



Standards  
& Testing  
Agency

# **2022 Validity Framework**

## **Multiplication tables check**

### **May 2023**

# Contents

Summary	3
Who is this publication for?	3
Claim 1: The Multiplication tables check is representative of multiplication fluency in the curriculum	4
Claim 2: Assessment results provide a fair and accurate measure of pupil performance	8
Claim 3: Pupil performance is comparable within and across schools	11
Claim 4: Differences in assessment difficulty from year to year are taken account of, allowing for accurate comparison of performance year on year	14
Claim 5: The meaning of test scores is clear to stakeholders	15

## **Summary**

The validity framework for the multiplication tables check (MTC) provides validity evidence gathered throughout every stage of the development of the assessment. It has been produced to help those with an interest in assessment to understand the validity argument that supports the assessment.

## **Who is this publication for?**

This publication is for test developers and others with an interest in assessment.

# **Claim 1: The MTC is representative of multiplication fluency in the curriculum**

## **1.1 The assessable area of the curriculum is clearly defined as a content domain**

The content domain is defined in the [multiplication tables check assessment framework](#) (Section 3, page 7).

The [national curriculum for mathematics](#) states: 'By the end of year 4, pupils should have memorised their multiplication tables up to and including the 12 multiplication table and show precision and fluency in their work' (Section: Lower key stage 2 – years 3 and 4, page 17). It goes on to state: 'Pupils should be taught to recall multiplication and division facts for multiplication tables up to  $12 \times 12$ ' (Section: Year 4 programme of study, page 25).

The content domain has been developed and reviewed by Standards and Testing Agency's (STA) expert mathematics Test Development Researchers (TDRs) and independent mathematics curriculum advisors.

STA first published the multiplication tables check assessment framework in 2018, and an updated version in 2022 ahead of the first statutory check. No concerns have been raised with STA about the content domain.

## **1.2 The non-assessable aspects of the content domain are understood and defined**

The MTC only assesses fluent recall of multiplication facts. The digital format allows this area of the content domain to be assessed through timed questions.

The non-assessable elements of the content domain are defined in table 1. The rationale for why any element of the content is not deemed assessable in a paper-based test is also provided.

<b>Element of content domain not assessed by the MTC</b>	<b>Rationale</b>	<b>How the element could be assessed</b>
Related division facts	The MTC tests the fluency of pupils' recall of multiplication facts. Related division facts would not be assessable due to the timed, fluent recall element of this assessment.	End of key stage 2 national curriculum assessments  Teacher assessment
Application of multiplication facts in mathematical context (e.g. word problems) or situational context (e.g. across all areas of 'their work')	The MTC tests the fluency of pupils' recall of multiplication facts. Application or situational context cannot be assessed effectively within the timed construct of this assessment.	End of key stage 2 national curriculum assessments  Teacher assessment
Strategies employed	A timed, digital test which is designed to assess recall is not an appropriate format to assess the strategies pupils use to calculate their times tables facts.	Teacher assessment

**Table 1: Non-assessable elements of the content domain**

### **1.3 The assessable content in the digital check is an accurate reflection of the curriculum requirements for the recall of multiplication facts**

The aim of the MTC is to test 'fluent' recall of multiplication facts and not to assess mathematical problem-solving. The digital format allows this area of the content domain to be assessed through timed questions. In order to check whether pupils have fluent recall of multiplication tables, each check includes a breadth of items. This includes items of different difficulties across different multiplication tables, with an emphasis on multiplication tables taught at key stage 2.

Evidence from quantitative STA studies on timing limits for the MTC indicated 6 seconds as the most suitable timing. This minimises the possibility of pupils working out answers while giving them enough time to recall and input. Evidence from qualitative research studies with teachers and pupils provided further support for the 6 second cut-off point, with participants indicating that pupils needed at least 4 seconds for correct recall of the answers and extra time to input the response.

## **1.4 Assessment items have been rigorously reviewed and validated by a range of appropriate stakeholders**

The MTC assesses the fluent recall of multiplication tables in accordance with curriculum expectations. The individual test items that make up each check (form) are drawn directly from the multiplication tables, which are a pre-defined set of questions.

Items were collated into equivalent forms in line with the test specification and trialled in a technical pre-test (TPT) in 2018 by year 4 pupils in a representative sample of schools. Data collected from this trial were analysed by the STA's expert psychometrics team and used to inform construction of the live test forms.

Items and forms were reviewed by independent mathematics curriculum advisers before and after the 2018 trial as part of the form construction process.

Following these trials and reviews, items and forms were subject to a large-scale pilot in 2019 and a national voluntary roll-out in 2021. Data collected from these were analysed by the STA's psychometrics team and used to evidence ongoing equivalence between forms.

## **1.5 Test items and item responses have been suitably interrogated to ensure only the desired construct is being assessed and construct-irrelevant variance is minimised**

Test items and item responses were evaluated through a timing trial, user research and in a TPT.

The 6 second time limit was established following a quantitative trial that involved over 1,000 children across 50 schools, with the results evidencing 6 seconds as the most suitable cut-off point. Evidence from qualitative research studies with teachers and pupils provided further support for the 6 second cut-off point. User research was carried out in 150 school visits, with qualitative feedback on pupil interactions gathered through observations and interviews. This evidence fed directly into the iterative development of the assessment.

A TPT was carried out in June 2018. In this trial, approximately 1,000 pupils saw each item and detailed psychometric analysis of their responses was produced. This included classical and item response theory (IRT) analysis estimating the difficulty and discrimination of every item.

Following the 2022 live administration, pupils' final responses were analysed in detail. This research was undertaken to explore possible construct irrelevant variance by examining the errors pupils produced. This work indicated that construct irrelevant keystroke and input errors arising from the digital delivery of the test comprised a negligible number of responses.

## **1.6 The range of questions included samples from a suitable range of multiplication table calculations.**

A breadth of items is included in each of the multiple forms, with items selected from the 121 calculations that make up the 2 to 12 multiplication tables. Each of the forms has multiplication tables that are introduced in years 2, 3 and 4, with an emphasis on multiplication tables taught at KS2. Each multiplication table has a maximum and minimum representation in each form. These limits are outlined in the assessment framework.

All forms were reviewed by a mathematics curriculum advisor, who confirmed that coverage was appropriate and in line with the assessment framework.

## **Claim 2: Assessment results provide a fair and accurate measure of pupil performance.**

### **2.1 Item-level data have been used to ensure only items that are functioning well are included in the check**

All test materials were trialled in a TPT in June 2018. In this trial, approximately 1,000 pupils, from a stratified sample of schools, saw each item. This trial provided STA with sufficient item-level data to be confident about how an item would perform in a live test.

TDRs and psychometricians used a computer algorithm and item-level data to construct equivalent forms assessing fluency and maximising test information across the ability range, while minimising the standard error of measurement. TDRs and psychometricians considered the test specification constraints and their own expertise to make a final decision on form construction.

### **2.2 Qualitative data have been used to ensure only items that are effectively measuring the desired construct are included in the check**

From January 2017 to December 2018, over 150 schools were visited as part of user research informing the iterative development of the MTC. Various aspects of the item interface were investigated, including the layout, inter-item pauses and input methods.

User feedback contributed to an iterative design and build process to ensure pupils were able to enter their responses effectively. Feedback on the item design has continued to be gathered via an exit questionnaire, which all schools had the opportunity to fill in following completion of the check. In this way, minor adaptations can be made as necessary, continuing to optimise measurement of the desired construct.

### **2.3 A range of items that are age-appropriate and cover the full ability range has been included in the final check**

All the calculations used in the MTC are drawn from the year 4 curriculum expectation that children will be fluent in their multiplication tables up to  $12 \times 12$ .

The assessment is administered at the end of lower key stage 2 (year 4), which aligns with the timing of the curriculum expectation.

In order to check whether pupils have fluent recall of the multiplication tables, a breadth of items is included in each form. This includes items of different difficulties and from



across the different multiplication tables, with an emphasis on multiplication tables taught at key stage 2. All forms are constructed to ensure they measure across the ability range, in line with the assessment framework.

## **2.4 Appropriate evidence has been used to ensure the check does not disproportionately advantage or disadvantage any subgroups**

Access arrangements have been developed to enable as many children as possible to take part in the assessment, without being disadvantaged by the digital interface. The development of access arrangements for the MTC was informed by the following:

- A literature review of access arrangements for digital assessments.
- Exploratory interviews with a range of stakeholders, including teachers and experts in visual impairment, dyscalculia, dyslexia, physical disability and digital assessment.
- Item review by a mathematics curriculum expert, who reviewed the items and considered whether they were suitable for inclusion in a trial.
- Targeted user research testing initial access-arrangement prototypes, including review by experts and school visits:
  - Experts across various fields of special educational needs, disabilities and temporary injuries, including dyslexia, dyscalculia, dyspraxia, deafness, visual impairments, multi-sensory impairment, and access arrangements were consulted. There was also an emphasis on speaking with people who had knowledge of technology and accessibility.
  - As part of the development of the access arrangements, 32 mainstream schools and 2 special schools were visited, 50 staff were interviewed, and prototypes were tested with approximately 500 children.
- a. Following the targeted research, feedback on access arrangements and their implementation, evidence has continued to be gathered. This has included user research in the beta phase of development, the national pilot in 2019, and the exit questionnaire following the 2021 and 2022 administrations, enabling further and ongoing optimisation of the user experience.

In addition to research into access arrangements, differential item functioning analysis has been carried out in association with gender. This evidence is outlined in relation to claim 3.1.

## **2.5 Pupil responses have been interrogated to ensure pupils are engaging with the questions as intended**

Following the 2022 live administration, pupils' final responses for every question within each form were analysed in detail. The results indicated that the majority of pupils answered correctly and that a relatively small portion of the responses were left unanswered, with omissions making up an expected proportion of the responses. In addition, the error analysis outlined in claim 1.5 indicated that a negligible quantity of the responses could be attributed to be construct-irrelevant errors arising from the digital delivery of the test.

## **2.6 Creditworthy responses are marked accurately**

The pupil's final answer is the number entered into the answer window when either the 'Enter' key is pressed or the time has elapsed.

Only numerical responses can be entered, and only the final response is marked. This is completed automatically by the system.

Following the close of the live window the scoring of every final response for every item completed is checked.

## **Claim 3: Pupil performance is comparable within and across schools**

### **3.1 Potential bias towards particular subgroups has been managed and addressed when constructing the check**

The MTC has only a single item type, with all calculations presented in a context-free format. Therefore, the potential for bias towards particular subgroups is considered to be low. Even so, test developers interpreted a wide range of qualitative evidence to ensure the MTC does not disproportionately advantage or disadvantage children based on their gender, digital literacy or special needs. This included user research, teacher panels and teacher questionnaires following trials.

Following the 2016 TPT, differential item functioning (DIF) analysis was carried out in relation to gender. The purpose of this was to identify differences in item performance based on membership of girls'/boys' groups. None of the items exhibited more than negligible DIF, and even this did not consistently favour one gender over another.

### **3.2 Appropriate systems are in place to ensure the security of the service and its data**

The MTC service can only be accessed by approved users through DfE Sign-in. The DfE Sign-in approver can add the 'multiplication tables check' service to their own or other user accounts. School employees can create a DfE Sign-in account and request the account to be linked to an organisation. The School DfE Sign-in approver must approve this request before they are able to view services for their school.

To enable pupils to take the official check, schools are required to generate a school password and individual PIN for each pupil. These expire at 4pm daily. PINs can only be generated by DfE Sign-in users with approved access to the MTC service.

Results are held securely within the MTC service and are available to schools to view after the end of the check administration period.

An annual IT health check is carried out on the MTC service before the service opens for the academic year. The health check provides an audit of the service and seeks out any potential vulnerabilities or issues so that risks can be identified rapidly and remediation action taken to ensure the digital service is secure.

### **3.3 Guidance on administration is available, understood and implemented consistently across schools**

The [MTC administration guidance](#) and supporting videos were published alongside the opening of the MTC service. In 2022, the guidance was accessed approximately 51,000 times and the videos watched approximately 5,000 times.

Specific [Multiplication tables check: IT guidance - GOV.UK \(www.gov.uk\)](#) is produced in the autumn term, to help schools prepare devices for check administration. The supporting guidance is regularly linked in STA's bi-weekly assessment update emails to schools.

Feedback on the guidance provided is gathered via helpdesk queries and the exit questionnaire that follows the live administration. This information feeds into the ongoing development of the materials provided to support implementation.

Implementation is monitored through the data collected by the service.

### **3.4 Appropriate arrangements are available to ensure access whilst maintaining assessment validity**

Access arrangements are adjustments that can be made to support pupils with specific needs to access the check and ensure they are able to demonstrate their attainment. Access arrangements are included to increase access without providing an unfair advantage to a pupil. The support given must be based on normal classroom practice and must not change the check questions. The answers must be the pupil's own. Access arrangements address accessibility issues rather than specific SEND and the evidence that led to their development is outlined in association with claim 2.4.

STA provides guidance about the range of access arrangements available to enable pupils with specific needs to take part in the check within the MTC administration guidance. The range of access arrangements available in the digital MTC includes:

- implementing an audible time alert
- applying a different font size
- applying a different colour contrast
- allowing an input assistant
- implementing a 'next' button between questions
- implementing an audio version
- removing the on-screen number pad.

Headteachers and teachers must consider whether any of their pupils will need access arrangements before they administer the official check. They can use the 'Try it out' area with pupils to enable them to determine the most appropriate access arrangement/s for individual pupils. Schools can also contact the national curriculum assessments helpline for specific advice about how to meet the needs of individual pupils.

A small number of pupils may not be able to access the check, despite the provision of additional arrangements.

### **3.5 Multiple forms presented in the MTC are comparable**

Trialling data from the TPT were used to estimate item parameters in a single group, IRT model. These item characteristics were used to construct multiple forms according to strict psychometric and content specifications. A computer algorithm was used to find the optimal distribution of items in multiple forms that met the desired criteria. Once the forms were constructed, the resulting test characteristic curves and information functions were examined to confirm the equivalence of the forms. True score estimates associated with 5 points on the ability scale were compared to confirm equivalence.

## **Claim 4: Differences in assessment difficulty from year to year are taken account of, allowing for accurate comparison of performance year on year**

This claim will be completed when the MTC has been statutory for more than one year to enable the comparison to be made.

## **Claim 5: The meaning of test scores is clear to stakeholders**

### **5.1 Appropriate guidance is available to ensure stakeholders understand the reported scores**

Schools receive a total score out of 25 for each pupil who completed the check. There is no pass mark or expected standard threshold for the MTC.

Results are published at national and local authority level.

The [MTC information for parents' guidance](#) advises that results will be used to identify pupils who need additional support and will be shared with the parents as they would with all national curriculum assessments.

### **5.2 Helpdesk queries regarding check scores are monitored to ensure stakeholder understanding**

STA gathers management information on all enquiries received by the national curriculum assessments helpline. This information is regularly monitored and used to highlight key trends in enquiries; these are passed on to individual teams for consideration.

STA always considers common stakeholder enquiries when drafting its regular assessment update emails to schools and making annual guidance updates. In 2022, the national curriculum assessment helpline received 68 enquiries categorised as 'multiplication tables check pupil results' (2% of total queries).



Standards  
& Testing  
Agency

© 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:

Any enquiries about this publication should be sent to us at

[assessments@education.gov.uk](mailto:assessments@education.gov.uk)

Download: [www.education.gov.uk/assessment](http://www.education.gov.uk/assessment)

ISBN: 978-1-83870-467-4

Follow us on Twitter:

@educationgovuk

Like us on Facebook:

facebook.com/educationgovuk