Supplemental Material S1: Standard errors of contextual effects estimates

When using contextual models (2)–(5) in manuscript, the standard errors of γ_1^{12}, γ_2^{12}, γ_1^{13}, γ_3^{13}, γ_2^{23}, and γ_3^{23} can be calculated as

$$se(\hat{\gamma}_1^{12}) = \sqrt{\frac{\sigma_{22}^*}{\sigma_{11}^*\sigma_{22}^* - \sigma_{22}^*\sigma_{22}^*}} = \sqrt{\frac{1}{\sigma_{11}^* - \sigma_{22}^*}} = \sqrt{\frac{1-\rho_1}{\sigma_x^2 IJK(1-\eta_2^2)(1-\eta_3^2)}},$$

$$se(\hat{\gamma}_2^{12}) = \sqrt{\frac{\sigma_{11}^*}{\sigma_{11}^*\sigma_{22}^* - \sigma_{22}^*\sigma_{22}^*}}$$

$$= \sqrt{\frac{(1-\eta_3^2)\rho_2 IJ[(1-\eta_2^2)(\rho_1-\rho_2)I + (1-\rho_1)] + (\rho_1-\rho_2)I[(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I - (\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2-2)(1-\rho_1)] + (1-\rho_1)^2}{IJK(1-\eta_2^2)(1-\eta_3^2)\sigma_x^2[IJ\eta_2^2(1-\eta_3^2)\rho_2 + (\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]}},$$

$$se(\hat{\gamma}_1^{13}) = \sqrt{\frac{\sigma_{33}^*}{\sigma_{11}^*\sigma_{33}^* - \sigma_{33}^*\sigma_{33}^*}} = \sqrt{\frac{1}{\sigma_{11}^* - \sigma_{33}^*}} = \sqrt{\frac{(1-\rho_1)[I(\rho_1-\rho_2)+(1-\rho_1)]}{\sigma_x^2 IJK(1-\eta_3^2)[(1-\eta_2^2)(\rho_1-\rho_2)I + 1-\rho_1]}},$$

$$se(\hat{\gamma}_3^{13}) = \sqrt{\frac{\sigma_{11}^*}{\sigma_{11}^*\sigma_{33}^* - \sigma_{33}^*\sigma_{33}^*}}$$

$$= \sqrt{\frac{(1-\eta_3^2)\rho_2 IJ[(1-\eta_2^2)(\rho_1-\rho_2)I + (1-\rho_1)] + (\rho_1-\rho_2)I[(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I - (\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2-2)(1-\rho_1)] + (1-\rho_1)^2}{IJK(1-\eta_3^2)\eta_3^2\sigma_x^2[(1-\eta_2^2)(\rho_1-\rho_2)I + (1-\rho_1)]}},$$

$$se(\hat{\gamma}_2^{23}) = \sqrt{\frac{\sigma_{33}^*}{\sigma_{22}^*\sigma_{33}^* - \sigma_{33}^*\sigma_{33}^*}} = \sqrt{\frac{1}{\sigma_{22}^* - \sigma_{33}^*}} = \sqrt{\frac{I(\rho_1-\rho_2)+1-\rho_1}{\sigma_x^2 IJK\eta_2^2(1-\eta_3^2)}},$$

$$se(\hat{\gamma}_3^{23}) = \sqrt{\frac{\sigma_{22}^*}{\sigma_{22}^*\sigma_{33}^* - \sigma_{33}^*\sigma_{33}^*}} = \sqrt{\frac{\eta_2^2(1-\eta_3^2)\rho_2 IJ + (\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]}{IJK\sigma_x^2[\eta_2^2(1-\eta_3^2)\eta_3^2]}} = se(\hat{\gamma}_3).$$

If the two-level model —which implies $K = 1$, $\sigma_3^2 = 0$, $\rho_2 = 0$, and $\eta_3^2 = 0$—is used to investigate contextual effect, then substituting these values into the above equations provides the standard errors of estimates as

$$se(\hat{\gamma}_{1|2}) = \sqrt{\frac{1-\rho_1}{IJ(1-\eta_2^2)\sigma_x^2}},$$

$$se(\hat{\gamma}_{2|2}) = \sqrt{\frac{I(1-\eta_2^2)\rho_1 + (1-\rho_1)}{IJ(1-\eta_2^2)\eta_2^2\sigma_x^2}}.$$

Supplemental Material S2: Sample size determination formulas based on desired statistical power when estimating contextual effects γ_2^{12}, γ_3^{13}, γ_3^{23}, and $\gamma_{2|2}$

Using the results provided in Supplemental Material S1, the formulas for lower limits of the sample sizes I, J, and K required to obtain ψ to test $H_0 : \gamma_2^{12} = 0$ can be derived as

$$I > \frac{(1-\rho_1)(\sqrt{I_1+I_2+I_3}+I_4)}{2(1-\eta_2^2)(1-\eta_3^2)[-(\rho_1-\rho_2)[\rho_2(J-1)+\rho_1](z_\psi+z_{1-\alpha/2})^2+\Delta_2^2 JK\sigma_x^2[\eta_2^2(1-\eta_3^2)\rho_2 J+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(\rho_1-\rho_2)]]},$$

$$I_1 = [(1-\eta_3^2)\rho_2 J[(1-\eta_3^2)\rho_2 J+2(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(\rho_1-\rho_2)]+(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2)^2(\rho_1-\rho_2)^2](z_\psi+z_{1-\alpha/2})^4,$$

$$I_2 = 2(1-\eta_2^2)(1-\eta_3^2)\Delta_2^2 JK\sigma_x^2[(1-\eta_3^2)(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)\rho_2 J+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2(\rho_1-\rho_2)](z_\psi+z_{1-\alpha/2})^2,$$

$$I_3 = (1-\eta_2^2)^2(1-\eta_3^2)^2(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2\Delta_2^4 J^2K^2\sigma_x^4,$$

$$I_4 = (\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2+2)(\rho_1-\rho_2)(z_\psi+z_{1-\alpha/2})^2+(1-\eta_3^2)J[\rho_2(z_\psi+z_{1-\alpha/2})^2-(1-\eta_2^2)(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)\Delta_2^2 K\sigma_x^2],$$

$$J > \frac{\sqrt{J_1+J_2+J_3}+J_4}{2(1-\eta_2^2)\eta_2^2(1-\eta_3^2)\rho_2\Delta_2^2 IK\sigma_x^2},$$

$$J_1 = \rho_2^2[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]^2(z_\psi+z_{1-\alpha/2})^4,$$

$$J_2 = [-2(1-\eta_2^2)\rho_2\Delta_2^2[(\rho_1-\rho_2)I+(1-\rho_1)][(1-\eta_2^2)(\eta_2^2\eta_3^2+\eta_3^2-\eta_2^2)(\rho_1-\rho_2)I-(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(1-\rho_1)]]K\sigma_x^2(z_\psi+z_{1-\alpha/2})^2,$$

$$J_3 = (1-\eta_2^2)^2(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2\Delta_2^4[(\rho_1-\rho_2)I+(1-\rho_1)]^2K^2\sigma_x^4,$$

$$J_4 = \rho_2[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)](z_\psi+z_{1-\alpha/2})^2-(1-\eta_2^2)(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)\Delta_2^2[(\rho_1-\rho_2)I+(1-\rho_1)]K\sigma_x^2,$$

$$K > \frac{[(1-\eta_3^2)\rho_2 IJ[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]+[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(1-\rho_1)]](z_\psi+z_{1-\alpha/2})^2}{(1-\eta_2^2)(1-\eta_3^2)\Delta_2^2 IJ\sigma_x^2[\eta_2^2(1-\eta_3^2)\rho_2 IJ+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]}.$$

Here, the effect size Δ_2 is defined in a manner similar to (8) in manuscript. Similar formulas to test $H_0 : \gamma_3^{13} = 0$ can be derived as

$$I > \frac{(1-\rho_1)(\sqrt{I_1+I_2+I_3}+I_4)}{2(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)[(\rho_2(J-1)+\rho_1)(z_\psi+z_{1-\alpha/2})^2-\eta_3^2\Delta_3^2 JK\sigma_x^2]},$$

$$I_1 = [(1-\eta_3^2)^2\rho_2^2 J^2+[2(1-\eta_3^2)(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)\rho_2(\rho_1-\rho_2)]J+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2(\rho_1-\rho_2)^2](z_\psi+z_{1-\alpha/2})^4,$$

$$I_2 = [-2(1-\eta_3^2)^2\eta_3^2\rho_2\Delta_3^2 J^2-2(1-\eta_3^2)\eta_3^2(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(\rho_1-\rho_2)\Delta_3^2 J]K\sigma_x^2(z_\psi+z_{1-\alpha/2})^2,$$

$$I_3 = (1-\eta_3^2)^2\eta_3^4\Delta_3^4 J^2K^2\sigma_x^4,$$

$$I_4 = -[(1-\eta_3^2)\rho_2 J+(\eta_2^2\eta_3^2-\eta_2^2-\eta_3^2+2)(\rho_1-\rho_2)](z_\psi+z_{1-\alpha/2})^2+(1-\eta_3^2)\eta_3^2\Delta_3^2 JK\sigma_x^2,$$

$$J > \frac{[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(1-\rho_1)](z_\psi+z_{1-\alpha/2})^2}{(1-\eta_3^2)I[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)][\eta_3^2\Delta_3^2 K\sigma_x^2-\rho_2(z_\psi+z_{1-\alpha/2})^2]},$$

$$K > \frac{[(1-\eta_3^2)\rho_2 I[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]J+(\rho_1-\rho_2)I[(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I-(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2-2)(1-\rho_1)]+(1-\rho_1)^2](z_\psi+z_{1-\alpha/2})^2}{(1-\eta_3^2)\eta_3^2\Delta_3^2 IJ\sigma_x^2[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]}.$$

Here, the effect size Δ_3 is defined in a manner similar to (8) in manuscript. Similar formulas to test $H_0 : \gamma_3^{23} = 0$ in the model can also be derived as

$$I > \frac{(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(1-\rho_1)(z_\psi+z_{1-\alpha/2})^2}{\eta_2^2(1-\eta_3^2)\eta_3^2\Delta_3^2 JK\sigma_x^2-(z_\psi+z_{1-\alpha/2})^2[\eta_2^2(1-\eta_3^2)\rho_2 J+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(\rho_1-\rho_2)]},$$
$$J > \frac{(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)](z_\psi+z_{1-\alpha/2})^2}{\eta_2^2(1-\eta_3^2)I[\eta_3^2\Delta_3^2 K\sigma_x^2-\rho_2(z_\psi+z_{1-\alpha/2})^2]},$$
$$K > \frac{[\eta_2^2(1-\eta_3^2)\rho_2 IJ+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]](z_\psi+z_{1-\alpha/2})^2}{\eta_2^2(1-\eta_3^2)\eta_3^2\Delta_3^2 IJ\sigma_x^2}.$$

Finally, when the two-level model (i.e., $K = 1$, $\sigma_3^2 = 0$, $\rho_2 = 0$, and $\eta_3^2 = 0$) is used, similar formulas to test $H_0 : \gamma_{2|2} = 0$ can be derived as

$$I > \frac{(1-\rho_1)(z_\psi+z_{1-\alpha/2})^2}{(1-\eta_2^2)[\eta_2^2\Delta_2^2 J\sigma_x^2-\rho_1(z_\psi+z_{1-\alpha/2})^2]},$$
$$J > \frac{[\rho_1 I(1-\eta_2^2)+1-\rho_1](z_\psi+z_{1-\alpha/2})^2}{(1-\eta_2^2)\eta_2^2\Delta_2^2 I\sigma_x^2}.$$

Supplemental Material S3:Sample size determination formulas based on desired width of confidence interval when estimating contextual effects γ_2^{12}, γ_3^{13}, γ_3^{23}, and $\gamma_{2|2}$

A $100(1-\alpha)\%$ confidence interval for γ_p is expressed as

$$\hat{\gamma}_p - z_{1-\alpha/2} se(\hat{\gamma}_p) \leqq \gamma_p \leqq \hat{\gamma}_p + z_{1-\alpha/2} se(\hat{\gamma}_p),$$

and so the width of a confidence interval L can be calculated as

$$L = 2z_{1-\alpha/2} se(\hat{\gamma}_p).$$

Note that this is also the width of the confidence interval for Δ_c because $\Delta_c = \gamma_c$ here.

When a desired width of the confidence interval is specified as L', from standard errors provided in (16)-(17) in manuscript, the relation $L \leq L'$ can be rewritten, resulting in formulas for lower limits of the sample sizes I, J, and K required in order to satisfy this relation in estimating γ_2 from the model (1). For this, $L \leq L'$ gives

$$I > \frac{4(1-\rho_1)z_{1-\alpha/2}^2}{(1-\eta_2^2)[\eta_2^2(1-\eta_3^2)JKL'^2 - 2z_{1-\alpha/2}^2(\rho_1-\rho_2)]},$$
$$J > \frac{4z_{1-\alpha/2}^2[(1-\eta_2^2)(\rho_1-\rho_2)I + (1-\rho_1)]}{(1-\eta_2^2)\eta_2^2(1-\eta_3^2)IKL'^2},$$
$$K > \frac{4z_{1-\alpha/2}^2[(1-\eta_2^2)(\rho_1-\rho_2)I + (1-\rho_1)]}{(1-\eta_2^2)\eta_2^2(1-\eta_3^2)IJL'^2}.$$

Similar formulas for estimating γ_3 can be derived as

$$I > \frac{4(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(1-\rho_1)z_{1-\alpha/2}^2}{\eta_2^2(1-\eta_3^2)[-\eta_3^2JKL'^2 - 2z_{1-\alpha/2}^2[\rho_2(J-1)+\rho_1]] - 2\eta_3^2(\rho_1-\rho_2)z_{1-\alpha/2}^2},$$
$$J > \frac{4(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)z_{1-\alpha/2}^2[I(\rho_1-\rho_2)+(1-\rho_1)]}{\eta_2^2(1-\eta_3^2)I(\eta_3^2KL'^2 - 2\rho_2 z_{1-\alpha/2}^2)},$$
$$K > \frac{4z_{1-\alpha/2}^2[\eta_2^2(1-\eta_3^2)\rho_2 IJ + (\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]}{\eta_2^2(1-\eta_3^2)\eta_3^2 IJL'^2}.$$

From standard errors provided in Supplemental Material S1, the formulas for lower limits of the sample sizes I, J, and K necessary to satisfy $L \le L'$ in estimating γ_2^{12} can be derived as

$$I > \frac{(1-\rho_1)(\sqrt{I_1+I_2}+I_3)}{2(1-\eta_2^2)(1-\eta_3^2)[-J[\eta_2^2(1-\eta_3^2)\rho_2 J-(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2)(\rho_1-\rho_2)]K\sigma_x^2 L'^2+4(\rho_1-\rho_2)z_{1-\alpha/2}^2(\rho_2 J-\rho_2+\rho_1)]},$$

$$I_1 = (1-\eta_2^2)(1-\eta_3^2)JK\sigma_x^2L'^2[(1-\eta_2^2)(1-\eta_3^2)(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2)^2JK\sigma_x^2L'^2+8z_{1-\alpha/2}^2[(1-\eta_3^2)(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)\rho_2 J+(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2)^2(\rho_1-\rho_2)]],$$

$$I_2 = 16z_{1-\alpha/2}^4[(1-\eta_3^2)\rho_2 J[(1-\eta_3^2)\rho_2 J+2(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(\rho_1-\rho_2)]+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2(\rho_1-\rho_2)^2],$$

$$I_3 = -(1-\eta_2^2)(1-\eta_3^2)(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2)JK\sigma_x^2L'^2-4z_{1-\alpha/2}^2[(1-\eta_3^2)\rho_2 J+(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2+2)(\rho_1-\rho_2)],$$

$$J > -\frac{\sqrt{J_1+J_2+J_3}+J_4}{2(1-\eta_2^2)\eta_2^2(1-\eta_3^2)\rho_2 IK\sigma_x^2L'^2},$$

$$J_1 = (1-\eta_2^2)^2(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2[I(\rho_1-\rho_2)+(1-\rho_1)]^2K^2\sigma_x^4L'^4,$$

$$J_2 = -8(1-\eta_2^2)\rho_2 z_{1-\alpha/2}^2[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(\eta_2^2\eta_3^2+\eta_3^2-\eta_2^2)(\rho_1-\rho_2)I-(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(1-\rho_1)]K\sigma_x^2L'^2,$$

$$J_3 = 16\rho_2^2 z_{1-\alpha/2}^4[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]^2,$$

$$J_4 = (1-\eta_2^2)(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)],$$

$$K > \frac{4z_{1-\alpha/2}^2[(1-\eta_3^2)\rho_2 I[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]J+(\rho_1-\rho_2)I[(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2+2)(1-\rho_1)]+(1-\rho_1)^2]}{(1-\eta_2^2)(1-\eta_3^2)IJ\sigma_x^2[\eta_2^2(1-\eta_3^2)\rho_2 IJ+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]L'^2}.$$

Similar formulas for estimating γ_3^{13} can be derived as

$$I > \frac{(1-\rho_1)(\sqrt{I_1+I_2}+I_3)}{2(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)[\eta_3^2JK\sigma_x^2L'^2-4z_{1-\alpha/2}^2[\rho_2(J-1)+\rho_1]]},$$

$$I_1 = (1-\eta_3^2)J(\eta_3^2K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2)[(1-\eta_3^2)J(\eta_3^2K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2)-8(\eta_2^2\eta_3^2-\eta_3^2+\eta_2^2)(\rho_1-\rho_2)z_{1-\alpha/2}^2],$$

$$I_2 = 16(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)^2(\rho_1-\rho_2)^2z_{1-\alpha/2}^4,$$

$$I_3 = (1-\eta_3^2)J(4\rho_2 z_{1-\alpha/2}^2-\eta_3^2K\sigma_x^2L'^2)+4(\eta_2^2\eta_3^2-\eta_3^2-\eta_2^2+2)(\rho_1-\rho_2)z_{1-\alpha/2}^2,$$

$$J > \frac{4z_{1-\alpha/2}^2[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(1-\rho_1)]}{(1-\eta_3^2)I[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)](\eta_3^2K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2)},$$

$$K > \frac{4z_{1-\alpha/2}^2[(1-\eta_3^2)\rho_2 IJ[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]+[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(1-\rho_1)]]}{(1-\eta_3^2)\eta_3^2IJ\sigma_x^2L'^2[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]}.$$

Similar formulas for estimating γ_3^{23} can also be derived as

$$\begin{aligned}
I &> \frac{4(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(1-\rho_1)z_{1-\alpha/2}^2}{\eta_2^2(1-\eta_3^2)J(\eta_3^2K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2)-4(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)(\rho_1-\rho_2)z_{1-\alpha/2}^2},\\
J &> \frac{4(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)z_{1-\alpha/2}^2[I(\rho_1-\rho_2)+(1-\rho_1)]}{\eta_2^2(1-\eta_3^2)I(\eta_3^2K\sigma_x^2L'^2-4\rho_2 z_{1-\alpha/2}^2)},\\
K &> \frac{4z_{1-\alpha/2}^2[\eta_2^2(1-\eta_3^2)\rho_2 IJ+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]}{\eta_2^2(1-\eta_3^2)\eta_3^2 IJ\sigma_x^2L'^2}.
\end{aligned}$$

Finally, when the two-level model (i.e., $K=1$, $\sigma_3^2=0$, $\rho_2=0$ and $\eta_3^2=0$) is used, similar formulas for estimating $\gamma_{2|2}$ in the model (5) can be derived as

$$\begin{aligned}
I &> \frac{4(1-\rho_1)z_{1-\alpha/2}^2}{(1-\eta_2^2)(J\sigma_x^2L'^2\eta_2^2-4\rho_1 z_{1-\alpha/2}^2)},\\
J &> \frac{4z_{1-\alpha/2}^2[\rho_1 I(1-\eta_2^2)+(1-\rho_1)]}{(1-\eta_2^2)\eta_2^2 I\sigma_x^2L'^2}.
\end{aligned}$$

Supplemental Material S4: On the relationship between derived formulas for evaluating contextual effects and those for experimental research

Although in manuscript the issue of contextual effect is the focus, let us consider, for theoretical interest, a situation in which investigating the contextual effect is not of interest and the researcher wants only to evaluate the relation between individual-level outcomes and a single predictor by using one of the following equations:

$$Y_{ijk} = \gamma_1^1 X_{ijk} + (e_{ijk} + e_{jk} + e_k),$$

$$Y_{ijk} = \gamma_2^2 X_{.jk} + (e_{ijk} + e_{jk} + e_k),$$

$$Y_{ijk} = \gamma_3^3 X_{..k} + (e_{ijk} + e_{jk} + e_k).$$

By using the similar procedure described in Appendix, the standard errors of $\hat{\gamma}_1^1$, $\hat{\gamma}_2^2$, and $\hat{\gamma}_3^3$ can be calculated as

$$se(\hat{\gamma}_1^1) = 1/\sqrt{\sigma_{11}^*} = \sqrt{\frac{(1-\rho_1)[I(\rho_1-\rho_2)+(1-\rho_1)]f}{IJK\sigma_x^2[(1-\eta_3^2)\rho_2 IJ[(1-\eta_2^2)(\rho_1-\rho_2)I+(1-\rho_1)]+[I(\rho_1-\rho_2)+(1-\rho_1)][(1-\eta_2^2)(1-\eta_3^2)(\rho_1-\rho_2)I+(1-\rho_1)]]}},$$

$$se(\hat{\gamma}_2^2) = 1/\sqrt{\sigma_{22}^*} = \sqrt{\frac{[I(\rho_1-\rho_2)+(1-\rho_1)]f}{IJK\sigma_x^2[\eta_2^2(1-\eta_3^2)\rho_2 IJ+(\eta_2^2+\eta_3^2-\eta_2^2\eta_3^2)[I(\rho_1-\rho_2)+(1-\rho_1)]]}},$$

$$se(\hat{\gamma}_3^3) = 1/\sqrt{\sigma_{33}^*} = \sqrt{\frac{f}{IJK\sigma_x^2\eta_3^2}}.$$

The first model may be used in experimental research, especially when using a multisite randomized trial (MRT; Raudenbush & Liu, 2000) with group assignment conducted at the level of individuals ($X_{ijk} = -1$ for the control group and $X_{ijk} = 1$ for the experimental group[*1]) is applied to evaluate an experimental effect through γ_1^1. In MRT, if the design is balanced and the experimental and control group sizes are equal in each level-2 unit (class), then $\eta_3^2 = \eta_2^2 = 0$. Thus, from this relation the formulas for lower limits of the sample sizes I, J, and K required to obtain ψ to test $H_0 : \gamma_1^1 = 0$ can be derived as

$$I = \frac{(1-\rho_1)(z_{1-\alpha/2}+z_\psi)^2}{\Delta_1^2 JK\sigma_x^2},$$

$$J = \frac{(1-\rho_1)(z_{1-\alpha/2}+z_\psi)^2}{\Delta_1^2 IK\sigma_x^2},$$

$$K = \frac{(1-\rho_1)Z(z_{1-\alpha/2}+z_\psi)^2}{\Delta_1^2 IJ\sigma_x^2}.$$

[*1] Here, the dummy variable taking -1 or 1 is assumed because X_{ijk} is standardized, meaning that the mean difference between groups is expressed as $2\gamma_1^1$.

Here, Δ_1 is an effect size and that can be defined in a manner similar to (8). These results imply a formula for evaluating the total required sample size IJK:

$$IJK = \frac{(1-\rho_1)(z_{1-\alpha/2} + z_\psi)^2}{\Delta_1^2 \sigma_x^2}.$$

This equation is essentially the same as that derived in Usami (2014), which assumes an unbalanced design [*2].

Next, the second model may be used when level-2 CRT is applied (i.e., classes within each school are randomly assigned to the experimental group ($X_{.jk} = -1$) or the control group ($X_{.jk} = 1$)) to evaluate an experimental effect through γ_2^2. In level-2 CRT, if the design is balanced and the sizes of the experimental and control groups are the same in each level-3 unit (school), then $\eta_3^2 = 0$ and $\eta_2^2 = 1$. Thus, from this relation, the formulas for lower limits of the sample sizes I, J, and K required to obtain ψ to test $H_0 : \gamma_2^2 = 0$ can be derived as

$$I = \frac{(1-\rho_1)(z_{1-\alpha/2} + z_\psi)^2}{\Delta_2^2 JK\sigma_x^2 - (\rho_1 - \rho_2)(z_{1-\alpha/2} + z_\psi)^2},$$
$$J = \frac{[I(\rho_2 - \rho_1) + (1-\rho_1)](z_{1-\alpha/2} + z_\psi)^2}{\Delta_2^2 IK\sigma_x^2},$$
$$K = \frac{[I(\rho_2 - \rho_1) + (1-\rho_1)](z_{1-\alpha/2} + z_\psi)^2}{\Delta_2^2 IJ\sigma_x^2}.$$

To the best of the author's knowledge, sample size determination formulas based on level-2 CRT in three-level model have not been derived.

Finally, the third model may be used when (level-3) CRT is applied (i.e., schools are randomly assigned to the experimental group ($X_{..k} = -1$) or control group ($X_{..k} = 1$)) to evaluate an experimental effect through γ_3^3. In (level-3) CRT, if the design is balanced and the experimental and control group sizes are equal, then $\eta_3^2 = 1$ and $\eta_2^2 = 0$. Thus, from this relation, the formulas for lower limits of the sample sizes I, J, and K required to obtain ψ to test $H_0 : \gamma_3^3 = 0$ can be derived as

$$I = \frac{(1-\rho_1)(z_{1-\alpha/2} + z_\psi)^2}{\Delta_3^2 JK\sigma_x^2 - (z_{1-\alpha/2} + z_\psi)^2[\rho_2(J-1) + \rho_1]},$$
$$J = \frac{[I(\rho_1 - \rho_2) + (1-\rho_1)](z_{1-\alpha/2} + z_\psi)^2}{I[\Delta_3^2 K\sigma_x^2 - \rho_2(z_{1-\alpha/2} + z_\psi)^2]},$$
$$K = \frac{[I[\rho_2(J-1) + \rho_1] + (1-\rho_1)](z_{1-\alpha/2} + z_\psi)^2}{\Delta_3^2 IJ\sigma_x^2},$$

[*2] Since X_{ijk} is standardized here, a term $P(1-P)$ (where P denotes the proportion of samples assigned to an experimental group) that appears in Usami (2014) to indicate the variance of dummy (group assignment) variables, equals 1, resulting to the derived formula.

yielding results mathematically the same as those derived in previous research (e.g., Heo & Leon, 2008; Usami, 2014) [*3].

Thus, it is clear that the formulas found in investigating contextual effects can also serve for investigating experimental effects by manipulating the values of η_3^2 and η_2^2 according to the experimental design (i.e., MRT, level-2 CRT, and (level-3) CRT). Formulas based on the desired width of confidence interval of experimental effect can also be easily derived in the same manner.

[*3] It should be noted that in Heo and Leon (2008) the number of units at the highest level (level-3) is set as $2K$, not K. Additionally, the derivation of results in Heo and Leon (2008) implicitly assumes a balanced design ($P = 1/2$). So, replacing $P(1 - P)$, which appears in the denominator of the formulas in Usami (2014), with 1/4 provides the same results as in Heo and Leon (2008). Thus, the formulas derived in Heo and Leon (2008) and Usami (2014) are mathematically equivalent to those derived here.