Using research-level statistics to ‘prove’ the impact of teaching initiatives

Darcy Fawcett, darcy@sounddata.co.nz

Director, Sound Data.

Today’s resources:

- Preliminary survey https://bit.ly/SD_NZAI2025_Survey
- Online dashboards https://sounddata.co.nz/dashboard/

SOUND DATA

Removing Noise From Decision Making

How can you ‘prove’ the impact of initiatives on learning?

- You must compare assessment data of those who experienced the initiative with those who didn’t
- The initiative group must make *significantly* more progress than the comparison group *and* this effect must be large enough to be worth attending to.

Experimental method

- Compare the performance of two equivalent groups
  - control and experiment group
- Often very difficult to arrange and can be unethical

Quasi-experimental method

- Compare naturally occurring groups
  - Compare the performance of latest cohort with historical performance
- Assumptions
  - latest cohort is typical
  - The historical record is large enough to describe the population

Data – basics

Summary statistics

Continuous/scale data

- parametric and non-parametric

Categorical

- Ordinal and nominal data

Summary statistics

Measures of centre (averages)

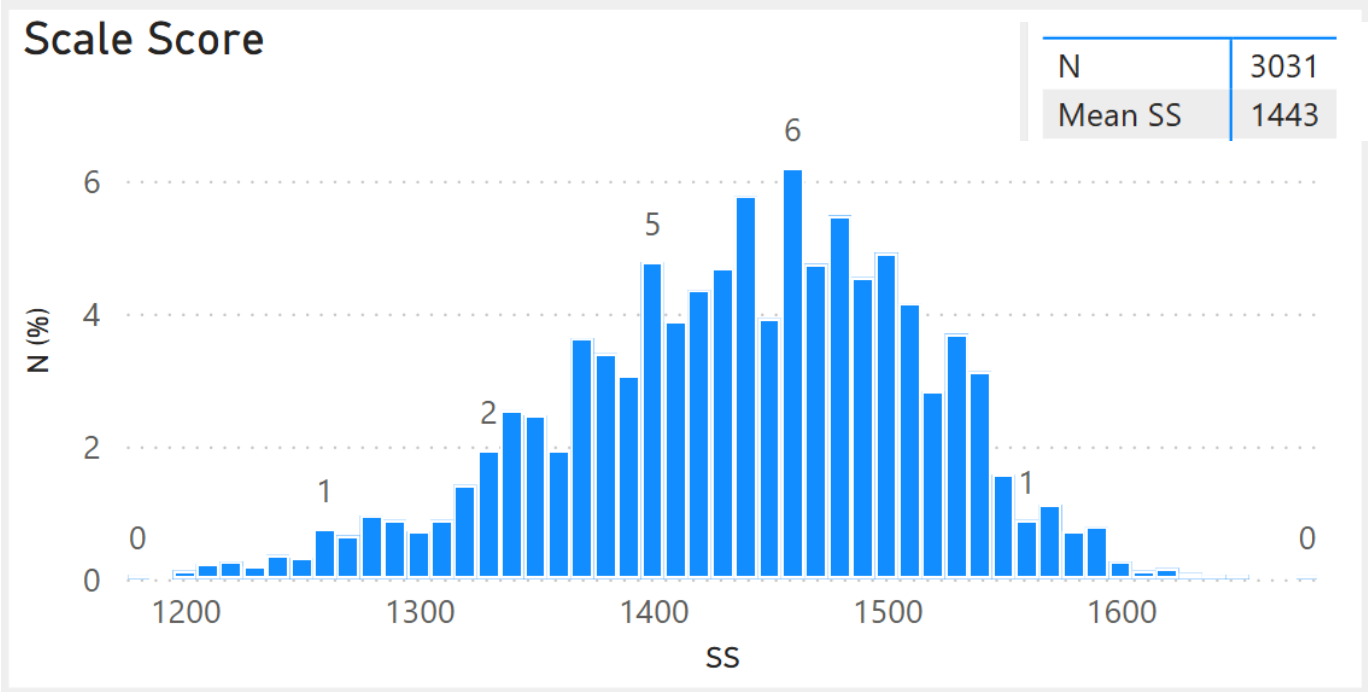
- mode
  - most frequently occurring value
- Medium
  - the middle of ranked data
- mean
  - $\mu = \bar{x} = \frac{x_1+x_2+x_3+\cdots}{n}$

Measures of spread (distribution)

- range
  - $R = x_{max} - x_{min}$
- interquartile range
  - $IQR = x_{Q3} - x_{Q1}$
- standard deviation
  - $\sigma = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots}{n}}$

Continuous/scale data - parametric

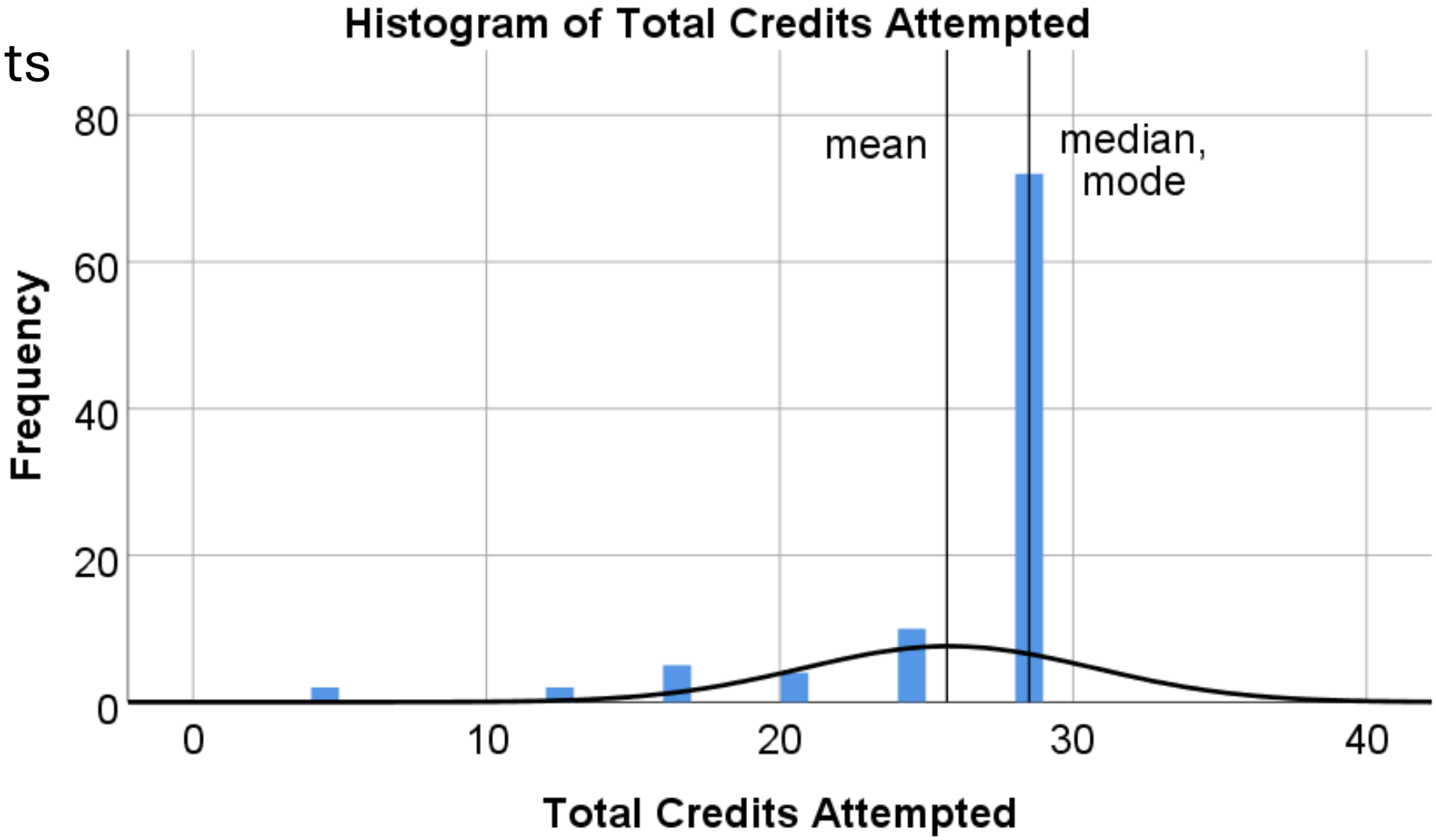
- **normally distributed**
- e.g. e-asTTle, PAT and STAR scale scores, IQ
- What are the median and mode?



- e-asTTle
  - National mean scale score: 1500
  - Standard deviation: 100

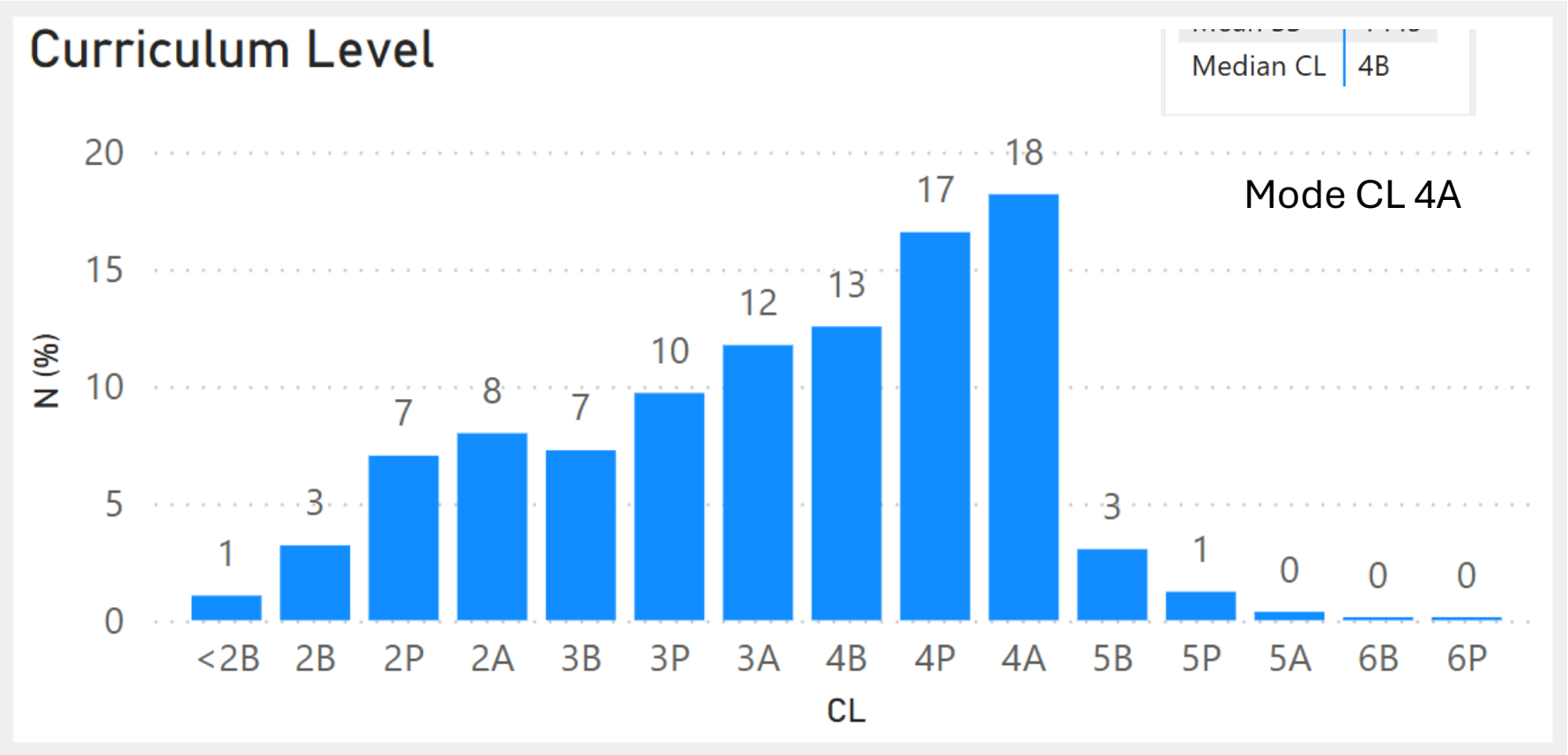
Continuous/scale data – non-parametric

- **not** normally distributed
- E.g. NCEA credits



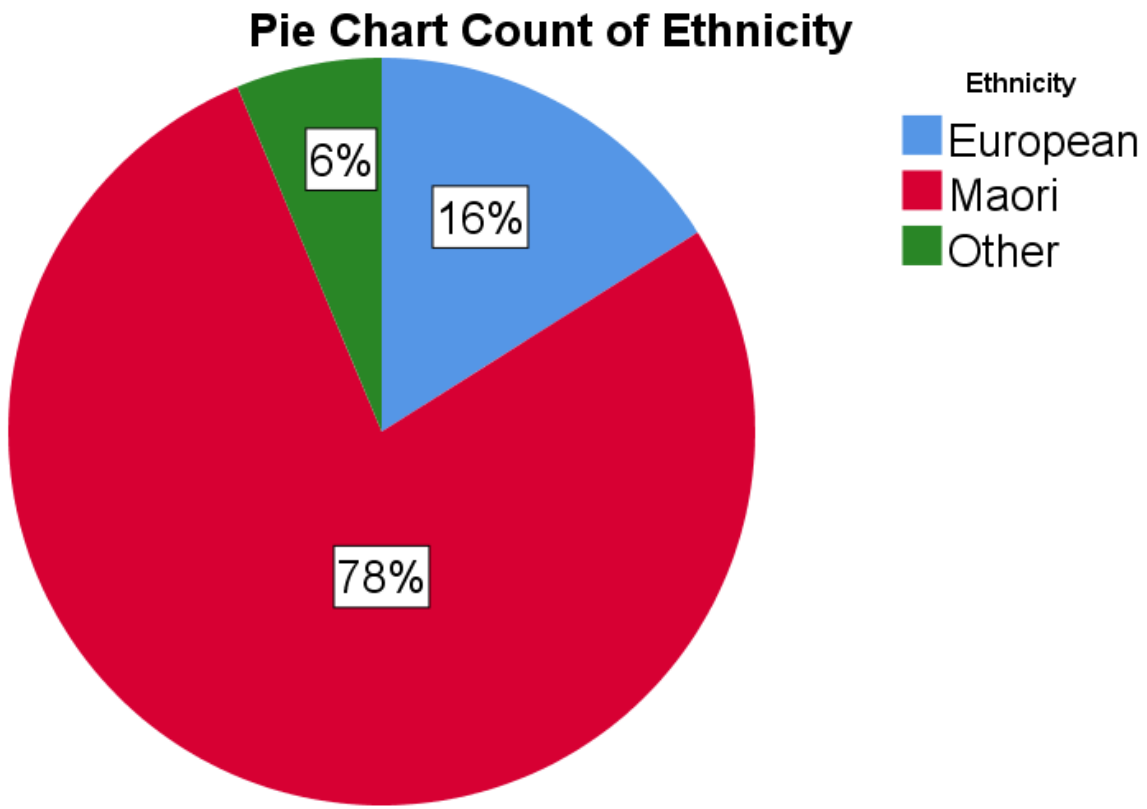
Categorical Ordinal data

- Categories that have order
- e.g. curriculum levels, achievement standards …



Categorical nominal data

- Categories that have no order
  - Count/frequency or percentage
- Summary: mode
- e.g. Gender, ethnicity, cohor



Statistical tests for significance

p: the probability that your observations are a random occurrence

Test result	probability
similar to / no significant difference	$p > 0.10$
notable difference	$0.10 \geq p > 0.05$
significant difference	$0.05 \geq p > 0.01$
highly significant difference	$0.01 \geq p$

The p-value acts as a filter for changes that are similar in size to the normal vaiations that occur from year to year

Effect size / Strength of association (only use if significant relationship)	R	G	D
None/minimal	0.00 – 0.09	0.00 – 0.24	0.00 – 0.09
Weak/small	0.10 – 0.29	0.25 – 0.49	0.10 – 0.19
Moderate	0.30 – 0.49	0.50 – 0.74	0.20 – 0.30
Strong/large	0.50 – 1.0	0.75 – 1.00	0.30 – 0.49
Very strong/large			0.50 – 1.00

But these trigger points are context dependent and should be varied depending on data properties and theoretical assumptions (Cohen 1969; Kraft 2020).

Associated measure for **strength of association** or **effect size**

- Large populations can show significant relationships for very small shifts. Use related tests to show the strength of the association or effect size
  - The relative size of the skew on the graph
- Which test you use depends on the nature of the variables E.g.,
  - Nominal by nominal (e.g Crammer's V) or
  - ordinal by nominal (e.g. Rosenthal's R, Goodman and Kruskal's Gamma)
  - Scale by nominal (e.g. Rosenthal's R, Cohen's D)

Statistical tests

https://stats.oarc.ucla.edu/other/mult-pkg/whatstat/

Variables		Tests for significance (and effect size)	
Independent	Dependent = 1	Independent samples	Related samples
0 variables - Theoretical	**Categorical, Cat = 2** Pass rates	**Cat2: One sample binomial test; Ord3+: One sample median test (Rosenthal's *R*)** - e.g., Year Level Achieved, School vs National % - e.g., Curriculum Level, School vs National %	
1 variable **2 levels** - Cohort - non/Māori - Gender **3+ levels** - Prioritised Ethnicity - Year - Room	**Ordinal, Ord = 2** UE, Below/At+	**Chi-Square Linear-by-Linear Association test (Gamma)** e.g., UE; this year vs previous years.	**McNemar's test (Phi, AR)** - e.g., Below/At+; T1 vs T4
	Ordinal, Ord ≥ 3 **Y**ear **L**evel **Q**ual. **A**chievement **S**tds **C**urriculum **L**evels **Non-Parametric Scale** Credits	1IV2L: **Wilcoxon-Mann-Whitney U rank test (R)** - e.g., YLQ; male vs female. 1IV3+L: **Kruskal-Wallis Anova by ranks (Eta)** - e.g., AS; Prioritised Ethnicity. 1IV3+L: **Anom of transformed ranks.** - e.g., progress on CL; School vs all rooms.	1IV2L: **Wilcoxon signed ranks test (Cohen's D)** - e.g., Curriculum Levels; T1 vs T4
	Parametric scale PAT scale scores e-asTTle s. scores PaCT s. scores	1IV2L: **Independent samples t-test (R)** - e.g., Scale score, Māori vs non-Māori. 1IV3+L: **Analysis of Variance by means (Eta).** - e.g., progress on s.sc; School vs Room1. 1IV3+L: **Analysis of means, Anom.** - e.g., progress on s.sc; School vs all rooms.	1IV2L: **Paired samples t-test (D)** - e.g., Scale score; T1 vs T4

The independent samples t-test

- 1 independent variable with 2 levels: ty v py
- 1 parametric scale dependent variables: scale scores, TX or progress

Cohort longitudinal comparison (all YL)

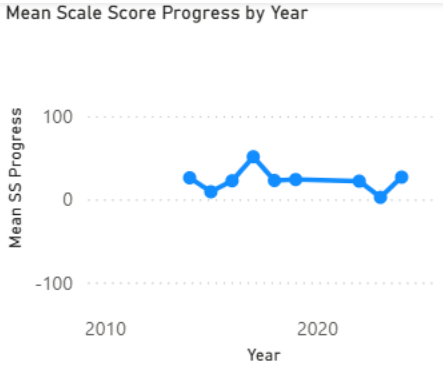
Reading | T1 to T4

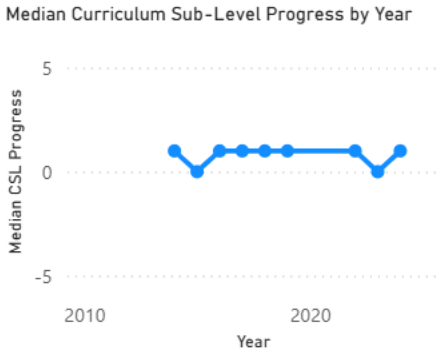
Gender: All | Ethnicity: All | ESOL: All | Year: All

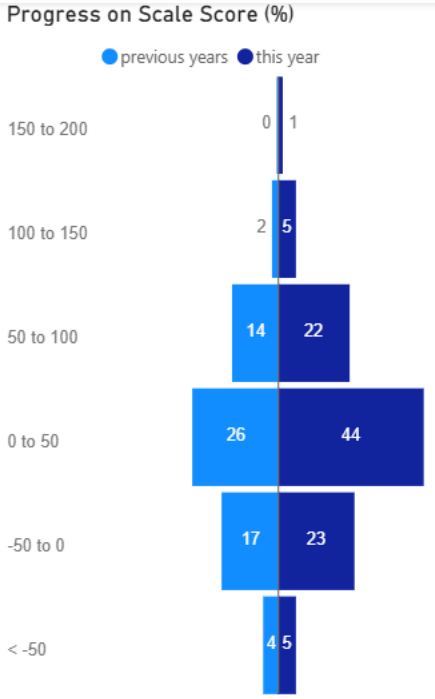
Cohort	N
previous years	1258
this year	274
Total	**1532**

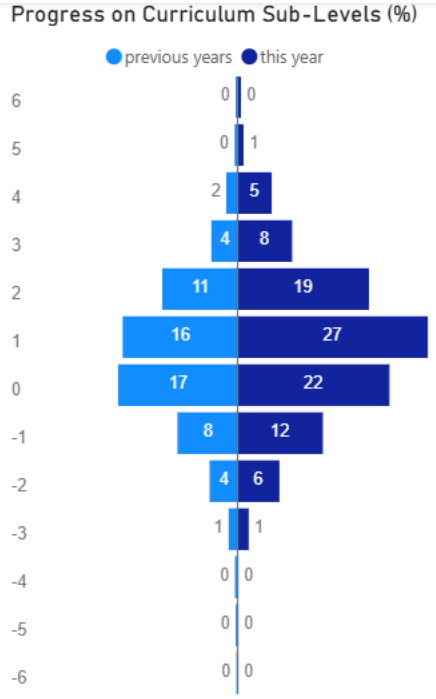
Search

Year	NSN	Student Name	Δ SS	Δ CSL
2014	100055	Student, 100055	57	2
2014	100060	Student, 100060	17	1
2014	100083	Student, 100083	40	0
2014	100089	Student, 100089	-29	-1
2014	100094	Student, 100094	72	2
2014	100098	Student, 100098	39	0
2014	100127	Student, 100127	13	0
2014	100131	Student, 100131	-17	0
2014	100141	Student, 100141	-15	-1









YearLevel	Dem	Statement	n (py)	n (ty)	mean Δ (py)	mean Δ (ty)	med Δ (py)	med Δ (ty)	p	ES
All	All	small highly significant improvement	1258	274	18.1	26.7	1.0	1.0	0.005	0.19
All	Asian	NO SIGNIFICANT DIFFERENCE, improvement	249	47	24.0	27.1	1.0	1.0	0.661	0.07
All	ESOL	INVALID TEST, n < 20, improvement	61	18	8.3	35.6	0.0	1.0	0.039	0.56
All	Female	moderate significant improvement	638	142	18.7	28.8	1.0	1.0	0.018	0.22
All	Male	NO SIGNIFICANT DIFFERENCE, improvement	620	132	17.5	24.4	1.0	1.0	0.118	0.15
All	Maori	NO SIGNIFICANT DIFFERENCE, improvement	178	31	6.4	14.8	0.0	1.0	0.379	0.17
All	MELAA	INVALID TEST, n < 20, improvement	61	18	8.3	35.6	0.0	1.0	0.039	0.56
All	Other	INVALID TEST, n < 20, reduction	16	6	22.8	20.7	1.0	0.5	0.922	-0.05
All	Pacific	NO SIGNIFICANT DIFFERENCE, improvement	199	46	16.5	23.9	0.0	1.0	0.325	0.16

Did students make more progress this year? Do any subgroups stand out?

Year Level and demographic
(All, Y7, Y8, M, F, Māori, Pacifica, Asian, MELAA, Pakeha)

number of students
(previous years or this year)

Mean progress on Scale Score and median progress on curriculum sub-levels
(previous years or this year)

YearLevel	Dem	Statement	n (py)	n (ty)	mean Δ (py)	mean Δ (ty)	med Δ (py)	med Δ (ty)	p	ES
All	All	small highly significant improvement	7047	693	88.5	98.3	2.0	3.0	0.003	0.12
All	Asian	NO SIGNIFICANT DIFFERENCE, improvement	1359	126	85.0	88.3	2.0	2.0	0.657	0.04

Computer generated statistical conclusion
(all you need if you don't like maths!)

the p-value

The effect size

Māori cross-sectional comparison

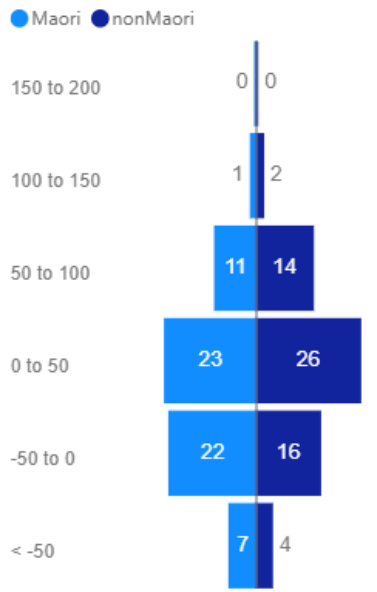
Reading	T1 to T4

How does the progress of Māori compare to males? Has this changed?

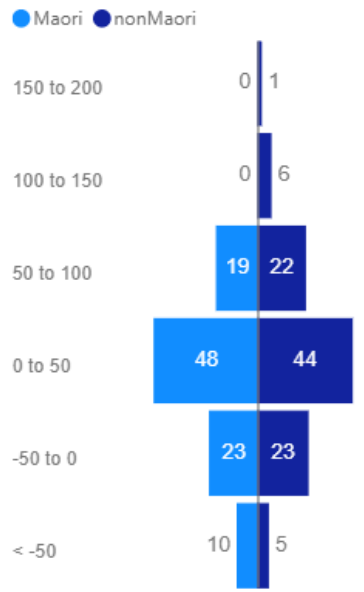
Year Level: All

YL	Cohort	Statement	n (M)	n (nM)	mean Δ (M)	mean Δ (nM)	med Δ (M)	med Δ (nM)	p	ES
All	this year	NO SIGNIFICANT DIFFERENCE, nonMaori > Maori	31	243	14.8	28.2	1.0	1.0	0.130	-0.29
All	previous years	moderate highly significant difference, nonMaori > Maori	178	1080	6.4	20.0	0.0	1.0	0.000	-0.30
Y07	this year	INVALID TEST, n < 20, nonMaori > Maori	15	116	11.9	32.9	1.0	1.0	0.091	-0.47
Y07	previous years	moderate significant difference, nonMaori > Maori	82	556	5.7	19.0	0.0	1.0	0.016	-0.29
Y08	this year	INVALID TEST, n < 20, nonMaori > Maori	16	127	17.5	23.9	1.0	1.0	0.609	-0.14
Y08	previous years	large highly significant difference, nonMaori > Maori	96	524	7.0	21.1	0.0	1.0	0.005	-0.31

Progress on SS (%) previous years



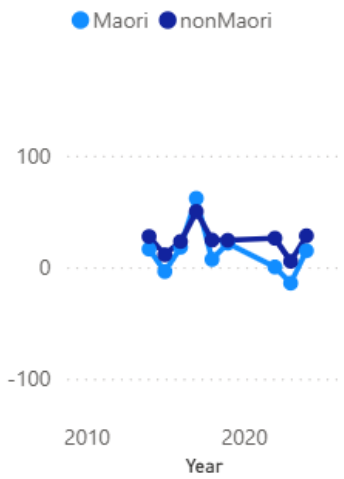
Progress on SS (%) this year



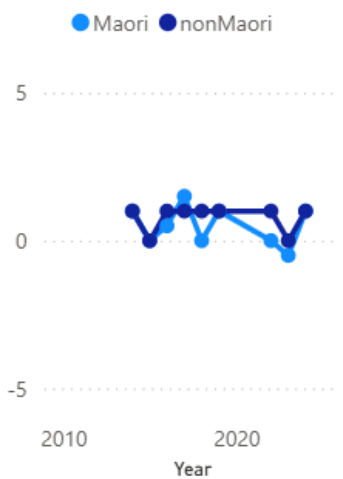
Search

Year	NSN	Student Name	Maori	SS Δ	CSL Δ
2016	100001	Student, 100001	nonMaori	-58	-3
2016	100002	Student, 100002	nonMaori	16	1
2024	100002	Student, 100002	nonMaori	71	2
2017	100003	Student, 100003	Maori	59	2
2019	100004	Student, 100004	nonMaori	129	3
2015	100005	Student, 100005	nonMaori	124	4
2017	100010	Student, 100010	nonMaori	7	0

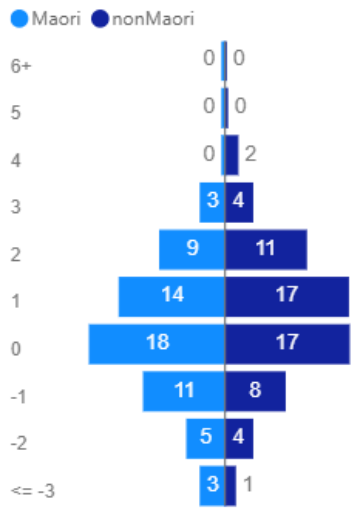
Mean SS Progress by Year



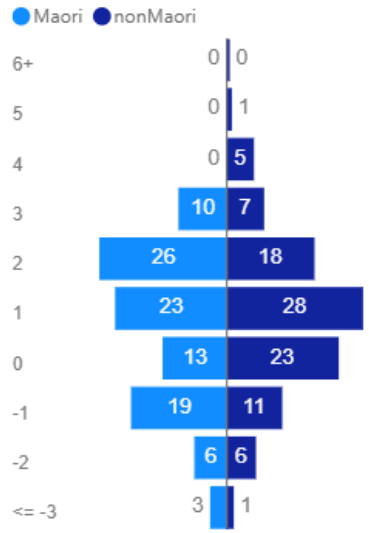
Median CSL Progress by Year



Progress on CSL (%) previous years



Progress on CSL (%) this year



Analysis of means, Anom

- 1 independent variable with 3+ levels: Many rooms
- 1 parametric scale dependent variable: Progress on scale score

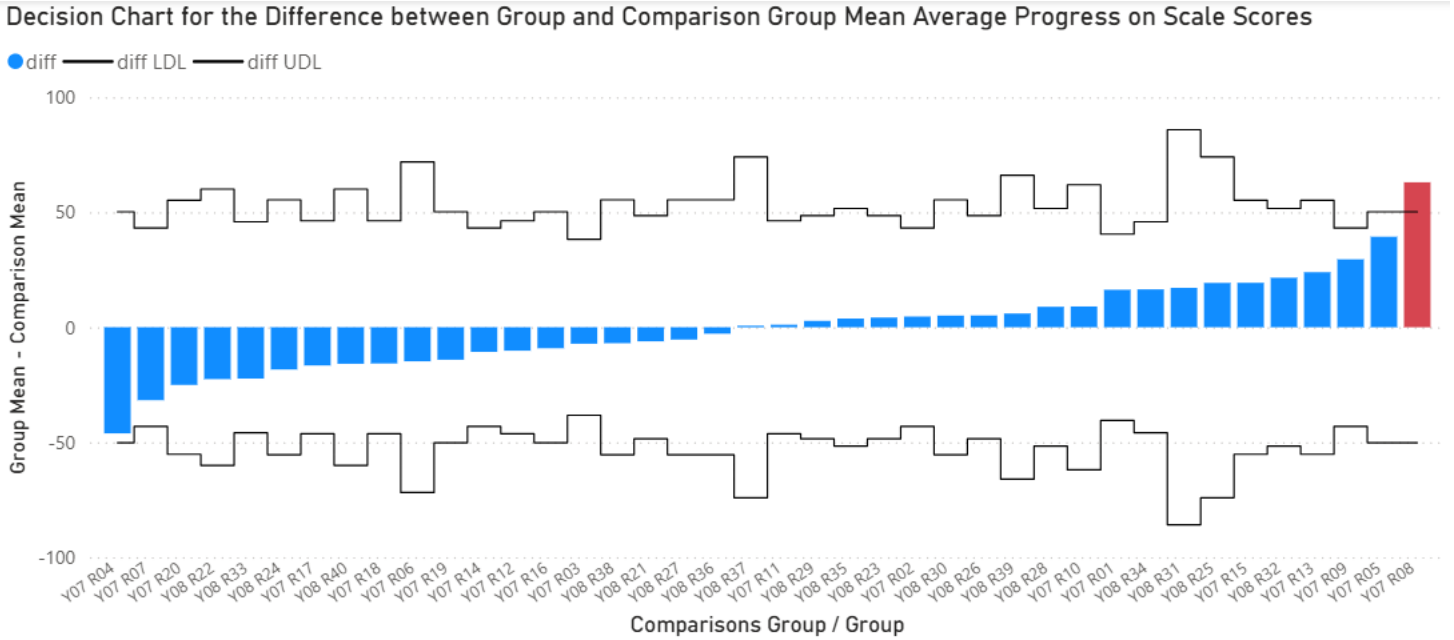
Comparative progress (Teaching Group)

Reading	T1 to T4

Group Type
- (●) Room
- ( ) YearLevel

Cohort
- ( ) previous years
- (●) this year

Do any Groups stand outside the normal range of progress for the Comparison Group?



Cohort	Comparison Group	Group	Statement	n (G)	mean ΔSS (G)	mean ΔSS (CG)	LDL	UDL	diff ΔSS	diff LDL	diff UDL	med ΔCSL (CG)	med ΔCSL (G)	diff ΔCSL
previous years	Y08	R38	Group mean significantly greater than Comparison mean	34	44.2	18.9	-3.6	41.4	25.3	-22.5	22.5	1	1	0
previous years	Y08	R21	Group mean similar to Comparison mean	29	9.0	18.9	-5.6	43.4	-9.9	-24.5	24.5	1	0	-1
this year	Y08	R21	Group mean similar to Comparison mean	9	17.1	23.2	-25.3	71.7	-6.1	-48.5	48.5	1	0	-1
previous years	Y08	R22	Group mean similar to Comparison mean	29	20.2	18.9	-5.6	43.4	1.3	-24.5	24.5	1	1	0
this year	Y08	R22	Group mean similar to Comparison mean	6	0.7	23.2	-36.9	83.3	-22.6	-60.1	60.1	1	0	-1
previous years	Y08	R23	Group mean similar to Comparison mean	30	15.9	18.9	-5.2	43.0	-3.0	-24.1	24.1	1	0	-1
this year	Y08	R23	Group mean similar to Comparison mean	9	27.4	23.2	-25.3	71.7	4.2	-48.5	48.5	1	2	1
previous years	Y08	R24	Group mean similar to Comparison mean	33	23.1	18.9	-4.0	41.8	4.2	-22.9	22.9	1	1	0
this year	Y08	R24	Group mean similar to Comparison mean	7	4.9	23.2	-32.2	78.6	-18.4	-55.4	55.4	1	0	-1
previous years	Y08	R25	Group mean similar to Comparison mean	25	3.5	18.9	-7.6	45.4	-15.4	-26.5	26.5	1	0	-1
this year	Y08	R25	Group mean similar to Comparison mean	4	42.5	23.2	-50.9	97.3	19.3	-74.1	74.1	1	1	0
previous years	Y08	R26	Group mean similar to Comparison mean	34	18.8	18.9	-3.6	41.4	-0.1	-22.5	22.5	1	0	-1
this year	Y08	R26	Group mean similar to Comparison mean	9	28.3	23.2	-25.3	71.7	5.1	-48.5	48.5	1	1	0

One sample median test

- 0 independent variable: theoretical comparison
- 1 ordinal dependent variable: change in Curriculum Sub-Levels

National Comparison

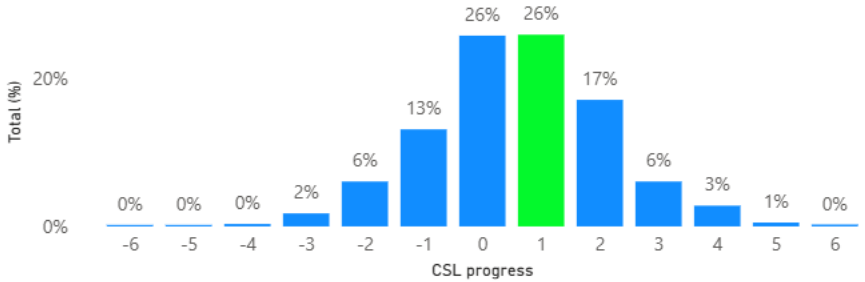
Reading	T1 to T4

Expected CSL Δ
1.0

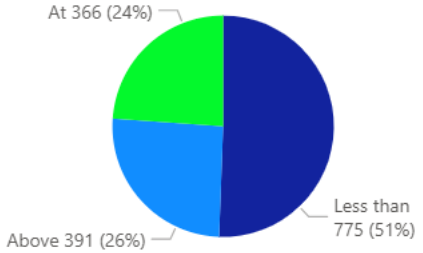
Gender: All
Ethnicity: All
Year Level: All
Cohort: All

Are students making the expected progress?
Do any sub-groups stand out

Curriculum Sub-Level progress (%)



Making the Expected Progress



YL	Cohort	Dem	Statement	N	Exp Δ	med Δ	med diff	p	ES
All	previous years	All	moderate highly significant difference, School < expectation	1258	1.0	1.0	-1.0	0.000	-0.27
All	previous years	Asian	moderate highly significant difference, School < expectation	249	1.0	1.0	-1.0	0.000	-0.20
All	previous years	Female	moderate highly significant difference, School < expectation	638	1.0	1.0	-1.0	0.000	-0.25
All	previous years	Male	moderate highly significant difference, School < expectation	620	1.0	1.0	-1.0	0.000	-0.30
All	previous years	Maori	moderate highly significant difference, School < expectation	178	1.0	0.0	-1.0	0.000	-0.36
All	previous years	MELAA	moderate highly significant difference, School < expectation	61	2.0	0.0	-1.0	0.004	-0.38
All	previous years	Other	invalid test, n < 20, School < expectation	16	2.0	1.0	-0.5	0.791	-0.25
All	previous years	Pacific	moderate highly significant difference, School < expectation	199	1.0	0.0	-1.0	0.000	-0.30
All	previous years	Pakeha	moderate highly significant difference, School < expectation	846	1.0	1.0	-1.0	0.000	-0.26
All	this year	All	small highly significant difference, School > expectation	274	0.0	1.0	0.0	0.000	0.12
All	this year	Asian	NO SIGNIFICANT DIFFERENCE, School > expectation	47	0.0	1.0	0.0	0.185	0.11
All	this year	Female	small highly significant difference, School > expectation	142	0.0	1.0	0.0	0.002	0.13
All	this year	Male	NO SIGNIFICANT DIFFERENCE, School > expectation	132	1.0	1.0	0.0	0.073	0.11
All	this year	Maori	NO SIGNIFICANT DIFFERENCE, School > expectation	31	0.0	1.0	0.0	0.678	0.05
All	this year	MELAA	invalid test, n < 20, School > expectation	18	0.0	1.0	1.0	0.039	0.27
All	this year	Other	invalid test, n < 20, School < expectation	6	1.0	0.5	0.0	1.000	-0.13
All	this year	Pacific	NO SIGNIFICANT DIFFERENCE, School > expectation	46	0.0	1.0	0.0	0.458	0.07
All	this year	Pakeha	small highly significant difference, School > expectation	180	0.0	1.0	0.0	0.001	0.14
Y07	previous years	All	moderate highly significant difference, School < expectation	638	2.0	1.0	-1.0	0.000	-0.45
Y07	previous years	Asian	moderate highly significant difference, School < expectation	123	2.0	1.0	-1.0	0.000	-0.44
Y07	previous years	Female	moderate highly significant difference, School < expectation	335	2.0	0.0	-1.0	0.000	-0.43
Y07	previous years	Male	moderate highly significant difference, School < expectation	303	2.0	1.0	-1.0	0.000	-0.46
Y07	previous years	Maori	large highly significant difference, School < expectation	82	2.0	0.0	-1.5	0.000	-0.50
Y07	previous years	MELAA	large highly significant difference, School < expectation	35	2.0	0.0	-2.0	0.000	-0.54

Linear regression

- **Single or multivariate Linear Regressions**
  - **Independent variables:** 1+ continuous independent variables e.g., scale scores; binary variables e.g., cohort, gender, non/Māori, etc
  - **Dependent variable**: 1 continuous variable e.g., IB Points
- Explains that increases in or presence of independents causes increase in continuous dependent
- Various conditions including **normal residuals**, **homoscedasticity** (variance of residuals should be equal), **multicollinearity** (low correlation between independents).
- Scale Score vs Scale Score or Scale Score vs credits would satisfy these conditions but Scale Score vs Ordinal NCEA Qualificions would not

Logistic ordinal regression

- Logistic ordinal regressions investigate the relationship between ordered responses and a set of explanatory variables.
  - Ordinal outcome variables: NCEA Level 3 and University Entrance Qualifications
  - A range of predictor variable types: parametric (PAT and e-asTTle scale scores), ordinal (effort scores) and binary (cohort, gender, ethnic heritages, …)
- Increase in or presence of independents cause increase in the probability of higher outcomes

$\hat{Y}_i = \frac{e^u}{1-e^u}$ where $\hat{Y}_i$ is the estimated probability that the *i*th case (*i* = 1, … n) is in one of the ordinal categories

where $u = \mathrm{A} + B_1X_1 + B_2X_2 + \ \ldots + B_kX_k$ with constant *A*, coefficients B_j , and predictors X_j , for *k* predictors (*j = 1, 2, 3, … k*).

The linear regression equation is the natural logarithm of the odds ratio to the predictors

$$ln\left(\frac{\hat{Y}_i}{1-\hat{Y}_i}\right) = \mathrm{A} + \sum B_jX_{ij}$$

where the goal is to find the *best* linear combination of predictors to maximise the likelihood of obtaining the observed outcome frequencies of the ordinal variable.

(Tabachnick & Fidell, 2014)

Assumptions and conditions

- **Optimisation technique**: Fisher's Scoring (stepwise)
- **Model Convergence Status (relative gradient convergence criterion)**: the maximum likelihood algorithm has converged
- **Score Test for the Proportional Odds Assumption**
- **Model Fit Statistics (Akaike Information Criterion**) i.e., the stepwise addition signifiantly improves the model
- **Likelihood ratio**: i.e., at least one coefficient is not equal to 0.
- **Adjusted / Pseudo R-sq (Nagelkerke)**
- **Gamma** (predicted vs observed)

The Ordinal Logistic Regression Model linking **NCEA Level 3** to Year 7 Term 1 e-asTTle Reading and demographics.

Actual vs Predicted Qualifications

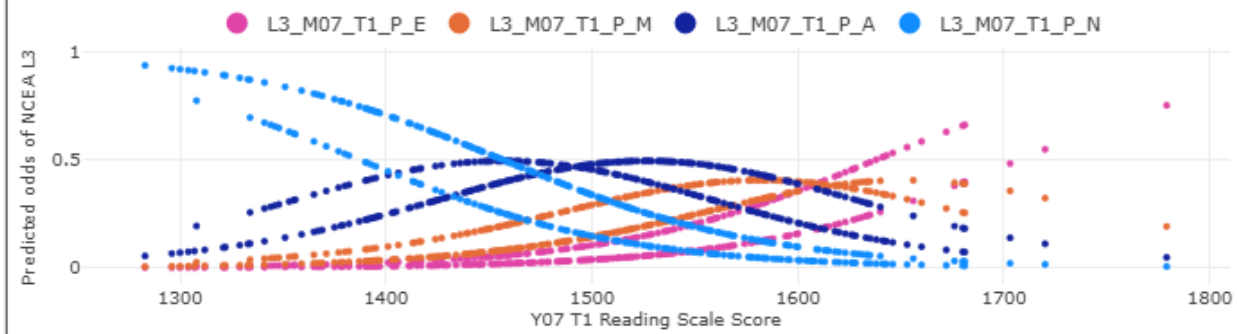
Actual \ Predicted	N	A	M	E	Total
N	13.5%	14.6%	1.2%		**29.3%**
A	8.1%	22.7%	4.9%		**35.7%**
M	1.9%	11.1%	7.9%	1.4%	**22.3%**
E		4.2%	4.2%	4.2%	**12.7%**
Total	**23.6%**	**52.5%**	**18.3%**	**5.6%**	**100.0%**

Correct predictions

Summary of Odds

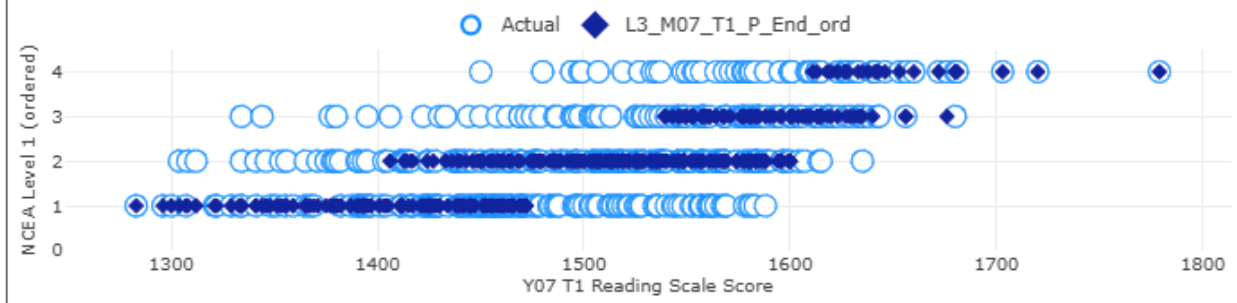
Predicted Grade	N	Min	Mean	Max
N	569	0.01	0.28	0.94
A	569	0.05	0.38	0.49
M	569	0.01	0.23	0.41
E	569	0.00	0.12	0.75

Predicted Odds of NCEA L3 Qualification using Y09 T1 e-asTTle Reading



Actual and Predicted NCEA L3 Qualification using Y09 T1 e-asTTle Reading

N = 1, A = 2, M = 3, E = 4



L3_Cohort

- ☐ previous years
- ☐ Kāhui Ako

Gender

- ☐ Female
- ■ Male

Reading Curriculum Level

2A	2B	2P	3A	3B	3P	4A	4B	4P	5A	5B	5P	6A	6B	6P
2.8%	0.2%	3.0%	5.4%	4.0%	4.2%	26.2%	10.0%	10.9%	3.9%	16.9%	8.4%	0.5%	2.3%	1.2%

Reading Scale Score

1244 — 1779

Ordinal logistic regression

Dependent	N	adj. R^2	p	B1_p	B1_odds	B1_LCL	B1_UCL	B2	B2_p	B2_odds	B2_LCL	B2_UCL	B2_var	B2_p	B2_odds	B2_LCL	B2_UCL
NCEA Level 3	1188	0.37	0.000	0.000	1.016	1.014	1.017	L3 Cohort	0.000	2.947	2.351	3.693	Female_Ord	0.002	1.419	1.142	1.764

The Ordinal Logistic Regression Model linking **University Entrance** to Year 7 Term 1 e-asTTle Reading and demographics.

Actual vs Predicted UE

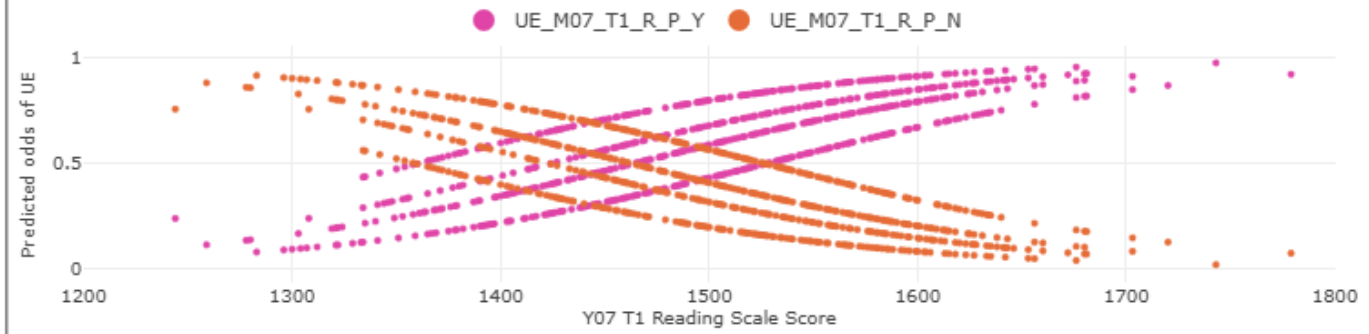
Actual \ Predicted	N	Y	Total
N	17.3%	21.0%	**38.2%**
Y	10.5%	51.3%	**61.8%**
Total	**27.8%**	**72.2%**	**100.0%**

Correct predictions

Summary of Odds

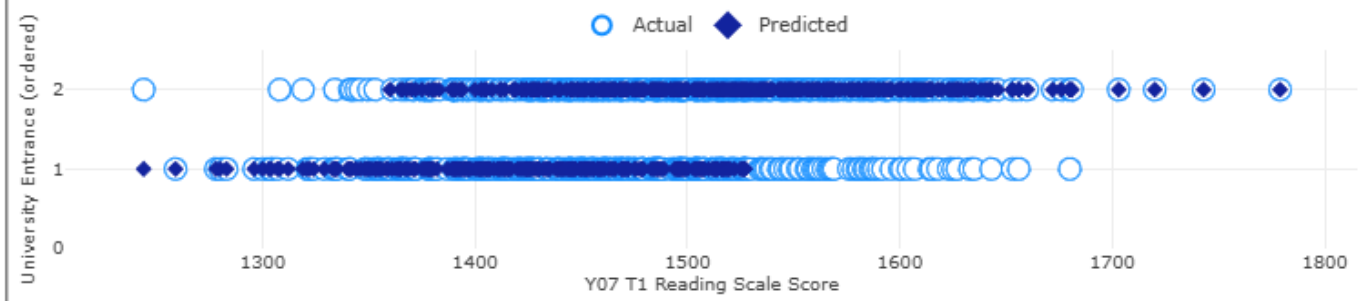
Predicted Grade	N	Min	Mean	Max
P_N	1188	0.02	0.38	0.92
P_Y	1188	0.08	0.62	0.98

Predicted Odds of UE Qualification using Y09 T1 e-asTTle Reading



Actual and Predicted UE Qualification using Y09 T1 e-asTTle Reading

N = 1, Y = 2



L3_Cohort

- [ ] previous years
- [ ] Kāhui Ako

Gender

- [ ] Female
- [ ] Male

Reading Curriculum Level

2A	2B	2P	3A	3B	3P	4A	4B	4P	5A	5B	5P	6A	6B	6P
3.5%	0.4%	2.6%	6.1%	3.5%	5.2%	25.3%	8.7%	12.7%	3.5%	15.5%	9.5%	0.4%	2.3%	0.8%

Reading Scale Score

1244 — 1779

Ordinal logistic regression

Dependent	N	adj. R^2	p	B1	B1_p	B1_odds	B1_LCL	B1_UCL	B2	B2_p	B2_odds	B2_LCL	B2_UCL	B3	B3_p	B3_odds	B3_LCL	B3_UCL
University Entrance	1188	0.23	0.000	Y07_eas_Rdg_SS_T1	0.000	1.010	1.008	1.012	L3 Cohort	0.000	2.795	2.129	3.670	Female_Ord	0.000	1.881	1.451	2.438

Logistic ordinal regression equations

$$\text{Prob(E v M, A, N)} = \hat{Y}_i = \frac{e^u}{1-e^u}$$

$$= \frac{e^{(-7.1110 + (0.0786 \times Y08\_PAT\_Mat\_SS\_T1))}}{1-e^{(-7.1110 + (0.0786 \times Y08\_PAT\_Mat\_SS\_T1))}}$$

$$\text{Prob(E , M v A, N)} = \hat{Y}_i = \frac{e^u}{1-e^u}$$

$$= \frac{e^{(-4.8481 + (0.0786 \times Y08\_PAT\_Mat\_SS\_T1))}}{1-e^{(-4.8481 + (0.0786 \times Y08\_PAT\_Mat\_SS\_T1))}}$$

Feedback data collected from 23 middle leaders

Dashboards are useful and easy to use!

How confident are you in your ability to use and interpret the DASHBOARD:	Confident or very confident
Independently?	61%
With coaching?	100%

How useful is the data in the DASHBOARD for:	Useful or very useful
understanding student achievement and progress?	91%
evaluating teaching/learning initiatives?	78%
informing future actions?	83%

QUESTION: Does a Wānanga day enhance student outcomes?

HERES WHAT HAPPENED IN 2019:

Wānanga Day and the MCAT
91027 Apply algebraic procedures in solving problems

Wānanga?
In essence students were given a day of supervised preparation for the 2019 MCAT. The MCAT has had a lot of media coverage over its difficultly through students and teachers being underprepared for the external.

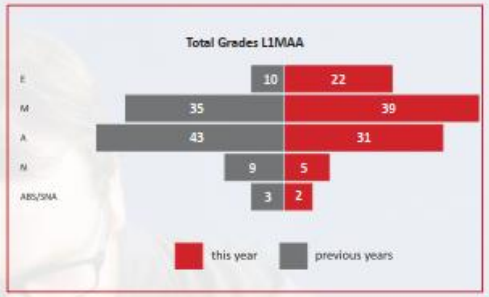
Teachers and NCEA students outraged over difficult Level 1 MCAT algebra exam – ***Stuff, 2016***
7 hacks for surviving NZ's toughest NCEA exam – ***Scoop, 2017***

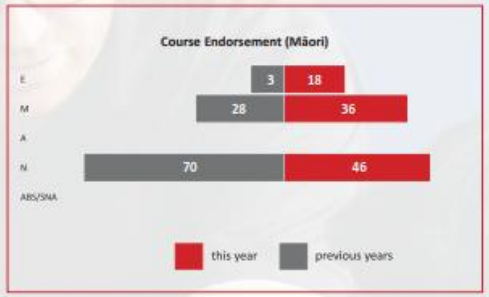
Methodology
After hearing positive results from the Art and Māori departments in instituting Wānanga days for 2019 the Mathematics department decided for all students sitting the external MCAT to have a day of organised preparation.

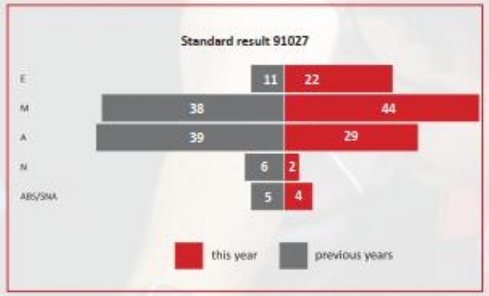
- [redacted]used the library and ran various presentations and activities for students the day prior to the external.

- Students where supplied lunch during a break.

Results Summary
- L1MAA students gained better overall grades in all topics. This difference was highly significant (p = <0. 0001).
- Course endorsement improved overall and this improvement was due to Māori students. This was large and significant (p = 0. 002, G = 0.475).
- MCAT showed significant improvement of grades earned including double the percentage of students gaining excellence (p = 0.016).







Conclusion
- Using a Wānanga day has enhanced student outcomes.

It is an opportunity to establish a learning focused relationship between students, and teachers. It is a time to fully focus on a standard without external distractions.
Although a significant advantage was established it was only for one standard, the results indicate that this practice should be promoted within the department and school. It would require a major change to the timetable.

GOAL: Was to significantly lift the rate of higher grades and endorsements at Level 1.

HERES WHAT HAPPENED IN 2019:

Highlights
- 90914 ***(IN)*** - Historically 19% Non achieved. 2019 - Large, highly significant improvement.

- 90916 ***(EX)*** - Historically 23% - SNA, 9% NA. 2019 - Similar to previous years.

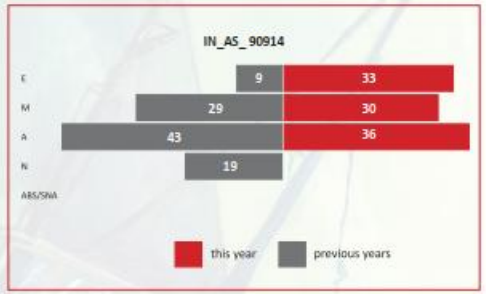
What did we do?
- The Wānanga day was a huge success. 100% of the students achieved their internal allowing us to move as a majority onto the external submissions (closure/peace of mind/admin done).
- Proposition had a clear direction. ***Unpack Arrival***
- Collaboration - Two classes - SM/WO

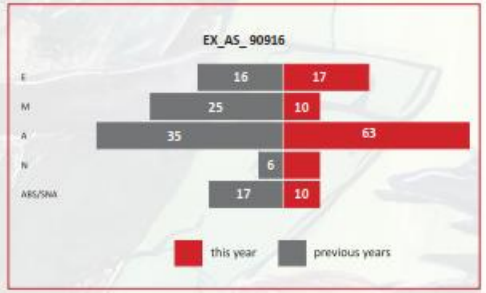
Work on ...
- We achieved our goal from 2019 for the internals by significantly improving both earned and quality of grade. However there was a shift in the externals both up and down.

Externals
SNA - DOWN 7%, N/A - DOWN 6%,
Achieved - UP 28%, Merit - DOWN 15%
Excellence - UP 1%

- As with the internals from 2019, the externals show a ***large portion*** of students gaining achievement. We have not yet fixed this issue!

- Externals are a ***school concern*** and our subject is no different.





Data presentations are useful and have an impact!

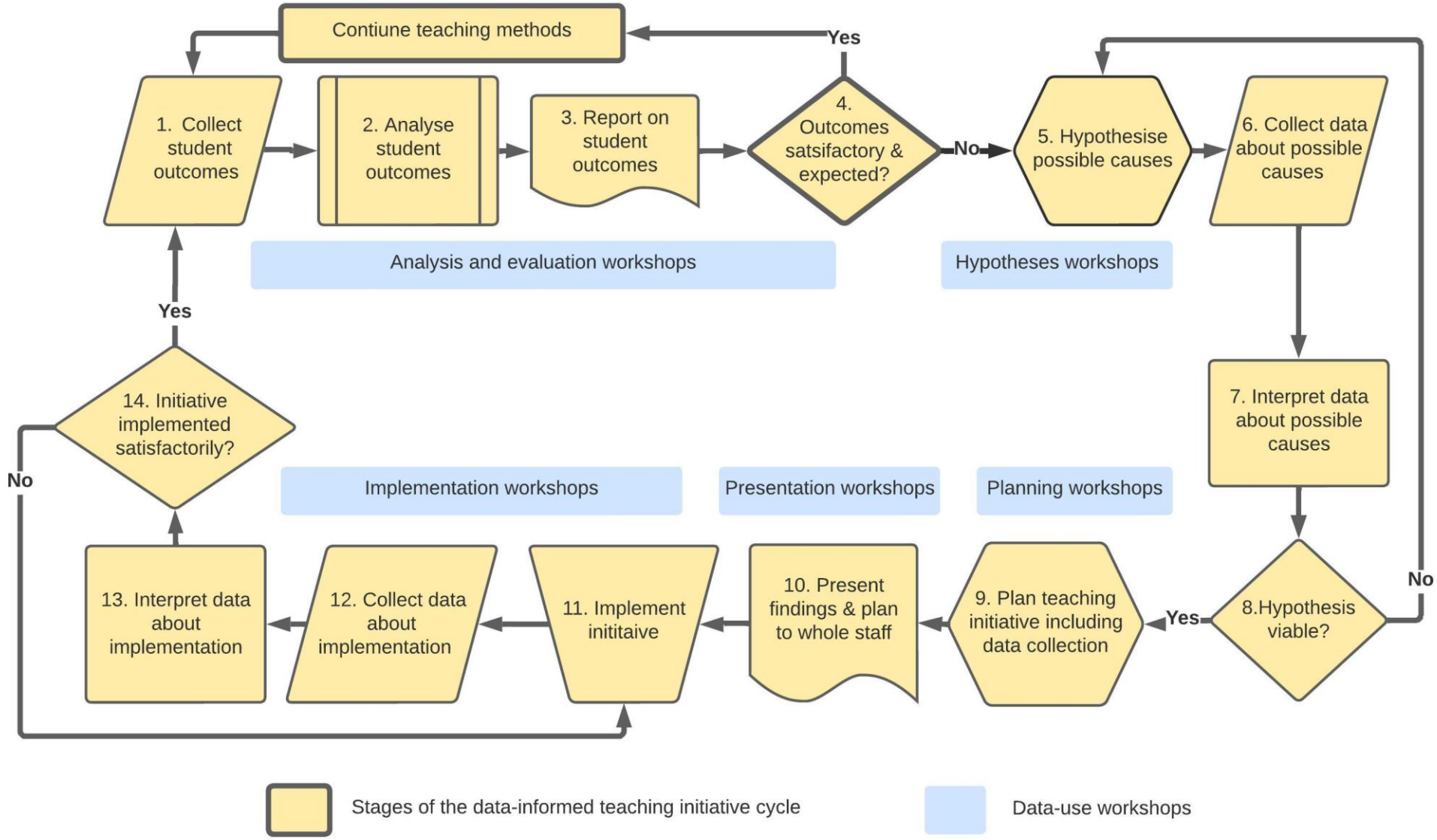
How useful were the poster and presentations for:	Useful or very useful
Clarifying your departmental challenges, goals and plans?	78%
Sharing your departmental challenges, goals and plans?	83%
Understanding the challenges, goals and plans of the other departments?	83%

Survey question	Impact or large impact
Do you think your POSTER PRESENTATION had an impact on your teaching colleagues?	61%

Leaders are confident and value data-informed teaching initiatives

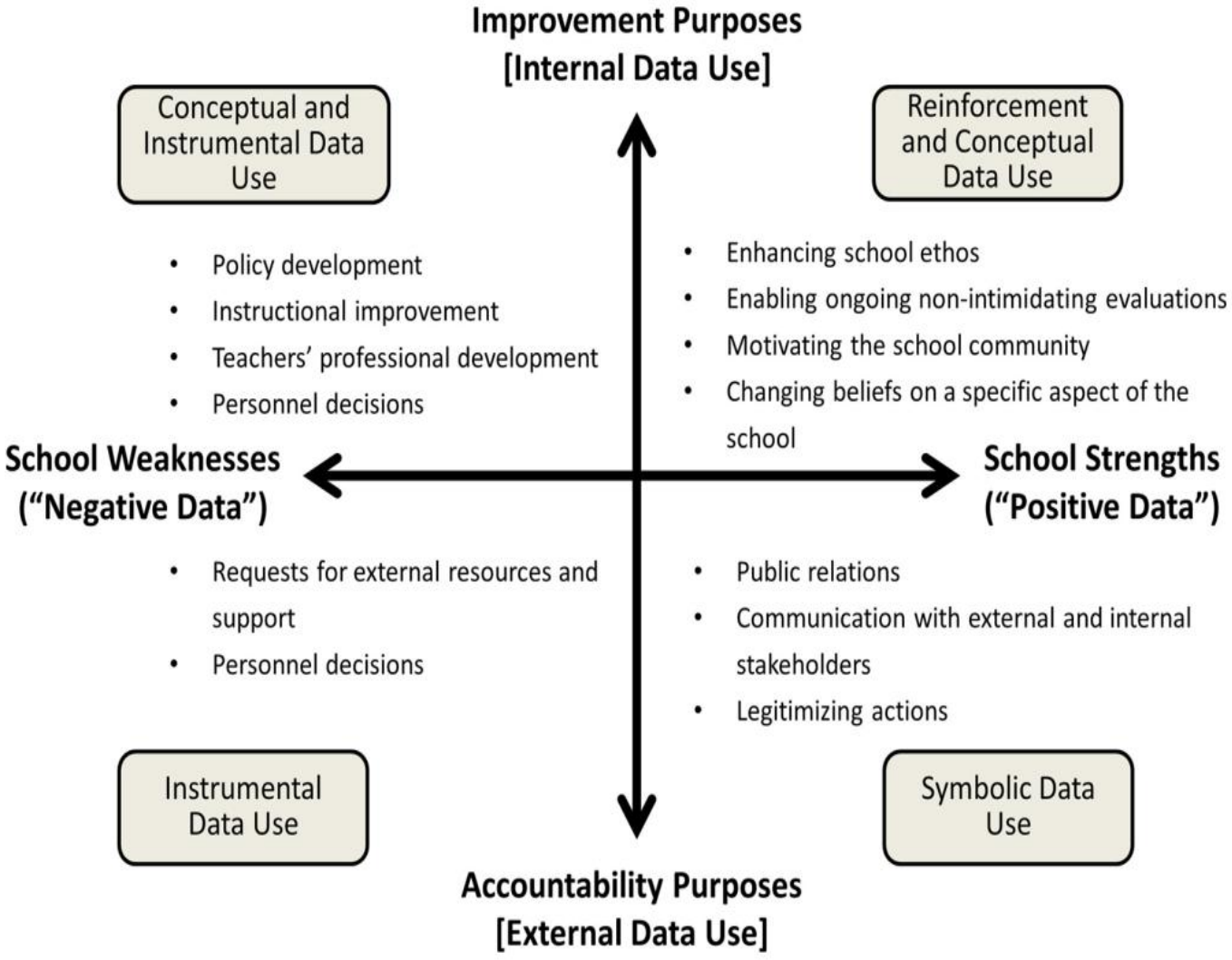
Survey Question	x or very x
How **confident** are you that your evidence-based initiative will lead to improvement?	78%
Were the outcomes of this process **worth** the effort you had to put in?	78%
How **likely** are you to recommend this process to colleagues at other schools?	78%
How **keen** are you to engage in this process again next year?	78%

The data-informed teaching initiative cycle



(Lai & Schildkamp 2013, Mandinach & Gummer, 2015, Schildkamp et al 2018)

The desirable uses of data in schools



(Aderet-German & Ben-Peretz, 2020)

- instrumental use – to directly influence actions and practice;
- conceptual use – to influence mental models and activity over time;
- persuasive or symbolic use – for a predetermined agenda
- reinforcement – to consciously highlight individual and school strengths.

The undesirable uses of data in schools

- non-use – ignoring or cursory data use i.e., ticking the box;
- interpreting data incorrectly – e.g., insufficient data literacy leading to inappropriate decisions or actions;
- narrowing the curriculum or teaching to the test;
- falsifying data – e.g., giving inappropriate assistance before or during assessments;
- educational triage and reshaping the data pool – focusing only on those on the threshold of passing and transferring less able students to non-examined classes;
- bullying or shaming teachers, leaders and schools

Data is misused when:

- Data use misrepresents reality to stakeholders;
- Data use has negative consequences for students;
- Data use removes the opportunity for the improvement of teaching.

(Booher-Jennings, 2005; Buly & Valencia, 2002; Ehren & Swanborn, 2012; Schildkamp & Ehren, 2013; Ford, 2018; Hardy & Lewis, 2017; Susan, 2016; Datnow & Park, 2018; Volante et al., 2020; Lockton et al., 2020; Bertrand & Marsh, 2021)