Supplementary Information

Ordered Probit Model

The full ordered probit model and priors are specified below with Interceptive Timing (IntT), age, steering, aiming, tracking and postural balance (eyes open and eyes closed) scores entered as predictors. The model was based on Kruschke (2015) and the model code is available online at https://github.com/OscartGiles/Ordered-Probit-Stan.

$$\boldsymbol{\beta} \sim N(0, K)$$

$$\boldsymbol{\mu} = \boldsymbol{X\beta}$$

$$\boldsymbol{C}_1 \equiv 1.5$$

$$\boldsymbol{C}_{t=2 \ldots K-1} \sim N(t + 0.5, K)$$

$$\boldsymbol{C}_{K-1} \equiv K - 0.5$$

$$\sigma \sim Cauchy^{+}(0, 100)$$

$$\boldsymbol{\theta}_{i,k} = \begin{cases} 1 - \phi\left(\frac{\mu_i - \boldsymbol{C}_1}{\sigma}\right), & k = 1 \\ \phi\left(\frac{\mu_i - \boldsymbol{C}_{k-1}}{\sigma'}\right) - \phi\left(\frac{\mu_i - \boldsymbol{C}_k}{\sigma}\right), & 1 < k < K \\ \phi\left(\frac{\mu_i - \boldsymbol{C}_{k-1}}{\sigma}\right), & k = K \end{cases}$$

$$\boldsymbol{y}_i \sim \textbf{Categorical}(\boldsymbol{\theta}_i)$$

Where N is the number of data points, K is number of levels in the attainment outcome, $i = 1 \ldots N$, $k = 1 \ldots K$, and $t = 1 \ldots K - 1$. $\boldsymbol{X}$ is an $N \times 7$ matrix of predictor variables where the first column is equal to 1. $\boldsymbol{\theta}$ is an $N \times K$ matrix, specifying the probabilities of obtaining each observed academic attainment score for the ith participant. ϕ is the cumulative normal function. $\boldsymbol{\mu}$ represents a continuous latent attainment outcome, and **y** is the observed attainment scores.

The first and last threshold value $\boldsymbol{C}_1$and $\boldsymbol{C}_{K-1}$ were fixed in order to identify the model. Thus all other model parameters must be interpreted with regards to this constraint. In addition, each threshold parameter was constrained to be greater than the last ($\boldsymbol{C}_k < \boldsymbol{C}_{k+1}$).

Effect size calculations

In the main text we provide an estimate of the effect size for each predictor in the model in terms of the equivalent change in age that would be required to produce the same change on the latent attainment score as the *typical range* of each of the sensorimotor measures (where the typical range was defined as 2 times the standard deviation of the motor measure of interest). The effect size can be formally defined as,

$$Equivilant\ age\ change = \frac{2 \times SD_j \times \beta_j}{\beta_{age}} \times 12$$

where SD_j is the estimated standard deviation for the jth sensorimotor measure (after controlling for age), β_j is the corresponding model coefficient and β_{age} is the coefficient for age. For clarity we illustrate this graphically in Figure S1 (see caption for details).

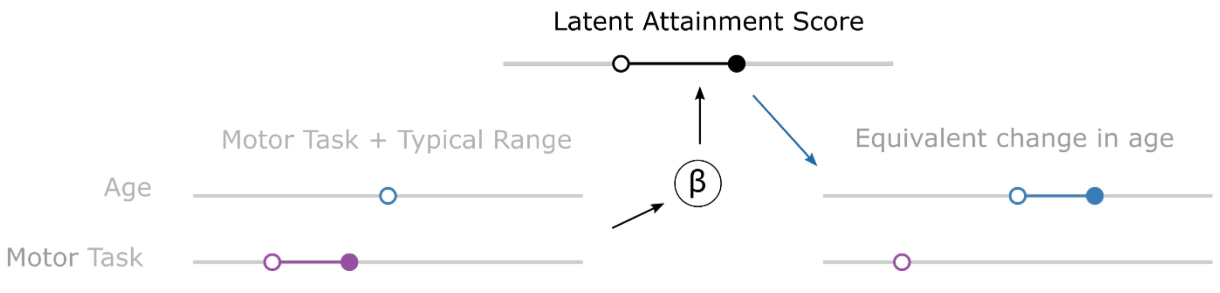

Figure S1: Illustration of how the effect size metric was calculated. The top line shows the latent Mathematics attainment score (μ_i) on a continuous scale. The model states that $\mu_i = X_i^T \beta$, where X is a design matrix specifying the predictor scores for each participant. As we change the values of the predictor variables, the predicted latent attainment score will change. Changing a motor task score by the *typical range* (left side; open to filled purple circle) results in a change in the predicted latent attainment score (open to filled black circle). Our effect size measure defines how much we would need to change the age predictor (right side; open to filled blue circle) in order to achieve the same change in the latent attainment score. In other words, how many months the typical range of the sensorimotor task predictor is worth.

Typical range of sensorimotor measures after controlling for age

We chose the *typical range* to be $2 \times SD$ as this is the difference between a score one SD above and below the mean. We therefore needed to estimate the SD for each motor task. However, we know that a substantial proportion in the variance in each motor task is explained by age. Thus we calculated the SD after controlling for age. For a single motor task we could calculate this by fitting a simple regression with age as a predictor and the motor task as the outcome variable. The SD then provides a measure of the variance not explained by age. Here we used a “seemingly unrelated regression” model which allowed for all the motor tasks to be modelled as output variables simultaneously. This is essentially the same as fitting multiple simple regressions between age and each motor task, except that the covariance between motor tasks is also estimated. The full model code is provided at https://github.com/OscartGiles/Hitting-the-target.

Understanding how the latent attainment score maps to the observed score

The latent attainment score is mapped to the observed data by a probit link function. For a given predicted latent attainment score (μ) the model provides a vector of probabilities for each possible ordered attainment outcome. For illustrative purposes, Figure S2a shows the probability distribution when $\mu = 5$, which we refer to here as μ_1 (orange bars) and when μ increases as a result of IntT increasing by the typical range, referred to as μ_2 (blue bars). We can see that in both cases an attainment score of 5 is most probable, but in the latter case higher scores have become more probable overall, while the probability of lower scores has decreased. Figure S2b shows the logarithm of the ratio between the two probability distributions shown in Figure S2a. Again, this shows that observed attainment scores above 5 are more probable when the latent attainment score is increased (positive values), while lower scores are less probable (negative values).

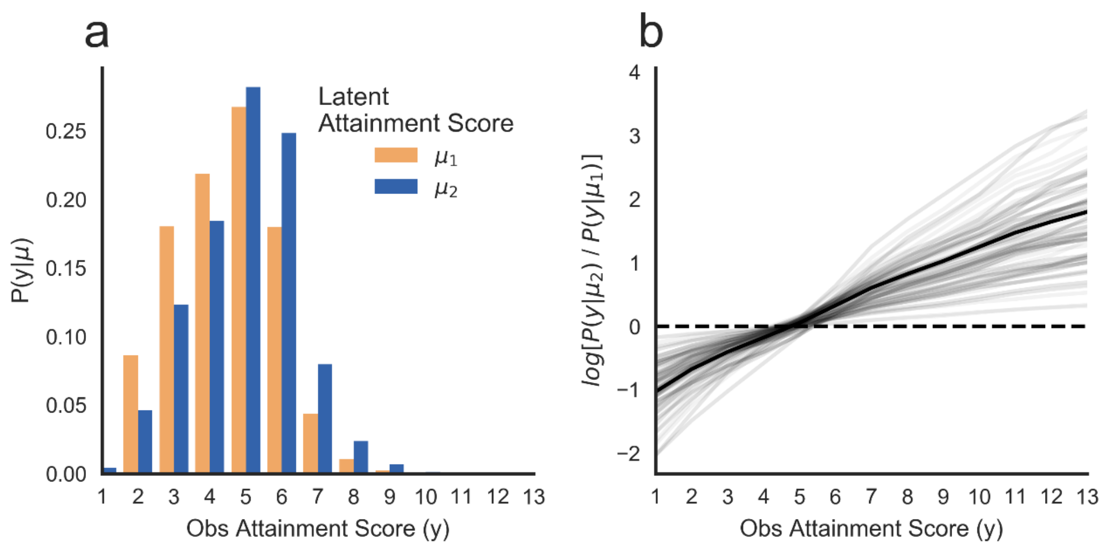

Figure S2: a) The probability of obtaining each possible observed Mathematics attainment outcome (y) when the latent Mathematics score is equal to 5 (μ_1; orange bars) and when the latent Mathematics score increases by the amount induced by the typical range of the interceptive timing metric (μ_2; blue bars). b) Log ratio of probability of each observed

Mathematics attainment score given μ_1 and μ_2. Dark line shows the posterior mean. Grey lines show 100 random samples from the posterior.

Graphical probes of model fit – Posterior predictive checks

To assess how well the model captures the data we simulated 16,000 data sets from the posterior (y_{rep}) and calculated the mean and standard deviation for each. The distribution of these test statistics are shown in Figure S3a and S3b respectively. The true mean and SD of the observed data is clearly plausible under the model simulations, suggesting this model captures these statistics well. We also calculated the mean score for each data point across all the expected score for each data point, $E(y_{rep})$. This is plotted again IntT in figure S4 (red dots) while the true Mathematics attainment scores are also plotted against IntT (blue dots). It's clear that the model captures the general pattern of observed relationship between interceptive timing and Mathematics attainment well.

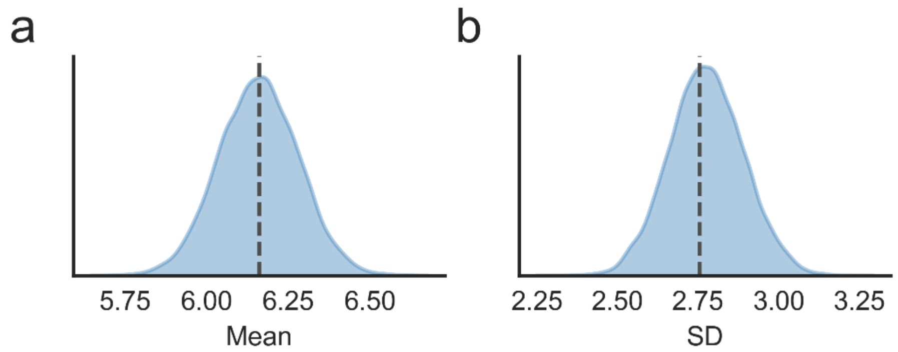

Figure S3: Distribution of the (a) mean and (b) standard deviation of test statistics for 16,000 simulated data sets (blue kernel density plots) alongside the true data sets (vertical black dashed line).

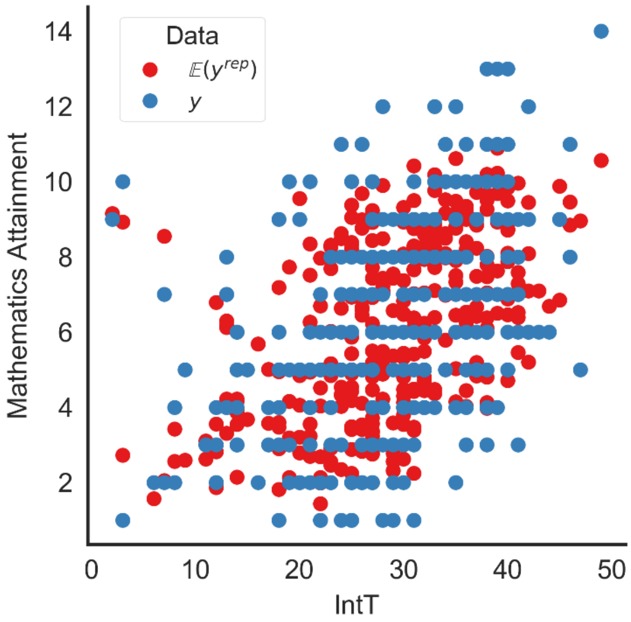

Figure S4: The expected value of the simulated data (y_{rep}) as a function of IntT score (blue dots). The observed data is also shown as a function of IntT score (red dots).

School Attainment Metrics:

Table S1 shows how the educational attainment code maps to the original code used by schools, as well as the school year and age at which children are expected to reach key attainment levels.

Table S1. Attainment score conversion table. A scale of 1 to K (where K was the highest observed score in the data) was used for the Bayesian Attainment Model. This scale maps to the UK nationally standardized scores. The school year and age at which children are expected to achieve these scores is shown.

Attainment Score	Government Code	Expected Year Group	Expected Age
1	1c		
2	1b		
3	1a		
4	2c		
5	2b	2	6-7
6	2a		
7	3c		
8	3b		
9	3a		
10	4c		
11	4b	6	10-11
12	4a		
13	5c		
14	5b	9	13-14
15	5a		

Table S2. In UK primary schools, mathematics is taught and assessed in two stages – Key stage 1 (years 1 and 2 when the children are 4-6 years) and Key stage 2 (years 3 to 6 when the children are 7-11 years). The table below is an extracted from:

https://www.gov.uk/government/collections/national-curriculum-assessments-test-frameworks

	Year	
Key Stage 1 The mathematics taught is very practical and related to everyday experiences. A variety of resources, such as coins, dice, dominoes, playing cards, beads and plastic bricks for counting.	1	number bonds, early skills for multiplication and solving simple problems; very practical mathematic related to everyday experiences.
	2	working on numbers through rehearsal and using addition and subtraction facts regularly; using number lines, tracks and 100 squares.
Key Stage 2 Shape, space, data handling, money and measures in addition to numeracy. Children are expected to read, write and order numbers on a number line (and place value cards, beads on a string etc).	3	puzzles, problems and investigations to practice, consolidate and extend understanding with an emphasis on real world situations.
	4	decimals (particularly with money and measurement); equivalent fractions introduced via diagrams and number lines used to teach fractions.
	5	Fractions, decimals and percentages; comparing, ordering and converting and solving problems in a meaningful context
	6	more complicated problems, including those that have decimals, fractions and percentages; expectation of working systematically, using the correct symbols and to check their results. They also learn about positive and negative numbers.