



Department  
for Education

# **Key stage 1 assessments data collection 2023**

**Technical specification – version 1.0**

**January 2023**

# Contents

Summary	4
Who is this publication for?	4
This specification is for:	4
Main points	4
Version history	5
1. Introduction	6
1.1. Statutory requirement	6
1.1.1. Data sharing	6
1.2. Structure	7
1.3. Scope	7
1.4. Changes for 2023	7
1.5. Coverage and timings	8
2. Structure and guidance	9
2.1. Outline data content	9
2.2. Process	9
2.2.1. Entry of teacher assessment results into school MIS	10
2.2.2. Creation of CTF key stage 1 file by school / setting for transmission to local authority	11
2.2.3. Loading key stage 1 results into the local authority's central management system	11
2.2.4. Creation of individual pupil results XML export file by the local authority for transmission to the department	12
2.3. Additional local authority software facilities	13
3. Data return formats	14
3.1. Special notes for XML returns	14
3.2. Example file structure for XML returns	14
4. Data validation rules	17
4.1. Header validation	17
4.2. Establishment level validation	18
4.3. Pupil level validation	19
5. Resubmission guidance	22
6. Valid values	23

6.1. Local authority codes	23
6.2. Assessment parameters	23

## Summary

This publication provides non-statutory guidance from the Department for Education. It has been produced to help suppliers of local authority and school management information systems to return end of key stage 1 teacher assessments to the Department for Education.

## Who is this publication for?

This specification is for:

- suppliers developing software for use in management information systems (MIS)
- local authorities collecting and submitting the key stage 1 (KS1) assessments data return 2023

## Main points

These specifications enable local authorities - and software suppliers working on their behalf - to prepare and submit data on key stage 1 assessments carried out in 2023.

It contains:

- changes to the previous year's data collection
- what pupil data to collect from each school
- methods for importing the data into local authorities' information systems
- how to structure the data in XML
- how to check the data against the validation rules

This guide must be read in conjunction with the current version of the [common basic data set \(CBDS\)](#). It defines common data items that schools use in MIS software and that we use in our data collections.

## Version history

Version	Change history	Author / date
0.1	Changes to the specification for the 2023 key stage 1 collection have been made as listed below: Roll over of dates All links updated Updated validation rule 1530 with new local authority codes	Marcia Merchant & Jon Leyshon 5 January 2023
0.2	Revised following working group QA	Marcia Merchant & Jon Leyshon 13 January 2023
1.0	Baselined	Jon Leyshon 24 January

# 1. Introduction

This document provides the specification for the return end of key stage 1 teacher assessments to the department for education. Returns are made using the department's COLLECT system and are in XML format.

## 1.1. Statutory requirement

The return of data to the department and its subsequent publication is underpinned by the following legislation:

- [the Education Act 1996 s.537A\(3\)](#)
- [the Education \(School Performance Information\)\(England\) Regulations 2007](#)

### 1.1.1. Data sharing

The UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018 (DPA 2018) provide certain safeguards regarding the use of personal data by organisations, including:

- the department
- local authorities
- schools

The UK GDPR and DPA 2018 detail the rights of those (known as) data subjects about whom data is held, such as:

- pupils
- parents
- teachers

This includes (amongst other information that we are obliged to provide):

- the right to know the types of data being held
- why it is being held, and
- to whom it may be communicated

For the purposes of data protection legislation, the terms 'process', 'processed' or 'processing' apply to any activity involving the personal data, such as:

- collecting
- storing
- sharing
- destroying
- etcetera – please note: this list is not exhaustive

The department provides suggested wording for [privacy notices](#) that schools and local authorities may wish to use. However, where the suggested wording is used, the school / local authority **must review and amend** the wording to reflect local business needs and circumstances. This is especially important, as data will be processed that is not solely for use within data collections.

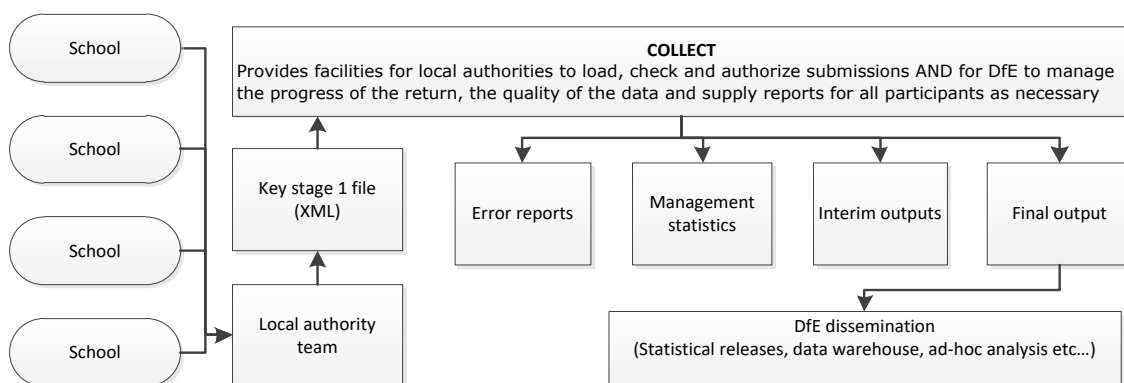
## 1.2. Structure

This requirements specification consists of one document containing a narrative, sample XML messages, and validation. Local authorities and suppliers must read this specification in conjunction with the [common basic dataset \(CBDS\)](#).

## 1.3. Scope

This data collection covers all local authorities in England, who are expected to provide data on all key stage 1 assessments carried out within their authority in 2023. The data consists of pupil level information.

The following diagram outlines the overall scope of the collection:



**Fig 1: high level data collection process**

## 1.4. Changes for 2023

Changes for the 2023 key stage 1 collection technical specification are:

- rolling dates over
- updating links
- updating validation rule 1530 with new local authority codes

## **1.5. Coverage and timings**

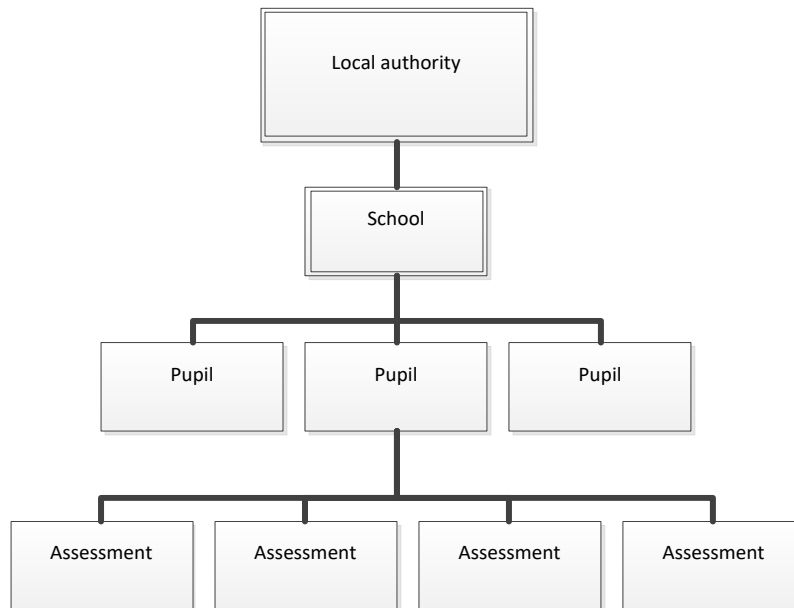
All local authorities are required to capture the data specified for the 2023 key stage 1 assessments. The data will then be collected via the department's COLLECT system between 12 June and 21 July 2023.



## 2. Structure and guidance

### 2.1. Outline data content

Each return from a local authority will consist of a header and one or more school modules (only one school per return as shown in Figure 2), each of which will contain a pupil module with a number of pupil records. Each pupil record will include identifiers and a series of assessment records.



**Fig 2: outline structure of data**

### 2.2. Process

Teachers will complete the end of key stage 1 assessment following the process set out in the [key stage 1 assessment and reporting arrangements 2023](#). The outcomes of those assessments will be reported to their local authority, who will then forward those results to the department.

This year, the department requires the teacher assessment (TA) outcomes in:

- reading
- writing
- mathematics
- science

Schools must report which pupils have been assessed using [the engagement model](#). This is a statutory teacher assessment tool for pupils working below the standard of the national curriculum and not engaged in subject-specific study.

The full list of valid assessment values for each of the above is as shown in assessment component files: [A\\_Comp\\_2023](#).

There are essentially four stages to the capture of key stage 1 pupil assessments once they have been made by teachers:

1. Entry by schools of individual pupils' key stage 1 results (including where the engagement model has been used) into their MIS.
2. Creation of a common transfer file (CTF) for submission to the local authority by secure means.
3. Loading of results (in all formats) into the local authority central database / processing system
4. Creation of school XML export files by local authority software for loading into key stage 1 COLLECT local authority.

Following moderation activities of the teacher assessment data by local authorities, in some cases data may be annulled. Code Q should only be made available in local authority software.

### **2.2.1. Entry of teacher assessment results into school MIS**

Assessment results are entered into the MIS for each subject for each pupil, based on [teacher assessments](#) of pupil attainment. The following subjects will be assessed and reported to the department:

- Reading (ENG REA)
- Writing (ENG WRI)
- Mathematics (MAT MAT)
- Science (SCI SCI)

For each of the above, an outcome should be recorded using the appropriate code in the list in A\_Comp. (Code Q should only be available in local authority software).

Pupils assessed against the engagement model should be recorded using 'EM'. Information about outcomes for these pupils should only be retained locally.

The school MIS software should check that a valid result has been entered for each teacher assessment. Any invalid or missing entries will be reported to the school when the data is entered and the CTF is created.

## **2.2.2. Creation of CTF key stage 1 file by school / setting for transmission to local authority**

Maintained schools, academies and participating independent schools must submit TA data for pupils at the end of KS1 to their geographical local authority by Tuesday 27 June.

This will normally be via a CTF and when using this route, the name of the CTF created by schools will be:

xxxnnnn\_KS1\_xxxLLLL\_yyy.XML where xxx represents the local authority number, nnnn is the school number and yyy is an incremented version number starting at 001.

The task / test results will not be included in the CTF unless the school selects an option to include. Ideally, the option will be defaulted not to include.

## **2.2.3. Loading key stage 1 results into the local authority's central management system**

i) The software will allow the local authority to import into its central database the CTF see section [2.2.2](#) and validate in accordance with the rules in sections [4.1](#) to [4.3](#). Any missing or invalid teacher assessment outcomes must be reported. The software will also warn if any UPNs are missing, or if any dates of birth are missing or incomplete.

ii) In addition, the software will allow the local authority to load into its central database a CSV file created by independent schools from a KS1 Excel independent spreadsheet (to be supplied to local authorities by the DfE for distribution if necessary) with the export filename "KS1\_yyyxxxx\_23.CSV", where yyy is the local authority no and xxxx is the school Estab number (- this number will always start with a 6). A copy of the spreadsheet can be provided when available, if required.

iii) Where any pupil identifying or contextual data are already present in the local authority central database there will be a prompt to ask if the existing data can be overwritten by the CTF or CSV import.

iv) Where a CTF contains any pupil's previous assessment data (for example, earlier year's FSP results) there will be a prompt to ask if the existing assessment data can be overwritten.

v) The local authority will have facilities:

- to remove and to add pupils
- to amend a pupil's identifying data or result levels that have been loaded, and for the new data to be re-validated in accordance with the rules in sections [4.1](#) to [4.3](#)

- to re-import a corrected CTF with an option for the operator to either replace all existing data or add as new data, and be validated in accordance with the rules in sections [4.1](#) to [4.3](#)

#### **2.2.4. Creation of individual pupil results XML export file by the local authority for transmission to the department**

i) On completion of the actions at [2.2.3](#), the software must allow the local authority to create separate XML files of individual pupil data, for each school (and local authorities can only load files relating to whole schools, not for individual pupils) as specified in the example XML file at section [3.2](#), even if still containing errors, for loading into the department's data collection system (COLLECT). Where no assessment result is reported for a subject, all assessment components for that subject must be excluded from the XML file. The XML file will be validated in accordance with the rules in sections [4.1](#) to [4.3](#). Any data failing the validation checks will be reported using the error codes and messages specified. In addition to local authority and school number, the file will contain the following data items for each pupil:

- UPN
- pupil surname
- pupil forename
- DoB
- gender
- each key stage 1 TA result as recorded

Task / test results are not to be included in the XML file even if these are included in the CTF submitted by the school to the local authority.

Please note – An approved 'SoftwareCode' will not be issued. COLLECT will use the 'SoftwareCode' tag value specified by suppliers for this purpose.

ii) The name of the XML file will be xxxLLLL\_KS1\_xxxDfE\_yyy.XML where xxx represents the local authority number, and yyy is an incremented version number starting at 001.

The local authority will be provided with options to include in the export XML:

- a) all schools' individual pupil results, even if these contain errors or have previously been exported (the software will inform the operator if a school is in error or has previously been exported)
- b) selected schools' individual pupil results, even if these contain errors or have previously been exported (the software will inform the operator if a school is in error or has previously been exported)

Please note: Whichever option above is chosen, all individual pupil results within the school are to be exported.

Where a CTF contains any pupil's previous assessment data (earlier years FSP results) there will be a prompt to ask if the existing assessment data can be overwritten.

### **2.3. Additional local authority software facilities**

The software should provide the following additional facilities to allow the local authority to:

- i) produce a list of schools to check that all expected schools returns have been processed
- ii) produce a report for each school or for all schools to show the number of boys and the number of girls and the total number of pupils included in each school's CTF submission
- iii) include in the department's XML file any independent schools from within its own local authority area whose results it has agreed to process
- iv) create separate departmental XML files for any independent schools from outside its own area whose results it has agreed to process
- v) exclude independent schools from its own local authority analysis
- vi) transmit results in batches without the need to resubmit results already transmitted

### 3. Data return formats

A submission file from a local authority / agent will be provided in a XML file structure as described in 3.2. Where an optional item is not included then the associated tag will also be omitted. Where a mandatory item is omitted then an error will be generated.

#### 3.1. Special notes for XML returns

The standard XML entity references must be used for the following special characters:

Character	Entity reference
Ampersand (&)	&amp;
Left angle bracket (<)	&lt;
Right angle bracket (>)	&gt;
Single quote / apostrophe (')	&apos;
Double quotes (“)	&quot;

For special characters such as é use a character reference such as &#233; this produces a generic XML file which can be viewed as such in a suitable browser. Details of all such characters are found within the XML standard documentation; see for example <http://www.w3.org/TR/xml/>, section 2.2.

#### 3.2. Example file structure for XML returns

CBDS ref	XMLFormatLayout	Notes
	<?xml version="1.0" encoding="UTF-8"?>	
	<KS1file>	
	<Header>	
N00600	<Collection>Key Stage 1 Transfer File</Collection>	Fixed value of "Key Stage 1 Transfer File"
N00609	<DateTime>2023-07-23T15:30:47</DateTime>	
N00602	<Year>2023</Year>	Fixed value of '2023'
N00216	<LEA>555</LEA>	Valid values listed in section 6.
N00605	<SoftwareCode>KEYPAS</SoftwareCode>	Suppliers decide on code used by their system
	</Header>	
	<School>	
N00279	<Estab>1234</Estab>	Valid values listed in section 6.
	<Pupils>	
	<Pupil>	A repeatable group of 1 or more

CBDS ref	XMLFormatLayout	Notes
N00001	<UPN>A123456789012</UPN>	
N00003	<Surname>SURNAME 1</Surname>	
N00004	<Forename>FORENAME 1</Forename>	
N00006	<DOB>2016-02-02</DOB>	
N00007	<Gender>M</Gender>	
	<Assessments>	
	<Assessment>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
N00134	<Subject>ENG</Subject>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
N00062	<Method>TA</Method>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
N00063	<Component>REA</Component>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
N00200	<ResultQualifier>NC</ResultQualifier>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
N00064	<Result>EXS</Result>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
	</Assessment>	<Assessment> : A repeatable group of 1 or more depending on the combination of <Subject>, <Method>, <Component> and <ResultQualifier> required. See A_Comp 2023.
	</Assessments>	
	</Pupil>	

CBDS ref	XMLFormatLayout	Notes
	</Pupils>	
	</School>	
	</KS1file>	



## 4. Data validation rules

This section specifies validation rules that are applied to the data by the department, using COLLECT. They must also be applied, as far as possible, by local authorities prior to loading the data onto the COLLECT system and any issues resolved. Those data items included in the specification but not appearing in the validation rules - that is, those not shown in the validation rules as "must be present" - must be supplied where available (for example, child's unique pupil number). The rules also show which validations are errors and which ones are queries. An error is a failed validation check that must be corrected. A query is one which must be investigated, and potentially corrected, as it identifies data that would usually be invalid.

Syntax:

- Validation checks are expressed as conditions that must be true (errors) or that should usually be true (queries). The error message shows what is displayed if the condition fails
- data items are indicated using the XML tag with the common basic data set number in parentheses, for example <UPN>
- data groups are indicated using the XML tag only, for example <Assessments>
- single dates or date ranges are sometimes defined in relation to <ReferenceDate>, which is from the XML header.

### 4.1. Header validation

Seq Number	Error/ Query	Validation Check	Error Message
001	Error	<Collection> must be present with a value of "Key Stage 1 Transfer File"	Document Name missing or invalid. The file must be recreated
101	Error	<SoftwareCode> must be present	Software Code is missing
110	Error	<Year> must be present and = 2023	Year is missing or invalid

## 4.2. Establishment level validation

Seq Number	Error/ Query	Validation Check	Error Message
301	Error	<LEA> must be provided and must be a valid value	Local Authority number missing or invalid
302	Error	<Estab> must be present with a valid value	Establishment No is missing
116	Query	COLLECT ONLY CHECK a query will be raised where the number of boys is fewer than the expected number of boys by more than 5	Number of boys loaded is more than 5 less than expected number
117	Query	COLLECT ONLY CHECK a query will be raised where the number of girls is fewer than the expected number of girls by more than 5	Number of girls loaded is more than 5 less than expected number
118	Query	COLLECT ONLY CHECK In independent schools only, moderation year needs to be one of 2018, 2019, 2022 & 2023 otherwise a query will be raised	Moderation required for 2023

### 4.3. Pupil level validation

Seq Number	Error/ Query	Validation Check	Error Message
1600	Error	<DOB> must be present and in a valid format	Pupil's date of birth is missing or in an invalid format. Format should be CCYY– MM-DD
1601Q	Query	<DOB> should be between 1/9/2014 and 31/8/2017	Pupil's Date of Birth is outside expected date range
1610	Error	<Gender> must be present and a valid value	Pupil Gender missing or invalid
1500	Error	<u>All schools except Independent schools</u> <UPN> (Unique Pupil Number) must be present.	UPN is missing
1510	Error	<UPN> must contain the correct check letter  To calculate the check letter: 1. Multiply the individual digits by their weights as follows: digit 2 by weight 2; digit 3 by weight 3; digit 4 by weight 4; digit 5 by weight 5; digit 6 by weight 6; digit 7 by weight 7; digit 8 by weight 8; digit 9 by weight 9; digit 10 by weight 10; digit 11 by weight 11; digit 12 by weight 12; digit 13 by weight 13. 2. Sum the individual results, divide the total by 23, and take the remainder. 3. Calculate the check letter from the result as follows: 0 = A; 1 = B; 2 = C; 3 = D; 4 = E; 5 = F; 6 = G; 7 = H; 8 = J; 9 = K; 10 = L; 11 = M; 12 = N; 13 = P; 14 = Q; 15 = R; 16 = T; 17 = U; 18 = V; 19 = W; 20 = X; 21 = Y; 22 = Z. For calculating the check letter (see above) any alphabetical character at digit 13 is accorded the same numerical value as listed in paragraph 3 above, A=0, B=1, C=2 etc.	UPN invalid (wrong check letter at character 1)
1520	Error	Each pupil <UPN> must be unique across all pupils in the school	More than one pupil record with the same UPN

Seq Number	Error/ Query	Validation Check	Error Message
1530	Error	If <UPN> is present Characters 2-4 of <UPN> must be a valid post April 1999 LA code or a recognised "pseudo LA" code (001-005, 201-213, 301-320, 330-336, 340-344, 350-359, 370-373, 380-384, 390-394, 420, 660-681, 701-708, 800-803, 805-808, 810-813, 815-816, 820-823, 825-826, 830-831, 835-841, 845-846, 850-852, 855-857, 860-861, 865-896, 908-909, 916, 919, 921, 925-926, 928-929, 931, 933, 935-938, 940-943)	UPN invalid (characters 2-4 not a recognised local authority code)
1540	Error	If <UPN> is present Characters 5-12 of <UPN> must be numeric	UPN invalid (characters 5-12 not all numeric)
1550	Error	If <UPN> is present Characters 13 of <UPN> must be numeric or A-Z omitting I, O and S	UPN invalid (character 13 not a recognised value)
1580	Error	<Surname> must be present	Pupil with last name missing
1590	Error	<Forename> must be present	Pupil with forename missing
113	Error	For each <Pupil> there must be only one <Assessment> for the same combination of <Subject><Component><ResultQualifier>	There are multiple assessments where the Subject, Component, and Result Qualifier combinations are the same.
223	Error	Where <Subject> = ENG and <Component> = REA and <ResultQualifier> = NC there should be a Stage Assessment <Result>	The reading result is missing, it must be one of A, Q, EM, PK1, PK2, PK3, PK4, WTS, EXS or GDS
224	Error	Where <Subject> = ENG and <Component> = WRI and <ResultQualifier> = NC there should be a Stage Assessment <Result>	The writing result is missing, it must be one of A, Q, EM, PK1, PK2, PK3, PK4, WTS, GDS or EXS
231	Error	Where <Subject> = MAT and <Component> = MAT and <ResultQualifier> = NC there should be a Stage Assessment <Result>	The mathematics result is missing, it must be one of A, Q, EM, PK1, PK2, PK3, PK4, WTS, GDS or EXS
232	Error	Where <Subject> = SCI and <Component> = SCI and <ResultQualifier> = NC there should be a Stage Assessment <Result>	The science result is missing, it must be one of A, Q, HNM or EXS

Seq Number	Error/ Query	Validation Check	Error Message
233	Error	If <Subject> concatenated with <Component> and <ResultQualifier> equals 'ENGREANC' or 'ENGWRINC' 'MATMATNC' then <Result> must be A, Q, EM, PK1, PK2, PK3, PK4, WTS EXS or GDS	Result is invalid, it must be one of A, Q, EM, PK1, PK2, PK3, PK4, WTS, GDS or EXS.
234	Error	If <Subject> concatenated with <Component> and <ResultQualifier> equals 'SCISCINC' then <Result> must be A, Q, HNM or EXS	Result is invalid it must be one of A, Q, HNM or EXS.
251	Error	Each <Pupil> record must contain an <Assessment> for each of the following: <Subject> = ENG and <Component> = REA <Subject> = ENG and <Component> = WRI <Subject> = MAT and <Component> = MAT <Subject> = SCI and <Component> = SCI	Each pupil record must contain a teacher assessment for English Reading, English Writing, Mathematics and Science

## 5. Resubmission guidance

When resubmissions are required the Serial No in the Header must be incremented by the provider's software. We have also assumed that the whole file will be resubmitted. Note that a resubmission will delete and replace all the data contained in the previous submission.

## 6. Valid values

Valid item values must be in the list or range specified in the [Common Basic Dataset](#).

### 6.1. Local authority codes

Local authority (previously referred to as 'LEA') codes for this collection only covers ENGLAND local authorities. Coverage also excludes local authority codes:

- 001, 003, 167
- codes starting with 6 and 7 except 702 (MoD schools, which is valid)
- XXX, MMM and NA

### 6.2. Assessment parameters

Assessment parameters - the following combinations of subject, component, assessment method and result type are required:

KS1-2023-ENG-REA-TA-NC

KS1-2023-ENG-WRI-TA-NC

KS1-2023-MAT-MAT-TA-NC

KS1-2023-SCI-SCI-TA-NC



Department  
for Education

© Crown copyright 2023

This publication (not including logos) is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

To view this licence:

visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/3](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3)

email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

write to Information Policy Team, The National Archives, Kew, London, TW9 4DU

About this publication:

enquiries [www.education.gov.uk/contactus](http://www.education.gov.uk/contactus)

download [www.gov.uk/government/publications](http://www.gov.uk/government/publications)



Follow us on Twitter:  
[@educationgovuk](https://twitter.com/educationgovuk)



Like us on Facebook:  
[facebook.com/educationgovuk](https://facebook.com/educationgovuk)