**SUPPLEMENTARY MATERIALS**

**School quality ratings are weak predictors of students’ achievement and wellbeing**

|  |  |
| --- | --- |
| **Supplementary Measures:** | **Page** |
| Measures of the school engagement | 2 |
| Measures of academic wellbeing | 3 |
|  |  |
| **Supplementary Tables:** | **Page** |
| **Table S1:** Representativeness of the current sample | 5 |
| **Table S2.** Sample sizes, means and standard deviations (SD) for Ofsted ratings | 6 |
| **Table S3.** Principal Component Analysis of Ofsted items | 8 |
| **Table S4.** Ofsted individual item loadings | 9 |
| **Table S5.** Analysis of variance with polynomial trend analysis and planned contrasts of GCSE scores between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate | 10 |
| **Table S6.** Results from multiple regression analysis predicting examination results at age 16 (GCSEs) from student covariates and Ofsted Headline Quality Rating. | 11 |
| **Table S7.** Analysis of covariance of GCSE scores between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate, accounting for covariates of prior achievement and socioeconomic status | 12 |
| **Table S8.** Analysis of variance with polynomial trend analysis of school engagement and wellbeing measures between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate | 13 |
|  |  |
| **Supplementary Figures:** | **Page** |
| **Figure S1.** Correlation matrix of Ofsted individual items | 17 |
| **Figure S2.** Scree plot | 19 |
| **Figure S3.** Spearman correlation coefficients (rs) for the relationship between measures of student wellbeing and school engagement and the Ofsted headline quality rating. | 20 |

# **SUPPLEMENTARY MEASURES**

At age 16, participants completed 8 questionnaires about school engagement and 6 questionnaires relating to academic wellbeing. A description of each of the questionnaires is included below. Measures were collected via web tests. All measures are self-report.

**Measures of school engagement**

**Teacher-student relations** – 6 items (Appleton, Christenson, & Reschly, 2006): This is a subscale of the Student Engagement Instrument and included items such as: “At my school, teachers care about students” and “My teachers are there for me when I need them” - rated on a 4-point scale from strongly disagree to strongly agree. The total score was created by taking the mean of the 6 items, requiring at least 3 to be present for an individual. The reported reliability of this subscale is good (alpha = .88).

**Control relevance of school work** – 4 items (Appleton, Christenson, & Reschly, 2006): This is a subscale of the Student Engagement Instrument and included items such as: “I feel like I have a say about what happens to me at school” and “When I do well in school, it’s because I work hard” - rated on a 4-point scale from strongly disagree to strongly agree. The total score was created by taking the mean of the 4 items, requiring at least 2 to be present for an individual. The reported reliability of this subscale is good (alpha = .80).

**Peer support for learning** – 3 items (Appleton, Christenson, & Reschly, 2006): This is a subscale of the Student Engagement Instrument and included items such as: “Students at my school respect what I have to say” and “Students at my school are there for me when I need them” rated on a 4-point scale from strongly disagree to strongly agree. The total score was created by taking the mean of the 3 items, requiring at least 2 to be present for an individual. The reported reliability of this subscale is good (alpha = .82).

**Family support for learning** – 3 items (Appleton, Christenson, & Reschly, 2006): This is a subscale of the Student Engagement Instrument and included items such as: “When something good happens at school, my family/carer(s) want to know about it.” and “My family/carer(s) want me to keep trying when things are tough at school.” rated on a 4-point scale from strongly disagree to strongly agree. The total score was created by taking the mean of the 3 items, requiring at least 2 to be present for an individual. The reported reliability of this subscale is good (alpha = .76).

**Homework behaviour** – 2 items (OECD Programme for International Student Assessment, 2000, 2003). These questions were taken from the PISA 2000, 2003 and 2006 student questionnaires. For homework behaviour, 2 items were selected: “I complete my homework on time” and “I do my homework while watching television” (reversed). These questions were rated on a 4-point scale from ‘never’ to ‘always’. A mean of these two items was taken as the total score for an individual, requiring both items to be present.

**Homework feedback** – 3 items (OECD Programme for International Student Assessment, 2000, 2003). These questions were taken from the PISA 2000, 2003 and 2006 student questionnaires. For homework feedback, 3 items were selected: “My teachers grade my homework”, “My teachers make useful comments on my homework” and “I am given interesting homework”. These questions were rated on a 4-point scale from ‘never’ to ‘always’. A mean of these three items was taken as the total score for an individual, requiring at least two items to be present (alpha = .74) .

**Attitudes to school** – 4 items (OECD Programme for International Student Assessment, 2000, 2003). These questions were taken from the PISA 2000, 2003 and 2006 student questionnaires. For this measure, four questions were asked relating to attitudes to the school, such as “School has done little to prepare me for adult life when I leave school” (reversed) and “School has taught me things which could be useful in a job”. These four questions were rated on a 4-point scale from strongly disagree to strongly agree. A mean of these items was taken as the total score for an individual, requiring at least 2 items to be present (alpha = .70).

**Peer victimisation** – 6 items (Mynard & Joseph, 2000). These questions were taken from the Multidimensional Peer-Victimization Scale which measures physical and verbal victimisation as well as social manipulation and attacks on property. Participants were asked to indicate whether and how often another student had victimised them, for example “How often during this school year has another student made fun of me for some reason” or “Hurt me physically in some way?” The response options were: ‘not at all’, ‘once’ or ‘more than once’. A mean of the items was used as a total score requiring at least half of the items to be present (alpha = .81).

**Academic wellbeing**

**Academic self-concept** – 10 items (Burden, 1998): These questions were taken from the ‘Myself-As-Learner Scale’ which was developed to measure academic self-concept in secondary-school aged learners. Participants were required to indicate the extent to which a series of statements describe them. These statements included things like “When I get stuck with my work I can usually work out what to do next” and “When I’m given new work to do, I usually feel confident I can do it”. There was a 5-point rating scale from ‘Very much like me’ to ‘Not like me at all’. A mean of the 10 items was taken as a total score requiring at least half to be present for each individual (alpha = .72).

**Future aspirations and goals** – 3 items (Appleton, Christenson, & Reschly, 2006): This is a subscale of the Student Engagement Instrument and included the following items: “I plan to continue my education following school”, “School is important for achieving my future goals” and “I am hopeful about my future.” rated on a 4-point scale from strongly disagree to strongly agree. The total score was created by taking the mean of the 6 items, requiring at least 3 to be present for an individual. The reported reliability of this subscale is good (alpha = .78).

**Life satisfaction in relation to school** – 4 items (Huebner, 1994). This is a subscale of the Multidimensional Students’ Life Satisfaction Scale. It included items tapping into life satisfaction, with a focus on the school environment such as “I like being in school” and “I enjoy school activities”. Participants were asked to indicate to what extent they agreed with these statements using a 6 point scale from ‘Strongly agree’ to ‘Strongly disagree’. The total score was created by taking a mean of the items, requiring at least half to be present (alpha = .82).

**Subjective happiness** – 4 items (Lyubomirsky & Lepper, 1999).This measure requires students to rate themselves on a 7 point scale for statements such as “In general, I consider myself to be”… (1) ‘not a very happy person’ to (7) ‘a very happy person’ or “Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. To what extent does this describe you?” From (1) ‘not at all’ to (7) ‘a great deal’. The total score was created by taking a mean of the four items, requiring at least half to be present. (alpha = .76)

**Grit** – 9 items (Duckworth & Quinn, 2009): This measure required participants to rate statements such as ‘I am driven to succeed’ on a 5-point scale from ‘very much like me’ to ‘not like me at all’. The total score was created by taking the mean of the 9 items, requiring at least 5 to be present (alpha = .91).

**Ambition** – 5 items (Duckworth & Quinn, 2009): This measure required participants to rate statements such as ‘I aim to be the best in the world at what I do’ and ‘I am ambitious’ on a 5-point scale from ‘very much like me’ to ‘not like me at all’. The total score was created by taking the mean of the 5 items, requiring at least 3 to be present (alpha = .79).

# **SUPPLEMENTARY TABLES**

**Table S1:** Representativeness of the current sample

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Achievement** | UK population | Current sample |
| 5 + GCSEs A\* - C grade | 75% | 81% |
|  |  |  |
| **Socioeconomic variables** |   |   |
| Mother employed | 49% | 48% |
| Father employed | 89% | 93% |

*Note:* We used the 2001 UK Census data for socioeconomic variables as this was taken at the time our variables were collected: <https://www.ons.gov.uk/census/2001censusandearlier/aboutcensus2001>; We used the 2011 GCSE statistics as this within the period the current sample took their GCSEs: <https://www.gov.uk/government/statistics/gcse-and-equivalent-attainment-by-pupil-characteristics-in-england-2010-to-2011>

|  |  |  |  |
| --- | --- | --- | --- |
| **Table S2.** Sample sizes, means and standard deviations (SD) for Ofsted ratings | **Total *N*** | Mean | *SD* |
| **Overall headline school quality measure:** |   |   |   |
| Overall effectiveness: how good is the school? | 4391 | 2.97 | 0.82 |
| **Individual ratings:** |   |   |   |
| The effectiveness of partnerships in promoting learning and well-being | 2903 | 3.29 | .657 |
| The schools capacity for sustained improvement | 1722 | 2.99 | .742 |
| Outcomes for individuals and groups of pupils/children | 3282 | 2.79 | .797 |
| Pupils achievement and the extent to which they enjoy their learning | 2903 | 3.00 | .897 |
| Pupils attainment | 2897 | 3.09 | .669 |
| The quality of pupils learning and their progress | 2895 | 3.12 | .683 |
| The quality of learning for pupils with special educational needs and/or disabilities and their progress | 2903 | 3.37 | .636 |
| The extent of pupils spiritual, moral, social and cultural development | 2903 | 3.32 | .610 |
| The extent to which pupils adopt healthy lifestyles | 2903 | 3.51 | .563 |
| The extent to which pupils feel safe | 2897 | 3.09 | .786 |
| Pupils attendance | 4385 | 3.15 | .693 |
| Pupils behaviour | 2903 | 3.52 | .604 |
| The extent to which pupils contribute to the school and wider community | 2903 | 3.27 | .707 |
| The extent to which pupils develop workplace and other skills that will contribute to their future economic well-being | 4391 | 2.86 | .702 |
| The quality of teaching | 2903 | 3.34 | .612 |
| The extent to which the curriculum meets pupils needs, including, where relevant, through partnerships | 2903 | 3.56 | .581 |
| The effectiveness of care, guidance and support | 1741 | 2.64 | .610 |
| The use of assessment to support learning | 4391 | 3.10 | .756 |
| The effectiveness of leadership and management in embedding ambition and driving improvement | 2897 | 3.30 | .675 |
| The effectiveness with which the school promotes equality of opportunity and tackles discrimination | 2389 | 3.06 | .695 |
| The effectiveness with which the school promotes community cohesion | 2897 | 3.20 | .737 |
| The effectiveness with which the school deploys resources to achieve value for money | 2897 | 3.19 | .672 |
| The effectiveness of the governing body in challenging and supporting the school so that weaknesses are tackled decisively | 1794 | 3.15 | .600 |
| The effectiveness of safeguarding procedures | 1722 | 3.04 | .604 |
| The effectiveness of the schools engagement with parents and carers | 1722 | 2.98 | .661 |
| The leadership and management of teaching and learning | 1722 | 2.98 | 0.66 |

**Table S3.** Principal Component Analysis of Ofsted items

|  |  |
| --- | --- |
|  | Eigenvalues |
| **Component** | **Total** | **% of Variance** | **Cumulative %** |
| 1 | 16.054 | 59.461 | 59.461 |
| 2 | 1.378 | 5.103 | 64.563 |
| 3 | 1.206 | 4.468 | 69.032 |
| 4 | .830 | 3.074 | 72.106 |
| 5 | .697 | 2.583 | 74.689 |
| 6 | .630 | 2.332 | 77.021 |
| 7 | .606 | 2.246 | 79.267 |
| 8 | .581 | 2.153 | 81.420 |
| 9 | .464 | 1.718 | 83.138 |
| 10 | .449 | 1.664 | 84.801 |
| 11 | .442 | 1.636 | 86.438 |
| 12 | .420 | 1.557 | 87.994 |
| 13 | .372 | 1.378 | 89.372 |
| 14 | .368 | 1.363 | 90.735 |
| 15 | .349 | 1.292 | 92.027 |
| 16 | .329 | 1.220 | 93.246 |
| 17 | .304 | 1.125 | 94.372 |
| 18 | .267 | .987 | 95.359 |
| 19 | .244 | .904 | 96.263 |
| 20 | .236 | .873 | 97.136 |
| 21 | .216 | .800 | 97.937 |
| 22 | .167 | .618 | 98.555 |
| 23 | .147 | .546 | 99.101 |
| 24 | .086 | .318 | 99.419 |
| 25 | .076 | .280 | 99.699 |
| 26 | .056 | .209 | 99.908 |
| 27 | .025 | .092 | 100.000 |

**Table S4.** Ofsted individual item loadings

|  |  |
| --- | --- |
| **Ofsted items** | **Component** |
| **1** | **2** | **3** |
| The effectiveness of partnerships in promoting learning and well-being  | .76 |  |  |
| The school’s capacity for sustained improvement  | .84 |  |  |
| Outcomes for individuals and groups of pupils/children  | .92 |  |  |
| Pupils achievement and the extent to which they enjoy their learning  | .89 |  |  |
| Pupils attainment  | .72 | -.49 |  |
| The quality of pupils learning and their progress  | .88 |  |  |
| The quality of learning for pupils with special educational needs and/or disabilities and their progress  | .80 |  |  |
| The extent of pupils spiritual, moral, social and cultural development  | .72 |  | .35 |
| The extent to which pupils adopt healthy lifestyles  | .63 |  | .31 |
| The extent to which pupils feel safe  | .76 |  |  |
| Pupils attendance  | .56 | -.49 |  |
| Pupils behaviour  | .79 |  |  |
| The extent to which pupils contribute to the school and wider community  | .73 |  | .36 |
| The extent to which pupils develop workplace and other skills that will contribute to their future economic well-being  | .84 |  |  |
| The quality of teaching  | .85 |  |  |
| The extent to which the curriculum meets pupils needs, including, where relevant, through partnerships  | .81 |  |  |
| The effectiveness of care, guidance and support  | .77 |  |  |
| The use of assessment to support learning  | .74 |  |  |
| The effectiveness of leadership and management in embedding ambition and driving improvement  | .88 |  |  |
| The effectiveness with which the school promotes equality of opportunity and tackles discrimination  | .83 |  |  |
| The effectiveness with which the school promotes community cohesion  | .62 |  |  |
| The effectiveness with which the school deploys resources to achieve value for money  | .92 |  |  |
| The effectiveness of the governing body in challenging and supporting the school so that weaknesses are tackled decisively...  | .70 | .37 |  |
| The effectiveness of safeguarding procedures  | .47 | .54 |  |
| The effectiveness of the school’s engagement with parents and carers  | .68 |  |  |
| The leadership and management of teaching and learning  | .82 |  |  |
|  |

**Table S5.** Analysis of variance with polynomial trend analysis and planned contrasts of GCSE scores between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate

|  |  |
| --- | --- |
| **ANOVA** |  |
|   | *SS* | *df* | *F* | *p* | *ɳ²* |
| *Between groups* |   |   |   |   |  |
| Combined | 296.37 | 3 | 67.47 | 1.22 x 10-42 |  |
| Linear | 295.73 | 1 | 201.96 | 7.68 x 10-45 |  |
| Quadratic | 0.55 | 1 | 0.38 | 0.54 |  |
| Cubic | 0.09 | 1 | 0.06 | 0.81 |  |
| *Within groups* | 6394.53 | 4367 |   |   |  |
| Total | 6690.9 | 4370 |   |   | 0.44 |
| **Planned contrasts** |  |
|  | Inadequate | Requires Improvement | Good | Outstanding  |  |
| **Inadequate***M*=8.17 (*SD* = 1.24) |   | 3.10\* | 6.49\*\* | 9.93\*\* |  |
| **Requires Improvement***M*=8.47 (*SD* = 1.23) | 0.30 |   | 6.35\*\* | 12.31\*\* |  |
| **Good***M*=8.77 (*SD* = 1.21) | 0.60 | 0.30 |   | 7.78\*\* |  |
| **Outstanding** *M*=9.11 (*SD* = 1.24) | 0.94 | 0.64 | 0.34 |   |  |

**Note:** *SS* = Sum of squares; *df* = degrees of freedom; *F* = test of overall ANOVA model; *p* = significance of F statistic; *ɳ²* = eta squared variance explained; Planned contrasts: *M* = mean; *SD* = standard deviation; the lower diagonal matrix shows the mean GCSE grade differences between Ofsted categories. The upper diagonal matrix presents the *t* static associated with the difference. \*\* = *p* <.001 \* = *p* <.05.

**Table S6.** Results from multiple regression analysis predicting examination results at age 16 (GCSEs) from student covariates and Ofsted Headline Quality Rating.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| **Predictors** | *B* | *SE* | *β* | *t* | *p* | *r* | *sr* |
| Socioeconomic status | .197 (.170-.224) | .014 | .156 | 14.298 | 1.87 x 10-44 | .444 | .143 |
| KS2 English score | .331 (.294-.368) | .019 | .262 | 17.236 | 5.25 x 10-64  | .713 | .173 |
| KS2 Mathematics score | .462 (.422-.502) | .020 | .376 | 22.647 | 9.41 x10105 | .759 | .228 |
| KS2 Science score | .232 (.189-.275) | .022 | .186 | 10.563 | 7.98 x 10-26 | .730 | .106 |
| Headline Quality Rating | .127 (.097-.157) | .015 | .085 | 8.234 | 2.80 x 10-16 | .211 | .083 |
| **Full model statistics:** |  |  | *F*(5, 3007) = 1379.093*p* = <.000001*R2* = .696 |  |  |

***Note***. Beta coefficients, standard errors and t statistics, p-values and correlations are presented for each of the predictors in the multiple regression model. *β* = standardised Beta coefficient; *SE* = standard error; t = unstandardized beta coefficient divided by the *SE*, *p* = significance of result; *r* = Pearson correlation between predictor and GCSE; *sr* = semi-partial correlation - unique prediction of predictor on GCSE corrected for other predictors in the model; *F* = test of overall ANOVA model; *R2*= variance explained by all the predictors in the model.

**Table S7.** Analysis of covariance of GCSE scores between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate, accounting for covariates of prior achievement and socioeconomic status

|  |  |
| --- | --- |
| **ANOVA** |  |
|   | *SS* | *df* | *F* | *p* | *ɳ²* |
| KS2 English | 131.22 | 1 | 297.29 | 1.35 x 10-63 | 0.03 |
| KS2 maths | 226.39 | 1 | 512.93 | 5.79 x 10-105 | 0.05 |
| KS2 Science | 49.22 | 1 | 111.51 | 1.28 x 10-25 | 0.01 |
| Socioeconomic status | 88.56 | 1 | 200.65 | 3.87 x 10-44 | 0.02 |
| Ofsted-rating | 30.22 | 3 | 22.82 | 1.33 x 10-14 | <.01 |
| *Error* | 1323.21 | 2998 |   |   |  |
| Total | 4355.16 | 3005 |   |   |  |
| **Pairwise comparisons** |
|  | Inadequate | Requires Improvement | Good | Outstanding  |
| **Inadequate***M*=8.55 (*SE* = 0.06) |   | 0.07 | 1.60 x 10-5 | 2.91 x 10-9 |
| **Requires Improvement***M*=8.72 (*SE* = 0.02) | 0.17 |   | 1.97 x 10-4 | 6.40 x 10-11 |
| **Good***M*=8.85 (*SE* = 0.02) | 0.30 | 0.13 |   | .001 |
| **Outstanding** *M*=8.96 (*SE* = 0.02) | 0.41 | 0.24 | 0.11 |   |

**Note:** *SS* = Sum of squares; *df* = degrees of freedom; *F* = test of overall model; *p* = significance of F statistic; *ɳ²* = eta squared variance explained; Pairwise comparisons: *M* = mean; *SE* = standard error; the lower diagonal matrix shows the mean GCSE grade differences between Ofsted categories, once accounting for student covariates. The upper diagonal matrix presents the significance of the results. Only the difference between Inadequate and Requires Improvement was not significant.

**Table S8.** Analysis of variance with polynomial trend analysis of school engagement and wellbeing measures between students attending schools rated as: Outstanding, Good, Requires Improvement and Inadequate

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   | *SS* | *df* | *F* | *p* | *ɳ²*  |
| Teacher-Student Relations | *Between Groups* | 3.92 | 3 | 2.42 | 0.06 |  |
| Linear | 3.90 | 1 | 7.21 | 0.01 |  |
| Quadratic | 0.01 | 1 | 0.01 | 0.91 |  |
| Cubic | 0.01 | 1 | 0.02 | 0.89 |  |
| *Within Groups* | 974.95 | 1802 |  |  |  |
| **Total** | 978.87 | 1805 |  |  | 0.00 |
| Control/Relevance of School Work | *Between Groups* | 2.12 | 3 | 1.55 | 0.20 |  |
| Linear | 1.61 | 1 | 3.54 | 0.06 |  |
| Quadratic | 0.35 | 1 | 0.76 | 0.38 |  |
| Cubic | 0.16 | 1 | 0.36 | 0.55 |  |
| *Within Groups* | 819.61 | 1802 |  |  |  |
| **Total** | 821.73 | 1805 |  |  | 0.00 |
| Peer Support for Learning | *Between Groups* | 3.88 | 3 | 2.25 | 0.08 |  |
| Linear | 3.16 | 1 | 5.50 | 0.02 |  |
| Quadratic | 0.68 | 1 | 1.18 | 0.28 |  |
| Cubic | 0.05 | 1 | 0.08 | 0.78 |  |
| *Within Groups* | 1031.82 | 1797 |  |  |  |
| **Total** | 1035.70 | 1800 |  |  | 0.00 |
| Homework Behaviour scale | *Between Groups* | 18.08 | 3 | 5.10 | 0.00 |  |
| Linear | 8.91 | 1 | 7.53 | 0.01 |  |
| Quadratic | 5.13 | 1 | 4.34 | 0.04 |  |
| Cubic | 4.04 | 1 | 3.41 | 0.06 |  |
| *Within Groups* | 2139.09 | 1809 |  |  |  |
| **Total** | 2157.16 | 1812 |  |  | 0.01 |
| Homework Feedback scale | *Between Groups* | 20.91 | 3 | 2.23 | 0.08 |  |
| Linear | 19.21 | 1 | 6.14 | 0.01 |  |
| Quadratic | 1.70 | 1 | 0.55 | 0.46 |  |
| Cubic | 0.00 | 1 | 0.00 | 1.00 |  |
| *Within Groups* | 5633.07 | 1801 |  |  |  |
| **Total** | 5653.98 | 1804 |  |  | 0.00 |
| Attitudes to School  | *Between Groups* | 2.23 | 3 | 2.25 | 0.08 |  |
| Linear | 1.91 | 1 | 5.81 | 0.02 |  |
| Quadratic | 0.17 | 1 | 0.52 | 0.47 |  |
| Cubic | 0.14 | 1 | 0.43 | 0.51 |  |
| *Within Groups* | 596.70 | 1810 |  |  |  |
| **Total** | 598.92 | 1813 |  |  | 0.00 |
| Family Support for Learning | *Between Groups* | 3.93 | 3 | 1.42 | 0.24 |  |
| Linear | 1.41 | 1 | 1.53 | 0.22 |  |
| Quadratic | 0.01 | 1 | 0.01 | 0.91 |  |
| Cubic | 2.51 | 1 | 2.71 | 0.10 |  |
| *Within Groups* | 1659.09 | 1797 |  |  |  |
| **Total** | 1663.02 | 1800 |  |  | 0.00 |
| Peer Victimisation | *Between Groups* | 27.18 | 3 | 0.86 | 0.46 |  |
| Linear | 5.31 | 1 | 0.50 | 0.48 |  |
| Quadratic | 6.19 | 1 | 0.59 | 0.44 |  |
| Cubic | 15.68 | 1 | 1.49 | 0.22 |  |
| *Within Groups* | 18924.74 | 1793 |  |  |  |
| **Total** | 18951.92 | 1796 |  |  | 0.00 |
| Academic Self-Concept  | *Between Groups* | 0.89 | 3 | 0.79 | 0.50 |  |
| Linear | 0.44 | 1 | 1.18 | 0.28 |  |
| Quadratic | 0.30 | 1 | 0.79 | 0.37 |  |
| Cubic | 0.15 | 1 | 0.40 | 0.53 |  |
| *Within Groups* | 635.54 | 1692 |  |  |  |
| **Total** | 636.43 | 1695 |  |  | 0.00 |
| Future Aspirations/Goals  | *Between Groups* | 3.19 | 3 | 1.13 | 0.33 |  |
| Linear | 2.72 | 1 | 2.90 | 0.09 |  |
| Quadratic | 0.04 | 1 | 0.04 | 0.84 |  |
| Cubic | 0.44 | 1 | 0.46 | 0.50 |  |
| *Within Groups* | 1683.47 | 1798 |  |  |  |
| **Total** | 1686.66 | 1801 |  |  | 0.00 |
| Life Satisfaction School  | *Between Groups* | 6.42 | 3 | 2.18 | 0.09 |  |
| Linear | 5.91 | 1 | 6.02 | 0.01 |  |
| Quadratic | 0.02 | 1 | 0.02 | 0.89 |  |
| Cubic | 0.49 | 1 | 0.50 | 0.48 |  |
| *Within Groups* | 1850.94 | 1884 |  |  |  |
| **Total** | 1857.36 | 1887 |  |  | 0.00 |
| Subjective happiness | *Between Groups* | 2.65 | 3 | 0.64 | 0.59 |  |
| Linear | 2.39 | 1 | 1.73 | 0.19 |  |
| Quadratic | 0.25 | 1 | 0.18 | 0.67 |  |
| Cubic | 0.01 | 1 | 0.01 | 0.92 |  |
| *Within Groups* | 2593.72 | 1886 |  |  |  |
| **Total** | 2596.36 | 1889 |  |  | 0.00 |
| GRIT | *Between Groups* | 0.18 | 3 | 0.18 | 0.91 |  |
| Linear | 0.00 | 1 | 0.00 | 0.98 |  |
| Quadratic | 0.05 | 1 | 0.15 | 0.70 |  |
| Cubic | 0.13 | 1 | 0.39 | 0.53 |  |
| *Within Groups* | 577.00 | 1712 |  |  |  |
| **Total** | 577.18 | 1715 |  |  | 0.00 |
| Ambition | *Between Groups* | 2.01 | 3 | 1.48 | 0.22 |  |
| Linear | 1.65 | 1 | 3.65 | 0.06 |  |
| Quadratic | 0.06 | 1 | 0.12 | 0.73 |  |
| Cubic | 0.30 | 1 | 0.67 | 0.41 |  |
| *Within Groups* | 765.86 | 1690 |  |  |  |
| **Total** | 767.87 | 1693 |  |  | 0.00 |

**Note:** *SS* = Sum of squares; *df* = degrees of freedom; *F* = test of overall ANOVA model; *p* = significance of F statistic; *ɳ²* = eta squared variance explained.

# **SUPPLEMENTARY FIGURES**

#

**Figure S1.** Correlation matrix of Ofsted individual items. **Part A** shows the intercorrelations among the 23 individual items. **Part B** shows how these individual items correlate with the Ofsted extracted component and the Ofsted Headline quality rating.

 

**Figure S2.** Scree plot showing eigenvalues for each principal component after performing PCA on individual Ofsted items. The dashed line represents the cut-off for principal component retention based on the Kaiser’s λ > 1 criterion.

**Figure S3.** Spearman correlation coefficients (*rs*) for the relationship between measures of student wellbeing and engagement and the Ofsted headline quality rating. The Ofsted measure was rated from 1 (inadequate) to 4 (outstanding). The total score for each of the student quality and engagement measures are in brackets. Details of how they are measured are in the Supplementary Measures section. After correcting for multiple testing (0.05 / number of correlations: 12 = .0035), only Homework behaviour was significant.