



**ImpactEd**  
Evaluation

# Changes in Behavioural Measures (24/25)

Commissioned by The Tutor Trust



## Executive Summary

On average, participants' school attendance decreased, but this was broadly in line with national benchmarks. Participants whose endline attendance was *after* receiving tutoring saw a smaller decrease in attendance rate than those whose endline attendance was *during* receiving tutoring, suggesting that Tutor Trust programme may have greater impact over the longer term. Pupil Premium (PP) pupils were more likely to experience a positive change in their attendance rate than their non- PP participating peers. On average, participants' behaviour improved. When Year 11s are removed from the sample, participants' behaviour remained stable. Generally, more disadvantaged pupils saw a greater improvement in their behaviour than their more advantaged peers. The rate of exclusions in participants increased, and disadvantaged participants saw a greater increase in their rate of exclusions than their more advantaged participating peers.

## Introduction

The Tutor Trust is an award-winning Northern charity that transforms lives through high-quality tutoring in English, maths, and science for children and young people from lower-income communities across Greater Manchester, Merseyside, Tees Valley and West Yorkshire. Operating with a mission of achieving educational equity, they deliver evidence-based, in-school tutoring interventions that help disadvantaged students reach their full potential regardless of postcode or background, while offering flexible employment opportunities for high-quality tutors who share their commitment to bridging the gap between privilege and opportunity.

In 2022/23, The Tutor Trust and ImpactEd Evaluation worked together to understand the impact of the TP (Tutoring Plus) and TRA (The Right Angle) projects' impact on attainment, social and emotional outcomes, and explore whether the combined tutoring-counselling model is more impactful on attainment than a single-stranded approach.

In this 2024/25 project, the two organisations have collaborated again to collect and analyse attendance, behaviour and exclusions data on pupils participating in the Tutor Trust's mainstream tutoring programme, comparing outcomes before, during and after programme participation. This report focuses on answering the following research questions:

- ▶ What change in attendance did pupils participating in the Tutor programme experience?
- ▶ What change in behaviour points did pupils participating in the Tutor programme experience?
- ▶ What change in exclusions did pupils participating in the Tutor programme experience?
- ▶ Did any subgroups of pupils benefit more from participating in the Tutor Trust programme?

This report focuses on pupils in Y7 – Y11 completing a Tutor Trust programme in 2024/25 in one of the following schools: Co-op Academy Leeds, King's Leadership Academy, Liverpool, Manchester Enterprise Academy Central, Mount St Mary's Catholic High School, Oasis Academy MediaCityUK, Oasis Academy Oldham and Ormiston Chadwick Academy.

## Methodology

This section outlines the methodological approach taken to collect and analyse data.

### Research Design

This research design balanced two key objectives: maximising sample size for data confidence and enabling before/during/after programme comparisons. Due to varying intervention completion times, baseline and endline data collection windows differed across participants.

#### Dividing Pupils into Cohorts

Pupils took part in their specific Tutor Trust programme received at different times across schools. To standardise evaluation windows, participants were divided into two cohorts based on when they received most of their tutoring: Spring term and Summer term. Pupils who completed all tutoring sessions by 22nd May 2025 were assigned to the Spring cohort; those whose majority of sessions occurred after this date formed the Summer cohort.

#### Baseline and Endline Windows

Initially, we planned to collect behavioural data one term before and one term after pupils received most of the programme. However, since many pupils received the programme in Summer Term, the term after would fall in Autumn 2025, which had not yet occurred.

We therefore maintained baseline collection one term before delivery, but adjusted endline timing: Spring cohort endline data were collected the term after delivery (Summer Term), whilst Summer cohort endline data were collected during the term of delivery (Summer Term). This approach means that the endline data captures changes that occur both during and after programme delivery, so analysing results by delivery term will produce different insights. The window dates are summarised below.

Delivery Term	Spring Cohort	
Window	Baseline	Endline
Window relative to delivery window	Term <b>before</b> most of the delivery	Term <b>after</b> most of the delivery

Academic Term	Autumn Term	Summer Term
Term Dates	02/09/2024 – 20/12/2024	28/04/2025 – 20/07/2025
Measures	Attendance, Behaviours, Exclusions	Attendance, Behaviours, Exclusions

Delivery Term	Summer Cohort	
Window	Baseline	Endline
Window relative to delivery window	Term <b>before</b> most of the delivery	Term <b>during</b> most of the delivery
Academic Term	Spring Term	Summer Term
Term Dates	01/01/2025 – 04/04/2025	28/04/2025 – 20/07/2025
Measures	Attendance, Behaviours, Exclusions	Attendance, Behaviours, Exclusions

## Data Collection, Timeline, Sample and Analysis

Behavioural data (attendance, behaviour points and exclusions) was collected through schools' Management Information Systems (MIS), and then pulled through to ImpactEd School Impact Platform (SIP). The sample size for behaviour data is smaller than for attendance and exclusion data because three schools were excluded from the analysis: two schools were excluded because their school record-keeping systems do not collect behaviour points data in a way that is captured by their MIS, and one school was excluded because there was insufficient sample size for standardisation. Data was collected on the following sample size at the following time points by delivery term.

	Baseline			Endline		
	Spring Delivery - Autumn Term	Summer Delivery - Spring Term	All	Spring Delivery - Summer Term	Summer Delivery - Summer Delivery	All
<b>Attendance</b>	<b>55</b>	<b>295</b>	<b>350</b>	<b>55</b>	<b>295</b>	<b>350</b>
<b>Behaviour</b>	<b>55</b>	<b>256</b>	<b>311</b>	<b>55</b>	<b>256</b>	<b>311</b>
<b>Exclusions</b>	<b>55</b>	<b>295</b>	<b>350</b>	<b>55</b>	<b>295</b>	<b>350</b>

### Analysis & Sampling

The total cohort size for the Tutor Trust's mainstream provision is 1170. To achieve a level 95% of confidence with a 5% margin of error (the standard for social science research), we would need a sample size of 290. We had a sample size of 350 pupils for attendance and exclusions data, and a sample size of 311 for behaviour data.

Please see tables below to understand the demographic composition of the sample, by school, year group, by Pupil Premium (PP), by Free School Meal (FSM) status, by English as an Additional Language (EAL) status, and by Special Educational Needs and Disabilities.

School Name	% of total pupils from school	Number of pupils
Co-op Academy Leeds	44.9	157

King's Leadership Academy, Liverpool	3.1	11
Manchester Enterprise Academy Central	10.9	38
Mount St Mary's Catholic High School	10.9	38
Oasis Academy MediaCityUK	10.9	38
Oasis Academy Oldham	17.7	62
Ormiston Chadwick Academy	1.7	6

Year Group	% of total pupils	Number of pupils
7	12.9	45
8	6.3	22
9	6.3	22
10	28.0	98
11	46.6	163

Pupil Premium Status	% of total pupils	Number of pupils
False	40.9	143
True	59.1	207

Free School Meal (FSM) Status	% of total pupils	Number of pupils
False	27.6	87
True	72.4	228

English as an Additional Language (EAL) Status	% of total pupils	Number of pupils
False	37.4	131
True	62.6	219

Special Educational Needs and Disabilities (SEND) Status	% of total pupils	Number of pupils
False	79.9	279
True	20.1	70

#### Reporting on 'baseline' and 'endline'

**Baseline data** are from the term before pupils finished the Tutor Trust programme:

- Spring cohort: Autumn term
- Summer cohort: Spring term

**Endline data** are from either during or after pupils finished the programme:

- Spring cohort: the term **after** (Summer term)
- Summer cohort: the term **during** (Summer term)

## Behaviour z-score

Behaviour points required sophisticated data cleaning since schools use different point systems: an incident worth one point in one school might warrant ten in another. To aggregate scores across schools, we used Z-score standardisation, converting each raw score by subtracting the participating pupils' mean and dividing by their standard deviation. This expresses each score as the number of standard deviations from the mean.

We used sample-level parameters (mean and standard deviation) from participating pupils as approximations for whole-school parameters. This approach is justified by the central limit theorem and is highly reliable when  $n \geq 30$ . Three schools were excluded from behavioural analysis due to unavailable data or insufficient sample size for standardisation.

## Exclusions methodology

Exclusion analysis included permanent exclusions, fixed-term exclusions, and suspensions, but excluded internal exclusions as these cannot be easily standardised across schools. For all participating pupils, we counted exclusion occurrences at baseline and endline, then calculated rates by dividing the number of exclusions by the number of pupils who could have experienced an exclusion. This produced baseline and endline exclusion rates for all pupils and by selected subgroups.

## Benchmarking

This report uses two benchmarks:

- Gov.uk data: represents all secondary pupils in England by term and FSM eligibility, not weighted by participating pupils' year group composition
- ImpactEd Evaluation's School Impact Platform (SIP): represents pupils in England by year group and term, weighted by participating pupils' year group composition. This data is collected on a half-termly basis, so we can provide a sample size by half-term:
  - Autumn: half-term 1 -  $n = 451,875$
  - Autumn: half-term 2 -  $n = 451,943$
  - Spring half-term 1 -  $n = 468,474$
  - Spring half-term 2 -  $n = 469,128$
  - Summer half-term 1 -  $n = 474,280$

- Summer half-term 2 –  $n = 442,116$

## Limitations

- Endline data combine two different timepoints: Spring cohort endline data is collected **after** programme completion, whilst Summer cohort endline data is collected **during** programme delivery. This means that 'endline' refers to two different points in time relative to delivery, and therefore the level of change we might expect to see would differ between cohorts.
- The behaviour points data sample size was smaller than the broader sample because three schools were excluded (either because behaviour points data were not available or there were insufficient data points to standardise).
- A large proportion (46.6%) are Year 11 pupils with Summer Term endlines, when they're often on study leave. This creates artificially low attendance figures and reduces opportunities for exclusions or behaviour points. All subgroup analyses will exclude Year 11s to avoid obscuring trends.
- We use two socio-economic disadvantage measures: Pupil Premium (our default measure) and Free School Meal eligibility (stated explicitly when used). We use FSM for attendance data because government benchmarks for secondary pupils compare attendance by FSM status (and not PP status).
- We calculate the mean of pupils' individual changes from baseline to endline which can obscure the stories of change for pupils at an individual level.

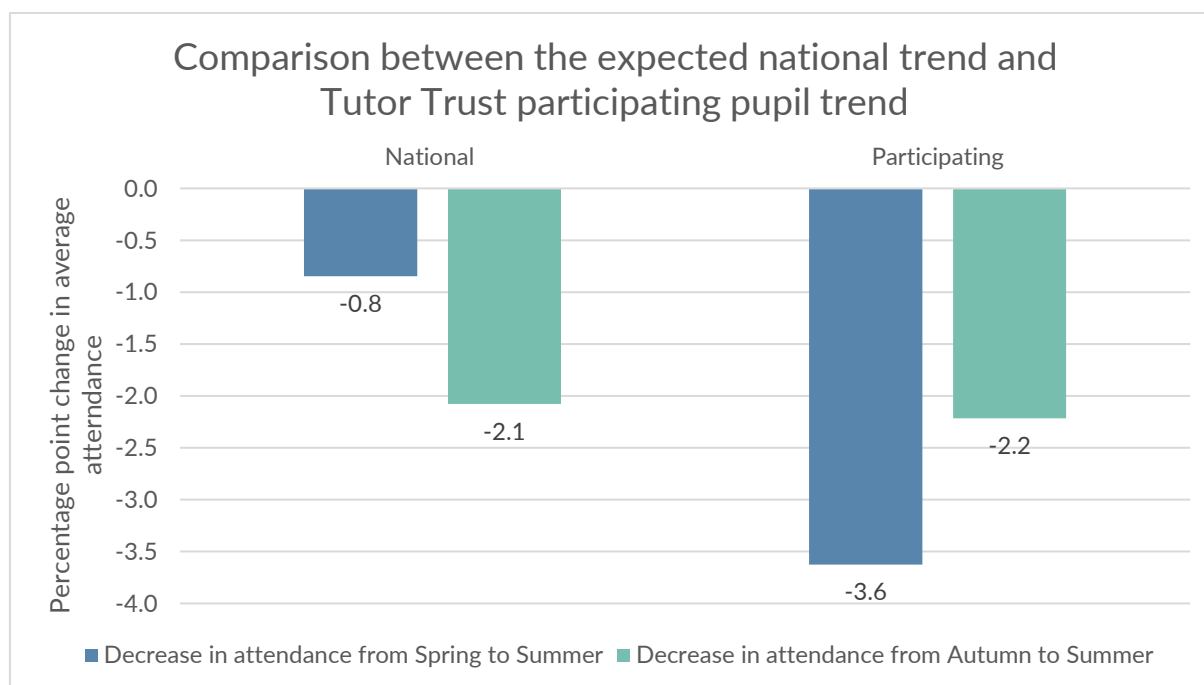
## Headline Findings

### Attendance

**Overall, across all non-Year-11s in the Spring and Summer cohorts, 32.1% of participating pupils' school attendance increased from the term before they received tutoring (either Autumn or Spring) to the Summer term.** When comparing disadvantaged groups of participants (PP, EAL, SEND) to their more advantaged counterparts (non-PP, non-EAL, and non-SEND), we saw that a **greater proportion of disadvantaged pupils' attendance increased compared to their more advantaged peers.** One specific story of positive change is a pupil whose attendance was 67.2% before engaging with the Tutor Trust and increased 26.0pp to 93.2% whilst they were engaging with the Tutor Trust.

**On average, participating pupils' school attendance decreased by 16.2 percentage points (pp) from the term before they received tutoring (either Autumn or Spring term) to the Summer term.** There are two key caveats to this finding. First, for the large majority of pupils, the Summer term occurred *during* programme delivery rather than after delivery, so this data is not representative of the programme's long-term impact. Second, this figure includes attendance data from Year 11 participants during the Summer term, when they are often on study leave, which creates artificially low attendance. Accordingly, this decrease was much smaller when Year 11 pupils were excluded from the sample (-3.4pp); this was a slightly larger decrease in attendance than those in the SIP national benchmark for Years 7-10 (-1.4pp). This benchmark is matched by year group only, not other demographics (PP, SEND, EAL), so comparisons should be interpreted carefully.

The Spring cohort (whose final attendance was measured *after* tutoring ended) showed a smaller decrease in attendance (-2.2pp) compared to the Summer cohort (whose final attendance was measured *during* tutoring, -3.6pp). This pattern is particularly notable because it **goes against the national trend.** National data shows that pupils' attendance drops more sharply from Autumn to Summer term than from Spring to Summer. However, participating pupils' appear to reverse this trend. The Spring cohort (Autumn baseline to Summer endline, measured *after* the programme) showed a smaller attendance decrease than the Summer cohort (Spring baseline to Summer endline, measured *during* the programme). This is the opposite of what we'd expect based on national patterns, which further suggests the programme's benefits are felt more over the long term.



**Pupils with English as an additional language saw smaller decreases than their non-EAL participating peers (-2.7pp vs -4.8pp).**

**Socio-economically disadvantaged pupils (FSM-eligible) experienced slightly larger decreases than their more advantaged peers (-3.9pp vs -1.4pp);** this mirrored national trends (-1.45pp vs -0.95pp). No major difference existed between SEND and non-SEND pupils. There was no clear pattern demonstrating that dosage impacted pupils' school attendance; details can be found in table 1 below. Change in attendance varies greatly by year group. Pupils in Year 11 showed the biggest decrease in attendance (-30.9pp), although this likely reflects artificially low attendance figures in Summer Term endlines due to Year 11s often being on study leave. Year 8 showed the smallest decrease in attendance (-2.5pp).

There are large ranges of change in attendance across all sub-groups of participating pupils. This variation suggests that more time should be taken to explore pupils' individual stories through a qualitative approach to better understand what drives these different changes in attendance.

Table 1: Group/sub-group (* indicates that Year 11s are excluded from group)	Sample size	Average baseline attendance (%)	Average endline attendance (%)	Change in average attendance (percentage point)	Benchmark *from SIP data **from government data	Range of change attendance for individuals (percentage point)	% of participants whose attendance decreased from baseline to endline	% of participants whose attendance did not change from baseline to endline	% of participants whose attendance increased from baseline to endline
All pupils	350	91.9	75.7	-16.2	-10.2*	+26.0 to -69.3	78.0	4.6	17.4
All non-Year 11 pupils	187	89.7	86.3	-3.4	-1.4*	+26.0 to -47.7	59.4	8.6	32.1
Not PP*	75	92.4	89.8	-2.6	-	+26.0 to -47.7	58.7	10.7	30.7
PP*	112	87.9	84.0	-3.9	-	+25.6 to -46.7	59.8	7.1	33.0
Not FSM*	60	92.3	91.0	-1.4	-1**	+26.0 to -33.2	51.7	13.3	35.0
FSM*	112	87.8	83.9	-3.9	-1.5**	+25.6 to -46.7	59.8	7.1	33.0
Not EAL*	61	89.5	84.7	-4.8	-	+20.9 to -47.7	60.7	9.8	29.5
EAL*	126	89.8	87.1	-2.7	-	+26.0 to -44.4	58.7	7.9	33.3
Not SEND*	141	91.4	88.0	-3.4	-	+25.6 to -44.4	60.3	9.2	30.5
SEND*	45	84.4	80.9	-3.5	-	+26.0 to -47.7	57.8	6.7	35.6
Dosage: 1 – 3 hours*	85	89.4	85.6	-3.8	-	+19.5 to -33.2	63.5	8.2	28.2
Dosage: 4 – 7 hours*	34	92.0	92.0	0.0	-	+25.6 to -36.5	38.2	14.7	47.1
Dosage: 8 – 10 hours*	17	85.3	73.8	-11.5	-	+26.0 to -25.4	88.2	0.0	11.8
Dosage: 10+ hours*	47	91.8	89.9	-1.9	-	+15.2 to -47.7	53.2	8.5	38.3
Spring cohort*	34	96.4	94.1	-2.2	-2.1*	+10.0 to -17.7	50.0	20.6	29.4
Summer cohort*	153	88.2	84.6	-3.6	-0.8*	+26.0 to -47.7	61.4	5.9	32.7

(i) Next steps:

- Tutor Trust should prioritize collecting pre- (one term before receiving tutoring) and post-intervention (the term after receiving tutoring) attendance data for all pupils in 2025/26. This is particularly important given that pupils' pre-/post-attendance decreased less than their pre-/during-attendance, which is unexpected considering the academic terms when these measurements were taken and suggests the intervention may have had a protective effect on attendance that warrants further investigation
- Considering the diversity in individual pupils' journeys in attendance, the Tutor Trust should explore pupils' experience through a qualitative approach to understand what drives the differences in pupils' change in attendance.

## Behaviour

Behaviour points analysis assumes negative behaviour points (awarded for poor behaviour), so decreases in behaviour scores indicate improved behaviour.

**On average, participating pupils saw a 0.3 behaviour point decrease**, from the term before they received tutoring (either Autumn or Spring term) to the Summer term, indicating improved behaviour. However, when Year 11s were excluded, there was no change in behaviour points (+0.0bp). **Socio-economically disadvantaged pupils showed improved behaviour (-0.1bp) compared to their more advantaged peers (+0.1bp)**. Pupils with English as an additional language maintained stable behaviour (+0.0bp) whilst their non-EAL peers showed slight deterioration (+0.1bp). Similarly, SEND pupils demonstrated small improvements (-0.1bp) compared to non-SEND peers who remained stable (+0.0bp). No clear trends emerged across different programme dosage levels, and delivery term showed no effect on behaviour changes.

Table 2: Group/sub-group (* indicates that Year 11s are excluded from group)	Sample size	Average baseline behaviour z-score	Average endline behaviour z-score	Change in average behaviour z-score)
All pupils	311	0.1	-0.1	-0.3
All non-Year 11 pupils	174	0.2	0.2	0.0
Not PP*	64	0.0	0.1	0.1
PP*	110	0.3	0.2	-0.1
Not EAL*	57	-0.1	0.0	0.1
EAL*	117	0.3	0.2	0.0
Not SEND*	132	0.0	0.0	0.0
SEND*	42	0.6	0.5	-0.1
Dosage: 1 – 3 hours*	75	0.3	0.1	-0.2
Dosage: 4 – 7 hours*	31	0.1	0.3	0.2
Dosage: 8 – 10 hours*	17	0.3	0.0	-0.3
Dosage: 10+ hours*	47	-0.1	0.2	0.3
Spring cohort*	34	-0.2	-0.1	0.2
Summer cohort*	140	0.3	0.2	0.0

(ii) Next steps

- Explore through qualitative research why less advantaged peers, on average, improve their behaviour more so than their more advantaged peers.

## Exclusions

An increase in exclusion rates indicates more pupils being excluded and represents a negative outcome.

**Exclusion rates among participating pupils increased by 2.6pp**, rising further when Year 11s were excluded (+5.4pp). However, this overall average masks significant disparities between groups, **pupils receiving more than 10 hours of tutoring saw decreased exclusion rates (-2.1pp)**, whilst all other dosage groups experienced increases. **Spring cohort pupils showed stability in their exclusion rates**, whereas Summer cohort pupils saw increases of 6.5pp. Pupils with English as an additional language saw increases of 8.7pp compared to decreases of 1.6pp for non-EAL pupils. SEND pupils experienced large increases (+20pp) compared to small increases for non-SEND pupils (+0.7pp).

### (iii) Next steps:

Table 3: Group/sub-group (* indicates that Year 11s are excluded from group)	Sample size	Baseline rate of exclusion (%)	Endline rate of exclusion (%)	Change in rate of exclusion (%)
All pupils	350	5.4	8.0	2.6
All non-Year 11 pupils	187	9.6	15.0	5.3
Not PP*	75	5.3	4.0	-1.3
PP*	112	12.5	22.3	9.8
Not EAL*	61	9.8	8.2	-1.6
EAL*	126	9.5	18.3	8.7
Not SEND*	141	6.4	7.1	0.7
SEND*	45	20.0	40.0	20.0
Dosage: 1 – 3 hours*	85	9.4	17.6	8.2
Dosage: 4 – 7 hours*	34	8.8	14.7	5.9
Dosage: 8 – 10 hours*	17	5.9	17.6	11.8
Dosage: 10+ hours*	47	8.5	6.4	-2.1
Spring cohort*	34	0.0	0.0	0.0
Summer cohort*	153	11.8	18.3	6.5

- Explore through a qualitative approach why pupils from less disadvantaged groups experience increased exclusion rates from baseline to endline.

## Tutor Trust's next steps for programme understanding

- ▶ Current findings show declining trends in attendance that broadly align with national averages, but the attendance rates are lower in the programme participants relative to the national averages. Related to this, whilst informative as a benchmark and reference point, it is important that the national averages are not treated as a direct comparison group because the pupils contributing to the national averages may differ in observable and unobservable ways to the programme participants. For example, the programme participants are likely selected for existing vulnerabilities that predispose them to poorer outcomes in these areas. To better understand, and isolate, the impact of the programme, **we propose building on the benchmarking and evaluation work already undertaken**. A quasi-experimental design (QED) could be used in a follow-up study to establish what would have happened to participants if they had not taken part in the intervention in relation to their attendance. A QED is widely used in education as a causal methodology. In the context of this evaluation, a matched control group of similar at-risk individuals who didn't receive the intervention would be created and their outcomes compared to the outcomes of the programme participants.
- ▶ Collect pre- and post- data for every participant (and control pupil). This would enable further analysis on how the Tutor Trust programmes impacts attendance in the long term.
- ▶ Whilst current findings demonstrate differences between sub-groups, it is important to undertake further work in understanding whether these changes are meaningful. We would propose conducting statistical significance testing on both pre-/post- analysis and between sub-groups. This would facilitate understanding on whether changes are not random and increase the confidence in our understanding of how the Tutor Trust impacts different groups.
- ▶ Collect qualitative data from teachers and pupils to gain a better understanding of individual pupils' journeys in attendance, behaviour and exclusions.