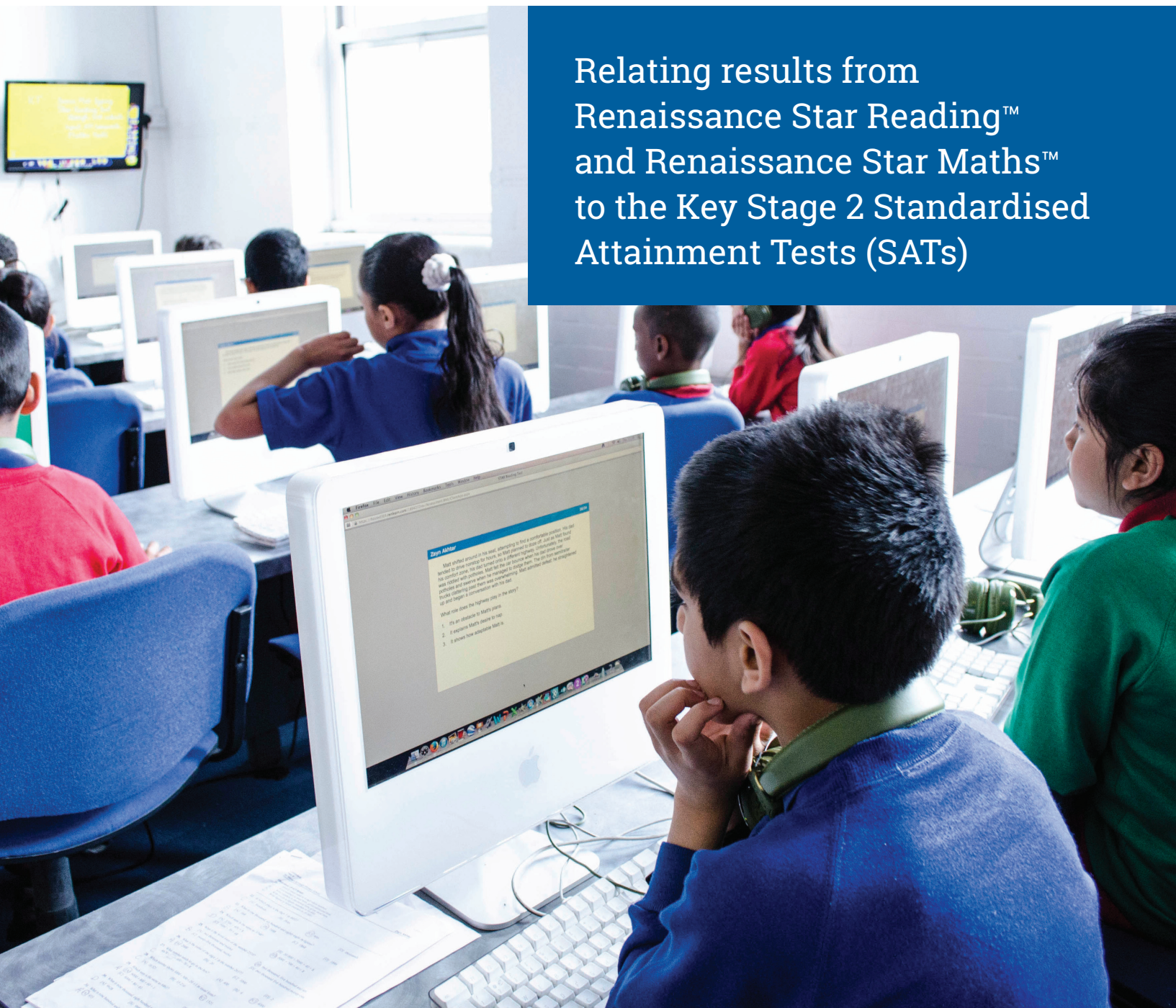


Relating results from Renaissance Star Reading™ and Renaissance Star Maths™ to the Key Stage 2 Standardised Attainment Tests (SATs)



Contents

Introduction	3
Main Findings	3
Study	3
Results	6
Appendix A: About Renaissance Star Assessments	8
Appendix B: Explaining equivalent Star scores	9
Appendix C: Case Study: Star Assessments at Battle & Langton School	10
References	12

Figures

Figure 1. Star Reading and Star Maths scores highly correlate with KS2 Scaled tests.....	4
Figure 2. Star Reading correlation chart with KS2 reading SATs Scaled scores.....	7
Figure 3. Star Maths correlation chart with KS2 maths SATs Scaled scores.....	6
Figure 4. Star Reading correlation chart with KS2 SPaG SATs Scaled scores.....	6

Tables

Table 1. Star Reading and Star Maths score equivalents for KS2..... SATs expected attainment designation	5
Table 2. SATs attainment estimates using Star Reading and Star Maths scores yield accurate results	8
Table 3. Higher standard estimating with Star	7
Table 4. Additional thresholds associating Star scores	9
with probability of meeting standards	

Introduction

Assessments that identify early any pupils at risk of missing academic standards are especially useful

The changes to assessment in the UK in recent years have been numerous: levels have been removed, the National Curriculum has been reformed, and new SATs and GCSE examinations have been introduced. As a result, teachers face a host of new challenges and opportunities. Schools are now presented with a more urgent need than ever to monitor progress and attainment reliably and objectively, identify pupils in need, and prepare for new national assessments and school performance measures.

Renaissance®, suppliers of Renaissance Star Assessments™, have been established in the UK since 1999, and have a presence in over 5,000 UK schools.

This linking study applied results from two interim assessments, Renaissance Star Reading™ and Renaissance Star Maths™, to help teachers assess whether pupils are likely to meet the Primary Standard or may be in need of additional support.

Main Findings

Results from the linking analysis revealed that Star Reading and Star Maths are accurate predictors of the Key Stage 2 SATs outcomes, meaning as a primary teacher you can use Star scores to:

1. Identify pupils likely to miss reading and maths progress targets in advance, enabling meaningful adjustments to teaching to be made well before the KS2 SATs.
2. Forecast the percentage of pupils on track to exceed, meet and fall short of the Primary Standard in reading.

Study

To determine whether Star Reading and Star Maths can predict pupil achievement in the Key Stage 2 SATs in reading, maths and SPaG, we began by linking the score scales for each assessment.

1. Please note that SPaG (Spelling, Punctuation and Grammar) reports will not be available in Renaissance software

Data collection

Using a secure data-matching procedure compliant with Department for Education (DfE) data policies, Renaissance provided the DfE with Star records from the 2015/16 academic year. The DfE then matched this data with the SATs data from that year. For the purposes of this study, only Star tests taken within 30 days of the SATs examination date were taken into account for the remainder of the study.

Sample characteristics

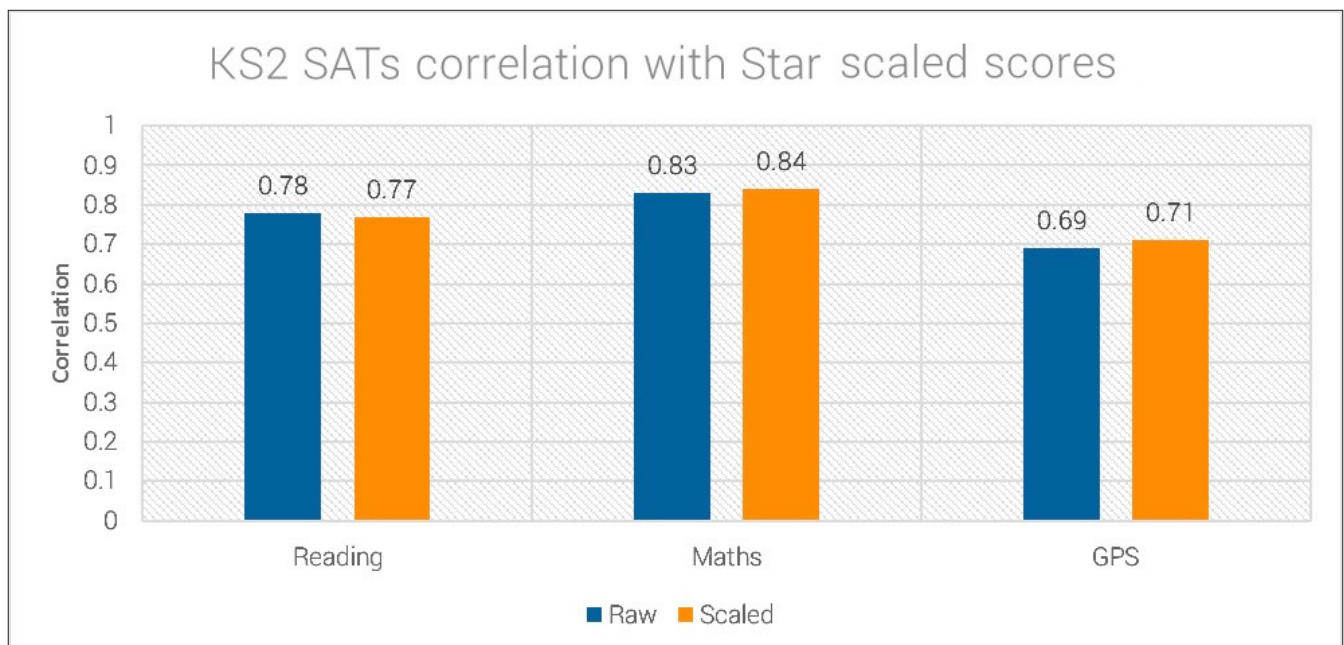
The linking analysis revealed that Star Reading and Star Maths are accurate predictors of Key Stage 2 SATs

Renaissance looked at SATs outcomes from Reading, Maths and SPaG (Spelling, Punctuation and Grammar). After matching records, Reading SATs outcomes from 12,104 pupils from 529 schools, Maths outcomes from 815 pupils from 29 schools, and SPaG outcomes from 12,212 pupils in 532 schools were used in the study. We then set aside scores from a subset of these pupils—10%—as a holdout sample to use only to evaluate the scale linkage.

Correlations

Before linking Star tests with the SATs results, we ensured there was a strong relationship between Star scaled scores and SATs scores. Generally a correlation of 0.7 indicates a strong relationship between score ranges. As seen in Figure 1, the correlations were well above this benchmark: 0.77 for scaled scores between Star Reading and the Reading SATs, 0.84 between Star Maths and the Maths SATs, and 0.71 between Star Reading and the SPaG SATs. These correlations show a very strong relationship between Star and SATs scaled scores.

Figure 1. Star Reading and Star Maths scores highly correlate with KS2 SATs



Scale linkage

Renaissance then linked the score scales for Star Reading/Star Maths and SATs in reading, maths and SPaG by applying equipercentile linking analysis (Kolen & Brennan, 2004). The concurrent sample (sans the holdout sample) was used in the linking (scores from all Star tests taken within 30 days before or after the SATs testing date), and the result was a table of Star scores for each possible SATs score.

Results

SAT scores and corresponding Star score equivalents

SATs results are reported in scaled scores ranging from approximately 80 -120, with a score of 100 signifying the expected attainment standard, and 110 marking the higher standard. A main purpose in linking Star Reading and Star Maths to the SATs was to identify Star scores approximately equivalent to the cut-off score that marks expected attainment. Table 1 displays these equivalent Star scores for each test.

Table 1. Star Reading and Star Maths score equivalents for KS2 SATs expected attainment

	SATs scaled score	SATs Raw Score	Star™ scaled score
Reading	100	21	510
	110	34	858
Maths	100	60	712
	110	98	861
SPaG	100	43	473
	110	61	818

Scatter plots

Using data from the linking sample, we plotted each pupil's scores on Star Reading/Star Maths and Key Stage 2 SATs, marking thresholds for the Star scores equivalent to expected standard and the higher standard.

Figure 2. Star Reading scatter plot with KS2 reading SATs raw scores



Figure 3. Star Maths scatter plot with KS2 maths SATs raw scores

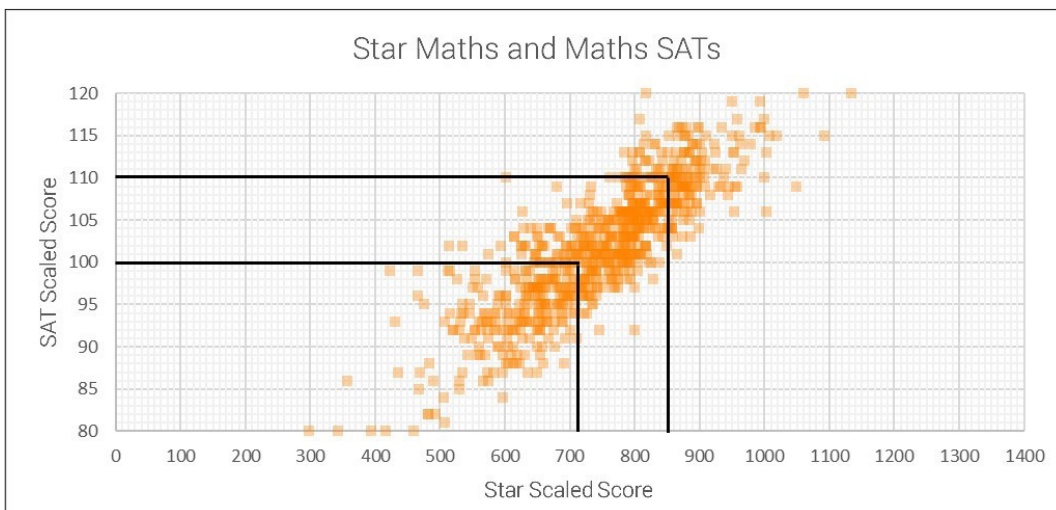


Figure 4. Star Reading scatter plot with KS2 SPaG SATs raw scores



Accuracy of scale linkage confirmed

In evaluating the accuracy of the scale linkage, we examined the differences between pupils' observed (actual) SATs scores and the Star equivalents. To do so, we applied the linking results (i.e. our table of SATs scores for each possible Star score) to the holdout sample, consisting of the subset of concurrent scores not used in the linking. Results showed that our linking computation performed as intended.

Using the sample of actual SATs scores, we were able to compare how the Star estimates aligned with the observed SATs scores. Tables 2 & 3 display classification diagnostics about whether pupils were correctly or incorrectly classified as on track to meet expectations according to their Star scores. On average, pupils were correctly classified (i.e., overall classification accuracy) 79% of the time for reading, 89% of the time for maths and 82% of the time for SPaG. When examining if pupils were on track to achieve higher expectations, overall accuracy was 87% for reading, 83% for maths and 84% for grammar

For Area Under the ROC Curve (AUC), a summary measure of diagnostic accuracy, reading equalled 0.87, maths equalled 0.97, and SPaG equalled 0.87 (also displayed in table 2). For higher expectations, AUC equalled 0.9 for reading, 0.87 for maths and 0.87 for SPaG. High AUCs are convincing evidence that an assessment can accurately predict another assessment result or outcome.

Table 2. SATs attainment estimating using Star Reading and Star Maths scores yields accurate

Measure	Reading	Maths	SPaG
Overall classification accuracy (percentage of correct classifications)	79%	89%	82%
Area under ROC curve	0.87	0.97	0.87

Table 3. Higher standard estimating with Star yields accurate

Measure	Reading	Maths	SPaG
Overall classification accuracy (percentage of correct classifications)	87%	83%	84%
Area under ROC curve	0.9	0.87	0.87

Other diagnostic accuracy measures studied:

- ✓ **Sensitivity** represents the percentage of pupils meeting the standard that were correctly forecasted, which for reading was 84%, for maths was 89%, and for SPaG was 87%.
- ✓ **Specificity** represents the percentage of pupils not meeting the standard that were correctly forecasted, which for reading equalled 70%, for maths equalled 90%, and for SPaG equalled 69%.
- ✓ **Positive predictive values** indicate that when Star scores forecasted pupils to be meet the standard, they actually were proficient 83% of the time for reading, 97% of the time for maths and 87% of the time for SPaG.
- ✓ **Negative predictive values** indicate that when Star scores forecasted pupils to fall short of the standard, they actually weren't proficient 71% of the time for reading, 70% of the time for maths and 69% of the time for SPaG.
- ✓ **Proficiency status projection error**, the difference between actual and projected proficiency rates, indicates how well scores accurately predict proficiency. This was 0% for reading and grammar, and -7% for maths (negative scores indicate under-prediction while positive scores show over-prediction).

Appendix A:

About Renaissance Star Assessments

RENAISSANCE
Star Assessments™



The Star screening report enables you to see at a glance which pupils in a class are on track, and which require additional support

Star Assessments for reading, maths and early learning comprise short, quickly administered computer-adaptive assessments which measure progress and attainment.

Star Assessments provide results immediately on completion. A wealth of data is produced, including a scaled score ranging from 0 - 1400, a norm-referenced standardised score comparing attainment with a pupil's peers nationally, percentile rank, reading age and zone of proximal development.

Multiple tests reveal a range of growth measures, including Student Growth Percentile (SGP), which reports a pupil's growth compared to pupils with similar starting abilities nationally. Star can be administered as often as needed without losing validity, and so can be used to track granular progress with all pupils or with target groups.

Star Assessments' skill-based questions determine mastery within skill domains to ensure age-related expectations are being met.

These skills are charted on the Core Progress learning progressions that have been built for the national curriculum in collaboration with the National Foundation for Educational Research (NFER).

Star Assessments' advanced reporting capabilities enable staff to view and present this information at an individual, class or whole-school level. Reports can be filtered by characteristics such as SEN, EAL and pupil premium, helping you to ensure that all pupils are making progress.

Learn more at: www.renlearn.co.uk/star-assessments/

Appendix B:

Explaining equivalent Star scores

The equivalent Star scores listed Table 1 represent the point on the Star scale at which a pupil reaches a 50% probability of meeting the expected SATs standard– i.e. a pupil scoring exactly 510 in Star Reading has a 50% probability to meet the expected standard in the SATs Reading exam.

To help you estimate beyond this threshold, we have also identified Star scores associated with higher chances of meeting standards.

Table 4. Additional thresholds associating Star scores with probability of meeting

Reading		Maths		SPaG	
Probability of meeting SATs standard	Star Reading score	Probability of meeting SATs standard	Star Maths Score	Probability of meeting SATs standard	Star Reading Score
50%	510	50%	712	50%	473
60%	537	60%	720	60%	486
70%	579	70%	739	70%	534
80%	631	80%	760	80%	592
90%	707	90%	793	90%	678

Within Star, the default cutoff is set to 50% probability, as this results in the lowest overall prediction error rate. There are two kinds of incorrect predictions: False Negatives, pupils who are predicted to fail but pass, and False Positives, pupils who are predicted to pass but go on to fail in the SATs. You may find the scores above useful in giving you increased confidence that individual pupils are SATs-ready. However, know that choosing higher Star scores, while increasing probability of passing, also increases error, specifically the rate of False Negative predictions.

Case study:

Star Assessments at Battle & Langton School



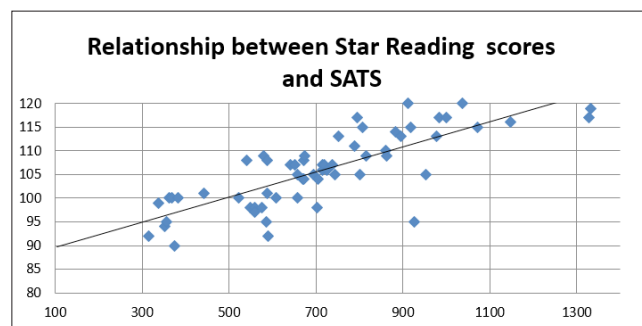
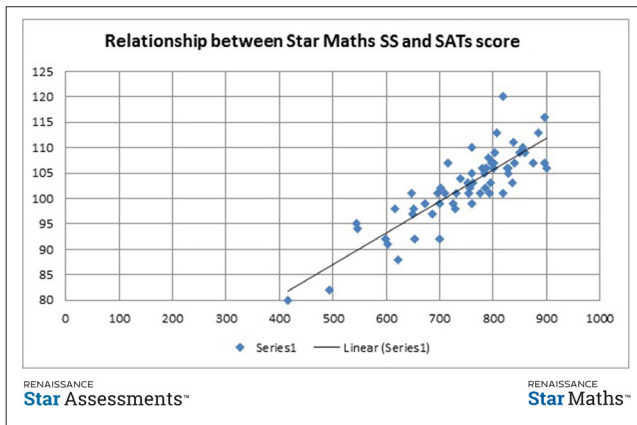
Gary Alexander is
Deputy Head Teacher
of Battle & Langton
CE Primary School, a
two-form entry primary
school on the south

"In 2014, I was keen to move on from old assessment procedures and took the abolition of National Curriculum levels as an opportunity for me to innovate practice. Although National Curriculum levels were familiar and comforting, there is no doubt in my mind that they needed to go. To start with, teacher assessment against National Curriculum levels took far too long, and generated so much work for teachers that it became an onerous task that dominated their thoughts. Accountability levels were rising (to levels far too high), and assessment information was being used to judge teachers rather than to support learning.

I stumbled across the power of Renaissance Star Assessments™ when looking at using Renaissance Accelerated Reader™ to support reading and organise my library. I was looking for something that would remove the burden of teacher assessment, whilst still giving strong and reliable information about learning. I needed the macro level information around group progress, the percentage of children 'on track' to achieve end of key stage expectations, and granular information about next steps in learning for individual children.

To begin with, we introduced Star to years 3-5 as a 6 month trial. Children completed the short and painless computer-adaptive tests, and teachers received information about learning instantly. It took us a while to begin to understand what all the information was telling us, and what parts we needed to look at first. Once we had a basic handle on the information, teachers began to realise that the information more often than not confirmed their own perception of how well that child was doing, even though they hadn't 'ticked a sheet' to say so. After the initial 6 month trial, the other teachers became agitated that they were still filling in 'APP style' sheets, whilst the trial year groups got all the information given to them with minimal workload. The time being saved was huge.

In the first full year we launched for years 2-6 and Star Assessments became our only formal assessment tool. Although teachers were still informally assessing day to day – how else would they teach – there was no formal recording of these judgements needed. About February of that year, I began to realise that either my school was incredible, or the default benchmark for reaching the expected standard was too low. I wish that 90% of my children had achieved the expected standard in their SATs as Star was predicting, but they didn't. But, of course, this was the first year of the new 'more challenging' SATs tests, and the Star data couldn't accurately predict anything.



After the results came in, and Primary Head Teachers across the country had finished weeping over the reading results, I did a correlation study of each test, comparing the SATs scores to the preceding Star scores. This proved that the correlation between the two tests was very strong (0.86 in Maths, 0.79 in Reading). This reassured me that Star could give a very accurate indication of KS2 performance.

The problem was that nobody knew where to draw the 'pass' line. I looked at the two sets of data and worked out what children needed to score in Star to be able to pass their SATs. This gave me a benchmark to work backwards through the years. Now we have a system that judges attainment very accurately, and can tell me at any point in the year which children are on track (given typical progress) to meet the expected standard at the end of the year. This gives me the ability to predict SATs scores this year with greater confidence.

In addition to robust attainment data, Star gives me relative progress information, which allows me to see instantly how children are progressing compared to all children using Star in the UK who started the year at a similar point. This is similar to the DfE 'value added' measure, in that it groups children according to prior attainment, and judges their progress against an average for the group. This way I can see how much progress my 'high flyers' are making when compared to all the other 'high flyers' in the system. So very powerful.

All in all, I believe we have as good an assessment system for reading and maths as is possible in these times. My teachers have more information than they've ever had on learning and progress, and have a lower workload. Children, parents and governors have robust information given to them when we need to, and I have more confidence in my judgements of how my school is performing."

References

Kolen, M. J., & Brennan, R. R. (2004). Test equating scaling and linking: Methods and practices. New York, NY: Springer Science+Business Media.

Renaissance Learning. (2016a). Star Maths: Technical manual. Wisconsin Rapids, WI: Author. Available by request to marketing@renlearn.co.uk

Renaissance Learning. (2016b). Star Reading: Technical manual. Wisconsin Rapids, WI: Author. Available by request to marketing@renlearn.co.uk

RENAISSANCE®

©Copyright 2017 Renaissance Learning, Inc. All rights reserved.

All logos, designs, and brand names for Renaissance's products and services, including but not limited to Accelerated Reader, and Renaissance are trademarks of Renaissance Learning, Inc., and its subsidiaries, registered, common law, or pending registration in the United States and other countries.

Brooklyn | Fremont | Hood River | London | Madison | Minneapolis | Toronto | Vancouver | Wisconsin Rapids