S1 Text. Internal Standardization of Scores using Age-Conditional Means and SDs.

For each scale, we removed tester effects from the raw score by running a regression of the raw scores on tester dummies using Ordinary Least Squares (OLS). We constructed the residuals of these regressions, which we standardized by age using non-parametric methods as follows. First, we computed the age-conditional mean using the fitted values of the regression in (1), estimated by kernel-weighted local polynomial smoothing methods:

$$\mathrm{Y_i} = \mathrm{f}(X_i) + \varepsilon_\mathrm{i} \qquad \forall\, i \qquad (1)$$

where $\mathrm{Y_i}$ is the residual of the raw score of child i in a given scale of a regression on tester dummies. X_i is the age of the child in days. Next, we regressed the square of the residuals in (1) on age of the child (in days) as shown in the kernel-weighted local polynomial regression in (2):

$$(\mathrm{Y_i} - \hat{\mathrm{f}}_i)^2 = \mathrm{g}(X_i) + v_\mathrm{i} \qquad \forall\, i \qquad (2)$$

Our estimate of the age-conditional standard deviation (SD) is the square root of the fitted values $\hat{\mathrm{g}}_i$ in (2). Finally, we computed the internally age-adjusted z-score, $\mathrm{ZY_i}$, by subtracting from the residual of the raw score the within sample age-conditional mean estimated in (1) and dividing by the within sample age-conditional SD obtained from (2). More specifically:

$$\mathrm{ZY_i} = \frac{\mathrm{Y_i} - \hat{\mathrm{f}}_i}{\sqrt{\hat{\mathrm{g}}_i}} \qquad \forall\, i \qquad (3)$$

This resulted in smooth normally distributed internally standardized scores, with mean zero across the age range.