

Understanding CEM Data

1. The data you receive (see anonymised profile)

2. What teachers receive and place in their electronic mark books

A Year 7 pupil:

Pupil name	Target Grade	Vocab	Maths	Non-Verbal	Skills
Jane Doe	7/8	111	124	122	104

A Year 12 student:

Pupil name	Target grade	Vocab	Maths	Non-Verbal
Ann Other	B/C	103	85	86

Note that the skills score is absent from the Y12 test. This only appears on Y7 and Y9 tests.

3. How can this data be useful for teachers and parents?

CEM data is not a replacement for individual face-to-face knowledge of a pupil, but it can be helpful when considering any of the following questions:

- (i) What is this pupils' progress like in each subject? Is it about as expected, high, or low given their baseline data?
- (ii) Why does this pupil behave and perform in the way that they do?
- (iii) What kind of action or advice might be useful in helping to raise this pupil's attainment in my subject?

- (iv) Should I request to alter the target grade in my subject for this pupil?
- (v) Might this pupil have an underlying special educational need?
- (vi) Is this pupil making sensible options choices?

4. How robust is this data?

No set of data drawn from a single test could ever claim to be perfect, but anyone who has worked with CEM data over a number of years will attest to its reliability.

The two key points are...

- (1) The underlying data set is very robust. CEM tests have been running since 1991 and just over 1200 schools nationally use them. Assuming an average school size of around 600 this produces a data set of just under 19 million sets of individual pupil data.
- (2) The tests are computer adaptive. That is, the next question a pupil is faced with is determined by the speed and accuracy of their previous answers. This makes the resulting bell curve of data much more flexible and finely tuned to individual variations in ability.¹

Having said this, one should always assume a standard deviation of up to 15 points depending on human factors such as mood, confidence and health on the day of the test. That is why it is useful to run such tests (as we do) in years 7, 9 and 12 and in Prep). That is also why CEM data should never be used in isolation.

5. What do the different types of score mean?

Vocab is based on tests that require pupils to locate synonyms and antonyms and select the most appropriate word to place into the gap in

¹ To give you some idea of how significant this point is, since 1991 only three individuals have managed to answer all the questions on their CEM test accurately.

a sentence. A pupil's vocab score will reflect how much they read for pleasure but can also be affected by dyslexia and (obviously) being a non-native English speaker.

Maths is fairly self-explanatory. This will, to an extent, be influenced by quality of Maths teaching to that point in time.

Non-verbal involves various visual tests, such as identifying a 3D shape that has been rotated through 180 degrees, or spotting the next step in a visual sequence. This section of the test is the most environment proof². This score also rises in significance the higher up the educational age range you go.³ That is, a low NV score can become much more of a problem in Sixth Form than it was in Years 7-11.

Skills (Y7 and Y9 only) involves proof reading and rapid checking for accuracy. On this section of the test, questions are on a timer. This tests attention to detail under timed conditions. The Majority of pupils score slightly lower on skills than on the average for their other scores. A reasonably good profile with a noticeably low skills score indicates either a serious tendency to meltdown when timed conditions kick in, or a tendency to laziness. A small minority of pupils score significantly higher on skills, indicating that they already work much better when placed under pressure and may underperform significantly when allowed to relax or given too much time.

6. What do the numbers mean?

The number scores are standardised, so 100 is always the national average. Then, with very slight variations year-on-year, the scores can be grouped into quartiles as follows...

87 and below: D quartile

88-100: C quartile

101-113: B quartile

² That is, the NV score is not as strongly influenced by previous quality of schooling, home life, parenting, or emotional stability and confidence.

³ One of the most important things to look for in a CEM profile is where the NV score is significantly higher than the other scores. This often indicates an intelligent pupil who underachieves due to poor schooling and/or 'chaotic' influences outside of school. To put it more crudely, a pupil with a significantly high NV score in relation to their other scores has been failed by some or all of the adults around them.

95% of all scores fall between 70 and 130. Therefore, 130+ is seen, nationally, as the key data measure for recognising gifted pupils. There is no cap on scores, but scores above 140 are exceptionally rare. Likewise, scores below 70 would indicate someone who probably will not cope within mainstream education.

7. Do people's scores tend to vary much over time?

The short answer is not much. Individual CEM profiles will always vary by a few numbers over the years, but they rarely alter significantly⁴ in a single area, and significant alteration in the overall score is very rare.

However, some scores are more variable than others, due to the influence of educational environment, working habits, and mental health, as follows:

Skills – most variable

Maths – second most variable

Vocabulary – third most variable

Non-verbal – least variable

8. Is this data more valuable for some subjects than others?

No. The main (and obvious) reason for this is that, broadly speaking, the higher a pupil's overall profile the better they will tend to perform in any subject. For example, with very occasional exceptions, an A quartile pupil will do much better than a C quartile pupil in both Maths and Art.

But, of course, individual areas of skill can have a greater or lesser influence on predictions in different subjects. So, for example, a Maths score will influence a science GCSE prediction slightly more than a History prediction.

Interestingly, whilst Vocab and Maths scores both range over a wide number of subjects in terms of their predictive power, only the NV score

⁴ 'Significant variation' can be taken to mean 10 points or more.

has absolutely universal scope. This is because the NV score points to the ability to see patterns in information, spot sequences, adapt your intelligence to unusual demands, and solve problems. A pupil with a very low NV score in relation to their other scores is more likely to fall apart quite easily when presented with unexpected contexts or challenges.

9. How can this data play a part in identifying possible underlying special educational needs?

Here it is important to sound a strong note of caution. There are more reliable ways of determining a special educational need than looking at CEM data. However, what is known as a 'spiky profile' is something that we are encouraged to flag up to SEN.

Most pupil's CEM profiles vary slightly from area to area, but not that much. Whatever their overall ability, the majority of pupils' Vocab, Maths and NV scores will be within about 15 points of one another. However, you do occasionally see a 'spiky profile', which is where there is a big divergence between one or more of these three scores. This can point to a special educational need and is therefore worth the SEN department being aware of as one piece in their general picture of each pupil.

The classic spiky profile would be an NVR score significantly higher than other scores, which is a classic indicator of dyslexia. Although, again, by itself it is not a sufficient indicator of dyslexia.

10. How does the School generate target grades?

CEM produce predicted grades by taking a mean of the individual CEM scores. However, note that the overall CEM score is a mean of the vocab, Maths and Non-Verbal scores with Non-Verbal weighted lower than the other scores⁵. The skills score is excluded from this calculation because, over the longer term, it is seen as highly variable according to

⁵ NV scores are the best indicator of abstract conceptual intelligence, but they are less reliable than Maths and Verbal scores in predicting examination results. Ironically, they are probably a better indicator of long-term career performance.

environment and quality of schooling⁶, and as more susceptible to whether the pupil is having a good or bad day when they take the test.

Broadly speaking, D quartile baseline scores would produce **raw**⁷ predicted grades from 1 up to 3; C quartile from 3/4 up to 4/5; B quartile from 5 to 6 and A quartile from 6/7 up.

Predictions for individual subjects do not tend to vary that much, but they can vary (and, obviously, this variation will be greater in proportion to the variation between individual baseline scores). For example, the pupil (Jane Doe) whose data is on page one has a prediction of 6/7 for every subject except Maths, where the prediction is a 7.

We then add one grade at GCSE and half-a-grade at A Level to whatever the CEM prediction is for each pupil in every subject. These initial predictions are passed to teachers and they can alter them up or down⁸ but only after consultation with senior leadership.

11. An exercise in using CEM data⁹

On the other side of this sheet are the CEM profiles of four pupils (without any target grades). Look at these profiles, for each one, consider the following questions:

- (i) What kind of pupil is this and what would their target grades look like?
- (ii) What potential challenges and pitfalls might this pupil face?
- (iii) What kind of tailored support might be most useful for this pupil?

⁶ To put it another way, it is not seen as a proper 'baseline' ability indicator in quite the same way as the other scores. This opens up a very interesting discussion as to why the ability to focus under pressure is seen as much more malleable than, say, the ability to read.

⁷ 'Raw' means with no value-added (teaching influence) factored in.

⁸ Which they rarely do, and mostly up when they do.

⁹ This is an artificial exercise because one would (and should) never normally use this data in complete isolation, but its purpose here is simply to show what you can glean from the data alone.

- (iv) What advice you might give this pupil regarding GCSE and/or A Level options?

PUPIL A (Year 7)

Pupil name	Target Grade	Vocab	Maths	Non-Verbal	Skills
???		94	94	104	113

PUPIL B (Year 9)

Pupil name	Target Grade	Vocab	Maths	Non-Verbal	Skills
???		78	135	120	123

PUPIL C (Year 9)

Pupil name	Target Grade	Vocab	Maths	Non-Verbal	Skills
???		103	119	114	94

PUPIL D (Year 12)

Pupil name	Target Grade	Vocab	Maths	Non-Verbal	Skills
???		107	87	57	N/A