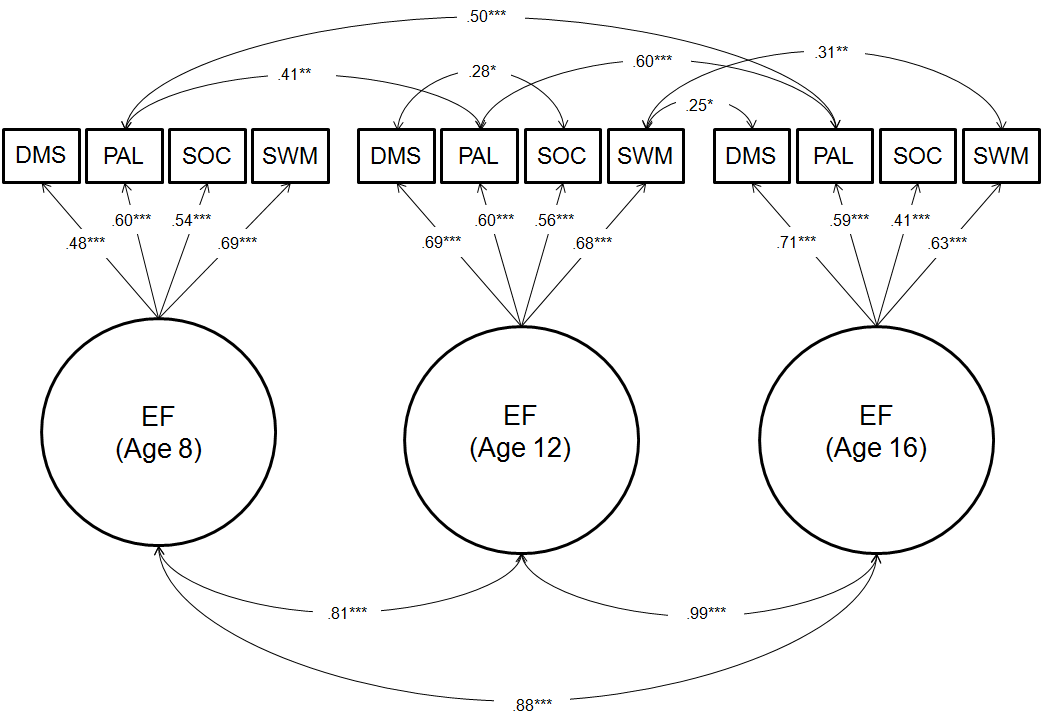
**Supplementary Materials**

**Methods**

**Measurement Model for EF Factors**



***Figure S1.*** Measurement model for the EF latent factors, estimated simultaneously at age 8, 12, and 16, with factors allowed to freely correlate. Loadings are standardized coefficients. Associations between measures are residual correlations. DMS – Delayed Match to Sample (percent correct over all delays); PAL – Paired Associates Learning (mean errors to success); SOC – Stockings of Cambridge (problems solved in the minimum number of moves); SWM – Spatial Working Memory (total errors). \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.

**Measurement Invariance of EF Factors**

We assessed factorial invariance for the latent EF factors over time (age 8, 12, and 16) in order to ensure they were roughly isomorphic. We did this by examining invariance at three levels: configural, metric (weak factorial), and scalar (strong factorial). Model fit for each of these steps is presented in Table S1. In the *configural model* (M1 in Table S1), we tested whether the same factor structure is present at each age by estimating the latent factors simultaneously, allowing the factors to freely co-vary over time. This model fit the data well, as reported in main text, with the pattern of loadings onto the EF factors being similar, thus supporting configural invariance. In the *metric model* (M2), we tested whether the factor loadings were equal across time. The model with all factor loadings constrained to equality over time fit the data relatively well (Table S1). However, a comparison of the alternative fit indices (CFI, RMSEA, and SRMR) across M1 and M2 suggested that the metric model (M2) fit slightly worse than the configural model (M1) based on field recommendations (Chen, 2007; Putnick and Bornstein, 2016). Modification indices suggested freeing the PAL loading on the EF factor at age 8. The resulting partial metric invariance model (M2a) did not fit worse than the configural model (M1), supporting partial metric invariance. Finally, we tested whether the intercepts of the scales were similar over time (*scalar invariance*; M3). Model fit was acceptable, and comparison of the scalar and partial metric models revealed that the scalar model (M3) did not fit worse than the metric mode (M2a), thus supporting scalar invariance. This pattern – that is, partial metric invariance and full scalar invariance – is permissible, as violations of metric invariance on mean-level analyses are minimal compared to scalar invariance (Steinmetz, 2013). Moreover, ignoring one non-invariant factor loading in the metric model has been shown not to bias parameter estimates (Guenole and Brown, 2014). These results suggest that the EF factors are largely invariant over time, thus enabling an examination of factor mean differences and construct relations over time (i.e., regression parameters), as reported in the main text.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S1:** Tests of measurement invariance for EF factors over time | | | | | | | |
| Model | χ2 (*df*) | CFI | RMSEA | SRMR | ΔCFI | ΔRMSEA | ΔSRMR |
| M1: Configural Invariance | 76.79 (45) | .94 | .061 | .057 | - | - | - |
| M2: Metric Invariance | 105.56 (53) | .90 | .073 | .098 | -.040 | .012 | .041 |
| M2a: Partial Metric Invariance | 88.78 (52) | .93 | .061 | .087 | -.010 | .000 | .030 |
| M3: Scalar Invariance | 93.64 (58) | .93 | .057 | .091 | .000 | -.004 | .004 |

**Measurement of Psychopathology Factors**

Estimation of the P, INT, and EXT factors described in the main text followed the procedure outlined in Wade *et al.* (2018). In short, teachers and caregivers provided ratings of psychopathology in eight domains (depression, overanxious, social anxiety, oppositional defiant, conduct problems, overt aggression, relational aggression, and ADHD) using the MacArthur Health and Behavior Questionnaire (HBQ; Essex *et al.*, 2002) at ages 8, 12, and 16 years. At age 8, reporters were the children’s teachers, and at age 12 and 16 both teachers and caregivers reported on problems, which were then combined into composite scores. Ratings were then subjected to latent bifactor models (Reise, 2012) in order to estimate general (P), internalizing-specific (INT), and externalizing-specific (EXT) factors at each age. In this model, all eight domains of psychopathoogy were set to load onto the latent general (P) psychopathology factor and their respective INT and EXT factors simultaneously. The latent factors were forced to orthogonality within-time. Measurement invariance on these factors was previously established and is reported in Wade *et al*. (2018). The factors were then extracted from this model and were used in the longitudinal path analysis reported in the main text.

**Results**

FigureS3_v2.tif

***Figure S2.*** Effect of placement stability on general psychopathology (A) and executive functioning (B) from age 8 to age 16. The y-axis is the factor score for each outcome, with a sample mean of zero. NIG – never-institutionalized group, FCG-S – foster care group with stable placements, FCG-U – foster care group with unstable placements, CAUG – care as usual group. At age 8, the NIG had lower P and higher EF compared to all other groups. By age 16, NIG continued to have higher EF than CAUG and FCG-U, but were not significantly different from FCG-S, who now outperformed both CAUG and FCG-U. At age 16, FCG-S also had lower P than CAUG, but continued to have higher P compared to NIG. \**p* < .05

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S2:** Bivariate correlations between placement variables and outcome variables | | | | | | | | | | | |
|  | Age 8 |  | Age 12 | | | |  | Age 16 | | | |
| EF |  | P | INT | EXT | EF |  | P | INT | EXT | EF |
| Age entered institution | .01 |  | -.15 | -.06 | -.10 | .03 |  | -.06 | -.14 | -.01 | .02 |
| Age entered foster care | -.06 |  | -.10 | -.23 | .13 | -.04 |  | .02 | -.01 | .01 | -.05 |
| % time in institution | -.05 |  | .23\* | .03 | .13 | -.09 |  | .39\*\*\* | .10 | .15 | -.16 |
| # placement disruptions | -.09 |  | .16 | .04 | .16 | -.11 |  | .34\*\*\* | -.01 | .15 | -.14 |
| \*\*\**p* < .001. \*\**p* < .01. \**p* < .05.  *Note.* These correlations are only among those who were ever raised in an institution | | | | | | | | | | | |

**References**

Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*. 2007; 14(3):464-504.

Essex MJ, Boyce WT, Goldstein LH, Armstrong JM, Kraemer HC, Kupfer DJ, MacArthur Assessment Battery Working Group. The confluence of mental, physical, social, and academic difficulties in middle childhood. II: Developing the MacArthur Health and Behavior Questionnaire. Journal of the American Academy of Child & Adolescent Psychiatry. 2002; 41(5):588-603.

Guenole N, Brown A. The consequences of ignoring measurement invariance for path coefficients in structural equation models. *Frontiers in Psychology*. 2014; 5:980.

Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review*. 2016; 41:71-90.

Reise SP. The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*. 2012; 47(5):667-96.

Steinmetz H. Analyzing observed composite differences across groups. *Methodology*. 2013; 9: 1-12.

Wade M, Fox NA, Zeanah CH, Nelson CA. Effect of foster care intervention on trajectories of general and specific psychopathology among children with histories of institutional rearing: a randomized clinical trial. *JAMA Psychiatry*. 2018; 75(11):1137-45.