Results to accompany:

Cho, S.-J., Preacher, K. J., & Bottge, B. A. (in press). Detecting intervention effects in a cluster randomized design using multilevel structural equation modeling for binary responses. *Applied Psychological Measurement*.

Supplementary Material: Two-Level Random Intercept MLM

Denote y_{1jki} and y_{2jki} as binary item responses for a person j ($j = 1, \ldots, J$) nested within a cluster k ($k = 1, \ldots, K$), and an item i ($i = 1, \ldots, I$), for pre-test (denoted by subscript 1) and post-test (denoted by subscript 2) scores, respectively. Assume that responses are coded 0 for an incorrect response or 1 for a correct response, for example. Total scores, $\sum_i^I y_{1jki} = y_{1jk.}$ and $\sum_i^I y_{2jki} = y_{2jk.}$, are used in MLM. In order to have an explicit modeling of individual-level and cluster-level effects, a pretest-score covariate $y_{1jk.}$ can be decomposed into $y_{1.k.}$ that varies at the cluster level and a component $(y_{1jk.} - y_{1.k.})$ that varies at the individual level (Neuhaus & Kalbfleisch, 1998).

A model at Level 1 (e.g., the student level) can be specified as follows:

$$y_{2jk.} = \beta_{0k} + \beta_{1j} \cdot (y_{1jk.} - y_{1.k.}) + e_{jk}, \tag{1}$$

where $y_{1.k.}$ is the mean of total scores for cluster k (for cluster-mean centering), β_{0k} is the effect of post-test score for cluster k adjusted for pre-test score for that person, β_{1j} is the effect of pre-test score for person j, and e_{jk} is the residual of a post-test total score at Level 1, assumed to follow $N(0, \sigma^2)$. A model at Level 2 (e.g., the teacher level) can be specified as follows:

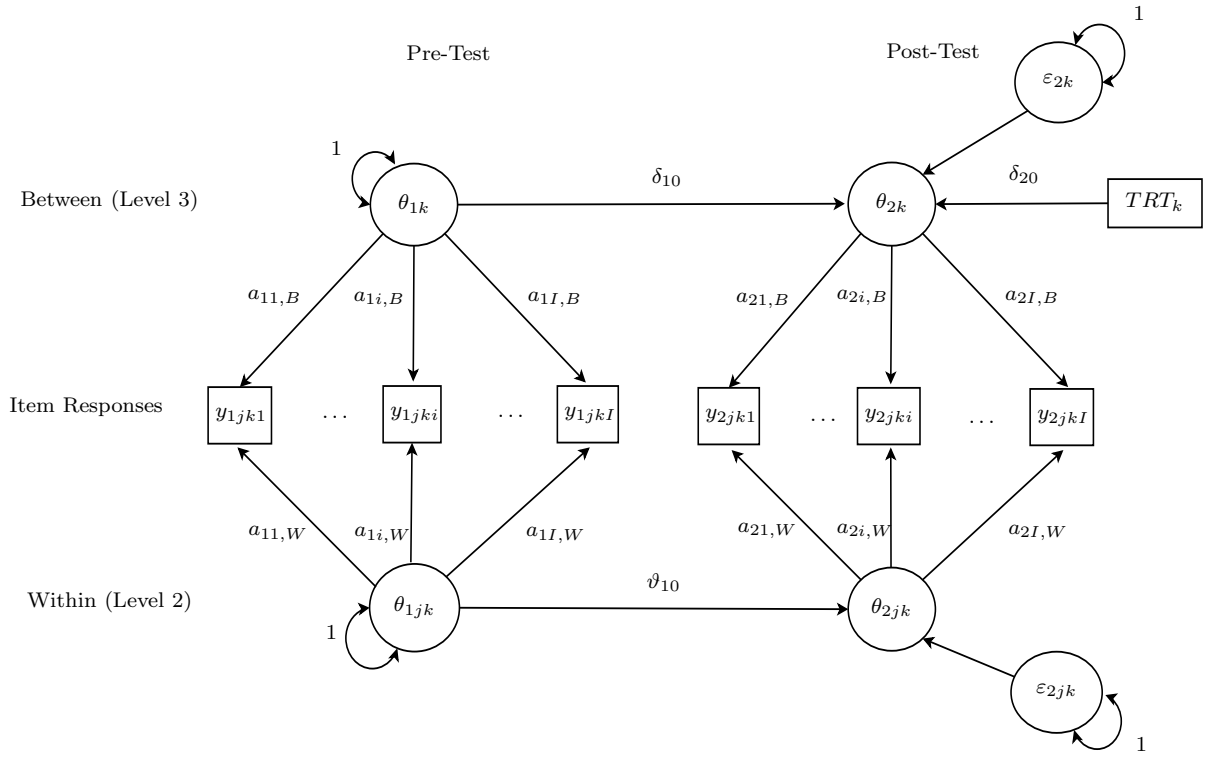
$$\beta_{0k} = \gamma_{00} + \gamma_{10} \cdot y_{1.k.} + \gamma_{20} \cdot TRT_k + u_k \tag{2}$$

and $\beta_{1j} = \gamma_{10}$, where TRT_k is a manifest treatment condition covariate with a value of 0 for members of the control group and a value of 1 for members of treatment group, γ_{00} is the average score at post-test for cluster k adjusted for the treatment condition, γ_{10} is the effect of the cluster mean pre-test score at Level 2, γ_{20} is the effect of the treatment condition, γ_{10} is the effect of pre-test score at Level 1, and u_k is the residual of the post-test total score

at Level 2, assumed to follow $N(0, \tau^2)$. Inserting the Level 2 model into the Level 1 model gives

$$y_{2jk.} = \gamma_{00} + \gamma_{10} \cdot (y_{1jk.} - y_{1.k.}) + \gamma_{10} \cdot y_{1.k.} + \gamma_{20} \cdot TRT_k + u_k + e_{jk}. \tag{3}$$

Supplementary Material: A Path Diagram of MSEM



In the figure, the squares and ellipses represent manifest and latent variables, respectively. At post-test, θ_{2jk} is explained by θ_{1jk} with a residual (ε_{2jk}) at Level 2 and θ_{2k} is explained by θ_{1k} and a treatment variable TRT_k with a residual (ε_{2k}) at Level 3. Each item at pre-test and post-test has its own item discrimination at Level 2 and Level 3, specified as a.

Supplementary Material: Mplus Code for MLM

```
TITLE: HLM

DATA: FILE IS data1.txt;

VARIABLE: NAMES ARE stuid clus support eai se LD

chisex chiethn1 y1-y20 yy1-yy20 pre prew preb post;

USEVARIABLES = clus post eai prew preb;

CLUSTER = clus;

BETWEEN = eai preb;

WITHIN = prew;

ANALYSIS: TYPE IS TWOLEVEL;

ESTIMATOR = Bayes;

PROCESSORS = 2;

MODEL:

%WITHIN%

post ON prew;

%BETWEEN%

post ON eai;

post ON preb;

OUTPUT: stdyx;
```

Supplementary Material: Mplus Code for MSEM

```
TITLE: MSEM

DATA: FILE IS data.txt;

VARIABLE: NAMES ARE stuid clus eai y1-y20 yy1-yy20;

USEVARIABLES = clus y1-y20 yy1-yy20 eai;

CATEGORICAL = y1-y20 yy1-yy20;

CLUSTER = clus;

BETWEEN = eai;

ANALYSIS: TYPE IS TWOLEVEL;

ESTIMATOR = Bayes;

PROCESSORS = 2;

MODEL:

%WITHIN%

f1w BY y1-y20*   (a1-a20);

f2w BY  yy1-yy20* (a21-a40);

f2w ON f1w;

[f1w-f2w@0];

f1w-f2w@1;

%BETWEEN%

f1b BY y1-y20*   (a41-a60);

f2b BY  yy1-yy20* (a61-a80);

f2b ON f1b;

f2b ON eai;

[y1$1 - y20$1];

[yy1$1 - yy20$1];

[f1b-f2b];

f1b-f2b@1;

y1-y20@1;

yy1-yy20@1;

MODEL PRIORS:

a1-a80 ~ N(0,1);

SAVEDATA:  save=fscores (30);

file is score.txt;
```

Table 1: Supplementary Material: Item Attributes and Item Parameter Estimates (95% Credibility Interval [CrI]) from MSEM.

Item	Attributes				Pre-test			Post-test		
	Operation	Denominator	Type	Stacks	$a_{1,W}$	$a_{1,B}$	b_1	$a_{2,W}$	$a_{2,B}$	b_2
1	addition	like	simple	2	1.20(0.93,1.52)	0.14(-0.29,0.56)	-2.10(-2.73,-1.50)	0.82(0.61,1.08)	0.17(-0.28,0.60)	-2.30(-3.02,-1.64)
2	addition	like	simple	2	1.37(1.05,1.76)	0.15(-0.30,0.59)	-2.61(-3.33,-1.94)	1.07(0.78,1.42)	0.32(-0.14,0.81)	-2.98(-3.97,-2.17)
3	addition	unlike	simple	2	3.02(2.43,3.82)	1.30(0.77,1.94)	-0.50(-1.88,0.65)	2.42(1.95,3.06)	1.22(0.69,1.87)	-0.98(-2.24,0.30)
4	addition	unlike	simple	2	2.59(2.11,3.21)	1.13(0.64,1.70)	-0.47(-1.52,0.48)	1.71(1.37,2.11)	0.84(0.41,1.33)	-0.56(-1.51,0.38)
5	addition	unlike	simple	2	2.92(2.30,3.61)	1.65(1.05,2.35)	0.63(-0.57,1.87)	1.98(1.58,2.47)	1.10(0.62,1.67)	0.06(-0.99,1.08)
6	addition	unlike	simple	2	2.92(2.36,3.66)	1.54(0.98,2.23)	0.03(-1.17,1.12)	2.57(2.02,3.28)	1.41(0.85,2.10)	-0.45(-1.84,1.01)
7	addition	unlike	mixed	2	2.60(2.09,3.17)	1.26(0.75,1.87)	0.14(-0.89,1.15)	2.18(1.72,2.78)	1.07(0.60,1.65)	-0.07(-1.19,1.17)
8	addition	unlike	mixed	2	2.32(1.84,2.89)	1.07(0.58,1.65)	0.35(-0.57,1.29)	2.08(1.68,2.63)	0.82(0.37,1.34)	0.14(-0.91,1.24)
9	addition	unlike	mixed	2	2.57(2.01,3.25)	1.26(0.72,1.89)	0.32(-0.67,1.33)	2.30(1.79,2.90)	1.41(0.86,2.08)	0.48(-0.80,1.66)
10	addition	unlike	mixed	2	2.60(2.02,3.41)	1.55(0.95,2.29)	0.76(-0.34,1.74)	1.88(1.48,2.39)	0.90(0.46,1.42)	0.26(-0.73,1.24)
11	addition	unlike	simple	3	2.67(2.13,3.34)	1.29(0.76,1.90)	-0.34(-1.47,0.69)	1.71(1.40,2.11)	0.86(0.43,1.36)	-0.33(-1.26,0.60)
12	addition	unlike	simple	3	2.92(2.28,3.67)	1.47(0.88,2.14)	0.22(-0.99,1.28)	1.89(1.51,2.34)	1.24(0.76,1.84)	0.21(-0.85,1.25)
13	addition	unlike	mixed	3	2.32(1.84,2.95)	1.31(0.77,1.94)	0.64(-0.31,1.58)	2.30(1.80,3.04)	1.15(0.65,1.74)	0.78(-0.46,2.17)
14	addition	unlike	mixed	3	1.96(1.55,2.47)	1.14(0.64,1.72)	0.80(-0.02,1.60)	1.85(1.43,2.37)	0.94(0.48,1.48)	0.87(-0.11,1.89)
15	subtraction	like	simple	2	1.06(0.80,1.37)	0.19(-0.21,0.59)	-1.97(-2.58,-1.37)	0.67(0.48,0.89)	0.18(-0.22,0.58)	-1.93(-2.52,-1.37)
16	subtraction	unlike	simple	2	2.73(2.18,3.43)	1.29(0.76,1.90)	0.19(-0.83,1.21)	1.71(1.38,2.18)	0.79(0.38,1.27)	0.03(-0.90,0.92)
17	subtraction	like	mixed	2	0.95(0.75,1.17)	0.16(-0.23,0.54)	-1.07(-1.60,-0.57)	0.50(0.37,0.65)	0.06(-0.30,0.41)	-0.93(-1.42,-0.45)
18	subtraction	unlike	mixed	2	1.10(0.82,1.42)	0.67(0.23,1.15)	1.05(0.49,1.60)	0.71(0.52,0.93)	0.61(0.22,1.07)	0.95(0.41,1.54)
19	subtraction	unlike	mixed	2	2.01(1.57,2.54)	1.08(0.57,1.66)	0.64(-0.12,1.47)	1.44(1.13,1.79)	0.81(0.39,1.30)	0.61(-0.21,1.41)
20	subtraction	unlike	mixed	2	1.13(0.79,1.55)	0.74(0.28,1.29)	1.67(1.03,2.33)	0.77(0.54,1.06)	0.67(0.22,1.21)	1.98(1.10,2.30)

Examples for the item attributes are as follows:
like denominators ($\frac{1}{4} + \frac{3}{4}$) or *unlike* denominators ($8\frac{2}{9} + 2\frac{1}{2}$)
simple fractions ($\frac{3}{8} + \frac{3}{4}$) or *mixed* numbers ($4\frac{1}{16} + \frac{1}{8} + \frac{1}{2}$)
two stacks ($\frac{\frac{3}{4}}{\frac{1}{2}}$) or *three stacks* (i.e., one more stack than in the two-stack example)

Table 2: Supplementary Material: RMSE of Estimates in MSEM

Conditions			Parameters								
K	n_j	ICC	$a_{2,W}$	$a_{2,B}$	b_2	$a_{1,W}$	$a_{1,B}$	b_1	δ_{00}	δ_{10}	μ
24	5	0.05	-	-	-	-	-	-	-	-	-
24	5	0.1	-	-	-	-	-	-	-	-	-
24	5	0.2	-	-	-	-	-	-	-	-	-
24	20	0.05	0.230	0.225	0.156	0.227	0.159	0.170	0.146	0.214	0.177
24	20	0.1	0.230	0.204	0.152	0.227	0.159	0.170	0.148	0.221	0.153
24	20	0.2	0.233	0.215	0.159	0.227	0.157	0.171	0.148	0.224	0.147
24	50	0.05	0.192	0.158	0.124	0.169	0.156	0.158	0.114	0.225	0.116
24	50	0.1	0.192	0.139	0.135	0.169	0.154	0.157	0.112	0.216	0.113
24	50	0.2	0.192	0.150	0.138	0.170	0.156	0.154	0.114	0.210	0.102
50	5	0.05	0.294	0.267	0.289	0.271	0.160	0.254	0.148	0.280	0.188
50	5	0.1	0.302	0.200	0.301	0.274	0.160	0.253	0.149	0.271	0.171
50	5	0.2	0.300	0.218	0.309	0.272	0.156	0.252	0.134	0.257	0.166
50	20	0.05	0.233	0.193	0.178	0.190	0.157	0.157	0.119	0.169	0.143
50	20	0.1	0.233	0.186	0.184	0.192	0.157	0.152	0.115	0.167	0.128
50	20	0.2	0.236	0.176	0.191	0.189	0.154	0.154	0.115	0.166	0.118
50	50	0.05	0.180	0.136	0.153	0.166	0.145	0.115	0.103	0.161	0.113
50	50	0.1	0.182	0.127	0.158	0.165	0.146	0.118	0.102	0.151	0.113
50	50	0.2	0.181	0.132	0.152	0.165	0.145	0.126	0.105	0.146	0.104
100	5	0.05	0.297	0.195	0.309	0.328	0.154	0.280	0.195	0.211	0.124
100	5	0.1	0.298	0.156	0.306	0.325	0.150	0.278	0.192	0.208	0.121
100	5	0.2	0.298	0.173	0.379	0.324	0.151	0.230	0.190	0.208	0.121
100	20	0.05	0.189	0.139	0.248	0.227	0.141	0.173	0.136	0.164	0.128
100	20	0.1	0.194	0.152	0.190	0.228	0.142	0.173	0.137	0.161	0.123
100	20	0.2	0.197	0.140	0.195	0.226	0.140	0.171	0.140	0.154	0.114
100	50	0.05	0.228	0.116	0.150	0.249	0.147	0.173	0.113	0.199	0.110
100	50	0.1	0.180	0.141	0.140	0.154	0.140	0.127	0.101	0.195	0.109
100	50	0.2	0.180	0.131	0.136	0.154	0.139	0.127	0.100	0.188	0.106