Predicting scale score gain

PAT Reading and PAT Maths

Australian Council *for* Educational Research

Contents

April 2024

Tables

Background

The ACER Progressive Achievement Tests in Reading and Mathematics, commonly known as PAT Reading and PAT Maths, are a set of assessments that allow teachers to measure students' abilities accurately and efficiently, to diagnose gaps, strengths, and weaknesses in student learning, and monitor student progress over time.

While it is well understood that a PAT scale score in the relevant learning area is a reliable indicator of student achievement, until now there has been no clear indicator for how much improvement, or gain, you might expect to observe for individual students from one assessment to the next.

All students are on their own individual learning trajectories and providing a single answer to the question, "how much improvement should we see?" is difficult. There are many factors affecting achievement and gain in Reading and Maths, including, but not limited to, the following:

- A student who achieves a score at the top of the described PAT scale may not have as much room to grow as a student lower down the scale.
- Students may be assessed at different intervals and therefore have different opportunities for teaching and learning between assessments.
- Students in the same classroom start at different points and learn at different rates.
- External and contextual factors may affect the reliability of one or more of a student's assessment results (for example, illness, fatigue, poor engagement with the test etc.)

Goal

This study was designed to illustrate and provide general advice on what scale score gain between assessments looks like in PAT Reading and Maths, using data collected from ACER's online assessment and reporting system. The study modelled the relationship between student year level, scale score starting point ('prior score'), and time between assessments. The intended result of the study was the production of a set of scale score gain predictions based on varying year level and scale score starting points. These data were to be based on multiple regressions that model the effect on gain of year level, prior scale score, and the time elapsed between assessments and to be utilised to represent predicted gain for students within online PAT Maths and PAT Reading reports.

Our conception of 'predicted gain' for the purpose of this study, can be understood as a response to the following question:

"For students in the same year level whose initial (prior) scale scores were the same or very similar, what is the typical scale score gain achieved after a given time interval?"

Importantly, in this context, a student's predicted gain does not reflect the student's 'expected' gain, or what they 'should' achieve after a given time interval, and does not account for curriculum standards, school goals, or individual student context. Equally, the gain prediction between subsequent assessments over a relatively short period should not form the basis of a value judgement about the quality of a student's broader learning progress. Instead, these gain predictions should be used as an additional piece of information to assist you in analysing and understanding your assessment data.

Data

The study data comprised PAT Maths 4th Edition and PAT Reading 5th Edition assessment results collected from ACER's online assessment and reporting system. Only completed test sittings from 2020 and 2021 were retained for the analysis and only from students who had completed two tests. In cases where students had completed three or more tests in the relevant time period, they were removed from the analysis. The time between tests was also considered, and the final data did not contain sittings more than 730 days (24 months) apart or fewer than 200 days (6.5 months) apart. Additional variables such as student gender, education system, and state or territory were not considered. The final dataset comprised results from 689 502 Maths tests and 391 223 Reading tests.

Table 1 *Data composition*

Year level	Maths tests	Reading tests
1	4167	797
2	36 327	13 795
3	81 767	28 212
4	97 467	31 951
5	99 127	31 345
6	150 318	72 521
7	46 407	38 604
8	60 916	61 887
9	59 632	59 525
10	53 374	52 586
Total	689 502	391 223

Models

Multiple linear and mixed effects models were fitted for each domain and at each year level. In the first model, a multiple linear regression was employed to model the relationship between the gain between scores (gainSS), the initial scale score (priorSS), and the time interval between assessments (daysBetween). The model aimed to investigate how changes in gain can be attributed to variations in both the initial scale score and the time elapsed between measurements.

Model 1

$$gainSS = \beta_0 + \beta_1 \times priorSS + \beta_2 \times daysBetween + \varepsilon$$

A mixed linear regression was employed in the second model to model the same relationships. In Model 2, the school was included in the model as a random effect variable to capture the variation in the intercept across different schools. This was to account for the potential clustering of data within schools.

Model 2

$$gainSS_{ij} = \beta_0 + \beta_1 \times priorSS_{ij} + \beta_2 \times daysBetween_{ij} + u_{0j} + \varepsilon_{ij}$$

Results

Results for year 8 students are used as exemplars below for both Reading and Maths.

Model 1

The linear regression analyses indicated that there is a significant relationship between both prior score and the time interval between assessments and scale score gain in domains. In Maths, when the predictors (prior scale score and time between tests) are at 0, students on average would gain 17.7 scale score points. For each scale score point on the initial assessment (prior score), the gain was reduced by 0.13 points, and for each day between tests, the gain increased by 0.004 points. For Reading, the average gain with the predictors at 0 was 32.2 scale score points, decreasing by 0.25 points for each prior scale score point, and increasing by 0.006 points for each day between assessments. The model was a good fit, but the predictors explain a relatively small proportion of the variability in gain; for Maths just 4% and for Reading 11%. Table 2 and Table 3 provide the results for year 8 Maths and Reading.

Table 2 *Effect of initial scale score and time between assessments on gain in year 8 Maths*

Variables	Estimate	Std error	t	Probability
Constant	17.652	0.373	47.386	<0.000
Initial scale score	-0.130	0.003	-49.582	<0.000
Interval between assessments	0.004	0.000	9.129	<0.000
Adjusted R^2= 0.040 F(2,60913) = 1271, p<0.000				

Table 3 *Effect of initial scale score and time between assessments on gain in year 8 Reading*

Variables	Estimate	Std error	t	Probability
Constant	32.211	0.402	80.17	<0.000
Initial scale score	-0.252	0.003	-89.33	<0.000
Interval between assessments	0.006	0.001	12.72	<0.000
Adjusted R^2= 0.116 F(2,61884) = 4069, p<0.000				

Model 2

In Model 2, the random effect of the school was included in a mixed effect model, along with the fixed effects used in the linear model. Using this model, the average gain observed for Maths at year 8 is 20.7 scale score points when the predictors are at 0. For each scale score point in the initial score, that gain is reduced by 0.15 points, for every day between assessments the gain increases by 0.002 points. The average gain varies by around 2.3 scale score points by school. The intraclass correlation, that is the proportion of variance between schools compared with the total variance in Maths, is approximately 9%. For Reading the average gain was 35.2 scale score points with predictors set to 0, and it was reduced by 0.3 scale score points for every prior scale score point and increased by 0.002 points for every day between

assessments. The average gain varies by around 2.4 scale score points by school. For Reading, the intraclass correlation was approximately 8%.

Table 4 and Table 5 provide the results for year 8 Maths and Reading.

Table 4 *Effect of initial scale score, time between assessments and school on gain in year 8 Maths*

Variables	Estimate	Std error	t
Intercept	20.702	0.461	44.813
Initial scale score	-0.150	0.003	-53.989
Interval between assessments	0.002	0.001	2.625
Random effects	**Variance**	**Std deviation**	
School (Intercept)	5.217	2.284	
Residual	53.126	7.289	

Table 5 *Effect of initial scale score, time between assessments and school on gain in year 8 Reading*

Variables	Estimate	Std error	t
Intercept	35.244	0.487	72.404
Initial scale score	-0.267	0.003	-91.685
Interval between assessments	0.002	0.001	2.596
Random effects	**Variance**	**Std deviation**	
School (Intercept)	5.581	2.364	
Residual	61.595	7.848	

It was found that adding school as a random effect was a significant addition to Model 1 in Reading and Maths, but with the school effect explaining less than 10% of the variance in gain in both Reading and Maths, it was not so important as to require the school to be included in the model to predict gain for PAT test-takers more generally.

The study results illustrate the variability of scale score gain between subsequent tests for students in different year levels and starting from different points. Table 6 shows the predicted scale score gain between two PAT Maths tests completed 12 months apart by students in years 2, 6, and 10, all of whom scored 125 on their initial test.

Table 6 *Example predicted gain by year level for students with the same prior PAT Maths scale score*

Year level	Prior scale score	Predicted gain	Predicted scale score
2	125	+0.5	125.5
6		+4.8	129.8
10		+3.7	128.7

Given the same scale score starting point, predicted gain over 12 months varies widely depending on students' year levels. Similarly, Table 7 shows the predicted 12-month scale score gains for students in year 6 whose initial scale scores differed significantly.

Table 7 *Example predicted gain by prior PAT Maths scale score for students in the same year level*

Year level	Prior scale score	Predicted gain	Predicted scale score
6	80	+13.0	93.0
	100	+9.4	109.4
	120	+5.7	125.7
	140	+2.1	142.1

Application

The reporting goal was to provide an indication of the range of scale scores that are typically achieved by students, given a particular scale score starting point and the time interval between assessments, at a particular year level. The study provided sufficient evidence that there is a significant relationship between students' prior scale scores and time between assessments, and the results may be used to predict scale score gain in Reading and Maths over time.

Within the online ACER Data Explorer, a student's predicted PAT Maths or PAT Reading scale score gain is shown along with an inner and outer range that represent a degree of uncertainty around the prediction. The outer reported range in predicted gain represents the 95% confidence interval, or approximately two standard errors on either side of the predicted value. This means that, on average, 95 times out of 100, when the same conditions are present, a student's gain will fall between the upper and lower limits of that confidence interval. The inner range represents the 68% confidence interval, or approximately one standard error on either side of the predicted value.

The inherent variability in test performance means that predicted gain does not perfectly reflect the true gain for all students. PAT scale scores are estimates and subject to a degree of uncertainty or error. While the range of gain reported represents a typical achievement for a student like the one for whom the report is generated, any individual student may achieve a score that differs from that range.

Predicted scale score 'decline'

In some cases, due to the relationship between student year level, prior scale score and the time interval between measurements, the study predicts that a typical student's scale score will decline. Again, it is necessary to differentiate here between the statistical prediction and the individual student. Where a student's scale score does, in fact, decline between two assessments, we would not assume that their learning progress has actually reversed. Rather, no progress has been measured and your professional judgement is required to understand and evaluate this result within the student's broader context, including through observations, classwork, and other formal and informal assessments.

Predicted scale score decline is not represented within the online reporting.

Predictions vs expectations

It is important to note that the scale score gain predictions do not account for system-, school-, or classroom-level expectations of student achievement or progress, curriculum standards, or other pedagogical benchmarks or performance targets. Your knowledge of students and your professional judgement are crucial when analysing student achievement and setting suitable targets for learning progress.

As mentioned above, our conception of predicted gain for this study can be thought of as a response to this question:

"For students in the same year level whose initial (prior) scale scores were the same or very similar, what is the typical scale score gain achieved after a given time interval?"

The following section provides guidance on how to use and understand predicted gain within the online ACER Data Explorer reporting.

Reporting

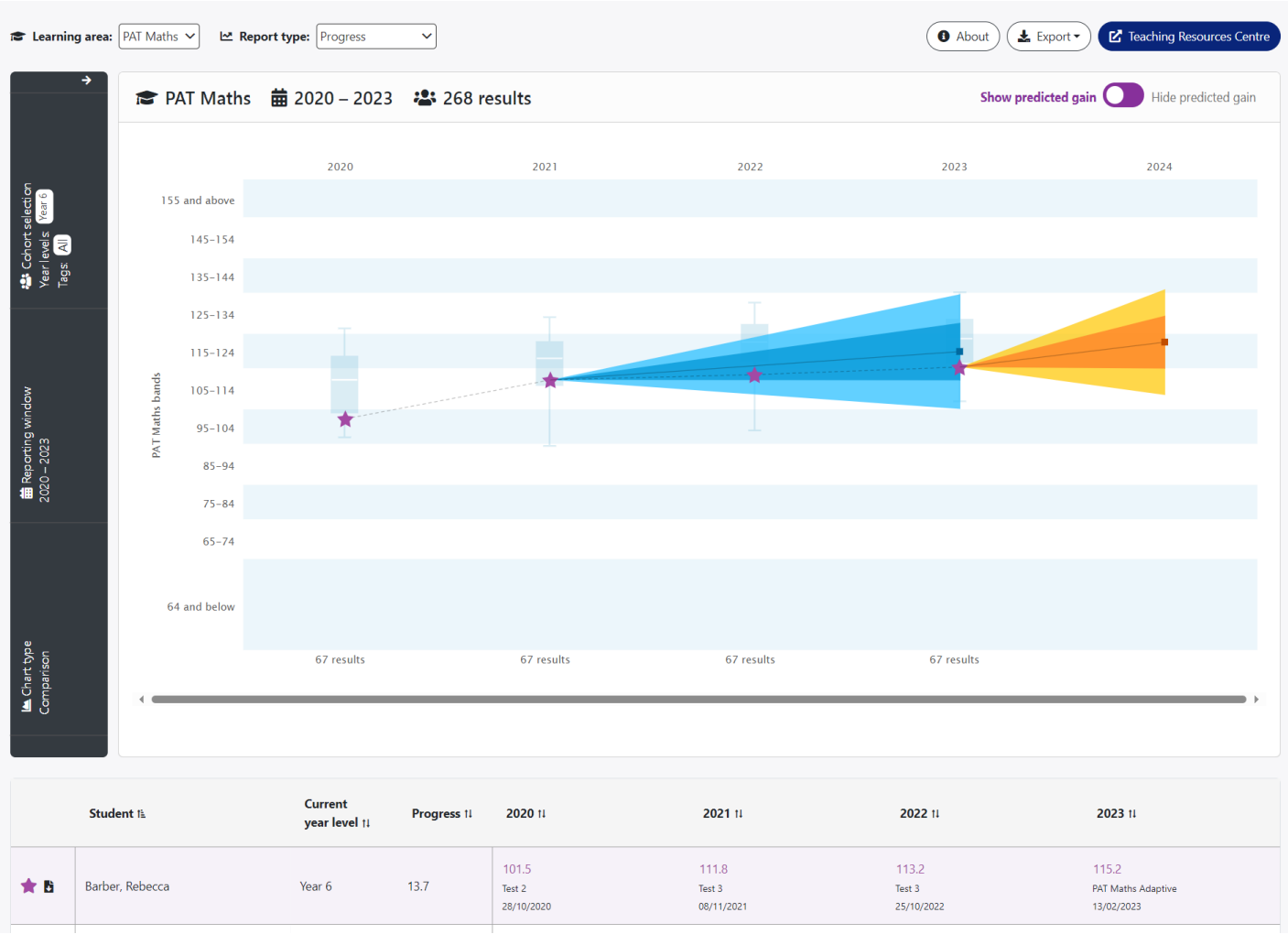
Students' predicted gain can be viewed within the ACER Data Explorer 'Progress' report. Highlight a student in the report table to view their historical scale score achievement. Each of the student's actual achieved scale scores are represented in the chart by purple stars.

Select 'Show predicted gain' to add this information to the report chart.

Predicted gain can be used for two distinct purposes:

1. Understanding the scale score gain a student is predicted to achieve 12 months after their most recent test (represented in orange).

2. Comparing a student's most recent achievement (purple star) against their previously predicted achievement (blue).

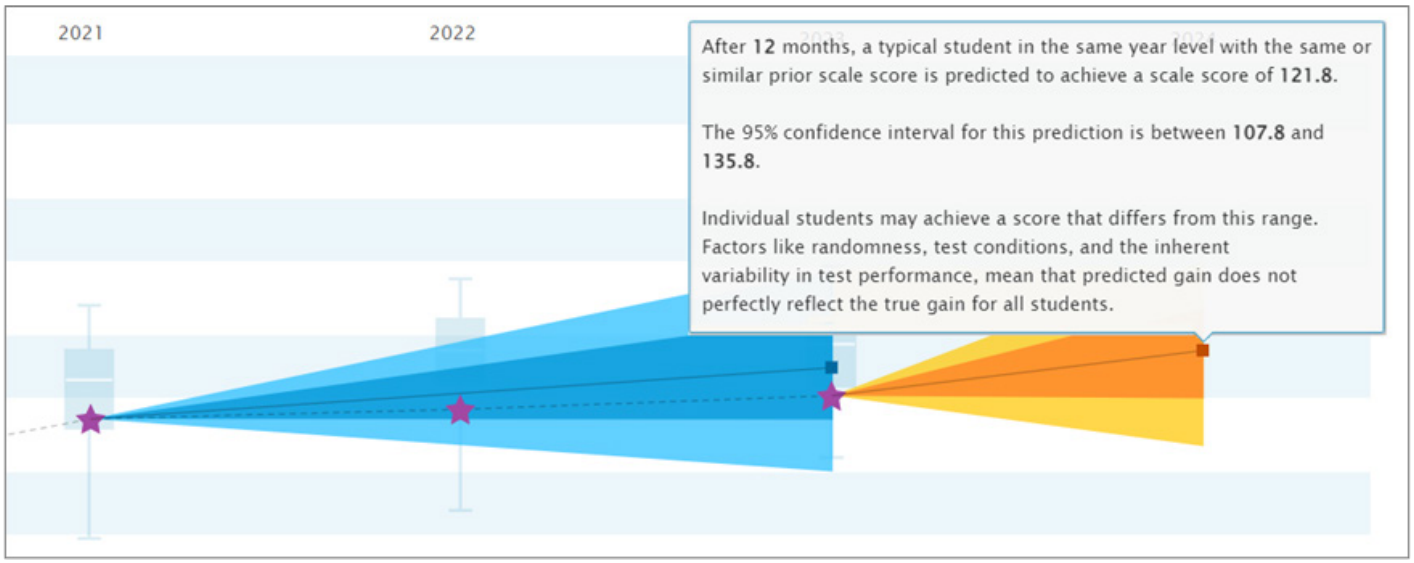


Future predicted gain

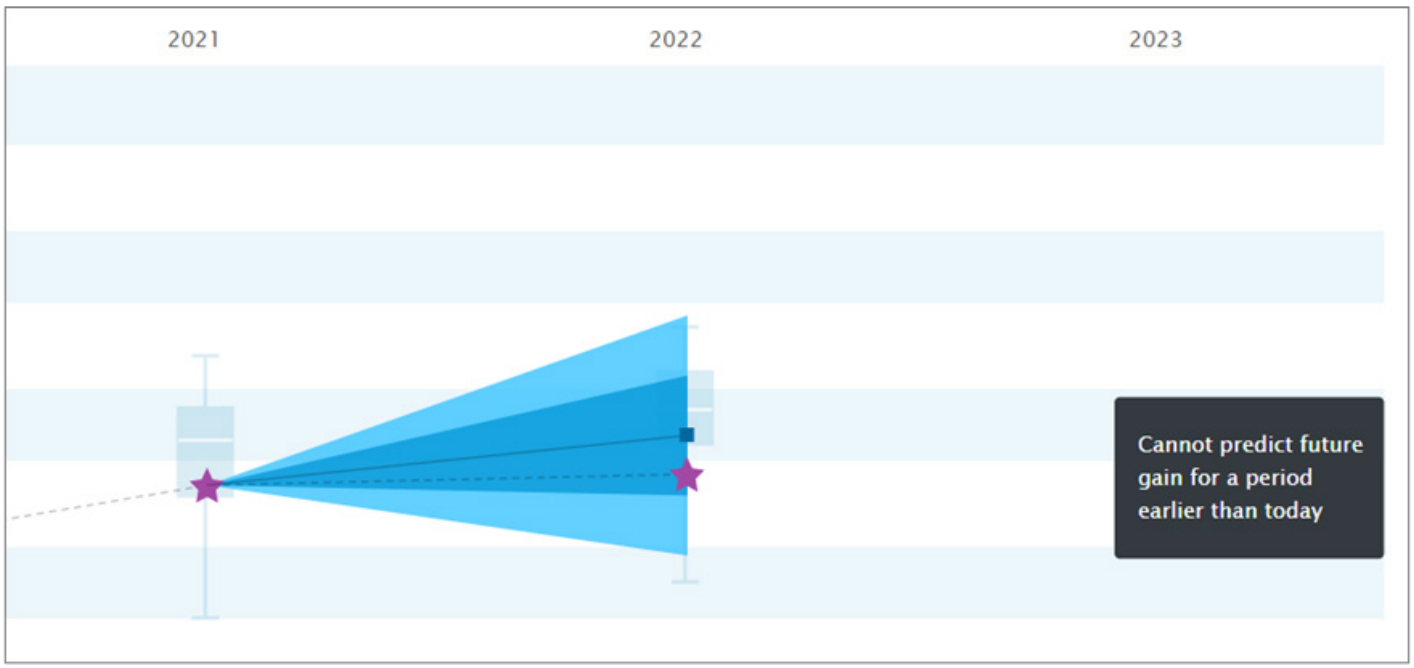
A student's 'future' scale score prediction represents their predicted achievement if they were to complete another PAT assessment in the learning area 12 months from the date of their most recent test. This predicted scale score is based on the 12-month gain of other students in the same year level who achieved the same or similar initial scale score.

This a statistical prediction and the ranges surrounding this single predicted value represent confidence intervals of the prediction. The inner, dark orange range represents one standard deviation either side of the prediction, or the 68% confidence interval. In other words, on average, when the same conditions are present, 68 times out of 100 a student's gain will fall within that range.

The broader, light orange range represents two standard deviations either side of the prediction, or the 95% confidence interval. On average, 95 times out of 100, when the same conditions are present, a student's gain will fall within that range.



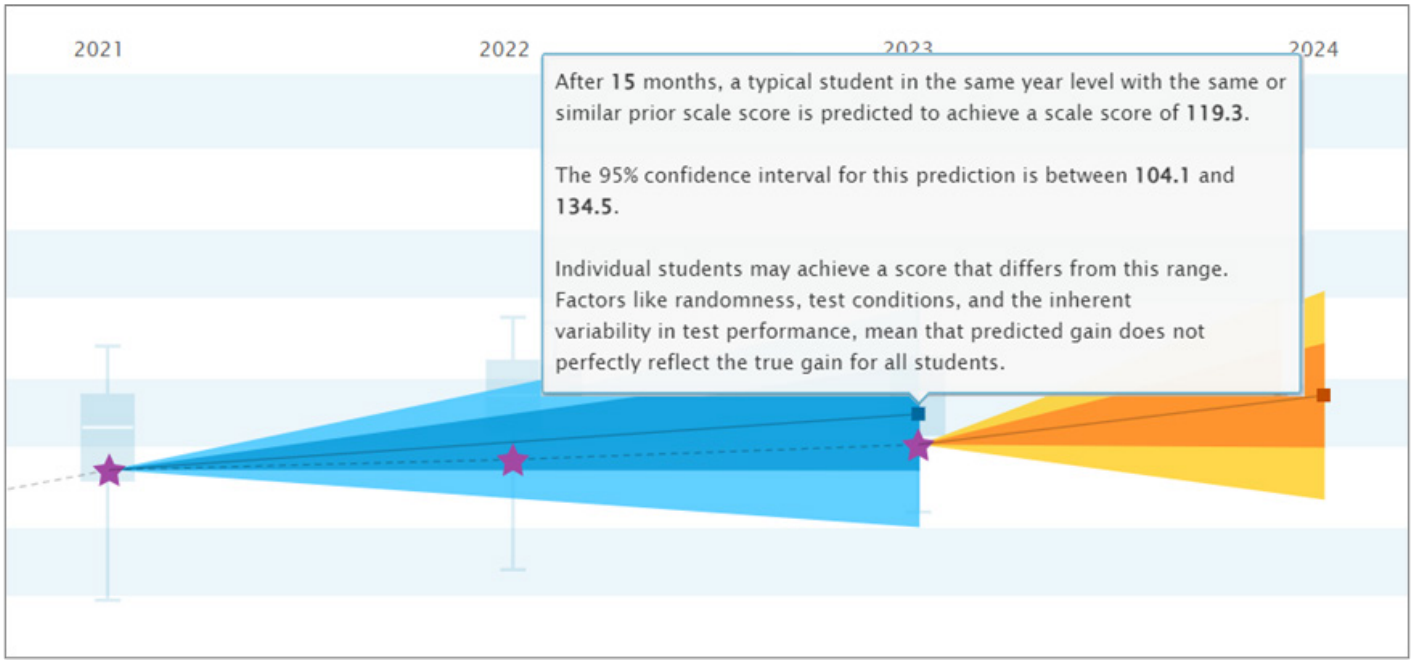
Note that 'future' predicted gain may not be displayed in the report if the student's most recently completed test is more than 12 months old. In other words, the period for which the 'future' prediction applies has already passed:



Past predicted gain

A student's 'past' scale score prediction allows you to compare your student's most recent actual achievement to their predicted achievement based on an earlier result.

For a 'past' prediction to appear in the report, the student must have completed a previous test in the same learning area 9–18 months prior to their most recent assessment. If the student does not have a previous result, or if they only have results fewer than 9 months prior or more than 18 months prior, no 'past' prediction will appear.



In the example above, the student's most recent previous assessment, from 2022, was completed fewer than 9 months prior to their 2023 assessment. As a result, their 'past' predicted gain has been calculated based on their result from an earlier test, completed 15 months prior, in 2021.

Predicted scale score ‘decline’

As mentioned above, in some cases, due to the relationship between student year level, prior scale score and the time interval between measurements, the study predicts that a typical student’s scale score will decline over a given time period.

The online reporting does not display predicted scale score decline but does display the confidence intervals for your reference.