

## Key elements of a School Self-Evaluation



Self-evaluations in many schools are weak. This is because they:

- Do not evaluate
- Where a school is not outstanding, do not explain why
- Are entirely positive
- Focus on areas which are not important
- Do not appear to understand what an organisational evaluation is.

The impression given by a well-created document is one of leadership competency. It is not the document per se which is important, but that the thinking represented within it is the thinking of the leadership of the school.

### What is a school self-evaluation?

Firstly, what it is not. It is not a Self-evaluation form, unless the school has created a form it completes annually. The 'SEF' is Ofsted terminology from several frameworks ago, when a 'Self Evaluation Form' was required by inspectors, to be completed annually. This document was lengthy, around 68 pages long when it was completed, and took an enormous time out of the headteacher's schedule. Therefore, it is a School Self Evaluation (SSE) which is now completed, and not a SEF.

A self-evaluation is a brief document (recommended maximum of eight pages – some schools have successful SSE documents which are three or four pages long). It looks at important areas for the school and judges how well they are doing, and whether improvement is needed. These might be:

- Areas which have proved to have weakness across the year, such as progress for a particular pupil group
- Areas which have not been evaluated for a time
- Areas of strength, so evidence can be gathered which proves this
- Areas in which the school believes it could do even better
- An analysis of areas where change is required externally, such as for a new government strategy, or as a result of an inspection.

### Being evaluative

Evaluate: 'form an idea of the amount, number, or value of; assess.; assign a value to'

Every sentence should be evaluative; that is, should contain an evaluative word, such as 'good', 'weak' or 'excellent'. This is a judgement based on evidence, and not an opinion. Where an evaluation becomes opinion, then the value given to this opinion will differ from one person to the next. An evaluation is a comparison of one piece of evidence with another, and not a value given against an idea in the mind of the writer. E.g. 'Teachers have high expectations' cannot be substantiated if outcomes at the end of school (KS2 SATs or GCSEs) are below average.

## Using evidence

A statement made without evidence, or without citing evidence, is assertion. Evidence should be gathered which is then analysed ('examine (something) methodically and in detail, typically in order to explain and interpret it.') Conclusions should then be drawn, in order to make an evidence-based judgement.

It is important to note that 'data' means 'a piece of information'. Whilst statistics are a form of data, not all data is statistics. Most information can be given a numerical value. For example, '92% of pupils said they were happy at school'. (NB 'data' is now generally used as both singular or plural)

Types of evidence:

- Quantitative (numerical; measurable by its quantity, rather than its quality) *Example: the % of pupils reaching the expected standard at a specific point*
- Qualitative (non-numerical; measured by its quality, rather than its quantity) *Example: colours, days of the week, brands*
- Primary (collected from an original source)
- Secondary (collected from a secondary source, not directly)
- Discrete (exact values or whole numbers that are not rounded) *Examples: number of pupils in a class, shoe size, the frequency an event happens*
- Continuous (measurements which are rounded) *Examples: time, length, mass, capacity*

It is important not to take a single source of evidence at face value, but to question it and seek corroboration. This is because, if incorrect conclusions are drawn, the school's resources will be allocated to the wrong place.

For example: '73% of pupils attain the expected standard in maths at the end of Y6'. No judgement can be made on this alone, without seeking answers to questions such as:

- How does this compare to the national average?
- How does this compare to data for this cohort in previous years?
- How does this compare to the school's previous Y6 cohort outcomes over the past three measurable years?
- What have been the changes to the cohort recently?
- What is the proportion of vulnerable pupils in the cohort, and what impact did this have?

Until answers are found for these questions, it is not possible to make an accurate judgement on whether 73% at expected in maths is good, weak or excellent.

## Use of language

This should be concise and evidential. It should draw conclusions, but not indicate the actions to be taken (this is for the School Improvement Plan). It should identify the most pressing items for the school to improve.

An example of clear evaluative writing:

'Pupils' attainment across all subjects is an area of concern. During an Evaluation of Overall School Effectiveness visit (Nov 2022), it was identified that the quality of teaching and learning in foundation subjects needed to be further developed as well as approaches to checking what pupils knew and

remembered against the intended learning. An analysis of foundation subject data for 2021-22 showed 29.2% of pupils were working at age related expectations or above across foundation subjects. Lower Prior attainers and disadvantaged pupils had been identified as the most vulnerable groups (Appendix 1).'

Every sentence should be causal: '(this document) considered this evidence, matched against this standardised or desired evidence, and drew this conclusion, making a judgement for how good the school's performance was in this area.'

## **Honesty**

The self-evaluation document is an important process which identifies the areas the school needs to assign its often-scarce resources to. These are:

- Finance
- Human resources (including time and seniority)
- Educational resources (including space and the look of the space)
- Teacher workload

Therefore, it is important to evaluate how well each of these were used in the previous year, whether they represent value for money and whether the strategy they were associated with needs to be ended or changed.

The self-evaluation document is for governors, stakeholders and staff to understand what are identified as the school's greatest strengths and weaknesses, and to be able to have confidence in the process.

The self-evaluation document is not an advertisement for the school, or in extreme cases, and encomium for the wonderful works apparently carried out by the headteacher. In these cases, a number of statements tend to be false. For example: 'The school's curriculum is ambitious', when outcomes are well below the national average.

The document must be honest in its appraisal of the school, its strengths and its weaknesses. There is no point to it if it is not, and at worst, it may cause real harm to the school. It will be one of the key documents requested by inspectors prior to their visit, and a document which lacks honesty is obvious. The best self-evaluations are open in looking at all information available in the areas being appraised, and does not ignore the weaker areas, or make over-inflated claims for apparently stronger areas.

## **Match to external verification**

The most important document to match to is an inspection report. The issues identified in the self-evaluation should be the same as those in the report. Inspectors are highly appreciative of a document which accurately identifies the same areas as the ones they find. There is no point in trying to disguise weaknesses in the school's practice, as inspectors will see the real picture within minutes when they visit.

Consistency: if the school's outcomes (progress and attainment) are not good, then leadership and management cannot be good. Teaching cannot be good, either – if it was, then outcomes, obviously,

would also be good. What is the test of good teaching? Good progress and attainment. The self-evaluation should be aligned with all other school documents, and the judgements within it should align with each other.

There are many other external reports and analyses which can be used to inform the self-evaluation. As well as SAT or exam outcomes, an school improvement partner/ adviser will issue visit reports, or there are externally-verified reports, such as the Wellbeing Award for Schools, investors in People or the Inclusion Quality Mark. Each produce a detailed report. Organisations such as the Fischer Family Trust or PiXL also produce detailed information specific to the school. All should be used to make judgements and cite relevant information.

### **What should be evaluated?**

The things which are most important to the school. Self-evaluation is about confirming the good, as well as identifying the weaknesses, and so all areas should be evaluated over time.

There is too much to evaluate in a single year. If this were to be attempted, it would represent days and days of leadership time, and produce a document of unreadable length, and consequently of little use.

If the school is claiming strength on its website, or in its brochure, then what is the evidence for the claims, or are they assertions rather than claims? Evidence needs to be gathered and evaluated.

In order to prioritise items, the school might look at:

- Progress towards the main objectives, set in the previous year's School Improvement Plan
- Progress towards any issues arising from inspection. These will usually have been developed significantly within the first year after inspection
- Any items arising, or which the school have been alerted to, as being below the standard desired
- Any items arising from external verification (SATs, reports, etc.)
- Any items of change – over the past year, or planned for the next (what is the current situation?)
- Items or areas which have not been looked at for some time
- Areas in which the school believes it can be even better
- Confirmation of the judgements in areas where the school already believes it is good.

Items may also be evaluated on a rolling programme. The danger of not doing that, is that some areas are never fully evaluated.

If every area is good, then the school is likely to be outstanding. Therefore, the most important evaluation is to explain why the school is not outstanding or, if already outstanding, how this should be maintained through identifying areas for further change.

Past evidence is relevant for three years. This usually includes the current year. 2019 SAT and GCSE outcomes are only valid, for example, until 2022. Evidence from before then, including Ofsted outcomes, lack relevance as things will have moved on.

*(cont)*

## The dangers of teacher assessment

Whilst external verification, such as SATs and GCSE outcomes, can give some level of reassurance, teacher assessments are notoriously unreliable. This is for many reasons:

- Historical. E.g. when KS1 SATs were first in place, there was an emphasis on assessing pupils as high as possible. Whilst this has been largely addressed, Y2 are still frequently assessed out of line with other primary year groups.
- Quality of teaching. Stronger teachers tend to be harsher/ more accurate assessors. Therefore, outcomes for their classes can be lower. Weaker teachers tend to assess higher, or to say every pupil in the class has achieved the task objective. Therefore, outcomes for their classes can appear higher.
- Interpretation. When assessing a specific criterion, some teachers will say it is achieved and the pupil is working at the expected standard when they see it in evidence once. Others will require seeing it three times, whilst some will want to see it evidenced all the time to award expected.
- Moderation. External moderation can skew outcomes. Different external moderation teams can give strikingly different interpretations, even in neighbouring schools, and particularly across boundaries. For example, in one Local authority area in 2022, external moderation teams insisted that no pupil was at the higher standard in KS2 writing, often because they had not included 'all the features of writing' in a specified piece, such as a diary entry. In this case, something had clearly gone awry with the process. Schools who had higher level outcomes of 25% in maths and reading, were still achieving 0% in writing after moderation. In another example, in a different local authority, at one point they had some of the highest KS1 outcomes in the country, and some of the lowest KS2. This was not reality, but how external moderation had become inaccurate.

Teacher assessment should give a straight line graph from entry to exit from the school. Ideally, two pupils in every group of thirty should move from below expected to expected, or from the expected level to above the expected standard from the start to end of school. The closer teacher assessment tends towards the mean (hopefully rising) straight line for attainment, the more trusted it can be.

## What represents improvement?

It is unlikely that any improvement which has been put into place will produce measurable outcomes within the first three or four months, particularly in terms of improved progress and attainment. It is likely, however, that some things will be in place, and that a reviewer can see that this is positive. Some improvements take years to come to fruition. For example, it is estimated that to move from RI to outstanding takes eight to twelve years of consistent improvement, with a stable leadership team in which there are the occasional change of members in order to bring new and fresh ideas. Therefore, a change should not be judged negatively if within a short period of time. What is more important is to report on the progress made towards the objective, and not whether it has been met.

It is important to assess improvement against a) SMART success criteria (SC) which were set and b) national expectations. If an SC says 'attainment rises in writing', and it has risen by 1% (or even 0.1%), then the objective is met. If the SC were to say 'attainment rises by at least 5%', then a rise of 1% would not meet the SC.

It is also important only to report on groups. This is defined as 'around five or more'. To report on less is to report on a statistically insignificant number. Less than five should be treated as individuals. A group of around five or more may be:

- A group of 5+ pupils (e.g. a class or a cohort)
- One pupil in one class, three in another, and two in a third (e.g. PP in a small school, or a particular ethnic group)
- One pupil this year, two last year, and three the year before (as above)

### **What activities might inform school self-evaluation?**

As well as external verification, such as inspection, internally, the school may engage in the following activities:

|                        |                                 |                   |
|------------------------|---------------------------------|-------------------|
| Book scrutiny          | Lesson observations             | Learning triads   |
| Learning walks         | Question analysis               | Questionnaires    |
| Pupil interviews       | Working in different classrooms | Planning analysis |
| Curriculum review      | Behaviour statistical analysis  | Audit             |
| Equipment review       | Opinion gathering               | Staff discussion  |
| 1:1 focused discussion | Governor visit                  | Etc.              |

### **Summary**

The best school self-evaluations:

- Inform the school improvement plans because they identify the most important areas for the school to improve
- Are open and honest, and identify weaknesses as well as strengths
- Rely solely on evidence for the identification of its judgements
- Have clarity, and are reasonably brief
- Do not attempt to evaluate everything in one year
- Are a valuable tool for governors and stakeholders to know how well the school is doing
- Are entirely aligned with inspection outcomes, and inform an inspection (or other external verification process) accurately from the outset

### **Variations for a Subject Leader Subject Self-Evaluation**

Most of the principles set out above apply. However, the self-evaluation should be briefer – perhaps two to four sides of A4. Evidence to base judgements on may be limited, as it is unlikely a subject leader will have sufficient time to carry out more than a limited number of evaluation activities. A subject leader should evaluate the evidence they have, and not guess at evidence they do not possess, even if this means an evaluation is not comprehensive. 'Don't know' or 'Unknown' is a perfectly acceptable part of a self-evaluation.