\ h \neq f \\ c_h \neq 0}}^{k-1} \gamma_{0h} \overline{c}_{h.j} + \underbrace{\sum_{\substack{h=1 \\ h \neq f \\ c_h = 0}}^{k-1} \gamma_{0h} \overline{c}_{h.j}}_{0} \right) \right)$$

$$= \gamma_{00} + \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{0f} \overline{c}_{f.j} \right) \right) + \underset{j}{E}\left(\underset{i|j}{E}\left(\sum_{\substack{h=1 \\ h \neq f \\ c_h \neq 0}}^{k-1} \gamma_{0h} \overline{c}_{h.j} \right) \right)$$

$$= \gamma_{00} + \gamma_{0f} \underset{j}{E}\left(\underset{i|j}{E}\left(\overline{c}_{f.j} \right) \right) + \sum_{\substack{h=1 \\ h \neq f \\ c_h \neq 0}}^{k-1} \gamma_{0h} \underset{j}{E}\left(\underset{i|j}{E}\left(\overline{c}_{h.j} \right) \right)$$

Because ICC of all contrast codes = 1, the cluster mean for any contrast code will simply be equal to the value of the contrast code. Thus, we can treat all instances of $\bar{c}_{f.j}$ as c_{fij}. Next, we plug in the value of the contrast code for the group(s) in f_1 and f_2.

$$\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)\Big|_{g\in f_1} = \frac{G_{f_2}}{G_{f_1}+G_{f_2}}$$

$$\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)\Big|_{g\in f_2} = \frac{-G_{f_1}}{G_{f_1}+G_{f_2}}$$

Where G_{f_1}, G_{f_2} are the total number of groups involved in f_1 and f_2, respectively.

$$\begin{aligned} E\left(y_{ij}\right)\Big|_{\substack{g\in f_1\\ ICC=1}} &= \gamma_{00} + \gamma_{0f}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{f.j}\right)\right)\Big|_{g\in f_1} + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right) \\ &= \gamma_{00} + \gamma_{0f}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right) + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right) \end{aligned}$$

$$\begin{aligned} E\left(y_{ij}\right)\Big|_{\substack{g\in f_2\\ ICC=1}} &= \gamma_{00} + \gamma_{0f}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{f.j}\right)\right)\Big|_{g\in f_2} + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right) \\ &= \gamma_{00} + \gamma_{0f}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right) + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right) \end{aligned}$$

Therefore:
The $f_1 - f_2$ mean difference is

$$\gamma_{00} + \gamma_{0f}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right) + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right) - \left(\gamma_{00} + \gamma_{0f}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right) + \sum_{\substack{h=1\\ h\neq f\\ c_h\neq 0}}^{k-1}\gamma_{0h}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right)\right)$$

$$= \gamma_{0f}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right) - \gamma_{0f}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right)$$

$$= \gamma_{0f}$$

γ_{0f} is the model-implied mean difference between focal groups(s) in f_1 and reference group(s) in f_2. This derivation did not require committing to balanced clusters or equal group sizes, and the contrast can compare any number of groups.

Second, when ICC of all contrast codes = 0, start with all the groups involved in the contrast.

$$E\left(y_{ij}\right)\Big|_{\substack{g\in f_1,f_2\\ ICC=0}}=\underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\underbrace{\gamma_{0h}}_{0}\bar{c}_{h.j}+\sum_{h=1}^{k-1}\gamma_{h0}\left(c_{hij}-\bar{c}_{h.j}\right)\right)\right)+\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j}+r_{ij}\right)\right)}_{0}$$

$$=\underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}+\gamma_{f0}\left(c_{fij}-\bar{c}_{f.j}\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\left(c_{hij}-\bar{c}_{h.j}\right)+\underbrace{\sum_{\substack{h=1\\h\neq f\\c_h=0}}^{k-1}\gamma_{h0}\left(c_{hij}-\bar{c}_{h.j}\right)}_{0}\right)\right)$$

$$=\gamma_{00}+\underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{f0}\left(c_{fij}-\bar{c}_{f.j}\right)\right)\right)+\underset{j}{E}\left(\underset{i|j}{E}\left(\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\left(c_{hij}-\bar{c}_{h.j}\right)\right)\right)$$

$$=\gamma_{00}+\gamma_{f0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)-\gamma_{f0}\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{f.j}\right)\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right)$$

$$=\gamma_{00}+\gamma_{f0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right)$$

From earlier we know:

$$\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)\Big|_{g\in f_1}=\frac{G_{f_2}}{G_{f_1}+G_{f_2}}$$

$$\underset{j}{E}\left(\underset{i|j}{E}\left(c_{fij}\right)\right)\Big|_{g\in f_2}=\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}$$

$$E\left(y_{ij}\right)\Big|_{\substack{g\in f_1\\ ICC=0}}=\gamma_{00}+\gamma_{f0}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right)$$

$$E\left(y_{ij}\right)\Big|_{\substack{g\in f_2\\ ICC=0}}=\gamma_{00}+\gamma_{f0}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\c_h\neq 0}}^{k-1}\gamma_{h0}\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{c}_{h.j}\right)\right)$$

Therefore:

The $f_1 - f_2$ mean difference is

$$\gamma_{00}+\gamma_{f0}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\c_h\neq 0}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\overline{c}_{h.j}\right)\right)$$

$$-\left(\gamma_{00}+\gamma_{f0}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right)+\sum_{\substack{h=1\\h\neq f\\c_h\neq 0}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(c_{hij}\right)\right)-\sum_{\substack{h=1\\c_h\neq 0}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j}{E}\left(\overline{c}_{h.j}\right)\right)\right)$$

$$=\gamma_{f0}\left(\frac{G_{f_2}}{G_{f_1}+G_{f_2}}\right)-\gamma_{f0}\left(\frac{-G_{f_1}}{G_{f_1}+G_{f_2}}\right)$$

$$=\gamma_{f0}$$

γ_{f0} is the model-implied mean difference between focal groups(s) in f_1 and reference group(s) in f_2. This derivation did not require committing to balanced clusters or equal group sizes, and the contrast can compare any number of groups.

Online Appendix C: Unweighted Effect Codes

$$y_{ij} = \beta_{0j} + \sum_{h=1}^{k-1} \beta_{hj} \left(e_{hij} - \bar{e}_{h.j} \right) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} + u_{0j}$$

$$\beta_{hj} = \gamma_{h0}$$

$$y_{ij} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) + u_{0j} + r_{ij}$$

With unweighted effect codes our focus is on mean differences for each group $h = 1 \ldots (k-1)$ relative to the unweighted mean of all k group means, so we need expected values of group means for group $h = 1 \ldots (k-1)$ and for the unweighted mean of all group means. Note that the mean of group k (the reference group) is never compared to anything in effect coding, so we do not need to consider the case when $g = k$.

First, when ICC of all effect codes = 1, start with the focal (group f):

$$E\left(y_{ij}\right)\Big|_{g=f,ICC=1} = \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\left(e_{hij} - \bar{e}_{h.j} \right)}_{0} \right)\right) + \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j} + r_{ij} \right)\right)}_{0}$$

$$= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} \right)\right)$$

$$= \gamma_{00} + \sum_{\substack{h=1 \\ h\neq f}}^{k-1} \gamma_{0h} \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{e}_{h.j} \right)\right)}_{0} + \gamma_{0f} \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\bar{e}_{f.j} \right)\right)}_{1}$$

$$= \gamma_{00} + \gamma_{0f}$$

Then, move on to the unweighted mean of the group means (here, $A[.]$ is the unweighted mean operator):

$$\underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(y_{ij} \right)\right)\Big|_{ICC=1} \right) = \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\left(e_{hij} - \bar{e}_{h.j} \right)}_{0} \right)\right)\right) + \underset{g}{A}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,g}{E}\left(u_{0j} + r_{ij} \right)\right)}_{0} \right)$$

$$= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} \right)\right)\right)$$

$$= \frac{1}{k} \sum_{g=1}^{k} \underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} \right)\right)$$

Next, remove γ_{00} from the summations

$$= \frac{1}{k} \sum_{g=1}^{k} \left(\gamma_{00} + \underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} \right)\right)\right)$$

$$= \gamma_{00} \underbrace{\sum_{g=1}^{k} \frac{1}{k}}_{1} + \frac{1}{k} \sum_{g=1}^{k} \left(\underset{j}{E} \left(\underset{i|j,g}{E} \left(\sum_{h=1}^{k-1} \gamma_{0h} \bar{e}_{h.j} \right) \right) \right)$$

$$= \gamma_{00} + \frac{1}{k} \sum_{g=1}^{k} \sum_{h=1}^{k-1} \gamma_{0h} \underset{j}{E} \left(\underset{i|j,g}{E} \left(\bar{e}_{h.j} \right) \right)$$

Next, separate group k (reference group) from the others

$$= \gamma_{00} + \frac{1}{k} \sum_{g=1}^{k-1} \sum_{h=1}^{k-1} \gamma_{0h} \underset{j}{E} \left(\underset{i|j,g}{E} \left(\bar{e}_{h.j} \right) \right) + \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h} \underbrace{\underset{j}{E} \left(\underset{i|j,k}{E} \left(\bar{e}_{h.j} \right) \right)}_{-1}$$

Next, switch the nested summations

$$= \gamma_{00} + \frac{1}{k} \sum_{h=1}^{k-1} \sum_{g=1}^{k-1} \gamma_{0h} \underset{j}{E} \left(\underset{i|j,g}{E} \left(\bar{e}_{h.j} \right) \right) - \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h}$$

$$= \gamma_{00} + \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h} \underbrace{\sum_{g=1}^{k-1} \underset{j}{E} \left(\underset{i|j,g}{E} \left(\bar{e}_{h.j} \right) \right)}_{1} - \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h}$$

$$= \gamma_{00} + \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h} - \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{0h}$$

$$= \gamma_{00}$$

Therefore:
$\gamma_{00} + \gamma_{0f}$ is the model-implied mean of generic focal group *f*.
γ_{00} is the model-implied unweighted mean of the group means.
γ_{0f} is the model-implied difference between the mean of group *f* and the unweighted mean of the group means.
This derivation did not require committing to balanced clusters or equal group sizes.

Second, when ICC of all effect codes = 0, start with the focal group (group *f*):

$$E\left(y_{ij}\right)\Big|_{g=f,ICC=0} = \underset{j}{E} \left(\underset{i|j}{E} \left(\gamma_{00} + \sum_{h=1}^{k-1} \underbrace{\gamma_{0h}}_{0} \bar{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right) + \underbrace{\underset{j}{E} \left(\underset{i|j}{E} \left(u_{0j} + r_{ij} \right) \right)}_{0}$$

$$= \underset{j}{E} \left(\underset{i|j}{E} \left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right)$$

$$= \gamma_{00} + \sum_{\substack{h=1 \\ h \neq f}}^{k-1} \gamma_{h0} \underset{j}{E} \left(\underset{i|j}{E} \left(e_{hij} - \bar{e}_{h.j} \right) \right) + \gamma_{f0} \underset{j}{E} \left(\underset{i|j}{E} \left(e_{fij} - \bar{e}_{f.j} \right) \right)$$

$$= \gamma_{00} + \sum_{\substack{h=1 \\ h \neq f}}^{k-1} \gamma_{h0} \left(\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(e_{hij} \right) \right)}_{0} - \underset{j}{E}\left(\underbrace{\underset{i|j}{E}\left(\bar{e}_{h.j} \right)}_{p_h - p_k} \right) \right) + \gamma_{f0} \left(\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(e_{fij} \right) \right)}_{1} - \underset{j}{E}\left(\underbrace{\underset{i|j}{E}\left(\bar{e}_{h.j} \right)}_{p_f - p_k} \right) \right)$$

$$= \gamma_{00} + \gamma_{f0} - \sum_{h=1}^{k-1} \gamma_{h0} \left(p_h - p_k \right)$$

Then move on to the unweighted mean of the group means (here, $A[.]$ is the unweighted mean operator):

$$\underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(y_{ij} \right) \right) \Big|_{ICC=0} \right) = \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \underbrace{\gamma_{0h}}_{0} \bar{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right) \right) + \underset{g}{A}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,g}{E}\left(u_{0j} + r_{ij} \right) \right)}_{0} \right)$$

$$= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right) \right)$$

$$= \frac{1}{k} \sum_{g=1}^{k} \underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right)$$

Next, remove γ_{00} from the summations

$$= \frac{1}{k} \sum_{g=1}^{k} \left(\gamma_{00} + \underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right) \right)$$

$$= \gamma_{00} \underbrace{\sum_{g=1}^{k} \frac{1}{k}}_{1} + \frac{1}{k} \sum_{g=1}^{k} \left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1} \gamma_{h0} \left(e_{hij} - \bar{e}_{h.j} \right) \right) \right) \right)$$

$$= \gamma_{00} + \frac{1}{k} \sum_{g=1}^{k} \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij} - \bar{e}_{h.j} \right) \right)$$

Next, separate group k (reference group) from the others

$$= \gamma_{00} + \frac{1}{k} \sum_{g=1}^{k-1} \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij} - \bar{e}_{h.j} \right) \right) + \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underset{i|j,k}{E}\left(e_{hij} - \bar{e}_{h.j} \right) \right)$$

Next, switch the nested summations and distribute expectations

$$= \gamma_{00} + \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\sum_{g=1}^{k-1} \underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij} \right) \right)}_{1} - \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{h0} \sum_{g=1}^{k-1} \underset{j}{E}\left(\underbrace{\underset{i|j,g}{E}\left(\bar{e}_{h.j} \right)}_{p_h - p_k} \right)$$

$$+ \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underbrace{\underset{i|j,k}{E}\left(e_{hij} \right)}_{-1} \right) - \frac{1}{k} \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underbrace{\underset{i|j,k}{E}\left(\bar{e}_{h.j} \right)}_{p_h - p_k} \right)$$

$$= \gamma_{00} + \frac{1}{k}\sum_{h=1}^{k-1}\gamma_{h0} - \frac{1}{k}\sum_{h=1}^{k-1}\gamma_{h0}\sum_{g=1}^{k}\underset{j}{E}\left(p_h - p_k\right) - \frac{1}{k}\sum_{h=1}^{k-1}\gamma_{h0}$$

$\underset{j}{E}$ disappears because ICC_X is 0

$$= \gamma_{00} - \frac{1}{k}\sum_{h=1}^{k-1}\gamma_{h0}\sum_{g=1}^{k}\left(p_h - p_k\right)$$

$$= \gamma_{00} - \frac{1}{k}\sum_{h=1}^{k-1}\gamma_{h0}k\left(p_h - p_k\right)$$

$$= \gamma_{00} - \sum_{h=1}^{k-1}\gamma_{h0}\left(p_h - p_k\right)$$

Therefore:

$\gamma_{00} + \gamma_{f0} - \sum_{h=1}^{k-1}\gamma_{h0}\left(p_h - p_k\right)$ is the model-implied mean of generic focal group *f*.

$\gamma_{00} - \sum_{h=1}^{k-1}\gamma_{h0}\left(p_h - p_k\right)$ is the model-implied unweighted mean of the group means.

γ_{f0} is the model-implied difference between the mean of group *f* and the unweighted mean of the group means.

This derivation did not require committing to balanced clusters or equal group sizes.

Online Appendix D: Weighted Effect Codes

$$y_{ij} = \beta_{0j} + \sum_{h=1}^{k-1} \beta_{hj}\left(e_{hij} - \overline{e}_{h.j}\right) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h}\overline{e}_{h.j} + u_{0j}$$

$$\beta_{hj} = \gamma_{h0}$$

$$y_{ij} = \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{0h}\overline{e}_{h.j} + \sum_{h=1}^{k-1} \gamma_{h0}\left(e_{hij} - \overline{e}_{h.j}\right) + u_{0j} + r_{ij}$$

With weighted effect codes our focus is on mean differences for each group $h = 1 \ldots (k-1)$ relative to the *grand mean of all observations*, which is also the *weighted average of the group means*. So, we need expected values of group means for groups $h = 1 \ldots (k-1)$ and for the grand mean of all observations. Note that the mean of group k (reference group) is never compared to anything in effect coding, so we do not need to consider the case when $g = k$.

First, when ICC of all effect codes = 1, start with the focal group (group f):

$$\begin{aligned}
E\left(y_{ij}\right)\Big|_{ICC=1,g=f} &= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1}\gamma_{0h}\overline{e}_{h.j} + \sum_{h=1}^{k-1}\gamma_{h0}\underbrace{\left(e_{hij} - \overline{e}_{h.j}\right)}_{0}\right)\right) + \underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(u_{0j} + r_{ij}\right)\right)}_{0} \\
&= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1}\gamma_{0h}\overline{e}_{h.j}\right)\right) \\
&= \underset{j}{E}\left(\underset{i|j}{E}\left(\gamma_{00}\right)\right) + \sum_{\substack{h=1 \\ h\neq f}}^{k-1}\gamma_{0h}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\overline{e}_{h.j}\right)\right)}_{0} + \gamma_{0f}\underbrace{\underset{j}{E}\left(\underset{i|j}{E}\left(\overline{e}_{f.j}\right)\right)}_{1} \\
&= \gamma_{00} + \gamma_{0f}
\end{aligned}$$

Then, move on to the grand mean of all observations, which is also the weighted mean of the group means (here, $A[.]$ is the weighted mean operator):

$$\begin{aligned}
\underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(y_{ij}\right)\right)\Big|_{ICC=1}\right) &= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1}\gamma_{0h}\overline{e}_{h.j} + \sum_{h=1}^{k-1}\gamma_{h0}\underbrace{\left(e_{hij} - \overline{e}_{h.j}\right)}_{0}\right)\right)\right) + \underset{g}{A}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,g}{E}\left(u_{0j} + r_{ij}\right)\right)}_{0}\right) \\
&= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1}\gamma_{0h}\overline{e}_{h.j}\right)\right)\right) \\
&= \sum_{g=1}^{k} p_g \underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00} + \sum_{h=1}^{k-1}\gamma_{0h}\overline{e}_{h.j}\right)\right)
\end{aligned}$$

Next, remove γ_{00} from the summations

$$=\sum_{g=1}^{k} p_g\left(\gamma_{00}+\underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1}\gamma_{0h}\bar{e}_{h.j}\right)\right)\right)$$

$$=\gamma_{00}\underbrace{\sum_{g=1}^{k}p_g}_{1}+\sum_{g=1}^{k}p_g\underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1}\gamma_{0h}\bar{e}_{h.j}\right)\right)$$

$$=\gamma_{00}+\sum_{g=1}^{k}p_g\sum_{h=1}^{k-1}\gamma_{0h}\underset{j}{E}\left(\underset{i|j,g}{E}\left(\bar{e}_{h.j}\right)\right)$$

Next, separate group k (reference group) from the others

$$=\gamma_{00}+\sum_{g=1}^{k-1}p_g\sum_{h=1}^{k-1}\gamma_{0h}\underset{j}{E}\left(\underset{i|j,g}{E}\left(\bar{e}_{h.j}\right)\right)+p_k\sum_{h=1}^{k-1}\gamma_{0h}\underbrace{\underset{j}{E}\left(\underset{i|j,k}{E}\left(\bar{e}_{h.j}\right)\right)}_{\frac{-n_h}{n_k}}$$

$$=\gamma_{00}+\sum_{g=1}^{k-1}p_g\sum_{h=1}^{k-1}\gamma_{0h}\underset{j}{E}\left(\underset{i|j,g}{E}\left(\bar{e}_{h.j}\right)\right)+p_k\sum_{h=1}^{k-1}\gamma_{0h}\underbrace{\underset{j}{E}\left(\underset{i|j,k}{E}\left(\bar{e}_{h.j}\right)\right)}_{\frac{-p_h}{p_k}}$$

Next, switch the nested summations

$$=\gamma_{00}+\sum_{h=1}^{k-1}\sum_{g=1}^{k-1}\gamma_{0h}\underset{j}{E}\left(\underset{i|j,g}{E}\left(\bar{e}_{h.j}\right)\right)-p_k\sum_{h=1}^{k-1}\frac{p_h}{p_k}\gamma_{0h}$$

$$=\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{0h}\underbrace{\sum_{g=1}^{k-1}p_g\underset{j}{E}\left(\underset{i|j,g}{E}\left(\bar{e}_{h.j}\right)\right)}_{p_h}-\sum_{h=1}^{k-1}p_h\gamma_{0h}$$

$$=\gamma_{00}+\sum_{h=1}^{k-1}p_h\gamma_{0h}-\sum_{h=1}^{k-1}p_h\gamma_{0h}$$

$$=\gamma_{00}$$

Therefore:
$\gamma_{00}+\gamma_{0f}$ is the model-implied mean of generic focal group *f*.
γ_{00} is the model-implied grand mean (weighted mean of the group means).
γ_{0f} is the model-implied difference between the mean of group *f* and the grand mean (weighted mean of the group means).
This derivation did not require committing to balanced clusters or equal group sizes.

Second, when ICC of all effect codes = 0, start with the focal group (group *f*):

$$\left.\underset{j}{E}\left(\underset{i|j,f}{E}\left(y_{ij}\right)\right)\right|_{ICC=0}=\underset{j}{E}\left(\underset{i|j,f}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\underbrace{\gamma_{0h}}_{0}\bar{e}_{h.j}+\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)+\underbrace{\underset{j}{E}\left(\underset{i|j,f}{E}\left(u_{0j}+r_{ij}\right)\right)}_{0}$$

$$
\begin{aligned}
&= \underset{j}{E}\left(\underset{i|j,f}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)\\
&= \underset{j}{E}\left(\underset{i|j,f}{E}\left(\gamma_{00}\right)\right)+\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j,f}{E}\left(e_{hij}-\bar{e}_{h.j}\right)\right)+\gamma_{f0}\,\underset{j}{E}\left(\underset{i|j,f}{E}\left(e_{fij}-\bar{e}_{f.j}\right)\right)\\
&= \gamma_{00}+\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,f}{E}\left(e_{hij}\right)\right)}_{0}-\underset{j}{E}\left(\underbrace{\underset{i|j,f}{E}\left(\bar{e}_{h.j}\right)}_{p_h-\left(\frac{p_h}{p_k}\right)}\right)\right)+\gamma_{f0}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,f}{E}\left(e_{fij}\right)\right)}_{1}-\underset{j}{E}\left(\underbrace{\underset{i|j,f}{E}\left(\bar{e}_{h.j}\right)}_{p_f-\left(\frac{p_h}{p_k}\right)}\right)\right)\\
&= \gamma_{00}-\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(p_h-\frac{p_h}{p_k}\right)+\gamma_{f0}\left(1-\underset{j}{E}\left(p_f-\frac{p_h}{p_k}\right)\right)\\
&= \gamma_{00}-\sum_{\substack{h=1\\h\neq f}}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(p_h-\frac{p_h}{p_k}\right)+\gamma_{f0}-\gamma_{f0}\,\underset{j}{E}\left(p_f-\frac{p_h}{p_k}\right)\\
&= \gamma_{00}+\gamma_{f0}-\sum_{h=1}^{k-1}\gamma_{h0}\left(p_h-\frac{p_h}{p_k}\right)
\end{aligned}
$$

Then move on to the unweighted mean of the group means (here, $A[.]$ is the weighted mean operator):

$$
\begin{aligned}
\underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(y_{ij}\right)\right)\bigg|_{ICC=0}\right) &= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\underbrace{\gamma_{0h}}_{0}\bar{e}_{h.j}+\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)\right)+\underset{g}{A}\left(\underbrace{\underset{j}{E}\left(\underset{i|j,g}{E}\left(u_{0j}+r_{ij}\right)\right)}_{0}\right)\\
&= \underset{g}{A}\left(\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)\right)\\
&= \sum_{g=1}^{k}p_g\,\underset{j}{E}\left(\underset{i|j,g}{E}\left(\gamma_{00}+\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)
\end{aligned}
$$

Next, remove γ_{00} from the summations

$$
\begin{aligned}
&= \sum_{g=1}^{k}p_g\left(\gamma_{00}+\underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)\right)\\
&= \gamma_{00}\underbrace{\sum_{g=1}^{k}p_g}_{1}+\sum_{g=1}^{k}p_g\,\underset{j}{E}\left(\underset{i|j,g}{E}\left(\sum_{h=1}^{k-1}\gamma_{h0}\left(e_{hij}-\bar{e}_{h.j}\right)\right)\right)\\
&= \gamma_{00}+\sum_{g=1}^{k}p_g\sum_{h=1}^{k-1}\gamma_{h0}\,\underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij}-\bar{e}_{h.j}\right)\right)
\end{aligned}
$$

Next, separate group k (reference group) from the others

$$= \gamma_{00} + \sum_{g=1}^{k-1} p_g \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij} - \bar{e}_{h.j} \right) \right) + p_k \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underset{i|j,k}{E}\left(e_{hij} - \bar{e}_{h.j} \right) \right)$$

Next, switch the nested summations and distribute expectations

$$= \gamma_{00} + \sum_{h=1}^{k-1} \gamma_{h0} \underbrace{\sum_{g=1}^{k-1} p_g \underset{j}{E}\left(\underset{i|j,g}{E}\left(e_{hij} \right) \right)}_{p_h} - \sum_{h=1}^{k-1} \gamma_{h0} \sum_{g=1}^{k-1} p_g \underset{j}{E}\left(\underbrace{\underset{i|j,g}{E}\left(\bar{e}_{h.j} \right)}_{p_h - \left(\frac{p_h}{p_k} \right)} \right)$$

$$+ p_k \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underbrace{\underset{i|j,k}{E}\left(e_{hij} \right)}_{\frac{-p_h}{p_k}} \right) - p_k \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\underbrace{\underset{i|j,k}{E}\left(\bar{e}_{h.j} \right)}_{p_h - \left(\frac{p_h}{p_k} \right)} \right)$$

$$= \gamma_{00} + \sum_{h=1}^{k-1} p_h \gamma_{h0} - \sum_{h=1}^{k-1} \gamma_{h0} \sum_{g=1}^{k} p_g \underset{j}{E}\left(p_h - \frac{p_h}{p_k} \right) - p_k \sum_{h=1}^{k-1} \gamma_{h0} \underset{j}{E}\left(\frac{p_h}{p_k} \right)$$

$\underset{j}{E}$ disappears because ICC is 0

$$= \gamma_{00} + \sum_{h=1}^{k-1} p_h \gamma_{h0} - \sum_{h=1}^{k-1} \gamma_{h0} \sum_{g=1}^{k} p_g \underset{j}{E}\left(p_h - \frac{p_h}{p_k} \right) - \sum_{h=1}^{k-1} p_h \gamma_{h0}$$

$$= \gamma_{00} - \sum_{h=1}^{k-1} \gamma_{h0} \sum_{g=1}^{k} p_g \left(p_h - \frac{p_h}{p_k} \right)$$

$$= \gamma_{00} - \sum_{h=1}^{k-1} \gamma_{h0} \left(p_h - \frac{p_h}{p_k} \right) \underbrace{\sum_{g=1}^{k} p_g}_{1}$$

$$= \gamma_{00} - \sum_{h=1}^{k-1} \gamma_{h0} \left(p_h - \frac{p_h}{p_k} \right)$$

Therefore:

$\gamma_{00} + \gamma_{f0} - \sum_{h=1}^{k-1} \gamma_{h0} \left(p_h - \frac{p_h}{p_k} \right)$ is the model-implied mean of generic focal group *f*.

$\gamma_{00} - \sum_{h=1}^{k-1} \gamma_{h0} \left(p_h - \frac{p_h}{p_k} \right)$ is the model-implied grand mean (weighted mean of the group means).

γ_{f0} is the model-implied difference between the mean of group *f* and the grand mean (weighted mean of the group means).

This derivation did not require committing to balanced clusters or equal group sizes.

Online Appendix E: Demonstrating the Relationship Between ICC$_X$ and Group Mean Differences

Here, we specify hypothetical population-level within- and between-cluster effects of a three-category predictor. These effects are held constant, and we choose five values of ICC$_X$ that range between 0 and 1. Under each of the five conditions, we derive group mean differences on *y*. This example was conducted with dummy codes, though our conclusions apply to any coding scheme.

To alter ICC$_X$ while maintaining simplicity in our demonstration, we assume that each dataset is comprised of 300 clusters each of size 100. We split the dataset into thirds (100 clusters in each third) and alter the composition of those 100 clusters such that cluster compositions deviate from population compositions by a specified margin.

Suppose we have a three-group categorical predictor, expressed with two dummy codes, and that in the population, 30% of individuals belong to the reference group, 30% of individuals belong to group 1, and 40% of individuals belong to group 2. Also suppose that, at the population level, the following parameters are true:

$\gamma_{00} = 20 \qquad \gamma_{10} = 2 \qquad \gamma_{20} = 6 \qquad \gamma_{01} = -1 \qquad \gamma_{02} = -3$

Condition 1: ICC$_{d1}$ = ICC$_{d2}$ = 0. Each cluster composition is equal to the population composition; all clusters have a 0.30/0.30/0.40 composition, meaning there are 30 members of group zero, 30 members of group one, and 40 members of group two, in each cluster.

$$E(\bar{d}_{1j})|_{g0} = \frac{\sum_{j=1}^{300}(n_{0j})(\bar{d}_{1j})}{N_0} = \frac{\sum_{j=1}^{300}(30)\left(\frac{30}{100}\right)}{9000} = 0.30 \qquad E(\bar{d}_{2j})|_{g0} = \frac{\sum_{j=1}^{300}(n_{0j})(\bar{d}_{2j})}{N_0} = \frac{\sum_{j=1}^{300}(30)\left(\frac{40}{100}\right)}{9000} = 0.40$$

$$E(\bar{d}_{1j})|_{g1} = \frac{\sum_{j=1}^{300}(n_{1j})(\bar{d}_{1j})}{N_1} = \frac{\sum_{j=1}^{300}(30)\left(\frac{30}{100}\right)}{9000} = 0.30 \qquad E(\bar{d}_{2j})|_{g1} = \frac{\sum_{j=1}^{300}(n_{1j})(\bar{d}_{2j})}{N_1} = \frac{\sum_{j=1}^{300}(30)\left(\frac{40}{100}\right)}{9000} = 0.40$$

$$E(\bar{d}_{1j})|_{g2} = \frac{\sum_{j=1}^{300}(n_{2j})(\bar{d}_{1j})}{N_2} = \frac{\sum_{j=1}^{300}(40)\left(\frac{30}{100}\right)}{12000} = 0.30 \qquad E(\bar{d}_{2j})|_{g2} = \frac{\sum_{j=1}^{300}(n_{2j})(\bar{d}_{2j})}{N_2} = \frac{\sum_{j=1}^{300}(40)\left(\frac{40}{100}\right)}{12000} = 0.40$$

Plugging these values in to solve for $E(y_{ij})$ for each group, we obtain:

$$E(y_{ij})|_{g0} = 20 + (-1)[0.30] + (-3)[0.40] + 2[0 - 0.30] + 6[0 - 0.40]$$
$$E(y_{ij})|_{g0} = 15.5$$

$$E(y_{ij})\,|_{g1} = 20 + (-1)[0.30] + (-3)[0.40] + 2[1 - 0.30] + 6[0 - 0.40]$$
$$E(y_{ij})\,|_{g1} = 17.5$$

$$E(y_{ij})\,|_{g2} = 20 + (-1)[0.30] + (-3)[0.40] + 2[0 - 0.30] + 6[1 - 0.40]$$
$$E(y_{ij})\,|_{g2} = 21.5$$

$$E(y_{ij})\,|_{g1} - E(y_{ij})\,|_{g0} = 17.5 - 15.5 = 2$$
$$E(y_{ij})\,|_{g2} - E(y_{ij})\,|_{g0} = 21.5 - 15.5 = 6$$

Note that group mean differences are exactly equal to the within-effects of the predictor.

Condition 2: ICC$_{d1}$ = 0.022, ICC$_{d2}$ = 0.018. Each cluster composition deviates up to +/- 0.1 from the population composition. More specifically, cluster compositions are as follows: 100 clusters have a 0.20/0.30/0.50 composition, 100 clusters have a 0.40/0.20/0.40 composition, and 100 clusters have a 0.30/0.40/0.30 composition. Even though all clusters do not have the same composition, overall *population proportions* still work out to 0.30, 0.30, and 0.40.

$$E(\bar{d}_{1j})\,|_{g0} = \frac{\sum_{j=1}^{100}(20)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{40}{100}\right)}{9000} = 0.289$$
$$E(\bar{d}_{2j})\,|_{g0} = \frac{\sum_{j=1}^{100}(20)\left(\frac{50}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{30}{100}\right)}{9000} = 0.389$$

$$E(\bar{d}_{1j})\,|_{g1} = \frac{\sum_{j=1}^{100}(30)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{40}{100}\right)}{9000} = 0.322$$
$$E(\bar{d}_{2j})\,|_{g1} = \frac{\sum_{j=1}^{100}(30)\left(\frac{50}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{30}{100}\right)}{9000} = 0.389$$

$$E(\bar{d}_{1j})\,|_{g2} = \frac{\sum_{j=1}^{100}(50)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{40}{100}\right)}{12000} = 0.292$$
$$E(\bar{d}_{2j})\,|_{g2} = \frac{\sum_{j=1}^{100}(50)\left(\frac{50}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{30}{100}\right)}{12000} = 0.417$$

Plugging these values in to solve for $E(y_{ij})$ for each group, we obtain:

$$E(y_{ij})|_{g0} = 20 + (-1)\left[0.289\right] + (-3)\left[0.389\right] + 2\left[0 - 0.289\right] + 6\left[0 - 0.389\right]$$
$$E(y_{ij})|_{g0} = 15.632$$

$$E(y_{ij})|_{g1} = 20 + (-1)\left[0.322\right] + (-3)\left[0.389\right] + 2\left[1 - 0.322\right] + 6\left[0 - 0.389\right]$$
$$E(y_{ij})|_{g1} = 17.533$$

$$E(y_{ij})|_{g2} = 20 + (-1)\left[0.292\right] + (-3)\left[0.417\right] + 2\left[0 - 0.292\right] + 6\left[1 - 0.417\right]$$
$$E(y_{ij})|_{g2} = 21.371$$

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = 17.533 - 15.632 = 1.901$$
$$E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = 21.371 - 15.632 = 5.74$$

Condition 3: ICC$_{d1}$ = 0.119, ICC$_{d2}$ = 0.103. Cluster proportions deviate up to +/- 0.20 from overall population proportions. Specific cluster compositions are as follows: 100 clusters have a 0.10/0.30/0.60 composition, 100 clusters have a 0.50/0.10/0.40 composition, and 100 clusters have a 0.30/0.50/0.20 composition.

$$E(\bar{d}_{1j})|_{g0} = \frac{\sum_{j=1}^{100}(10)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(50)\left(\frac{10}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{50}{100}\right)}{9000} = 0.256$$

$$E(\bar{d}_{2j})|_{g0} = \frac{\sum_{j=1}^{100}(10)\left(\frac{60}{100}\right) + \sum_{j=1}^{100}(50)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(30)\left(\frac{20}{100}\right)}{9000} = 0.356$$

$$E(\bar{d}_{1j})|_{g1} = \frac{\sum_{j=1}^{100}(30)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(10)\left(\frac{10}{100}\right) + \sum_{j=1}^{100}(50)\left(\frac{50}{100}\right)}{9000} = 0.389$$

$$E(\bar{d}_{2j})|_{g1} = \frac{\sum_{j=1}^{100}(30)\left(\frac{60}{100}\right) + \sum_{j=1}^{100}(10)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(50)\left(\frac{20}{100}\right)}{9000} = 0.356$$

$$E(\bar{d}_{1j})|_{g2} = \frac{\sum_{j=1}^{100}(60)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{10}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{50}{100}\right)}{12000} = 0.267$$

$$E(\bar{d}_{2j})|_{g2} = \frac{\sum_{j=1}^{100}(60)\left(\frac{60}{100}\right) + \sum_{j=1}^{100}(40)\left(\frac{40}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{20}{100}\right)}{12000} = 0.467$$

Plugging these values in to solve for $E(y_{ij})$ for each group, we obtain:

$$E(y_{ij})|_{g0} = 20 + (-1)[0.256] + (-3)[0.356] + 2[0 - 0.256] + 6[0 - 0.356]$$
$$E(y_{ij})|_{g0} = 16.028$$

$$E(y_{ij})|_{g1} = 20 + (-1)[0.389] + (-3)[0.356] + 2[1 - 0.389] + 6[0 - 0.356]$$
$$E(y_{ij})|_{g1} = 17.629$$

$$E(y_{ij})|_{g2} = 20 + (-1)[0.267] + (-3)[0.467] + 2[0 - 0.267] + 6[1 - 0.467]$$
$$E(y_{ij})|_{g2} = 20.996$$

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = 17.629 - 16.028 = 1.601$$
$$E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = 20.996 - 16.028 = 4.968$$

Condition 4: ICC$_{d1}$ = 0.408, ICC$_{d2}$ = 0.580. Cluster proportions deviate up to +/- 0.50 from overall population proportions. Specific cluster compositions are as follows: 100 clusters have a 0/0.70/0.30 composition, 100 clusters have a 0.80/0.20/0 composition, and 100 clusters have a 0.10/0/0.90 composition.

$$E(\bar{d}_{1j})|_{g0} = \frac{\sum_{j=1}^{100}(0)\left(\frac{70}{100}\right) + \sum_{j=1}^{100}(80)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(10)\left(\frac{0}{100}\right)}{9000} = 0.178$$

$$E(\bar{d}_{2j})|_{g0} = \frac{\sum_{j=1}^{100}(0)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(80)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(10)\left(\frac{90}{100}\right)}{9000} = 0.10$$

$$E(\bar{d}_{1j})|_{g1} = \frac{\sum_{j=1}^{100}(70)\left(\frac{70}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right)}{9000} = 0.589$$

$$E(\bar{d}_{2j})|_{g1} = \frac{\sum_{j=1}^{100}(70)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(20)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{90}{100}\right)}{9000} = 0.233$$

$$E(\bar{d}_{1j})|_{g2} = \frac{\sum_{j=1}^{100}(30)\left(\frac{70}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{20}{100}\right) + \sum_{j=1}^{100}(90)\left(\frac{0}{100}\right)}{9000} = 0.175$$

$$E(\bar{d}_{2j})|_{g2} = \frac{\sum_{j=1}^{100}(30)\left(\frac{30}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(90)\left(\frac{90}{100}\right)}{9000} = 0.75$$

Plugging these values in to solve for $E(y_{ij})$ for each group, we obtain:

$$E(y_{ij})|_{g0} = 20 + (-1)[0.178] + (-3)[0.1] + 2[0 - 0.178] + 6[0 - 0.1]$$
$$E(y_{ij})|_{g0} = 18.566$$

$$E(y_{ij})|_{g1} = 20 + (-1)[0.589] + (-3)[0.233] + 2[1 - 0.589] + 6[0 - 0.233]$$
$$E(y_{ij})|_{g1} = 18.136$$

$$E(y_{ij})|_{g2} = 20 + (-1)[0.175] + (-3)[0.75] + 2[0 - 0.175] + 6[1 - 0.75]$$
$$E(y_{ij})|_{g2} = 18.725$$

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = 18.136 - 18.566 = -0.43$$
$$E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = 18.725 - 18.566 = 0.159$$

Condition 5: $ICC_{d1} = ICC_{d2} = 1$. All variability is at the between-cluster level, meaning that all clusters are either composed entirely of reference group, entirely of group 1, or entirely of group 2. Specifically, there are 100 clusters with a 1.00/0/0 composition, 100 clusters with 0/1.00/0 composition, and 100 clusters with 0/0/1.00 composition.

$$E(\bar{d}_{1j})|_{g0} = \frac{\sum_{j=1}^{100}(100)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{100}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right)}{10000} = 0$$

$$E(\bar{d}_{2j})|_{g0} = \frac{\sum_{j=1}^{100}(100)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{100}{100}\right)}{10000} = 0$$

$$E(\bar{d}_{1j})|_{g1} = \frac{\sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(100)\left(\frac{100}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right)}{10000} = 1$$

$$E(\bar{d}_{2j})|_{g1} = \frac{\sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(100)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{100}{100}\right)}{10000} = 0$$

$$E(\bar{d}_{1j})|_{g2} = \frac{\sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{100}{100}\right) + \sum_{j=1}^{100}(100)\left(\frac{0}{100}\right)}{10000} = 0$$

$$E(\bar{d}_{1j})|_{g2} = \frac{\sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(0)\left(\frac{0}{100}\right) + \sum_{j=1}^{100}(100)\left(\frac{100}{100}\right)}{10000} = 1$$

Plugging these values in to solve for $E(y_{ij})$ for each group, we obtain:

$$E(y_{ij})|_{g0} = 20 + (-1)[0] + (-3)[0] + 2[0-0] + 6[0-0]$$
$$E(y_{ij})|_{g0} = 20$$

$$E(y_{ij})|_{g1} = 20 + (-1)[1] + (-3)[0] + 2[1-1] + 6[0-0]$$
$$E(y_{ij})|_{g1} = 19$$

$$E(y_{ij})|_{g2} = 20 + (-1)[0] + (-3)[1] + 2[0-0] + 6[1-1]$$
$$E(y_{ij})|_{g2} = 17$$

$$E(y_{ij})|_{g1} - E(y_{ij})|_{g0} = 19 - 20 = -1$$
$$E(y_{ij})|_{g2} - E(y_{ij})|_{g0} = 17 - 20 = -3$$

Note that group mean differences are exactly equal to the between-effects of the predictor.

This exercise demonstrates that the expected value of *y* across groups is largely dependent upon the expected values of cluster means across those groups; further, the expected values of cluster means across groups differ as a function of the ICC of each coding variable. As ICC of the coding variables increases, expected values of cluster means across groups diverge further from their population values. Said differently, as more variability is introduced into the cluster compositions, the expected value of a given cluster mean – having conditioned on a particular group – will diverge further from its corresponding population proportion. As a consequence of this, derived group mean differences move away from the true within-cluster effects and toward the between-cluster effects.

Under more commonly-encountered data conditions wherein the ICC of a given coding variable is neither 0 nor 1, expected mean differences on *y* across groups are driven partially by within-effects and partially by between-effects. To the extent that ICC of the coding variable is low, group mean differences are driven by within-effects. To the extent that ICC of the coding variable is high, group mean differences are driven by between-effects. Thus, raw group mean differences on *y* are equal to neither the within-cluster slope nor the between-cluster slope, and are not of practical utility in multilevel contexts (when $ICC_X \neq 0$ or 1). Note that in order for the group mean difference to be equal to a slope coefficient, the ICC of only *that particular contrast* must be 0 or 1, and ICC with respect to the other coding variables does not matter.

The results of these derivations are akin to those presented by Raudenbush and colleagues (2002; 1995) regarding single-level OLS regression models fit to multilevel data. Here, the single-level regression coefficient – denoted β_t – corresponding to a predictor with a true within-cluster effect β_W and a true between-cluster effect β_B, will be a mix of these two true effects that is pulled toward one or the other as a function of ICC$_X$:

$$\hat{\beta}_t = ICC_X\hat{\beta}_B + (1 - ICC_X)\hat{\beta}_W \tag{D.1}$$

Our results suggest that a similar pattern emerges in categorical predictor settings, such that group mean differences, as estimated in a naïve single-level setting, will vary as a function of ICC of the relevant coding variable.

The overall group mean differences derived here, akin to β_t, are different from the conflated slope estimates, $\hat{\gamma}_{10}$ and $\hat{\gamma}_{20}$, observed in the UN Model. Whereas each group mean difference β_t superficially resembles a regression coefficient that one would obtain from a single-level OLS regression model, and varies as a function of ICC$_X$, the UN multilevel model yields conflated slope estimates $\hat{\gamma}_{10}$ and $\hat{\gamma}_{20}$ that vary as a function of ICC$_X$ *as well as* other extraneous characteristics of the data, including cluster size, ICC$_Y$, and correlations among the dummy codes. With this distinction in mind, we reiterate that conflated slope estimates hold no practical or substantive meaning.

Online Appendix F: *R* Code for Empirical Example

```
library(dplyr)
library(lme4)
library(lmerTest)
data <- read.csv("xwdata_9var.csv", header = T)
data$PID <- as.factor(data$PID)

############################################################################
# VRQ: A verbal reasoning score from tests pupils took when they entered secondary school
# ATTAIN: Attainment score of pupils at age 16
# PID: Primary school identifying code
# SEXF: Pupil’s gender (0 = boy and 1 = girl)
# SC: Pupil’s social class scale (continuous score from low to high social class)
# SID: Secondary school identifying code
# FED: Father’s education (0,1)
# CHOICE: Choice of secondary school that they attend (1=first choice, ... 4=fourth choice)
# MED: Mother’s education (0,1)
############################################################################

# descriptives
by_school <- data %>% group_by(data$PID)
n_groups(by_school)
summary(group_size(by_school))
sd(group_size(by_school))
summary(data$ATTAIN)

# constructing the categorical predictor
# 1 = mom no, dad no
# 2 = mom yes, dad no
# 3 = mom no, dad yes
# 4 - mom yes, dad yes
data$PED <- NA
data$PED <- ifelse(data$MED == 0 & data$FED == 0, 1,
            ifelse(data$MED == 1 & data$FED == 0, 2,
            ifelse(data$MED == 0 & data$FED == 1, 3,
            ifelse(data$MED == 1 & data$FED == 1, 4, NA))))

table(data$PED)/3435

# constructing the dummy codes
data$d1 <- data$d2 <- data$d3 <- NA
data$d1 <- ifelse(data$PED==2, 1, 0)
data$d2 <- ifelse(data$PED==3, 1, 0)
data$d3 <- ifelse(data$PED==4, 1, 0)

# constructing the effect codes
data$e1 <- data$e2 <- data$e3 <- NA
data$e1 <- ifelse(data$PED==1, -1,
            ifelse(data$PED==2, 1, 0))
data$e2 <- ifelse(data$PED==1, -1,
            ifelse(data$PED==3, 1, 0))
data$e3 <- ifelse(data$PED==1, -1,
            ifelse(data$PED==4, 1, 0))

# constructing the contrast codes
data$c1 <- data$c2 <- data$c3 <- NA
data$c1 <- ifelse(data$PED==1, -1, 1/3)
data$c2 <- ifelse(data$PED==2 | data$PED==3, -1/2,
            ifelse(data$PED==4, 1, 0))
data$c3 <- ifelse(data$PED==2, -1,
            ifelse(data$PED==3, 1, 0))

# compute cluster means
data$d1.mean <- ave(data$d1,data$PID)
data$d2.mean <- ave(data$d2,data$PID)
data$d3.mean <- ave(data$d3,data$PID)

data$e1.mean <- ave(data$e1,data$PID)
data$e2.mean <- ave(data$e2,data$PID)
```

```
data$e3.mean <- ave(data$e3,data$PID)

data$c1.mean <- ave(data$c1,data$PID)
data$c2.mean <- ave(data$c2,data$PID)
data$c3.mean <- ave(data$c3,data$PID)

# compute CWC level-1 variables
data$d1.cwc <- data$d1 - data$d1.mean
data$d2.cwc <- data$d2 - data$d2.mean
data$d3.cwc <- data$d3 - data$d3.mean

data$e1.cwc <- data$e1 - data$e1.mean
data$e2.cwc <- data$e2 - data$e2.mean
data$e3.cwc <- data$e3 - data$e3.mean

data$c1.cwc <- data$c1 - data$c1.mean
data$c2.cwc <- data$c2 - data$c2.mean
data$c3.cwc <- data$c3 - data$c3.mean

######################### UN MODEL ##########################################
un.d <- lmer(ATTAIN ~ d1 + d2 + d3 + (1|PID), data=data)
un.e <- lmer(ATTAIN ~ e1 + e2 + e3 + (1|PID), data=data)
un.c <- lmer(ATTAIN ~ c1 + c2 + c3 + (1|PID), data=data)
anova(un.d, un.c, un.e)

######################## CWC(M) MODEL #######################################
cwcm.d <- lmer(ATTAIN ~ d1.cwc + d2.cwc + d3.cwc + d1.mean + d2.mean + d3.mean + (1|PID),
data=data)
cwcm.e <- lmer(ATTAIN ~ e1.cwc + e2.cwc + e3.cwc + e1.mean + e2.mean + e3.mean + (1|PID),
data=data)
cwcm.c <- lmer(ATTAIN ~ c1.cwc + c2.cwc + c3.cwc + c1.mean + c2.mean + c3.mean + (1|PID),
data=data)
anova(cwcm.d, cwcm.c, cwcm.e)

######################## UN(M) MODEL #######################################
unm.d <- lmer(ATTAIN ~ d1 + d2 + d3 + d1.mean + d2.mean + d3.mean + (1|PID), data=data)
unm.e <- lmer(ATTAIN ~ e1 + e2 + e3 + e1.mean + e2.mean + e3.mean + (1|PID), data=data)
unm.c <- lmer(ATTAIN ~ c1 + c2 + c3 + c1.mean + c2.mean + c3.mean + (1|PID), data=data)
anova(unm.d, unm.c, unm.e)


################# TESTING OVERALL EFFECTS OF THE PREDICTOR ####################

### deviance tests ###

# null model - no predictors
null.model <- lmer(ATTAIN ~ (1|PID), data=data)

# level 1 CWC predictors only
cwc.d <- lmer(ATTAIN ~ d1.cwc + d2.cwc + d3.cwc + (1|PID), data=data)
cwc.e <- lmer(ATTAIN ~ e1.cwc + e2.cwc + e3.cwc + (1|PID), data=data)
cwc.c <- lmer(ATTAIN ~ c1.cwc + c2.cwc + c3.cwc + (1|PID), data=data)

# level 2 cluster means only
m.d <- lmer(ATTAIN ~ d1.mean + d2.mean + d3.mean + (1|PID), data=data)
m.e <- lmer(ATTAIN ~ e1.mean + e2.mean + e3.mean + (1|PID), data=data)
m.c <- lmer(ATTAIN ~ c1.mean + c2.mean + c3.mean + (1|PID), data=data)

anova(null.model, cwc.d) # joint significance of level-1 fixed effects
anova(null.model, m.d) # joint significance of level-2 fixed effects
anova(null.model, cwcm.d) # joint significance of all fixed effects at both levels

anova(null.model, cwc.e) # joint significance of level-1 fixed effects
anova(null.model, m.e) # joint significance of level-2 fixed effects
anova(null.model, cwcm.e) # joint significance of all fixed effects at both levels

anova(null.model, cwc.c) # joint significance of level-1 fixed effects
anova(null.model, m.c) # joint significance of level-2 fixed effects
anova(null.model, cwcm.c) # joint significance of all fixed effects at both levels
```

```
### multivariate Wald-style F-tests ###
## using package lmerTest

# 3df joint test of the three level 1 fixed effects
L.l1 <- diag(length(fixef(cwcm.d)))[2:4, ]
contest(cwcm.d, L.l1)
contest(cwcm.e, L.l1)
contest(cwcm.c, L.l1)

# 3df joint test of the three level 2 fixed effects
L.l2 <- diag(length(fixef(cwcm.d)))[5:7, ]
contest(cwcm.d, L.l2)
contest(cwcm.e, L.l2)
contest(cwcm.c, L.l2)

# 6df joint test for all fixed effects at both levels
L.all <- diag(length(fixef(cwcm.d)))[2:7, ]
contest(cwcm.d, L.all)
contest(cwcm.e, L.all)
contest(cwcm.c, L.all)
```

Online Appendix G: Interpreting Unweighted Effect Codes in Empirical Example

With unweighted effect codes, slopes are interpreted as the difference between the mean of the focal group *f* and the unweighted mean of all group means in the sample. Our derivations in Appendix C indicate that this interpretation can be carried forward to multilevel settings in a similar way. The within-cluster slope of an effect code is interpreted as the expected difference between the within-cluster mean of group *f* and the unweighted mean of all groups. Similarly, the between-cluster slope can be interpreted as the difference between a cluster composed entirely of group *f* and the unweighted mean of all groups.

The CWC(M) Model

The within-cluster slope of the first effect code is $\gamma_{10} = 0.053$. Within schools, on average, we expect that children with an educated mother will have academic achievement scores that are about 0.053 points greater than the unweighted mean. Next, the within-cluster slope of the second effect code is $\gamma_{20} = 0.249$, implying that within schools, children with an educated father will have academic achievement scores that are about 0.249 points greater than the overall unweighted mean. Finally, the within-cluster slope of the third effect code is $\gamma_{30} = 0.248$, indicating that we expect children with two educated parents to have academic achievement scores that are about 0.248 points higher than the unweighted mean, within schools, on average. Here, we see that relative to the unweighted mean, having only an educated father or having two educated parents leads to similar gains in academic achievement among students.

The between-cluster slope of the first effect code is $\gamma_{01} = –0.764$. This slope implies that as we go from the unweighted mean of all groups to a school composed entirely of children with an educated mother, we expect mean academic achievement score to decrease by about 0.764 points. Next, the between-cluster slope of the second effect code is $\gamma_{02} = 3.057$. As we go from the unweighted mean of all groups to a school composed entirely of children with an educated father, we expect the mean academic achievement score to increase by 3.057 points. Finally, the between-cluster slope of the third effect code is $\gamma_{03} = 0.889$. As we go from the unweighted mean of all groups to a school composed entirely of children with two educated parents, the mean academic achievement score is expected to increase by 0.889 points. These slopes indicate that relative to the unweighted mean, attending a school in which many children have educated fathers yields the greatest gains in academic achievement.

Note that the slopes associated with effect codes bear interpretations similar to those of dummy codes, but instead of using the reference group (here, children with no educated parents) as the comparison point, we instead use the unweighted mean of the group means from the entire sample. Because children with no educated parents have been designated as the reference group in our effect coding system, we do not obtain any estimates with respect to this group specifically.

The UN(M) Model

The within-cluster slopes yielded by the UN(M) Model are identical to those yielded by the CWC(M) Model and can be interpreted identically. At level 2, the contextual effects yielded by the UN(M) Model can be interpreted as follows. First, the contextual effect associated with the first effect code is $\gamma_{01}^{c} = –0.817$. Holding an individual student’s parental education constant, we expect that as we move from the unweighted mean of all groups to a school where all children have an educated mother, the mean academic achievement score will decrease by 0.817 points. Thus, holding constant the effect of having an educated mother at the individual level, attending a school where many children have educated mothers predicts a reduction in academic achievement relative to the unweighted sample mean. Next, the contextual effect of the second effect code is γ_{02}^{c} = 2.808. Holding an individual student's parental education constant, we expect academic achievement scores to increase by 2.808 points upon moving from the unweighted mean to a school where all children have an educated father. Finally, the contextual effect of the third effect code is γ_{03}^{c} = 0.641. Holding an individual student's parental education constant, upon moving from the unweighted mean of all groups to a school where all children have two educated parents, we expect academic achievement scores to increase by 0.641 points. Beyond any effects that may be present at the individual level, it appears that attending a school where many children have an educated father results in the greatest academic achievement gains relative to the unweighted sample mean.