**Supplemental Materials**

These materials contain supplementary analyses, reported in Tables S1 – S4 and Figure S1, as well as Mplus code illustrating the analyses, presented in Figures S2-S4.

**Supplemental Analyses**

Tables S1 – S4 summarize a number of sensitivity analyses for research question four. The sensitivity analyses all used a multi-group model to assess measurement invariance, as a more general alternative to the MIMIC model. For each domain, the multi-group model was fitted using five different subsamples, all of which were balanced on sample size. The five different subsamples and the sample sizes per country are listed below.

1) All: the same sample reported in the main paper (*N* = 480)

2) Age 4-5: only children aged 4-5 (*N* = 453).

3) ECCD: only children who had received formal early child care via a Save the Children program (*N* = 245)

4) ECCD and Age 4-5: combined the conditions from analyses 2 and 3 (*N* = 243)

5) Without Vietnam: used the same subsamples from analyses 1, but omitted Vietnam (*N* = 480).

All analyses additionally balanced on gender. Analyses that required balance on age required omitting the Ethiopian sample since there were insufficient children below age 6 in this sample. Additionally, for the Social Emotional domain, the comparison involving the full sample (``All``) required fixing the residual loadings on the Personal Awareness subtask in Uganda and Bolivia to avoid Heywood cases. As shown in the tables, the results of all of the secondary analyses corroborated the conclusion made in the main paper: the IDELA domains are not measurement invariant over countries.

Figure S1 summarizes the partial invariance model in terms of country-specific deviations on the item thresholds. To obtain the figure, we re-specified the MIMIC model with the factor means fixed at their estimated values from the partial invariance model, and estimated all of the item thresholds freely. The country-specific item thresholds are reported as deviations from the mean threshold estimated overall countries. The figure also indicates the items for which the 95% confidence intervals on threshold deviations the included zero in all countries. Based on examination of the figure, it was not apparent that the invariant items corresponded to any particular subtask or domain, or that any one country was disproportionately responsible for the dissimilarities.

*Table S1. Summary of Measurement Invariance Analyses, Emergent Literacy Domain*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Domain | χ2 (df) | RMSEA  (90% CI) | TLI | χ2 diff (df) | p-value |
| **All** |  |  |  |  |  |
| Configural | 1395.646 (1155) | .021 (.016, .025) | .995 |  |  |
| Scalar | 2205.019 (1351) | .036 (.034, .039) | .983 | 756.600 (196) | <.001 |
| **Age 4-5** |  |  |  |  |  |
| Configural | 1091.261 (924) | .020 (.015, .025) | .994 |  |  |
| Scalar | 1768.886 (1071) | .038 (.035, .041) | .979 | 612.830 (147) | <.001 |
| **ECCD** |  |  |  |  |  |
| Configural | 1255.149 (1155) | .019 (.008, .026) | .994 |  |  |
| Scalar | 1826.645 (1351) | .038 (.033, .042) | .977 | 569.181 (196) | <.001 |
| **ECCD and Age 4-5** |  |  |  |  |  |
| Configural | 1024.509 (924) | .021 (.011, .029) | .992 |  |  |
| Scalar | 1503.249 (1071) | .041 (.036, .045) | .986 | 484.417 (147) | <.001 |
| **Without Vietman** |  |  |  |  |  |
| Configural | 1165.886 (924) | .023 (.019, .027) | .994 |  |  |
| Scalar | 1722.081 (1071) | .036 (.032, .039) | .986 | 526.396 (147) | <.001 |

*Notes*. Χ2 (df) denotes the chi-square test of model fit and its degrees of freedom. RMSEA denotes the root mean square error of approximation and (90% CI) its 90% confidence interval. TLI denotes the Tucker Lewis Index. Χ2 –diff (df) denotes the chi-square difference test and is degrees of freedom, and its *p*-value is reported in the last column. Analyses that selected on Age did not include Ethiopia. ECCD denotes early child care and development programs.

*Table S2. Summary of Measurement Invariance Analyses, Emergent Numeracy Domain*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Domain | χ2 (df) | RMSEA  (90% CI) | TLI | χ2 diff (df) | p-value |
| **All** |  |  |  |  |  |
| Configural | 1866.641 (1501) | .023 (.019, .026) | .992 |  |  |
| Scalar | 2651.937 (1708) | .034 (.031, .037) | .982 | 768.544 (207) | <.001 |
| **Age 4-5** |  |  |  |  |  |
| Configural | 1550.514 (1201) | .025 (.022, .029) | .989 |  |  |
| Scalar | 2037.266 (1356) | .033 (.030, .036) | .982 | 516.871 (155) | <.001 |
| **ECCD** |  |  |  |  |  |
| Configural | 1654.198 (1510) | .020 (.012, .026) | .994 |  |  |
| Scalar | 2320.914 (1717) | .038 (.034, .042) | .979 | 663.218 (207) | <.001 |
| **ECCD and Age 4-5** |  |  |  |  |  |
| Configural | 1377.186 (1201) | .025 (.017, .031) | .989 |  |  |
| Scalar | 1800.899 (1356) | .037 (.032, .041) | .976 | 465.661 (155) | <.001 |
| **Without Vietman** |  |  |  |  |  |
| Configural | 1505.855 (1201) | .023 (.019, .027) | .983 |  |  |
| Scalar | 2125.998 (1356) | .034 (.032, .037) | .984 | 601.412 (155) | <.001 |

*Notes*. Χ2 (df) denotes the chi-square test of model fit and its degrees of freedom. RMSEA denotes the root mean square error of approximation and (90% CI) its 90% confidence interval. TLI denotes the Tucker Lewis Index. Χ2 –diff (df) denotes the chi-square difference test and is degrees of freedom, and its *p*-value is reported in the last column. Analyses that selected on Age did not include Ethiopia. ECCD denotes early child care and development programs.

*Table S3. Summary of Measurement Invariance Analyses, Motor Domain*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Domain | χ2 (df) | RMSEA  (90% CI) | TLI | χ2 diff (df) | p-value |
| **All** |  |  |  |  |  |
| Configural | 34.275 (26) | .026 (.000, .047) | .998 |  |  |
| Scalar | 1031.281 (78) | .160 (.151, .168) | .933 | 950.237 (52) | <.001 |
| **Age 4-5** |  |  |  |  |  |
| Configural | 26.674 (20) | .027 (.000, .052) | .998 |  |  |
| Scalar | 676.780 (59) | .152 (.142, .162) | .948 | 599.482 (39) | <.001 |
| **ECCD** |  |  |  |  |  |
| Configural | 31.442 (26) | .029 (.000, .061) | .998 |  |  |
| Scalar | 665.835 (78) | .175 (.163, .188) | .932 | 603.678 (52) | <.001 |
| **ECCD and Age 4-5** |  |  |  |  |  |
| Configural | 28.079 (20) | .041 (.000, .073) | .998 |  |  |
| Scalar | 477.900 (59) | .171 (.157, .185) | .933 | 409.679 (39) | <.001 |
| **Without Vietman** |  |  |  |  |  |
| Configural | 31.007 (20) | .032 (.000, .054) | .997 |  |  |
| Scalar | 931.811 (59) | .172 (.163, .182) | .911 | 834.643 (39) | <.001 |

*Notes*. Χ2 (df) denotes the chi-square test of model fit and its degrees of freedom. RMSEA denotes the root mean square error of approximation and (90% CI) its 90% confidence interval. TLI denotes the Tucker Lewis Index. Χ2 –diff (df) denotes the chi-square difference test and is degrees of freedom, and its *p*-value is reported in the last column. Analyses that selected on Age did not include Ethiopia. ECCD denotes early child care and development programs. The Sclar

*Table S4. Summary of Measurement Invariance Analyses, Social Emotional Domain*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Domain | χ2 (df) | RMSEA  (90% CI) | TLI | χ2 diff (df) | p-value |
| **All** |  |  |  |  |  |
| Configural | 426.211 (291) | .031 (.025, .037) | .985 |  |  |
| Scalar | 1038.175 (387) | .059 (.055, .064) | .947 | 577.408 (96) | <.001 |
| **Age 4 -5** |  |  |  |  |  |
| Configural | 297.867 (234) | .025 (.015, .033) | .990 |  |  |
| Scalar | 715.976 (306) | .054 (.049, .060) | .953 | 401.640 (72) | <.001 |
| **ECCD** |  |  |  |  |  |
| Configural | 318.907 (233) | .039 (.027, .049) | .975 |  |  |
| Scalar | 663.176 (321) | .066 (.059, .073) | .927 | 355.613 (88) | <.001 |
| **ECCD and Age 4-5** |  |  |  |  |  |
| Configural | 426.211 (291) | .031 (.025, .037) | .985 |  |  |
| Scalar | 1038.175 (387) | .059 (.055, .064) | .947 | 577.408 (72) | <.001 |
| **Without Vietman** |  |  |  |  |  |
| Configural | 322.816 (235) | .028 (.020, .035) | .989 |  |  |
| Scalar | 780.873 (307) | .057 (.052, .062) | .954 | 425.644 (72) | <.001 |

*Notes*. Χ2 (df) denotes the chi-square test of model fit and its degrees of freedom. RMSEA denotes the root mean square error of approximation and (90% CI) its 90% confidence interval. TLI denotes the Tucker Lewis Index. Χ2 –diff (df) denotes the chi-square difference test and is degrees of freedom, and its *p*-value is reported in the last column. Analyses that selected on Age did not include Ethiopia. ECCD denotes early child care and development programs.

Figure S1. Item thresholds estimated treating the country means in Partial Invariance model (see Figure 4 from main document) as fixed values. For vertical axis reports the estimated deviation of each country’s threshold from the overall threshold for that item (*δjc*). Error bars denote approximate 95% confidence intervals using a Normal reference distribution. Invariance is indicated as “yes” if the confidence intervals on the item threshold deviations includes zero for all countries.

Bay 1:Users:peterhalpin:Dropbox:IDELA Measurement Analysis:CD:figures:all_dif.pdf

**Mplus Code**

The following figures present example Mplus code illustrating the bi-factor model from Equation 1 of main document (Figure S2), as well as its configural (Figure S3) and scalar (Figure S4) specifications discussed in connection with Equation 2 of main document. The examples use the Early Literacy domain of the IDELA. Code for the other domains is similar. Code for the overall models (i.e., all four domains simultaneously) is a straightforward combination of the domain-level models.

Figure S2: Mplus code for the bi-factor model.

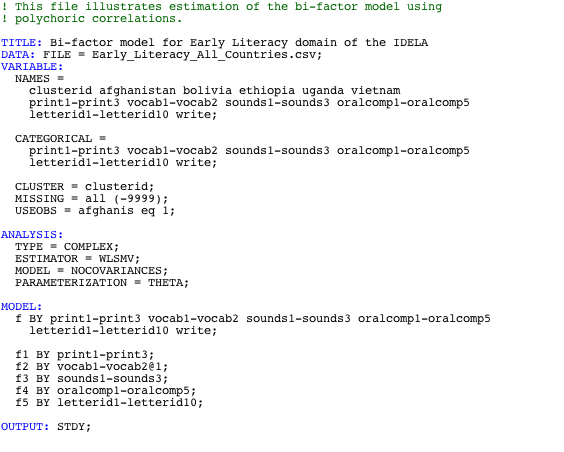


Figure S3. Mplus code for the configural invariance model.



Figure S4. Mplus code for the scalar invariance model.

