



City
&
Guilds

Digital Learning

BEST PRACTICE INITIAL AND DIAGNOSTIC ASSESSMENT
HOW BEST PRACTICE PRINCIPLES OF ASSESSMENT
INFORMED THE RATIONALE, DESIGN AND
DEVELOPMENT OF CITY & GUILDS' INITIAL
AND DIAGNOSTIC ASSESSMENT TOOL

CONTENTS

Purpose of the paper.....	3
Introduction and background.....	3
Purpose of City & Guilds' initial and diagnostic assessment	3
Validity.....	5
Reliability.....	6
Authenticity.....	7
Accessibility and transparency.....	8
Efficiency and cost effectiveness.....	8
Feedback.....	10
e-assessment.....	13
Conclusion.....	13

Purpose of the paper

The purpose of this paper is to identify and reflect upon principles of assessment best practice and to demonstrate how these can be embodied in an approach to initial, diagnostic and formative assessment. This is informed and illustrated by reference to the rationale, design and development of City & Guilds' initial and diagnostic assessment tool for mathematics and English, which forms part of our [Maths and English e-Toolkit](#).

Introduction and background

Effective and reliable initial and diagnostic assessment of mathematics and English is increasingly important in the post-16 and adult education sector. As well as pedagogical considerations, there are also the following policy and funding drivers within England:

- In 2013 the Government introduced a requirement for post-16 learners on study programmes of more than 150 hours a year to continue studying maths and English if they have not already achieved GCSEs at grade A*-C in these subjects. GCSE grading, particularly at the lower levels, does not provide information about a candidate's strengths and weaknesses. These grades cannot therefore, be used to identify exact skills level, nor what skills learners need to concentrate on in order to progress. For learners arriving at a college with lower level GCSE grades, accurate initial and diagnostic assessment is required to assist providers in ensuring learners are placed on an appropriate programme of learning.
- The Skills Funding Agency (SFA) requires centres to carry out thorough initial assessment of their learner's English and mathematics skills. Learners must be enrolled onto courses at the level above that which they have been assessed to be operating. SFA further require that learners sit appropriate diagnostic assessment to inform their learning agreement and the basis of their programme. These requirements are subject to Ofsted inspection.

Similar expectations are also placed on providers in Northern Ireland and Wales.

Purpose of City & Guilds' initial and diagnostic assessment tool

In planning the development of our initial and diagnostic assessment tool, City & Guilds undertook research into best practice in initial and diagnostic assessment and into our customers' needs.

Based on our research, we identified the following key objectives for the tool:

- high quality, accurate assessment
- clear and transparent reporting
- rich, detailed and purposeful feedback
- ability to support learners working towards a range of qualifications in maths and English, including but not limited to, Functional Skills, Essential Skills, Mathematics/English Skills, Principles of Using Mathematics/English and GCSE
- a tool for identifying the maths and English needs of those undertaking vocational programmes.

As an awarding organisation, we drew upon our expertise in assessment and applied this to the needs of initial and diagnostic assessment. An important consideration is that the needs of initial and diagnostic assessment are not the same as those of summative assessment, which is intended to be delivered at the end of a programme of study. For this reason, the approach we took to initial and diagnostic assessment differs from the approach we take to summative assessment.

The following factors and considerations informed our approach:

- The purpose of practice tests is to prepare candidates to take summative assessment. The purpose of initial and diagnostic assessment is to measure prior attainment and support further learning.
- Summative assessments, particularly at the higher levels, often test selection and application of a number of skills together in solving problems. With this approach it is not always possible to identify with accuracy the reason *why* a candidate has answered incorrectly. A more direct approach, with questions that isolate and test individual skills, is necessary in order to assess strengths and gaps for initial and diagnostic assessment purposes.
- Summative assessments are intended to measure overall achievement against a framework of skills and use an overall pass mark. An aggregated whole test score can only give an approximation of ability. A better approach for initial and diagnostic assessment is to measure against each curriculum area separately in order to focus on strengths and weaknesses and to provide rich and purposeful feedback.

All assessment developed by City & Guilds is informed by eight principles of assessment. Although these were originally developed for summative assessment they are of equal relevance and are applicable to initial and diagnostic assessment. The table below summarises some key considerations in relation to this:

VALIDITY:

- Assessment must have a direct relation to the curriculum / standards.
- Candidates must be able to demonstrate achievement in relation to the curriculum / standards.
- Items must avoid assessment of skills other than that which is being assessed.

RELIABILITY:

- Assessment must give the same result for the same response each time it is used.
- Assessment must credit the full range of correct responses.
- Feedback must be consistent and reliable.

AUTHENTICITY:

- Context is relevant and purposeful.
- Items should be purposeful and match real life application of the skills.

EFFICIENCY:

- Assessment effectively supports learning.
- Does the assessment approach effectively support the needs of individual candidates?
- Do items and tasks maximise opportunities for feedback?

ACCESSIBILITY / TRANSPARENCY:

- The purpose of the assessment is clear.
- All items / tasks are clear and unambiguous.
- Language used is direct and unambiguous (unless there are assessment-related reasons otherwise).
- Outcomes and next steps for further learning are clear.

COST EFFECTIVENESS:

- The product offers value to practitioners, employers and learners.

ADEQUACY OF FEEDBACK:

- Assessment is followed by timely feedback.
- Feedback is supportive and recognises achievement.
- Feedback is constructive and sufficiently detailed to support further learning.

The next section of this paper will look at how these principles have been addressed in the development of City & Guilds' Maths and English initial and diagnostic assessment tool.

Validity


Maths and English provision in England, Northern Ireland and Wales is based on national frameworks (these include, the Functional Skills standards, national standards for Adult Numeracy and Literacy and GCSE subject content). All of these frameworks originally derive from and are aligned with the National Curriculum in England.

The team that developed the content for City & Guilds' initial and diagnostic assessment tool brought a wealth of experience in the development and interpretation of these frameworks. It included former national policy leads for National Curriculum Maths, English and Functional Skills standards and assessment, National Curriculum test developers and awarding organisation chief examiners.

Drawing on this range of expertise and experience we have been able to develop assessments that accurately assess the standards. In addition, we have created a tool that supports maths and English skills application broadly and is not specific to any single group of qualifications.

A common criticism of mathematics assessment, especially at lower levels, is that the level of language required to articulate the question can act as a barrier, introducing an element of reading which has adverse impact on the validity of the assessment. For this reason, a guiding principle in the development of the initial and diagnostic assessment tool was to use design and graphics as much as possible to set the context and basis for questions.

Taking the below examples, wholly written versions of these questions might read:




How many cabins does a group of 30 people need?

 cabins

A group of 30 people decide to rent holiday cabins.

Each cabin has room for a maximum of 4 people.

How many cabins does the group of 30 need to rent in total?



How much is the perfume now?

☐ £6
☐ £10
☐ £20
☐ £24

A bottle of perfume cost £30.

It is reduced by 20%.

How much does the perfume cost now?

Even on an individual question basis, it is apparent how the written form of the question adds to the reading demands of the assessment. If we consider this over the course of a 40-item test we can begin to see how this might adversely affect length and increase literacy demand. Additionally, the use of graphics further adds validity by ensuring the questions more closely match how such information is displayed in the real world.

Reliability

The initial tests are entirely computer-marked, ensuring results and feedback are accurate and consistent.

The initial tests are adaptive, each learner's pathway through the test adapts to whether questions are answered correctly or incorrectly. Adaptive assessment enables us to measure prior attainment against a number of different levels within single maths and English tests, each comprising around 40 items. In order to ensure that this measurement is reliable, the following considerations have been built into the design of the adaptive algorithm:

- **Strength and gaps profiling.** The initial tests give a measure of performance against a number of curriculum subdomains (six for mathematics, seven for English). We know that learners do not display 'flat' profiles of attainment across a level so each curriculum subdomain is assessed separately. The adaptivity does not assume that because a learner is at a certain level for a given subdomain (eg

spelling) they are at the same level for any other subdomain (eg punctuation, grammar, etc).

- **Benchmarking.** All learners take the same set of benchmarking questions. These form the basis of the adaptivity and underpin the skills profiling measurement.
- **Best practice in adaptive testing.** In line with best practice for adaptive assessment, the tests' adaptivity is based on grouped clusters of questions for each curriculum subdomain and level. Systems which adapt on the basis of single questions are less reliable as these provide the learner with only one chance to demonstrate attainment of the skill / level. This does not take into account the possibility that the learner may know the correct answer but enter it incorrectly or not know the answer but make a lucky guess. Using a greater number of items mitigates against this and provides a more robust and reliable measure.

Authenticity

In developing the tests, care has been taken to ensure that question contexts are purposeful and relevant to work and everyday life.

SCHEDULED TIME	FLIGHT NUMBER	COMING FROM	STATUS	TERMINAL
09:00	LY8053	CHICAGO	LANDED 09:54	3
09:00	LY8053	NEW YORK	LANDED 10:49	3
09:00	MH9410	NEW YORK	LANDED 10:49	3
09:00	MH9424	CHICAGO	LANDED 09:56	3
09:05	AA6171	WASHINGTON	LANDED 10:17	5
09:05	AA6151	WASHINGTON	LANDED 10:20	5
09:05	BA266	WASHINGTON	LANDED 10:17	5
09:05	IB4654	WASHINGTON	LANDED 10:17	5
09:15	AA6151	NEWARK	LANDED 09:38	5
09:15	AT5488	NEWARK	LANDED 09:38	5
09:25	AF3673	NEW YORK	LANDED 10:12	4
09:25	DLS	NEW YORK	LANDED 10:12	4
09:25	KL6005	NEW YORK	LANDED 10:12	4
09:25	VS4005	NEW YORK	LANDED 10:12	4
09:35	AA6159	DALLAS/FT WORTH	LANDED 09:45	5
09:35	AC860	HALIFAX	LANDED 09:49	3
09:35	AT5492	DALLAS/FT WORTH	LANDED 09:45	5
09:45	AA6135	NEW YORK	LANDED 09:41	5

What time did flight AA6151 arrive?

☐ 10.17
☐ 10.20
☐ 09.38
☐ 09.15

NEXT

We have used professionally produced graphics so that stimulus material appears authentic and has an engaging, high quality look and feel.

BEAUTIFUL DESIGNS, OUTSTANDING VALUE

We collaborate with some of the best talent in the industry to create our unique designs. Then we commission innovative craftsmen to create our award winning furniture using only sustainable timber. We want you – our valued customer – to be satisfied, so we never compromise on quality, even at our unbeatable prices.

Find one word the writer uses to give the impression that the furniture can only be bought from that design company.

Write your word in the box:

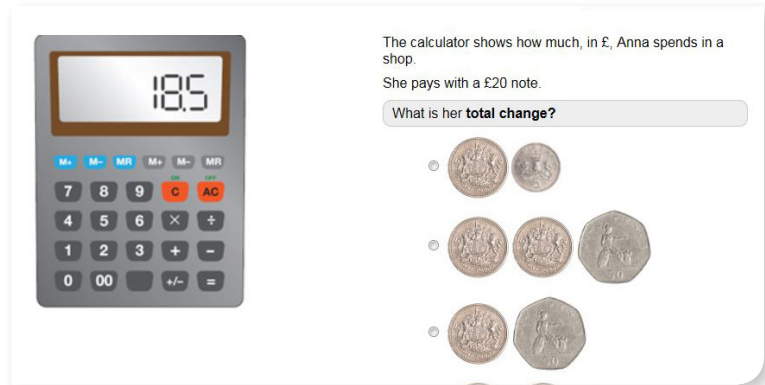
NEXT

How many miles has this car travelled?

☐ two thousand, four hundred and eighty
☐ twenty thousand, four hundred and eight
☐ twenty-four thousand and eighty
☐ twenty-four thousand, eight hundred

NEXT

Use of authentic-looking source material means that questions can be focused and concise, and reduces the need to use text to frame the context (eg 'This is an arrivals board at an airport...', 'This is an advertisement in a magazine...').



Accessibility and transparency

We have already seen how the test questions have been designed to be focused and clear, supported by the use of realistic, high quality graphics. The previous examples also illustrate other design features which have been used to clarify and focus the question requirements:

- As a general rule, the question is given to the right of the screen and the context and information needed to answer the question is displayed to the left. This models the way in which we read written information, from left to right.
- The question 'stem' (ie the question itself) is articulated clearly, using as few words as possible.
- The stem is emphasised by always appearing in a grey highlighted area, positioned in the same place on each screen.
- In line with best practice in e-assessment, screens maximise use of white space to make these easily accessible and free from clutter and distraction.

Efficiency and cost effectiveness

The initial tests are designed to take an average of about 30 minutes each to complete and are intended to be taken by learners in single sittings. This is to ensure learners stay focused, without the need to pause and resume tests. The initial tests use adaptivity to give a reliable measure of learner attainment at four levels. The tests' adaptivity does not affect the overall assessment length – all learners take 36 questions for maths and 42 for English.

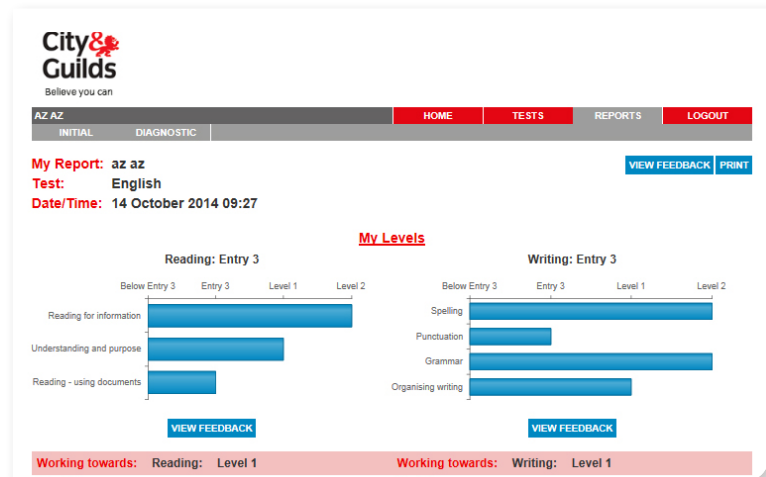
Following each 30 minute initial test detailed performance feedback is provided. The reporting features provide clear results which can be easily understood by both learners and centre staff regardless of their training levels or familiarity with national frameworks and the requirements of the standards.

DIAGNOSTIC TESTS

ASSIST LESSON PLANNING AND MEASURE SKILLS ACQUISITION

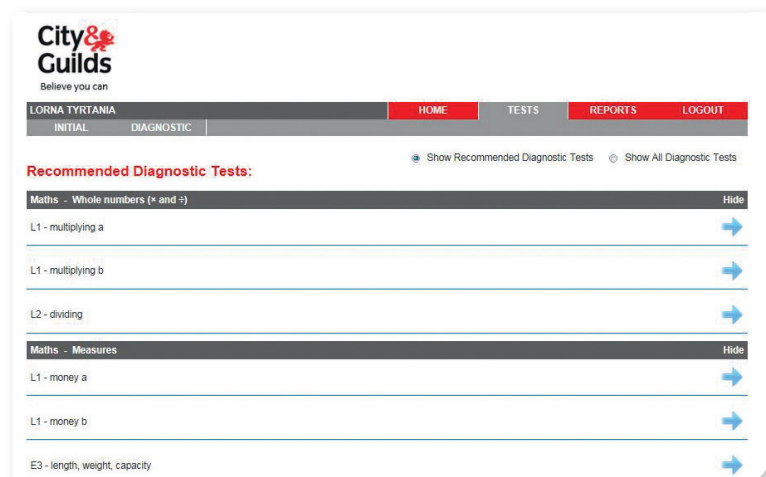
The report screen (shown) clearly displays the learner's:

- current 'secure' level (ie the lowest level achieved for any curriculum subdomain)
- skills profile ('spiky profiling' of level of attainment for each curriculum area assessed)
- 'work towards' level
- next steps.



Based on the outcomes of the initial tests, the system selects and automatically routes the learner to diagnostic tests relating to *only* those skills they need to develop and practise in order to progress. The diagnostic tests are designed to be short and manageable and can be worked through at the learner's own pace. The diagnostic tests can be delivered at the start of, or throughout a programme of learning to assist lesson planning and measure skills acquisition.

Once the learner has progressed to the next level, or should a need be identified to do further diagnostic tests, learners can self-serve from the entire bank of diagnostic tests, as directed by a tutor, without the need for the tutor or centre administrator to assign these tests to each learner individually via the system. The learner can simply select 'Show All Diagnostic Tests' on the screen above to access all tests in the bank.



Feedback

Initial and diagnostic assessment are both types of formative assessment in that they form the basis of further learning. The main purpose of formative assessment is to generate high quality, 'rich' feedback that is useful to tutors and learners.

The following principles for good quality feedback have been identified:

1. It clarifies what good performance is (goals, criteria and expected standards).
2. It facilitates the development of self-assessment in learning.
3. It provides high quality information to students about their learning.
4. It encourages teacher and peer dialogue around learning.
5. It encourages positive motivational beliefs and self-esteem.
6. It provides opportunities to close the gap between current and desired performance.
7. It provides information to teachers that can be used to help shape teaching.

(Nicol, David J.; Macfarlane-Dick, Debra (2006). "Formative assessment and self-regulated learning: a model and seven principles of good practice". *Studies in Higher Education*).

Feedback is often described as a conversation or dialogue that a learner has with tutors, peers and employers about their performance. Through the exchange of concepts, ideas, questions and performance feedback, positive adaptation of a learner's actions (ie learning) takes place.

The initial and diagnostic tests give both high level and detailed feedback in a way which is modelled on and supportive of this 'conversational' process. The tests are presented in a supportive and accessible way to engage and involve learners in the assessment and subsequent learning.

On completion of the initial and diagnostic tests (pictured right) learners are provided with clear and accessible feedback on their performance. This recognises their achievement and advises on next steps.

The top screenshot shows the 'Diagnostic Tests' page for Lorna Tyrantania. It includes a navigation bar with 'HOME', 'TESTS', 'REPORTS', and 'LOGOUT'. The page is divided into 'INITIAL' and 'DIAGNOSTIC' sections. The 'Diagnostic Tests' section explains that these tests are designed to check Maths and English skills and provide a 'TAKE TEST' button.

The bottom screenshot shows the 'My Report' page for Lorna Tyrantania. It displays test results for 'L2 - Getting an overview' in the 'Reading for information' domain. The report includes a 'VIEW FEEDBACK' button and a recommendation to view feedback and practice techniques.

As well as summary reporting, on completion of each initial and diagnostic test, the system delivers a test transcript. This shows how the learner performed on each question and gives rich, detailed feedback covering:

- the skill being assessed (at item level for initial test and test and item level for diagnostic)
- techniques for working out the correct answer
- suggestions for further learning
- guidance on test-taking techniques.

Feedback
Test Level feedback:

The questions in this test focused on how you can use the organisational features of texts to find information.

Large documents such as text books, reports and even some articles and essays often use specific organisational features to help you locate specific information quickly and find your way around the document easily. Organisational features include things such as a contents page, index, footnotes and references.

The questions in this test looked specifically at using an index page to find information.

Most non-fiction books will have an index page at the back. The index page lists the content of the book with the page numbers it appears on in alphabetical order. It can help you find information quickly by pointing you to the information you're looking for without having to skim through the whole book.


Some questions ask you to look for something specific. You can save time by identifying the key words in the question and then scanning the document to find them. When you have, that's probably where you will find the answer. This means you don't have to read the whole document again.

When you think you've found what you are looking for, always read on because there may be a better answer later in the document.

In a test if you find two details and both seem right, go back to the question and see which one you think fits best.


This question is asking you to use organisational features, like index pages, to find information.

What is the key word in the question?



When this photograph was taken, this puppy weighed 1 kg.
It was 5% of its adult weight.

What is its adult weight?

80 kg  Incorrect

Feedback

The puppy weighed 1 kg.
So 5% = 1kg.
We need to find 100%

$$\begin{array}{c} \text{5\%} = 1 \text{ kg} \\ \times 2 \\ \text{10\%} = 2 \text{ kg} \\ \times 10 \\ \text{100\%} = 20 \text{ kg} \end{array}$$

So the adult weight (100%) is 20kg

Feedback for the diagnostic tests provides a whole test commentary, giving item-by-item guidance on the skills being assessed and test-taking techniques.

1.

Feedback
Test Level feedback:

The questions in this test are asking you to find specific information by reading a text in detail.

In a document with a lot of details, you will never be able to remember everything on the first read. So once you've got the general idea of what the document is about, go back and **scan** the document for the **specific** information you are looking for.

Looking at **headings**, **sub-headings**, opening sentences and specific words can help you do this

When you think you've found what you are looking for, always read on because there may be a better answer later in the document.

In a test if you find two details and both seem right, go back to the question and see which one you think fits best.

In this question you are looking for detail about what the Inclusive Fitness Initiative provides.

The section titles and opening sentence do not help you so next scan the document for the words Fitness Initiative. The only place it is mentioned is towards the end of the first paragraph, so that's almost certainly where you'll find the answer.

By reading carefully you will see that the Inclusive Fitness Initiative focuses mainly on **gym equipment**. This means option C, which is about swimming, and option D which is about sign language, are not correct.

Now you've narrowed it down to option A and option B which are both about gym equipment.

2.

[▲ Feedback](#)

In this question you are looking for detail about when the leisure centres open and close.

Scan the document for where you might find the information.

Each leisure centre has a line on opening times.

Compare the opening times to find which is open earliest and which is open latest.

When using written information about times it's important to know the difference between am (before midday) and pm (after midday). If you are not sure about this, keep looking out for examples from your everyday life until you feel confident.

3.

[▲ Feedback](#)

For true and false questions, focus on each statement in turn and then scan for where you are most likely to find the answer.

Start with the true/false statement. Scan the document for key words from the statement. The first section is about Jane Smithdown Centre. Read the information there and work out whether the statement is true or false.

Work through each statement in this way.

All tests and reports are saved on the system and can be accessed and referred to by the learner and tutor to form the basis of further learning, help monitor progress and assist revision.

FORMATIVE ASSESSMENT:
HIGH QUALITY, 'RICH'
FEEDBACK THAT IS USEFUL
TO TEACHERS AND LEARNERS

e-assessment

The assessment platform used to develop the tests offers a range of different question types and formats. In designing the tests we were careful to avoid assumptions about learners' familiarity with ICT and e-assessment but balance this against the positive benefits that can be delivered through variety and interactivity.

Where possible, on screen item types were chosen or developed to match or model how the skills tested are used in real life.

Conclusion

The need for accurate and effective initial and diagnostic assessment in the planning and delivery of maths and English learning programmes is beyond question.

We have listened to customers to understand this need and, as a leading awarding organisation, have applied our own expertise and understanding of assessment best practice to produce a high quality initial and diagnostic assessment tool designed to support maths and English learning in a wide range of contexts.

Our initial and diagnostic assessment tool forms part of our Maths and English e-Toolkit, designed to support centres in the effective delivery of maths and English learning programmes across their organisation.

Put the three sentences below in the correct order by dragging and dropping them into the boxes below.

Then, before I knew it, it was all over.

As it got closer, I got more and more nervous.

The time of the interview was fast approaching.

Sentence 1

Sentence 2

Sentence 3

Staff had known in advance of the consequence, then they would have acted. In future, it is vital that staff are told that they inform a

Employee Weekly Timesheet

Employee Name: Anna Jones

Week 1

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Start time	9am	8am	—	—	9am	8am	—
Finish time	1pm	1pm	—	—	1pm	3pm	—
Hours worked	4	5	0	0	4	7	0
Total hours worked							
Total pay for the week							£

Anna works at a supermarket.
She is paid £9.25 per hour.

Write the missing information in the blue boxes and work out her **total pay**.

BOOK A FREE, NO OBLIGATION DEMO

Contact us to arrange a demo of our Maths and English e-Toolkit, including a preview of our initial and diagnostic assessment tool and samples of our teaching and learning resources.

Visit: www.cityandguilds.com/mathsandenglishetoolkit

Call: 01924 206709

Email: digitalsales@cityandguilds.com