

Reception baseline: criteria for potential assessments

This document provides the criteria that will be used to evaluate potential reception baselines. Suppliers will need to provide evidence against these criteria in order to be accepted onto the approved list of reception baselines. Details of the process that suppliers will need to follow in order to submit a potential reception baseline for evaluation will be published in June 2014. The first deadline for submission of evidence will be in November/December 2014.

Section A of this document provides the criteria against which potential reception baselines will be evaluated.

Section B of this document contains questions which we will ask of suppliers when they submit their evidence but which will not be used to evaluate the potential reception baselines.

Section A

Purpose

The purpose of the reception baseline is to support the accountability framework and help assess school effectiveness by providing a score for each child at the start of reception which reflects their attainment against a pre-determined content domain and which will be used as the basis for an accountability measure of the relative progress of a cohort of children through primary school.

The assessment must demonstrate that it meets the purpose defined by the Department.

If a reception baseline is accepted onto the approved list, the relationship with key stage 1 assessments will be reviewed once children have reached the end of year 2. Reception baselines with a poor relationship will be considered for removal from the approved list.

Content domain

The assessment must have an age-appropriate content domain that is suitable for the range of children s

The Provider must provide evidence that the assessment measures what it purports to measure (validity)

Reliability

Data from the assessment must demonstrate that it is able to provide a sufficiently stable measure that is replicable [eg 95% of second administrations should be within two standard errors of measurement (SEM) of the first administration].

Data from the assessment must demonstrate that it is able to provide a sufficiently stable measure regardless of who is administering and scoring the assessment [eg values of Cohen's Kappa of above 0.8].

Data from the assessment must demonstrate sufficient le

