

A WORLD WITHOUT MAPS?
ASSESSMENT IN TECHNICAL EDUCATION

A report to the Gatsby Foundation

SIMON FIELD

September 2021



GATSBY

GATSBY IS A FOUNDATION SET UP
BY DAVID SAINSBURY
TO REALISE HIS CHARITABLE OBJECTIVES.
WE FOCUS OUR SUPPORT ON A LIMITED
NUMBER OF AREAS:

PLANT SCIENCE RESEARCH
NEUROSCIENCE RESEARCH
SCIENCE AND ENGINEERING EDUCATION
ECONOMIC DEVELOPMENT IN AFRICA
PUBLIC POLICY RESEARCH AND ADVICE
THE ARTS

WE ARE PROACTIVE IN DEVISING PROJECTS
TO ACHIEVE OUR AIMS. WE ARE ENTHUSIASTIC
ABOUT SUPPORTING INNOVATION. WE ARE
ANALYTICAL AS WE BELIEVE IT IS IMPORTANT
TO UNDERSTAND THE OPPORTUNITIES
AND PROBLEMS WE TACKLE. WE TAKE A
LONG-TERM VIEW AS WE DO NOT THINK MUCH
CAN BE ACHIEVED BY SHORT, ONE-OFF
PROJECTS. WE ARE ALWAYS EAGER TO FORM
PARTNERSHIPS WITH ORGANISATIONS WHO
SHARE OUR GOALS.

The Gatsby Charitable Foundation
The Peak, 5 Wilton Road, London SW1V 1AP
T +44 (0)20 7410 0330 F +44 (0)20 7410 0332

www.gatsby.org.uk

Copyright © Gatsby Charitable Foundation 2021

CONTENTS

Acknowledgements	iv
Executive summary	
A world without maps? Assessment in technical education. Summary and implications for England	1
1 Introduction: the aims of this paper	4
2 Living with an unmapped world: tools for navigation	6
3 The design of assessments: three dilemmas	13
4 Why the objectives of assessment matter	24
5 How can we be sure of occupational competence in certifications?	33
Annex A. Assessment in Swiss apprenticeship	38
Annex B. A BTEC certificate for Art and Design and a Danish certificate for a Health Care Cook	42
References	44

BOXES, FIGURES AND TABLES

Box 1. Contrasting assessment with certification	5
Box 2. In Italy, assessments that vary by region and institution	6
Box 3. Could a PISA-type international study of technical competence in occupations be undertaken?	7
Box 4. Validity and reliability in assessment: an illustration using the example of French medical education	8
Box 5. An approach to continuous improvement in training providers in Western Australia	9
Box 6. In Finland, radical flexibility in pathways and the implications for assessment	10
Box 7. Recognition of prior learning as a special form of assessment: the example of France	11
Box 8. Written examinations for trade apprentices in Canada and electricians in New Zealand	14
Box 9. Work-embedded assessment projects in the Netherlands	16
Box 10. Avoiding plagiarism in periodic assessment	17
Box 11. Assessing placements in technical programmes: some comparisons across countries	18
Box 12. A blend of periodic modular and synoptic final assessment in Luxembourg	19
Box 13. What happened to synoptic assessment in English apprenticeship?	20
Box 14. In Germany, employers are systematically involved in apprenticeship assessments	25
Box 15. What is the 'right' pass rate for final assessments for apprentices? England compared with Germany and Norway	26
Box 16. In Sweden and Denmark, pathways to higher education are key to the attractiveness of technical programmes	28
Box 17. The United States is distinctive in making use of free-standing assessments of employability skills	29
Box 18. Is there scope to make more use of an international standard for assessments and certification? ISO/IEC 17024:2012	36
Figure 1. A promising start in quantifying assessment: percentage of apprenticeship assessment plans in England that use different assessment tools	7
Figure 2. The UK has weak employment protection relative to many OECD countries	27
Table 1. The relative advantages of standardised and work-embedded tasks in technical assessment	15
Table 2. England 2019/20 apprenticeship standards	26

ACKNOWLEDGEMENTS

The author is very grateful to Daniel Sandford Smith and Jenifer Burden of the Gatsby Charitable Foundation for their helpful and constructive guidance and advice throughout this exercise. Helpful input was also received from Paul Newton of Ofqual, Seb Ross of DfE and other colleagues convened in two seminars on this topic organised by the Gatsby Charitable Foundation. My thanks also to Sophie Field for her contribution of international examples.

DISCLAIMER

The views and opinions expressed in this report are those of the author and do not necessarily state or reflect those of the Gatsby Charitable Foundation.

EXECUTIVE SUMMARY

A WORLD WITHOUT MAPS? ASSESSMENT IN TECHNICAL EDUCATION. SUMMARY AND IMPLICATIONS FOR ENGLAND

Although assessment is a critical element in technical education, our knowledge of the topic is remarkably weak. Technical assessment is hard to study, given the diversity of occupations, of competences required for individual occupations, and of assessment tools. This fragmentation makes it difficult to summarise or measure, and to provide the evidence base to guide good practice. As a result, the landscape of technical assessment may often seem to be like a world without maps. But despite all the challenges, technical education can be studied: different types of assessment can be identified and their use measured, informing the design of technical assessments. The aim of this paper is to explore some of the key issues in technical assessment in England, drawing on a wide range of international, as well as UK experience.

Several choices are made in the design of assessments: whether to employ a largely standardised approach, or to rely more extensively on variable work-embedded assessment tasks; whether assessment should be undertaken only at the end point, or periodically throughout learning; whether the approach should be synoptic (looking at occupational competence in the round) or more atomistic; how to set passing thresholds, and if and how to establish grades above pass. In part, these choices may be determined by the varying objectives of different assessments, which might include demonstrating occupational competence to employers, or proving the capacity to pursue further learning, or motivating students. So the function of assessment determines its design: an intellectually demanding knowledge-based exam may impress a university, but not help to persuade an employer of an individual's ability to tackle the messy realities of day-to-day working life.

So how can the evidence reviewed here help to design assessments in England? This paper suggests some conclusions:

- **Balancing authentic work tasks and standardisation in assessment.** Technical assessments need to balance two competing demands. On the one hand, there are powerful arguments that assessment should concern itself with authentic work tasks, as these are the best tests of occupational competence. On the other hand, such tasks are variable and context-dependent, making it difficult to ensure assessment reliability – with all candidates having to meet the same standards. One solution is to make overall assessment results depend both on performance in authentic work tasks, and on more standardised tasks – for example in France the final marks in a professional baccalaureate depend both on performance during a placement with an employer, and on national written examinations. Moreover when assessment does involve variable work tasks, the assessment criteria can be carefully standardised, so that among other matters the ability to use a key tool might be a required part of every assessment.

- **Implementing synoptic assessment.** 'Synoptic' assessment involves looking at occupational competence in the round, rather than just assessing a list of the competences required for an occupation. Typically it addresses higher level 'meta' competences involved in marshalling knowledge and skills to solve problems, and respond to unexpected challenges. Such competences are essential in nearly all jobs, and vital in the context of changes in technology and in the work environment. Synoptic assessment is challenging, as typically it means looking at the capacity of a candidate to undertake a complete work project in a realistic context. In the Netherlands, a central element of final assessment in technical education is performance on work projects which may be undertaken over a period of weeks. In England, while synoptic assessment is widely promoted, its full implementation remains elusive.
- **Granting a role to periodic assessment.** An end-point assessment is administratively neat, and provides a good frame for a synoptic assessment, but it can sometimes usefully be augmented by assessments undertaken periodically during learning, especially in longer programmes – most German apprenticeships involve a 'half-time' assessment which contributes to the final result. Periodic assessment can have formative as well as summative uses; some elements of a training programme (like placements with an employer) require assessment at the time; and records of partial accomplishment, embodied in periodic assessments, can help those who drop out of programmes to return to education at a later date.
- **Implementing grading.** In assessment, the use of grades above pass (such as 'merit' and 'distinction') has often been championed on the grounds that such grading recognises and encourages excellence. Grading in technical assessment is common in many countries and encouraged in England. However, for excellence to be effectively recognised and rewarded, it needs to be measured consistently and accurately, and this is challenging. While the evidence base is limited, such evidence as we do have, from the UK and Australia, suggests that it has proved difficult to establish clear and meaningful grading criteria, and consistent grading practice. This implies that grading should be used cautiously in technical education.
- **Quality assuring certification.** Technical certifications need to offer employers a guarantee of occupational competence. Certifications usually depend not only on a candidate passing an assessment, but also on other requirements such as completion of an approved training programme, and/or relevant working experience. Quality assurance of certification should therefore treat assessment as part of a package, alongside other requirements, which collectively provide convincing evidence of competence. For example, in Swiss apprenticeship, requirements on the final assessment sit alongside quite demanding expectations on the staff in enterprises who supervise and train apprentices, as well as requirements on off-the-job education and training. In English apprenticeship, by contrast, while there are quite demanding requirements on end-point assessments, some training requirements, such as for 20% of programme time spent in off-the-job training, are very weakly enforced.

- **Delivering independence in assessments.** Independent elements in assessment, separate from local training providers, represent an important means of removing local biases and increasing consistency in standards. So in Korea, the award of technical qualifications takes place after an internal assessment undertaken by a training institute, and an external assessment undertaken by the awarding body. However while it is important to ensure independence in the sense of avoiding any personal stake on the part of assessors on the outcome, the perspective of some interest groups should be reflected in assessments. There are good reasons to allow employers collectively to influence the shape of assessments, as they are best placed to judge the competences necessary for the job, and the level of competence that might reasonably be expected in a newly qualified person.
- **Improving the evidence base.** The evidence base on technical assessment can and should be improved. Assessment tools (such as work observations) can be classified, and their use in different technical qualifications quantified. The prevalence of elements such as periodic, synoptic, and standardised assessments can be measured, alongside data on pass rates, grades, and resits. Such data can be disaggregated to show how different assessment bodies are using assessment tools, supporting data-driven quality assurance. Research on the effectiveness of different assessment tools can be undertaken, guiding quality improvement.

I INTRODUCTION: THE AIMS OF THIS PAPER

Technical assessment has received limited attention

Assessment, in the form of written examinations, is at the heart of much discussion of academic education. Issues such as grade inflation, 'teaching to the test', and test reliability and validity are hotly debated. Often, the shape of the examination, for example in English A-levels, comes to define the role and value of the qualification. In technical education this picture is reversed. Analysis and research on technical education is limited (Newton, 2018¹), so that a search for articles with technical or vocational assessment in their title in the 'Assessment in Education' journal yielded only two articles out of 400 over the last decade (Assessment in Education: Principles, Policy & Practice, 2021²). International comparative studies, such as OECD and CEDEFOP reviews of country systems of vocational education and training, and OECD reviews of evaluation and assessment, give little coverage to technical assessment (OECD, 2021³) (CEDEFOP, 2021⁴). One reason for the low profile of technical assessment is that it is extraordinarily diverse in respect of occupational target, types of competence, and assessment tools, and therefore resistant to summary description or measurement. It is nearly, if not entirely, a world without maps.

This paper draws heavily on international comparisons

This paper draws on international evidence and experience to look at technical assessment in England. Sometimes this is because other countries have pursued approaches which could be used in England, but more often it is to illustrate the possibility of different approaches to assessment, and therefore widen horizons. All countries face choices and dilemmas in managing and designing technical assessments; whether to use standardised assessment tasks, periodic or final assessments, synoptic or atomistic approaches; whether to use grading, how to set the passing threshold, how to involve employers, and how to ensure independence in the assessment. This paper aims to illuminate these choices.

The focus of the paper is on the upper secondary level

The primary focus of this paper is on upper secondary technical programmes at level 3 in the UK Regulated Qualifications Framework, alongside their international equivalents (mostly ISCED 3). This reflects the policy importance of this level of education, and allows sufficient comparability across countries. But there is some latitude in the coverage: in apprenticeship the assessment issues are similar across different levels, and higher level and tertiary technical assessments are not excluded from consideration. Given its close links with assessment, the paper also looks at certification (see Box 1). In 2020 and 2021, COVID-19 has had a large impact on assessments and technical education more broadly. However, although longer-term effects of the COVID-19 crisis on assessment are likely, most impact will be temporary. Consequently this paper does not seek to describe the effects of COVID-19 on assessment.

BOX 1. CONTRASTING ASSESSMENT WITH CERTIFICATION

Technical certification is typically intended to confirm that an individual is occupationally competent. As certification commonly follows success in an assessment, the two processes are often seen as one, but there are important distinctions. Often certification depends on more than just an assessment, such as completion of a learning programme, or requirements for employment experience or placements with an employer. It follows that the capacity of certification to confirm occupational competence depends not only on assessment but also on these other requirements, a point pursued in section 5 of this paper:

Certification typically provide a variety of additional information (see the two examples in Annex B). In the Danish example, grades in the three main subcomponents of the assessment are reported in the certificate. The T-level certificate will include separate grades for the core component, using A* to E, and for each occupational specialism, shown as pass, merit or distinction (DfE, 2020⁵).

Sometimes, more than one certificate is issued to reflect the different components of a programme and the connected assessments. For example in Germany, the graduating apprentice receives three certificates: one reflects an examination of professional competence, one a report of performance in the vocational school, and one a report on the apprentice's performance by their training employer (European Alliance for Apprenticeships, 2021⁶).

The remainder of this paper addresses the challenges faced by assessment and draws out the implications for England

Section 2 of this report describes the weak evidence base on technical assessment, and suggests ways in which it might be improved. It draws out the implications of fallible assessments, including the potential need to buttress assessment with requirements on programmes of study and other experience. Section 3 explores choices in the design of technical assessment systems. It describes the competing arguments for standardised assessments as compared with non-standardised work-embedded assessment tasks. It looks at whether periodic assessment should be employed in the course of a technical programme, or rely only on an end-point assessment (EPA). It notes the strong arguments for synoptic (as opposed to atomistic) assessment, and at the challenges involved in implementing a synoptic approach. Section 4 identifies three key objectives of summative technical assessment. It looks first at assessment as a means of demonstrating occupational competence, noting the advantages of assessments that involve employers in enhancing the credibility of the assessment to employers. Second, it examines assessment in relation to the wider competences that support further learning. Finally it explores how assessments may motivate students, especially by offering a grade above pass. Section 5 examines quality assurance and independence in assessment. It suggests that quality assurance is best conceived in relation to certification, whereby the quality both of learning and of assessment jointly confirm occupational competence. Finally it notes the different means used by countries to ensure independence in assessment.

2 LIVING WITH AN UNMAPPED WORLD: TOOLS FOR NAVIGATION

This section looks at why the evidence base on technical assessment is so weak, and at how it might be improved. It explores how the weak evidence base means that experiential requirements, such as employment experience or completion of a training programme, remain necessary conditions for many types of certification.

A WEAK EVIDENCE BASE

Technical assessment is more complicated than academic assessment

Most technical assessments are primarily concerned with measuring the capacity to undertake an occupation – occupational competence. As a result they are several times more complicated than most academic assessments. First, an extraordinary range of competences, from interpersonal to mathematics skills, from physical dexterity to attention to detail, are needed for every occupation, and there are many different occupations. Second, a diverse set of assessment tools are employed, ranging from observation of job performance to written exams. Third, occupational competence is most directly measured by observing real working practice, and this is highly variable. Fourth, technical assessment is often delegated to local teachers, trainers, and assessment bodies (see Box 2). Compare these challenges with an academic assessment – say mathematics – where a specific competence is assessed, as one of a handful of core academic skills, and where assessments are dominated by standardised and often national written examinations.

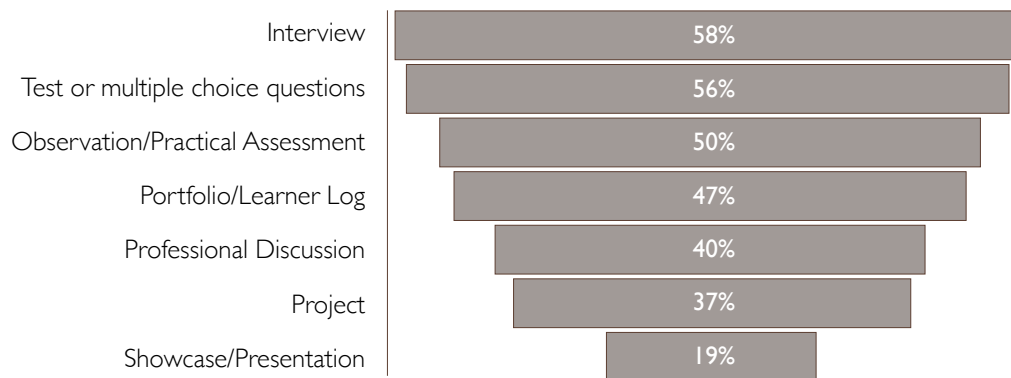
BOX 2. IN ITALY, ASSESSMENTS THAT VARY BY REGION AND INSTITUTION

In Italy, alongside regular school-based upper secondary technical education, there are regionally organised technical programmes (IFP). The assessments for these programmes vary from region to region, but some elements are common. There are three types of tests, theory, practical, and oral, developed by institutions in some regions, and at regional level in others. Sometimes theory tests are prepared regionally and practical tests at institution level. The final mark is a weighted sum of marks from the last year of study, the final examination and sometimes an employer appraisal from a placement. The weighting varies between regions. Although assessment procedures, tests and admission requirements to the final examination are defined nationally, assessment scales vary by region (Eurydice, 2020⁷).

The diversity of technical assessment makes it hard to map

Given its multi-level diversity, it is hard to quantify or summarise – to map – how technical assessments are conducted in a given country or sector. Summary metrics, including the numbers or proportions of assessments using specific tools, such as multiple-choice tests, or workplace observation, or portfolios of completed work, are rarely collected, although SDN has made a promising start in England (see Figure 1). Moreover, multi-level diversity makes international assessment instruments, that have proved so powerful in the context of general education, difficult to deploy (see Box 3). As a result of this largely unquantified diversity, it rarely makes sense to refer to the technical assessment ‘principles’ or ‘system’ employed in a country.

Figure 1. A promising start in quantifying assessment: percentage of apprenticeship assessment plans in England that use different assessment tools



This figure shows the results of an analysis of 208 apprenticeship standards, employing 637 assessment tools (all the plans and tools published at the time of the analysis). On average each assessment plan makes use of about three assessment tools. (Strategic Development Network, 2018⁸).

BOX 3. COULD A PISA-TYPE INTERNATIONAL STUDY OF TECHNICAL COMPETENCE IN OCCUPATIONS BE UNDERTAKEN?

The OECD PISA study, providing an assessment of literacy, maths and science knowledge and skills, has become a key benchmark in comparing the performance of school systems across the world. Could the success of PISA not be emulated by comparing technical training systems through a standardised assessment, so that, for example, electrician training in England could be compared with that in the Netherlands according to the measured skills of newly-qualified electricians in the two countries? This beguiling prospect has attracted attention over many years - see for example (Baethge, 2006⁹).

But the obstacles have proved formidable. Occupational competence involves a diverse range of soft and hard competences, and is often subject to cultural and contextual variation bearing on international comparability. In these respects the assessment challenge is much greater than an assessment – say – in mathematics. Moreover, while developing an international assessment of mathematics is demanding and resource intensive, these costs are balanced by the huge value of information about something which bears on the education of every child. No such returns are available from an assessment which might bear on just one out of several hundred occupations.

One special form of international assessment takes place through the World Skills Competition, in which national teams compete in a sports-type competition to display their prowess in a variety of technical domains. The assessment takes a form suitable for an international spectator competition and is therefore not designed to assess full occupational competence, but it plays an important role in popularising and giving status to technical competences (WorldskillsUK, 2021¹⁰).

There is little evidence on validity and reliability

There is little solid knowledge about what makes for an effective assessment, in terms of validity (accurately measuring the desired competences), and reliability (consistent measurement across different contexts) (see Box 4). We have little or no research evidence on whether observing trainees in the course of their ordinary work, or, alternatively, setting a separate practical test outside the usual working context is a more valid way of determining occupational competence. When portfolios of evidence of work achievements are used to measure occupational competence, we know little about how reliable these assessments may be, recognising that some individuals may have had help in preparing their portfolios, or have chosen the contents more wisely than others. Hard evidence on these matters would require complex and costly research studies, following up assessments longitudinally to explore how well they predict subsequent job performance, and retesting examinees with standardised tests to check assessment reliability. Even where such evidence exists, the diversity of assessments means that localised evidence on effectiveness yields little generalisable guidance on good practice. So not only is the evidence base weak, constructive development of that evidence base is also challenging.

**BOX 4. VALIDITY AND RELIABILITY IN ASSESSMENT:
AN ILLUSTRATION USING THE EXAMPLE OF FRENCH
MEDICAL EDUCATION**

Assessments are commonly appraised in respect of their validity and reliability (New Zealand Ministry of Education, 2021¹¹) (Darr, 2005¹²) (Darr, 2005¹³).

- *Validity* refers to the capacity of an assessment to accurately measure what it intends to measure. For a technical assessment this will mean the capacity to accurately measure the ability of the candidate to perform well in the target occupation.
- *Reliability* refers to consistency, so that the assessment applies the same standards to different individuals, in different contexts, with different assessors and assessment bodies, and does not change over time. Various quantitative measures of reliability are available.

These two characteristics are different, but not independent. A highly valid assessment must, by definition, be relatively reliable, since a very inconsistent measure cannot yield accurate predictions of occupational competence. However a reliable technical assessment may have low validity. For example in France, the six-year technical programme for doctors used to terminate in a single national written examination, the *épreuves classantes nationales* (ECN). Marks received in this examination had a big influence on the career of a doctor, as places in more or less prestigious hospitals and specialities are determined by the marks. While this national examination removes the risk of variation in marking between test centres – a critical point given the influence the examination has on a medical career – it is being reformed because it is deemed to have insufficient emphasis on medical skills as opposed to knowledge. The new assessment tools will include simulations to measure practical medical and interpersonal skills. So an assessment of

high reliability but questionable validity (the national examination) is being reformed in order to improve validity, but possibly also reducing reliability (Coussin, 2017¹⁴) (Ministère de Solidarité et de Santé, 2018¹⁵).

Some practical steps would improve the evidence base

Looked at across countries, there are few obvious examples of outstanding practice in data collection and research on assessment that could be imported to England. But better evidence on assessment could be collected. For example:

- A taxonomy of assessment techniques, including standardised/non-standardised, periodic/final, and synoptic/atomistic approaches might be developed. This taxonomy should be designed to be readily applicable in practice, rather than just theoretical, recognising that experience has shown, for example, that it is sometimes quite difficult to identify when an assessment is synoptic (see Box 13).
- Drawing on this taxonomy, a better understanding of assessment techniques could be developed, based on the systematic collection of data on assessment techniques across a range of different qualifications and programmes.
- Data on pass rates, and grade rates by field of study and training provider could be routinely collected, and backed up by information on resits and retakes. This is key information that can help to pinpoint any areas of doubtful or inconsistent practice, supporting quality assurance arrangements.
- More systematic research could be undertaken on different assessment tools to explore how well they predict later job performance. This could compare teacher assessments, employer assessments, written tests, and practical tests in relation to different occupational targets.

Better data support transparency

Quality assurance of assessment (discussed below in section 5) is, at present, rarely supported by data. Better data and evaluation would inform quality assurance by providing guidance on the assessment approaches that are most likely to be effective. Equally pass and grading rates for different subjects and different providers offer greater transparency to apprentices, employers, providers, and other actors regarding how difficult or easy it might be to pass, or obtain a higher grade. Any apparent inconsistencies in grading or pass rates become open to examination in the context of quality assurance.

BOX 5. AN APPROACH TO CONTINUOUS IMPROVEMENT IN TRAINING PROVIDERS IN WESTERN AUSTRALIA

The government of Western Australia provides guidance on a continuous improvement approach to assessment practice. This requires training providers to regularly review their assessment processes and outcomes to identify flaws and to pinpoint the scope for improvement. This involves a data collection strategy (for example client satisfaction surveys, and data from consultation with learners and industry bodies) to monitor assessments and support quality improvements. When assessment practice is modified to make improvements, the impact of the changes in assessment practice should be monitored (Government of Western Australia, 2016¹⁷).

Better data also helps quality assurance

Assessment bodies themselves, as part of internal quality assurance and improvement, can check that their assessments are providing reliable tests of occupational competence, by seeking feedback from employers on any gaps in the competences of newly-qualified persons, and addressing such gaps through modifications in assessment and through advice to training providers (see Box 5). Similarly, in England, the Institute for Apprentices and Technical Education (IfATE) emphasises the importance of continuous improvement in their quality assurance framework, and identifies the strongest end-point assessment (EPA) organisations as ones that see themselves as “learning organisations” constantly improving their performance through both internal quality assurance and feedback from stakeholders (IfATE, 2020¹⁶).

MANAGING THE UNCERTAINTIES OF ASSESSMENT***In principle, rigorous assessment should allow competences to be acquired by any route***

If an assessment successfully tests everything required for occupational competence, it should be irrelevant how a person has acquired such competence – formally or informally, through employment experience, or through different types of learning adapted to individual needs. These possibilities are fully exploited in recent radical reforms of the Finnish technical education system (see Box 6.) Unfortunately, in practice, assessment is a fallible instrument, and often we do not know how well our assessments are measuring occupational competence.

BOX 6. IN FINLAND, RADICAL FLEXIBILITY IN PATHWAYS AND THE IMPLICATIONS FOR ASSESSMENT

In 2018, Finland made large changes in its system of technical education. At upper secondary level, a fixed three-year programme was replaced by a more ‘personalised’ model for both adults and young people. Under this model, students may start the programme at any point during the year, and follow an individual study path of no fixed length, adapted to their individual needs and allowing for recognition of prior learning. Increasing emphasis is placed on virtual learning environments, and workbased learning. These different study paths all lead to the same final qualifications. Assessment, which previously involved some emphasis on examinations, will now give primacy to the demonstration of technical skills in the workplace (Finnish Ministry of Education and Culture, 2017¹⁸) (Karttunen, 2018¹⁹).

At first sight, the attractions of this model, in responding to individual student needs, encouraging lifelong learning, and meeting employer needs are substantial. However it is asking a lot of local providers to expect them to manage an infinite variety of learning pathways while maintaining consistent national standards. One profound challenge is to sustain assessment rigour, especially in terms of reliability, in this more fluid arrangement (see e.g. (Virolainen, 2018²⁰)).

Many countries allow flexibility in learning pathways

Some countries allow alternative learning pathways to final assessments and certifications, but without the complete flexibility now encouraged in Finland. The professional examination systems in Austria, Germany, and Switzerland, (see the description of master craftsperson qualifications in (Field, 2020²¹)), allow experienced workers to enter the examination with no constraints on how they prepare themselves for the assessment although there is a requirement for relevant employment experience. Similarly, many leading apprenticeship countries, including Austria, Canada, Germany, Norway, and Switzerland (but not England) allow experienced workers direct access to the final apprenticeship assessment (Field, 2018²²). In the Netherlands there are two parallel routes to upper secondary technical qualifications – apprenticeship, and a school-based route with extensive workbased learning (CEDEFOP, 2016²³). France offers many technical programmes – such as the *baccalauréat professionnel* – with an option of a workbased route (*apprentissage* or *alternance*) (CEDEFOP, 2018²⁴). These models depend not only on the robustness of final assessments, but on how the approved programmes are quality assured. They all contrast with England, where by and large each assessment is associated with a single type of programme – such as when an apprenticeship programme is associated with a specific EPA. The question this poses is whether in England, we might have sufficient confidence in the final assessment to permit more than one quality-assured programme as alternative pathways to the same final assessment and certification.

BOX 7. RECOGNITION OF PRIOR LEARNING AS A SPECIAL FORM OF ASSESSMENT: THE EXAMPLE OF FRANCE

(Colardyn and Bjornavold, 2004²⁵) use data on the use of RPL in European countries to describe five different methods used to assess and certificate prior learning. They identify:

- *examination*, often of a traditional written form;
- *declarative*, in which the candidate makes a presentation to a jury to explain and justify their competences;
- *observation*, in which the performance of the candidate at work is observed;
- *simulation*, in which a work problem is simulated, often using technology, to test work competences;
- *portfolios*, in which documentary evidence of employment experience and performance is presented.

For example in France, a 2002 law established an individual right to the recognition of professional experience (*validation des acquis de l'expérience*) in the acquisition of a qualification. This allows an individual to obtain part or all of the qualification based on professional experience. The candidate prepares an application documenting relevant professional experience, which is then examined by a panel including both academic and professional members. The panel may then either grant the full qualification, or alternatively set out the courses which need to be followed by the candidate to obtain it. The final qualification is the same as that which can be realised through a conventional learning programme (Chaparro, 2012²⁶).

Recognition of prior learning is a special type of assessment

Recognition of prior learning (RPL) is a special type of technical assessment, as it involves identifying competences regardless of how they have been acquired, but typically over a long period of time and through practical experience as well as formal education and training (see Box 7.) RPL has multiple benefits, particularly for adults with substantial career histories. It helps to make competences transparent, with benefits for both individuals and potential employers, and has strong equity benefits for groups such as migrants and informal workers who have skills but often lack recognised qualifications (Field, 2020²⁷). However experience in countries like Denmark and France with quite fully developed systems of RPL has shown that take-up is sometimes disappointing (Field et al., 2012²⁸) (Ministere de l'Education Nationale et de la Jeunesse, 2018²⁹).

Living in an unmapped world: implications for England

This section has described the thin evidence base on the reliability and validity of technical assessments, and how this hampers quality assurance. There are four implications for England:

- *2.1 Better data on assessment can be collected, and more analysis undertaken.*
It is challenging, but not impossible, to develop the evidence base on technical assessment. Assessment tools (such as work observations) can be classified, and their use in different technical qualifications quantified. The prevalence of elements such as periodic, synoptic and standardised assessments can be measured, alongside data on pass rates, grades, and resits. Such data can be disaggregated to show how different assessment bodies are using assessment tools, supporting data-driven quality assurance. Research on the effectiveness of different assessment tools can be undertaken, guiding quality improvement.
- *2.2 To demonstrate occupational competence in certification, assessment usually needs to be supported by requirements on the candidate's learning.*
Given the limited evidence base, assessment rarely offers a guarantee of occupational competence. It follows that for certification to represent a reliable indicator of occupational competence, final assessments usually need to be supported by other requirements, such as that the candidate has either passed through a training programme or has relevant employment experience. Such requirements provide added evidence that a full range of occupational competences have been acquired (see also section 5 and implications 5.1 and 5.2).
- *2.3 Consideration might be given to allowing different types of programme to lead to the same final assessment and certification.* England tends to follow a relatively rigid approach by associating each final assessment with just one type of programme (such as an EPA with an apprenticeship). In contrast, other European countries tend to allow more flexibility, allowing more than one type of programme, and therefore more than one pathway to certification and associated qualifications. England might consider adopting this approach more widely, recognising that decisions need to be made on a case-by-case basis, taking into account the rigour of the final assessment, and the evidence provided by each programme that occupational competence has been realised.
- *2.4 Policy and practice need to take account of the fallibility of assessment.*
The recognition that assessments are fallible needs to be built into the design of programmes and qualifications and the way in which they are used.

3 THE DESIGN OF ASSESSMENTS: THREE DILEMMAS

This section looks at three choices, and potential dilemmas, arising in the design of technical assessment systems. First, it examines the extent to which technical assessment tasks and procedures can and should be standardised, so that effectively the same test is applied to every candidate. Second, it looks at whether assessment should be undertaken only at the end point, or at different points throughout learning. Third, it considers whether assessments should be synoptic appraisals of occupational competence in the round, or more atomistic, separately addressing each of the elements which make up occupational competence.

STANDARDISED OR WORK-EMBEDDED ASSESSMENT TASKS?

Technical assessment involves both candidate tasks and procedures

Any assessment may be conceived as involving two elements: first, a set of tasks which candidates are expected to perform, and second, procedures for evaluating these candidates on the basis of those tasks. For example in England, the assessment plan for a pipe welder includes tasks such as welding operations and health and safety procedures, while the procedures include rules such as the requirement that for an overall pass, the candidate must pass every assessment module. The outcome depends both on the tasks and the procedures (IfATE, 2021³⁰).

Assessment procedures are normally as standardised as possible

If all candidates in an assessment have to undertake the same tasks, and are assessed using the same procedures, then the assessment should be reliable (see Box 4). Strenuous attempts are routinely made to ensure that assessment procedures, in the sense of criteria for assessment, persons involved in the assessment, rules for resits and retakes and so forth are as consistent as possible. Arrangements such as validation and external assessment, and other quality assurance measures, are often designed to reinforce procedural consistency and therefore reliability. In the Spanish Basque country, a calibration procedure is used to ensure that teachers use similar assessment and grading criteria. Every two years, groups of teachers are brought together to correct the same written assignment independently and then compare outcomes, discussing their proposed grades and seeking consensus. The results of this grading discussion are recorded for future use (CEDEFOP, 2015³¹).

Assessment tasks can also be standardised

Occupational competence is usually associated with a body of technical theory and knowledge, which can be assessed in a standardised written test. Thus an electrician needs to understand the physics of electricity. This theoretical or knowledge dimension is often classroom taught, and typically assessed through written examinations, which may be standardised (as in New Zealand, see Box 8). The practical dimension can more naturally be assessed through a test in which an electrician demonstrates that they can undertake – for example – rewiring a house (Misko, 2014³²). Practical tasks can also be standardised, by defining a set of tasks expected of all candidate electricians, and requiring candidates to undertake those tasks under controlled conditions, such as in a regional technical assessment centre, and in England this takes place through the AM2 test (National Electrotechnical Training, 2021³³). Standardised tasks may include (Alphaplus Consultancy Ltd., 2014³⁴) (IfATE, 2020³⁵):

- Traditional forms of written examination including multiple choice (see Box 8 for how this works in Canada and New Zealand).
- Practical examinations, in which all candidates are expected to perform the same, or a similar set of tasks, and are observed and/or interrogated about their approach to these tasks.
- Oral Interviews, involving a sequence of standardised questions.
- Role-play, sometimes to establish soft skills such as handling clients using a standardised set of role-play situations.

BOX 8. WRITTEN EXAMINATIONS FOR TRADE APPRENTICES IN CANADA AND ELECTRICIANS IN NEW ZEALAND

In Canada, apprenticeship is managed by the separate provinces, but a national examination is used to assess skills in different trades, and to certificate successful candidates not just for work in one province, but also across Canada. The Red Seal examination, which is based on national occupational standards, involves between 100 and 150 multiple choice questions, to be answered during a 4-hour examination. Around three quarters of candidates pass the exam. No practical or workbased tests are involved (Canadian Council of Directors of Apprenticeship, 2016³⁶).

In New Zealand, trainee electricians must undergo a practical assessment and a final computer-based examination. The practical assessment is undertaken in a decentralised way by different training providers, with results submitted to the Electrical Workers Registration Board for approval. The final examination for an electrician involves a multiple choice, computer-based test undertaken in an examination centre. Candidates must have undertaken an approved course with an approved training provider to be eligible to take the examination. Resits are allowed but after three failures within 3 months, some retraining will be required (Electrical Workers Registration Board, 2021³⁷).

Standardisation supports quality assurance of training

Standardised tasks in assessments not only measure candidate performance, they also provide data to support the quality assurance of training providers, for example by identifying training providers where students have unusually low marks. In Poland, national standardised examinations, with external examiners in regional centres, are used to quality assure the training provided in different regions and schools (Chłoń-Domińczak, 2019³⁸). Comparisons of results from non-standardised tasks would offer less reliability when it comes to pinpointing variations in training outcomes.

Simulation technology can be used both to train and assess students

In fields where working practice involves human subjects (as in healthcare), or expensive machinery (such as aircraft or CNC machines), technology-assisted simulation, where no persons or expensive machines are at risk, has large attractions. Simulation also allows practice and assessment in the handling of rare but critical events, such as medical emergencies in robot patients, or engine failures for pilots. A controlled set of challenges can be offered, both to train students, and subsequently to assess their skills. A substantial literature has emerged on

the use of such technology, recognising both its potential and its limitations as for example in simulating the use of interpersonal skills. Simulation technology may also facilitate standardisation in assessment, so that candidates face the same, or similar challenges in a final examination (Ahn and Nyström, 2020³⁹). In the reform of medical education in France, designed to enhance the assessment of practical skills, the intention is to use simulation technology, notably programmable robotic patients, extensively, not only to train but also to assess medical students (see Box 4 and (Ministère de Solidarité et de Santé, 2018¹⁵)).

Real world, work-embedded assessment tasks have many advantages

Often, assessment tasks are part of, or linked to, ordinary work activities. Such tasks may include ordinary work (assessed through observation), special projects undertaken in the context of workplace requirements, and the tasks reflected in portfolios of previous work. As illustrated in Table 1, such 'real world' work tasks have many advantages as the target of assessment. Their authenticity means that they offer direct and therefore convincing tests of occupational competence; they allow for the engagement of employers in assessment; and they can be less burdensome than standardised tasks, particularly when they are part of ordinary working practice. Moreover, as discussed below, they can be more readily designed as synoptic assessments, appraising the capacity to deploy higher level 'meta-skills' to solve novel problems. One example of this approach is in the Netherlands (see Box 9).

Table 1. *The relative advantages of standardised and work-embedded tasks in technical assessment*

	Standardised tasks	Work-embedded tasks
Confidence that the same standards are applied to all candidates	✓	More challenging
Suitability to cognitive aspects of competence	✓	More challenging
Realistic work tasks assessed.	More challenging	✓
Engagement of employers in assessment	More challenging	✓
Suitability to synoptic assessment and the assessment of meta-skills	More challenging	✓
Cost	National practical assessments are burdensome	Tend to be cheaper

But the variability of work-embedded assessment tasks create a challenge for reliability

Despite their advantages, work-embedded assessment tasks vary from workplace to workplace, and therefore from one candidate to another. This makes it, as (Stanley, 2017⁴⁰) puts it, "very difficult to address consistency" in the assessment. Similarly (Yu and Frempong, 2012⁴¹) recognise a tension between complete standardisation in assessment and work-embedded forms of assessment. The implication is that some technical qualifications accept risks in terms of consistency of assessment standards in return for the deepened credibility realised through assessment of authentic occupational practice.

There are ways of profiting from both standardised and work-embedded assessments

There is a compelling case for work-embedded assessment as the most credible test available for occupational competence. But it would be self-defeating if the effect were to permit so much variation that pure luck comes to determine the difficulty of the assessment task, and hence who passes the assessment. While the tension between these two perspectives is a challenge, there are ways of designing assessments so that this tension can be effectively managed, if not eliminated:

- Assessment tasks, even if work-embedded, may still be required to meet some standard requirements, for example to ensure that they always allow for the assessment of those skills identified as key parts of occupational competence. Moreover, the procedures used to assess the work-embedded tasks may be subject to standardisation. Practice in the Netherlands illustrates both points (see Box 9).
- Many countries blend standardised and work-embedded assessment tasks in a composite assessment. In the Czech Republic, students in technical upper secondary programmes are assessed through a combination of national exams and practical tests devised and managed by local technical schools (CEDEFOP, 2015³¹). In New Zealand, trainee electricians pursue practical assessments organised locally by their training provider, and a national examination to test knowledge (see Box 8). In Switzerland, assessment of the practical skills of an apprentice involves first, an examination related to a standardised set of tasks or project which are the same for all candidates in the occupational field, and are usually conducted at the same time; and second, an individual practical project completed at the workplace and agreed with the individual employer. The project is presented by the apprentice to the examiners who award the grade - see annex A and (International Labour Organisation, 2020⁴²).

BOX 9. WORK-EMBEDDED ASSESSMENT PROJECTS IN THE NETHERLANDS

In the Netherlands, the practical component of technical assessment at upper secondary level is linked to placements. An examination project, associated with a real professional activity, is chosen and approved (by the student, trainer at school and workplace trainer) at the beginning of the school year. The candidate must then carry out the project within the company (whilst treated as a regular employee) over a period of around six weeks. The student prepares a written report and a presentation, in which they are expected to demonstrate mastery of the required learning outcomes. Assessment and grading are undertaken by a minimum of two people to ensure impartiality – typically the school and workplace trainers. In their final year, students take part in three of four of these projects. This practical assessment is a component of a decentralised assessment system where the training providers themselves (or regional authorities) develop examination questions and tasks. However, since 2014 the certification exams in English, Dutch, and Mathematics which form part of the technical programmes have been centralised (CEDEFOP, 2015³¹).

PERIODIC OR FINAL ASSESSMENT?

Periodic and final assessments have competing attractions

Final assessments may be contrasted with periodic assessments undertaken at intervals throughout a learning programme and contributing to the final grade. (The terminology of 'continuous' assessment can be misleading since assessment processes are rarely truly continuous). A CEDEFOP survey found that, England apart, nine of the eleven European countries they surveyed tended to use final assessments in their initial technical education systems, with the exceptions being Spain and Finland (CEDEFOP, 2015³¹). Periodic and final assessments have competing advantages:

- *Periodic assessment* has a useful formative as well as summative function. It can also fit well with a modular approach in which individual elements of occupational competence are separately assessed, facilitating a record of partial learning. This can be used, as in Denmark, as a means of granting the possibility of later completion to those who might otherwise drop out with nothing to show for their efforts (Danish Ministry of Education, 2008⁴³). The case for an element of periodic assessment is therefore stronger in longer programmes, where students need formative feedback in the course of the programme, and the risk of dropout is higher.
- *Final assessments* are better placed to assess occupational competence as an integrated whole through a 'synoptic' assessment (see below). Periodic assessment with written assignments can be at risk of plagiarism, although such risks can be minimised (see Box 10).

BOX 10. AVOIDING PLAGIARISM IN PERIODIC ASSESSMENT

One of the challenges in periodic assessment is that if students are offered time to complete periodic assignments, they may resort to plagiarism, or get help from friends with relevant expertise. Periodic assessments are often more informal, with less scrutiny than might occur in a final examination. Written exams at the end of a programme are often used to deal with this problem, increasing reliability, but sometimes reducing the reliance on authentic work problems.

Various techniques have been proposed to deal with this problem – including assessment methods that relate to real world challenges, oral assessments involving face-to-face interaction between lecturers and their students, and task design in which students resubmit assignments after receiving feedback (Bjaelde, Lauridsen and Lindberg, 2018⁴⁴). One solution, adopted in a bachelor's degree on programming in England, is to randomise student assignments, while ensuring that they all reflect the curriculum (Bradley, 2016⁴⁵).

Formative assessment is important

Formative assessment uses information about learning progress to guide both teachers and learners and thereby support learning. Much evidence shows that it is a powerful pedagogical tool in general education (Black and Wiliam, 1998⁴⁶), suggesting, although direct evidence is limited, that the same might be true in technical education (University of Exeter, 2021⁴⁷). Norway requires half-yearly formative assessments both in the school and the workplace segments of their 2+2 system of apprenticeship. In the final two-year apprenticeship period of the programme, half-year assessments are undertaken by the training supervisor of the apprentice in the training company. The supervisor is expected to explain to the apprentice what competences they still need to acquire, and how they can be acquired (Norwegian Ministry of Education and Research, 2006⁴⁸).

Some programme elements call for immediate assessment

Some critical elements of a training programme can be more readily assessed at the time they take place, implying an element of periodic assessment. A placement in industry, in which the trainee gains practical working experience as part of a technical programme provides an opportunity for assessment that cannot easily be replicated at the end of the programme. Assessments of placements are common (see Box 11).

BOX 11. ASSESSING PLACEMENTS* IN TECHNICAL PROGRAMMES: SOME COMPARISONS ACROSS COUNTRIES

Placements, where students spend an extended period with an employer (in programmes other than apprenticeship) are a key part of technical programmes, as it is in this context that students gain a concrete understanding of the demands of the workplace, both in respect of practical technical skills, and in more subtle requirements such as handling working relationships. Assessing the learning outcomes from such placements therefore offers a strong indicator of occupational competence, and signals to the student the importance of what can be learnt in the workplace. But as these placements are at one remove from the main training provider for a technical programme, assessing the learning outcomes can be challenging.

The new T-level qualifications in England include a mandatory industry placement, and employers will be required to prepare an appraisal of the performance of students undertaking these placements (DfE, 2021⁴⁹). While these appraisals will not contribute to T-level grades, the Sainsbury review recommended that the T-level certification should include the employer's appraisal of the student's performance during the industry placement (Independent Panel on Technical Education, 2016⁵⁰) (DfE, 2020⁵¹). Some countries make assessments of placements a more formal part of overall assessment. For example in France, the 22-week placements that are part of the baccalauréat professionnel are subject to an assessment by the teachers

*In this report, a 'placement' refers to elements of technical programmes in which students spend substantial periods of time undertaking work with employers. Such placements can be distinguished from short exercises of work observation or job-shadowing.

from their vocational schools, and represents a varying (according to profession) but substantial contribution to their overall mark in the baccalauréat. In the Netherlands, a large part of the assessment of students in technical programmes is based on performance on a project undertaken over some weeks with the host employer, assessed through a presentation and discussion with the candidate (CEDEFOP, 2015³¹).

Mixed models are possible

Periodic assessment often supplements final assessment in a mixed model. In German apprenticeship, an assessment normally takes place halfway through the 3–4-year programme to measure the apprentice's acquisition of both theory and practical skills. This is used formatively to provide feedback on learning progress, but, increasingly, it is also used summatively, representing 30-40% of the final mark for the apprenticeship, depending on the profession (see the case of plumber assessment in Germany (BIBB, 2021⁵²)). In Swiss apprenticeship, some elements of periodic assessment, reflecting marks given by teachers in inter-company training courses, and in classroom-taught courses, contribute to the final mark in the overall assessment (see Annex A). In Luxembourg, technical assessment involves a unique mix of modular periodic assessment and a synoptic, work-embedded final assessment (see Box 12).

BOX 12. A BLEND OF PERIODIC MODULAR AND SYNOPTIC FINAL ASSESSMENT IN LUXEMBOURG

In Luxembourg, the technical education system has many similarities with that of Germany, with a dual system of apprenticeship at upper secondary level, alongside some school-based technical programmes. Summative assessment involves both periodic and final assessment (European Alliance for Apprenticeships, 2021⁵³):

- *Modular periodic assessment.* The programmes are organised in modules, each leading to a subset of competences for a specific occupation. Each module is assessed individually by the technical education teacher or the in-company trainer responsible for the associated teaching or training. The apprentice must pass a fixed proportion of mandatory modules before entering the final assessment.
- *Final synoptic assessment.* A 2008 reform replaced theoretical and practical final exams with an assessment based on an integrated project, which corresponds to a simulated or real working situation, undertaken over a period of up to 24 hours. The integrated projects are developed and assessed by teams of experts from employer organisations, and technical teachers from secondary schools (plus some additional assessors). Success in this final assessment leads to certification.

SYNOPTIC OR ATOMISTIC ASSESSMENT?

Occupational competence may be conceived in a holistic or atomistic fashion

'Occupational competence' has a concrete meaning for employers who know what they expect in a recruit. However the expression remains somewhat fluid, as professions evolve over time, so occupational competence needs to include the capacity to learn and adapt, as well as immediate job-readiness. In the UK, as in many English-speaking countries, occupational competence has been linked to defined occupational standards which set out what is expected of those working in the occupation (Allais, Molebatsi and Marock, 2014⁵⁴). But countries have pursued different approaches to occupational competence – notably in the UK, France and Germany – with implications for technical training and assessments (Lester and Religa, 2017⁵⁵). One key distinction is between atomistic and holistic approaches.

- Under an *atomistic* approach, occupational competence is defined in terms of a list of the types of knowledge, skills and behaviours required for the target occupation. This lends itself to an assessment that seeks to measure acquisition of each and every competence on this list (unless this requirement is relaxed to allow just a sample of these competences to be tested). This has the advantage of demonstrating a transparent link between assessment and the list of competences which constitute occupational competence.
- Under a *holistic* approach, occupational competence is associated with a professional identity, and the core abilities that maintain that identity over time, even while specific skill and task requirements come and go with changing technology and workplace organisation. This approach, and the central role it gives to occupational identity is influential in the German apprenticeship system. While in the past this approach was less salient in the UK (see (Fuller and Unwin, 2013⁵⁶)), the move towards a more holistic approach was central to the reforms introduced following the Richard review, even if some elements have proved challenging to implement (see Box 13).

BOX 13. WHAT HAPPENED TO SYNOPTIC ASSESSMENT IN ENGLISH APPRENTICESHIP?

...there needs to be a test that demonstrates that the apprentice can take the knowledge and expertise they have gained and apply it in a real world context to a new, novel problem. The final test and validation must be holistic, in that it seeks to test the full breadth of the relevant competencies not merely the incremental progression of the apprentice. That may take the form of a project or an assessment in front of an examiner. It should be performance and real world based, rather than just theoretical. It should be primarily at the end of an apprenticeship, not measuring progress during it. And the examiners should be neutral parties with no interest in the outcome, drawn from the ranks of employers as well as educators. (Richard, 2012⁵⁷) p. 8.

In implementing the findings of the Richard review, the government stated that "synoptic, end-point assessment means that an apprentice will need to be able to demonstrate the full range of skills, competency and knowledge required to meet the standard, putting different strands of learning together and applying these in different contexts" (DfE and BIS, 2013⁵⁸). Pursuing that principle, IfATE requires EPAs to use at least one 'synoptic' assessment

method, (for example work observation) defined to mean that the method is used to test some of the required knowledge, some of the skills and some behaviours required by the occupation (IfATE, 2020³⁵). However this means the assessment method can be used on separate occasions to assess different KSBs – for example observing a technical skill in a work context on one occasion, and observing interpersonal skills on a separate occasion, which is not a synoptic assessment. In the last 6 approved apprenticeship standards at levels 3 and 4 (approval dates from 9.11.2020 to 24.2.2021) there is only one mention of synoptic assessment in the six assessment plans (IfATE, 2021⁵⁹). Ofqual's review of the quality of assessment materials developed by assessment bodies does not mention synoptic assessment (Ofqual, 2020⁶⁰).

There are reasons for the implementation challenges. Synoptic assessment is hard to operationalise into a set of clear enforceable rules for assessments. As a result, the principles behind synoptic assessment can easily be marginalised relative to the apparently more concrete need for the assessment to cover all the KSBs. But the effect is that the whole rationale for the replacement of atomistic apprenticeship frameworks with holistic standards has become unclear.

Some professions with a clear mission demand a holistic conception

The resonance of holistic approaches to occupational competence depends partly on the profession: professions like teaching and nursing are associated with strong professional missions and values, independent of the changing set of competences necessary to deliver those missions. In other professions, like IT technicians, professional missions are harder to define. Regardless of the profession, the capacity to perform at work depends both on specific technical and soft skills and on the more holistic ability to deploy the right mix of skills to address evolving challenges.

Synoptic technical assessment often requires authentic work-embedded assessment tasks

In any job, workers need to deal with unexpected developments. This ranges from something as simple as dealing with an unexpected complaint from a customer to an adaptation to new technology over a period of years. In the face of these and other demands, a fixed skillset associated with a set of traditional tasks is not enough. Instead, workers need the high-level integrative capacity necessary to build on existing skills, research new information, develop the new skills needed to solve novel and open-ended problems, and perform effectively in changing work environments. In the UK in the last two decades the terminology of 'synoptic' assessment – meaning, in the technical sphere, holistic assessment of occupational competence – has emerged, and synoptic assessments have come to be expected, at least as an element of overall assessment, in both general and technical education. Synoptic technical assessment tends to rest on practical tasks showing how candidates address real world workplace challenges. This may involve either assessment of workplace performance or of specific problem-solving projects. Box 12 gives an example of how this works in Luxembourg. Such work-embedded assessments are difficult to standardise (see Table 1) and (Constantinou, 2020⁶¹). In England, implementation of synoptic assessment in apprenticeship has proved challenging (see Box 13).

Synoptic assessment may be required to measure high level 'meta-skills'

Some high-level competences, involving the capacity to deploy a range of knowledge and skills to solve practical problems, are more naturally addressed by synoptic assessment. Gaps in such competences were highlighted by the Sainsbury review, when it reported that "Employers repeatedly told us that, although applicants may look good on paper, they often struggle to apply knowledge and skills in the workplace" (Independent Panel on Technical Education, 2016⁵⁰). The terminology can be confusing, as such high-level competences are variably characterised, in overlapping but varying notions, as 'soft', '21st-century' 'meta-cognitive' and 'cross-curricular' skills and competences. One example is the identification of 'meta-skills', including self-management, social intelligence and innovation, as the key to future-proofing the skills system in Scotland. It is argued that such competences are not easily taught in a classroom, and they can most naturally be developed in a workbased context. By the same token, they are difficult to measure, and therefore assess, except in the context of regular work, or special projects that closely mimic the demands of ordinary work (Skills Development Scotland and Centre for Workbased Learning, 2018⁶²). Similarly, Bjaelde and colleagues describe how assessments of higher-level professional competences (often at tertiary level) require candidates to solve 'authentic' work problems with competences that may include, among others, teamwork, critical thinking and interpersonal skills (Bjaelde, Lauridsen and Lindberg, 2018⁴⁴).

Three dilemmas in the design of technical assessments: implications for England

How then do we balance the competing arguments for standardised and work-embedded non-standardised forms of assessment, or between periodic and final assessment, or synoptic and atomistic assessment? Three implications for England are set out below.

- *3.1 Technical assessments need to balance the strong arguments for authentic work-embedded assessment tasks, with the need for sufficient standardisation to preserve reliability.* In pursuit of this balance, technical assessment often involves a mix of standardised and work-embedded assessment tasks. When assessment tasks are work-embedded and variable, standardised criteria can be used to constrain and guide the choice of suitable assessment tasks, and standard procedures used to undertake the assessment. This approach is often reflected in England: for example the EPA for a level 3 knitting technician apprenticeship standard, requires (like all standards) employer approval at the gateway, and observation of work in practice, (both these elements being work-embedded). The multiple-choice test included in the EPA offers a more standardised element (IfATE, 2020⁶³).
- *3.2 The potential of periodic assessment needs to be reassessed.* An end-point assessment is administratively neat, and provides a good frame for a synoptic assessment, but it can often be usefully augmented by periodic assessment, especially in longer programmes. An element of periodic assessment may be particularly useful when elements of a programme such as a placement with an employer demand assessment at the time; when it is helpful to provide partial certification to those who drop out; and when it also usefully contributes to formative assessment. (Formative assessment is a pedagogical tool whose value is now widely promoted in general education, but with a potential that may still be insufficiently exploited in technical education). For England, this raises two specific options: first, that of some form of half-time assessment in longer

apprenticeship programmes, following the German model; second, including in T-level certification the employer's appraisal of student performance during the industry placement, as proposed in the Sainsbury review.

- 3.3 *Synoptic assessments need to be fully implemented in apprenticeship.* In England, one major thrust of the Richard review, and the subsequent reform that replaced frameworks by standards was to establish a more holistic approach to occupational competence, embodied in a synoptic end-point assessment. However implementation has proved challenging, with limited evidence of a synoptic approach remaining in the implementation of end-point assessments. To remedy this, one option would be to require all assessment plans to set out explicitly how they plan to implement synoptic assessment. A final assessment project, as proposed in the Richard review, employed in many European countries, and already used in many apprenticeship EPAs, would be one means of realising this routinely.

4 WHY THE OBJECTIVES OF ASSESSMENT MATTER

This section identifies three key objectives of summative technical assessment and explores their implications. It looks first at assessment as a means of demonstrating occupational competence, noting how involving employers in the assessment process may enhance its credibility. Secondly it examines the role of assessment in relation to the wider competences that support transitions to further learning. Finally it explores how assessments may motivate students, especially by offering a grade above pass.

Technical assessment can have diverse objectives

The design of any assessment should depend on what the assessment is intended to achieve – its objectives (Newton, 2010⁶⁴). It follows that an assessment designed to convince employers of a candidate's ability to do the job will be quite different from one where the measurement of academic ability and potential is a key aim. Assessment objectives therefore bear on the design choices – such as standardised versus work-embedded assessments – discussed in the previous section. Summative technical assessment usually has a mix of purposes, partly reflecting how different interest groups, such as learners, employers, and trainers, will use the assessment (Clayton, 1995⁶⁵).

Three main objectives can be identified

Three main objectives of summative technical assessment may be identified, and in the remainder of this section they will be considered in turn:

- To provide an effective test of occupational competence, that is credible to employers;
- To demonstrate the wider and deeper competences that underpin further learning, adaptation to change and career development;
- To motivate students, recognising and rewarding their learning, and signalling the mix of competences they need to acquire.

CREDIBLE TESTS OF OCCUPATIONAL COMPETENCE

Assessments need to be valid, and seen to be valid by employers

Good technical assessments must not only be valid, they must also be seen to be valid by employers, otherwise the associated qualification will have no labour market value for the qualification-holder. This raises the question of the extent to which assessments need to be embedded in working practice to be credible to employers, and how employer involvement in assessment may support credibility. Employers will naturally have most confidence in their own observation of a trainee facing a real work task; no abstract test or examination will be as convincing.

Employer involvement in assessment may enhance its credibility with other employers

Employer involvement in assessment has a double benefit – first, employers in the relevant sector are well-qualified to judge occupational performance, and second, the involvement of employers increases the credibility of the assessment in the eyes of other employers. This sometimes emerges directly in the organisation of assessments and associated certifications. In Estonia, technical schools organise the final assessment leading to a school-leaving certificate. But the occupational qualification linked to the school programme, and the associated assessment for

that qualification, is organised by awarding bodies (employers' or professional associations) set up by employer-based sector skills councils (CEDEFOP, 2015³¹).

In England, employers have a limited role in undertaking assessments

The major assessment role played by independent bodies in England is unusual by international standards. Typically, when a qualification is established by an awarding body in England, that body specifies both the content of the qualification and how it is to be assessed. A training provider then delivers the programme and the assessment under the supervision of the awarding body. This means that awarding bodies and training providers, but not employers, are directly involved in undertaking assessments. The Richard review of apprenticeships argued strongly for employer involvement in assessments, such that a representative of industry would be part of the team that undertakes assessments (Richard, 2012⁵⁷). While employers now develop the assessment plans through Trailblazer Groups, employers are not necessarily involved in undertaking assessments in the way that Richard envisaged.

BOX 14. IN GERMANY, EMPLOYERS ARE SYSTEMATICALLY INVOLVED IN APPRENTICESHIP ASSESSMENTS

In Germany, assessments for apprenticeship involve multiple stakeholders. The assessments are organised by the economic 'chambers' (sectoral employer organisations) and the chambers set out the framework for the assessment, subject to approval by federal government authorities. Chambers establish assessment boards to deliver the assessments – these boards contain at least three members, and two-thirds of the participants must be representatives of employers and employees. Members of assessment boards are unpaid volunteers. The assessment boards may request 'expert reports' from the trainers of the apprentice within companies – for example in respect of apprentice performance of practical tasks, and from the teachers in vocational schools, and take account of these reports in deciding whether an apprentice has passed the assessment, but they are not bound by these reports. (Federal Institute for Vocational Education and Training, 2017⁶⁶). Chambers issue certificates to successful candidates alongside the vocational school and individual employer certificates (see Box 1).

The optimal pass threshold for a technical assessment depends on the relative costs of different assessment errors

For the sake of simplicity, let us assume that there is a binary distinction between occupationally competent and non-competent candidates in an assessment. While assessments will seek to pass all and only those who are competent, in practice, as fallible instruments, they will sometimes pass those who are not competent, and fail those who are competent. If passing the non-competent is seen as the most worrying type of assessment error (as for airline pilots) the passing threshold should be set high, even if the cost is that some competent candidates fail the examination.

Employers may prefer a higher passing threshold than training providers

Different stakeholders will perceive the costs of these different types of assessment error differently, leading to different views on the optimal passing threshold. The candidate, the training provider, and the individual employer in the case of

apprenticeship will normally want the apprentice to pass, since a positive outcome will reflect their collective efforts. Conversely employers in the sector will want the pass threshold to be sufficiently high for them to have confidence in the skills of those with the qualification. The failure rates in the final assessments for apprenticeship in England, (where employers are not regularly involved in undertaking assessments), are significantly lower than in Germany, where employers are more fully involved (see Box 14) and in Norway (see Box 15).

In England, where employment protection is relatively weak, employers may care less about rigorous assessment

Relative to England, Germany and Norway make it more difficult to fire non-competent staff (see Figure 2). As a result, rigorous technical assessments (and underpinning them robust technical programmes) may be even more important for a German or Norwegian employer than for a UK counterpart, since the costs of recruiting an incompetent worker by mistake are higher than in the UK. (Minimum wage levels are also relevant to this calculation, but relative to the medium wage, there is little difference between the UK and Germany on this point, with no available comparable data for Norway (OECD, 2021⁶⁷)).

BOX 15. WHAT IS THE ‘RIGHT’ PASS RATE FOR FINAL ASSESSMENTS FOR APPRENTICES? ENGLAND COMPARED WITH GERMANY AND NORWAY

In England, in 2019/20, around 40% of all those starting apprentices on apprenticeship standards, failed to complete. The remaining 60% undertook EPAs, at which point just 1.5% of the original starters fail (see Table 2). So there are around 27 times as many prior dropouts as EPA failures. Some apprentices may drop out because they do not see their way to passing the EPA, but this effect may be limited: recent qualitative research suggests that fear of failure in the EPA is no more than a minor factor in apprentice dropout, which usually has multiple causes (Kantar Public, 2019⁶⁸). For apprenticeship standards, IfATE guidance allows a lot of flexibility to assessment plans (IfATE, 2020³⁵). For example, the advanced upholsterer standard allows for resits and retakes of each failed assessment module, but resits and retakes are not permitted as a means of seeking a higher grade (IfATE, 2020⁶⁹). Perhaps more significantly, apprentices only face an EPA once an employer considers them ready to do so through the ‘gateway’, so, from the individual employer’s point of view, the EPA may provide no more than formal confirmation of their view of an employee who, day to day, demonstrates occupational competence – at least in their current job.

Table 2. *England 2019/20 apprenticeship standards*

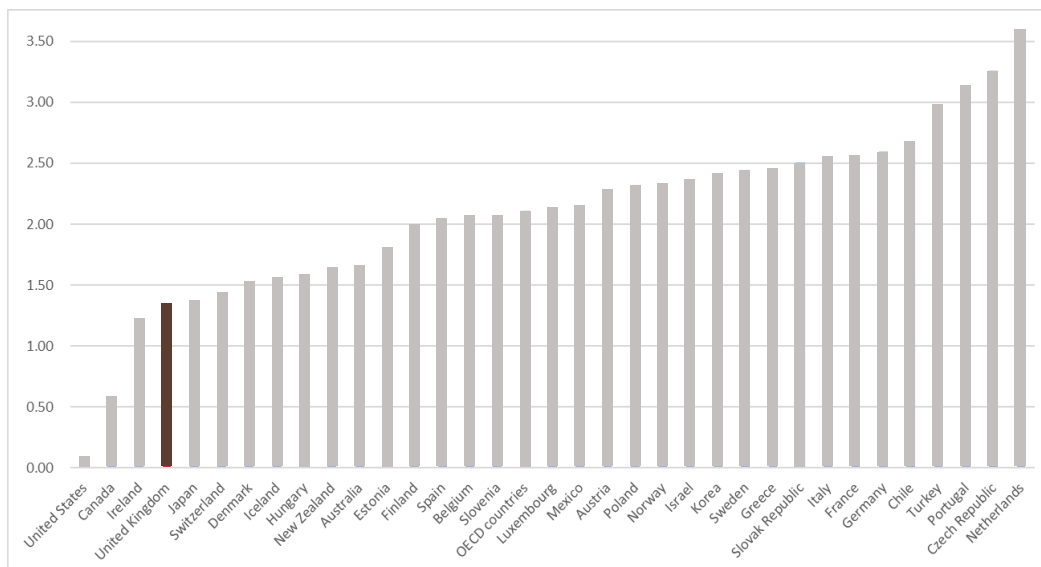
Retention rate (% of starters who reach EPA)	EPA pass rate (as a % of those who reach EPAs)	Achievement rate (% of starters who reach and pass EPAs)	Failures in the EPA expressed as a % of all starters.
60.2%	97.5	58.7%	1.5%

Source: (DfE, 2021⁷⁰)

These figures demonstrate that the final hurdle – the EPA – plays a small role, relative to dropout, in determining which starting candidates successfully complete their apprenticeship, as it only fails 1.5% of starting apprentices. This also implies that the shape of the EPA may have a limited incentive impact on students. This perspective has implications for the relative importance of efforts to improve assessment, as opposed, for example to measures designed to reduce dropout rates.

In Germany only around 10% of apprentices drop out during the apprenticeship, and in the final examination, around 90% pass on the first attempt, but 95% eventually pass, sometimes only on second or third attempts, after a permitted extension of their apprenticeship contract of up to one year (BMBF, 2019⁷¹) (BMBF, 2003⁷²). Similarly in Norway, in 2015, 80% of candidates who entered a vocational programme in 2011 passed the exam; 6.7% completed their apprenticeship but failed the exam; 12% failed to complete their apprenticeship and 1.3% were still undertaking their apprenticeship (Norwegian Centre for International Cooperation in Education, 2016⁷³). (In Canada, also, some 25% of candidate apprentices fail the Red Seal apprenticeship certification examination – see Box 8). These figures imply, firstly, that final assessment failure rates for apprentices are two or three times higher in Germany and Norway than in England, and, secondly, that while drop out is roughly twice as likely as final exam failure in Germany and Norway: the comparable ratio for England, at around 27, is much larger. So the low EPA failure rate/ high dropout rate pattern for apprentices in England is quite different from that observed in some other apprenticeship countries.

Figure 2. The UK has weak employment protection relative to many OECD countries



OECD index of difficulty of individual dismissal for workers on regular contracts 2019:
a higher index indicates greater difficulty in dismissing workers

Source: (OECD, 2021⁷⁴)

PATHWAYS TO FURTHER LEARNING

Technical programmes often need to provide the competences that support further learning

In recent decades, partly because of rising educational aspirations, and partly because of increasing skills requirements in the labour market, many technical programmes and qualifications have been modified to facilitate access to further learning, including higher education (Field and Guez, 2018⁷⁵). Consequently, technical assessments have needed to give attention not just to the immediate ability to perform on the job, but also to the foundational competences that enable individuals to go on learning. These include numeracy and literacy, but also wider elements of general education. In many countries, this is realised through an upper secondary technical track, which offers occupation-specific technical training alongside a substantial component of general education. While upper secondary education in England is organised differently, some of the challenges are similar. Many of those considering or entering technical pathways in England would also be interested in higher education options. The profound importance for technical upper secondary education of maintaining an effective pathway to higher education is vividly illustrated by recent experience in Sweden and Denmark (see Box 16).

European countries offer different routes of assessment and certification leading to higher education

For most countries with an upper secondary technical track, pathways are in place to allow students to progress from technical qualifications into other education and training programmes, including higher education. In some cases this is achieved through a qualification that integrates an academic and technical assessment, such as the French *baccalauréate professionnel*, in which the overall mark out of 20 includes marks for academic subjects such as maths and French alongside marks awarded for demonstrating occupational competence during a placement (see Box 11). In other cases it is realised by allowing, over and above an assessment of occupational competence, an 'add-on' more academic assessment and/or qualification. For example, in Switzerland, apprentices may opt, alongside or after their regular apprenticeship, to pursue a more general educational programme leading to the 'federal vocational baccalaureate' granting entry to universities of applied science (Swiss State Secretariat for Education, 2020⁷⁶). This separation allows not only for different content, but also for different approaches to assessment – so in Switzerland the technical assessment is very much employer-led, and based on performance in a working context, whereas the more academic qualification involves a national examination. This separation is one way of addressing the competing demands, and competing purposes, of these different types of assessment and associated certifications.

BOX 16. IN SWEDEN AND DENMARK, PATHWAYS TO HIGHER EDUCATION ARE KEY TO THE ATTRACTIVENESS OF TECHNICAL PROGRAMMES

In Sweden up to 2011, all students in upper secondary technical programmes pursued the general education courses that provided access to higher education. A 2011 reform meant that some of the more demanding general education courses, including in mathematics, became optional, although still required for entry to higher education. Following the reform, the proportion

of students following technical programmes dropped sharply, down from 53% in 2005-10 to 33% in 2016/17, with a particularly sharp fall in the number of strong performers. Survey evidence suggests that many Swedes now see technical programmes as an option for low performers, and not a good preparation for higher education. A recent OECD review has proposed further reform so that, by default, technical programmes would include the courses required for higher education, while allowing an opt-out (Kuczera and Jeon, 2019⁷⁷).

Denmark's point of departure was rather different in that it has historically had a form of dual apprenticeship in upper secondary education, but has recently found it difficult to sustain the interest of young people in the apprenticeship track. In response, Denmark has created a hybrid qualification (EUX) providing young people both with an apprenticeship and access to higher education. These programmes are academically demanding, and attract only a few per cent of those in the technical track. But they have attracted some strong performers who would otherwise not have considered an apprenticeship (Kuczera and Jeon, 2019⁷⁷).

'Employability' competences raise additional assessment challenges

Technical education and training programmes, as well as including general education, may also directly or indirectly develop other generic competences. These may include some of the meta-skills, discussed earlier, which allow individuals to address occupational competence in the round, but they also include what are commonly called employability skills. These include traits such as self-discipline, honesty and determination and interpersonal skills including teamwork. There is evidence that these skills have significant labour market returns and are often best developed through workbased learning rather than in classrooms (Lerman, 2013⁷⁸). Some of these skills, notably interpersonal skills, have been shown to be of increasing relative importance in England, most plausibly because they correspond to the elements of occupations that are least subject to automation (Adecco, 2017⁷⁹). In the United States, much attention has been given to these competences, and different tools developed to assess them (see Box 17).

BOX 17. THE UNITED STATES IS DISTINCTIVE IN MAKING USE OF FREE-STANDING ASSESSMENTS OF EMPLOYABILITY SKILLS

US high schools do not have the kind of systematic technical tracks found in many European countries, but, in a decentralised system which varies extensively from state to state and school to school, technical courses are sometimes offered within a broadly comprehensive approach to upper secondary education. Much emphasis is placed on demonstrating that those graduating from high school are 'career and college-ready'. The 'career-readiness' element of this is subject to diverse assessment tools. In a survey across US states, assessments of career readiness were classified into tests of academic, employability and technical skills. Technical skills were often assessed through industry recognised certifications of different types. But the US is distinctive in making extensive use of tests of employability independently of

specific technical domains; these include 'Work Keys' (used in 32 states), and ASVAB (developed by the Department of Defence) also used in 32 states. Very often students themselves, alongside school districts, have to share the costs of these assessments (Centre on Educational Policy, 2013⁸⁰).

INCENTIVES FOR LEARNING

Assessment can encourage both the mix and intensity of student learning

Expectations about assessments have both a qualitative and quantitative impact on student learning. Qualitatively, the salience of certain competences in an assessment signals the relative importance of these competences in the programme. One risk of technical assessments which privilege academic examinations is that they may remove the incentive to acquire those more practical competences and meta-skills which are hard to represent in a knowledge-based examination (see the example of medical education in France in Box 4).

Grading may also provide incentives to students

While the desire to pass an assessment is potentially motivating, this effect may be muted when pass rates are high, so that students of average competence have little fear of failure (see the example of English apprenticeship in Box 15). Graded assessments can in principle encourage students to work hard to realise a higher grade, although empirical doubts have been raised about whether this is always realised in general education (Grant and Green, undated⁸¹). Internationally, many countries allow for graded technical assessments, but it is not universal; in Estonia, occupational qualifications do not record a grade (CEDEFOP, 2015³¹). In the Netherlands, although assessment projects are individualised, their grading is standardised. The workplace and school trainers use a grading grid in which each of the learning outcomes included in the assessment project are graded as inadequate, adequate, or good. An overall grade is reached through dialogue between the assessors, rather than through a simple averaging of the marks awarded (CEDEFOP, 2015³¹).

In England, grading of technical assessments is now widely encouraged

In England, (Wolf, 2011⁸²) argued for grading of technical assessments (with marks above pass) as a means of supporting progress into further learning, while (Richard, 2012⁵⁷) proposed grading of apprenticeships as a way of improving their signalling value to employers. Grading has now become government policy, with the aim of both motivating students to pursue excellence and allowing them to signal excellence to employers and education institutions. In 2017, 14% of the nearly 16000 technical and technical qualifications in England offered grades above the passing grade (Newton, 2018¹). Apprenticeship standards must have at least one grade beyond pass, (distinction), and sometimes two (merit and distinction). T-levels will be graded as pass, merit, distinction, or a starred distinction (DfE, 2020⁵). BTEC qualifications have a range of levels – pass, merit, distinction (Pearson, 2020⁸³).

For grading to be of value it needs to apply consistent standards

For grading to have value as a signal to employers or education institutions, or as an incentive to students, it must apply consistent standards, but this is challenging. Thus:

- The pass/fail boundary refers to a commonly understood notion of the ability to do the job – occupational competence. ‘Merit’ and ‘distinction’ have no such common currency of meaning, increasing the risk that any given set of criteria for identifying ‘merit’ or ‘distinction’ will be applied inconsistently, both within and across qualifications (Stanley, 2017⁴⁰).
- Research has shown that methods used to grade technical qualifications are diverse and do not derive from any underlying principles, and a study of grading criteria across 18 technical qualifications demonstrated widely varying approaches (Newton, 2018⁴). In Australia, a sequence of attempts to define good practice in grading has not been successful (Newton, 2018¹). In apprenticeship EPAs, proportions of merits and distinctions varied substantially across standards and between assessment bodies; one assessment body awarded all 116 apprentices on one standard a distinction (IfATE, 2020¹⁶).

Why the objectives of assessment matter: implications for England

This section has identified three main purposes of technical assessment: to demonstrate occupational competence to employers, to establish that the candidate has the foundation skills for lifelong learning and career development, and to guide and motivate students by setting clear targets in the assessment process. For any individual technical assessment, while all three objectives are likely to be present, their balance will vary according to the function of the assessment and associated qualification. This has three implications for England:

- *4.1 Mechanisms for employer engagement in assessments need to be further developed.* Technical assessments and qualifications need to be credible to employers. This privileges forms of authentic assessment which are convincing to employers, and in which employers can more naturally engage. Employer involvement in the assessment, for good reasons, may also render the passing threshold more demanding. This also suggests that mechanisms to involve employers in technical assessments in England need further development.
- *4.2 The high pass rates for EPAs deserve review.* In England only 1.5% of starting apprentices, and 2.5% of EPA candidates, fail the EPA. This is a high pass rate relative to other countries, and raises questions about whether the EPA is sufficiently demanding to adequately support the credibility of the apprenticeship qualification. Conversely, if it is an appropriate pass rate, the resources devoted to EPAs and their quality assurance may be excessive, given the marginal impact of EPA quality on who graduates as an apprentice, relative to other factors such as dropout.

- *4.3 Useful and consistent grading standards (for grades above pass) have proved elusive, suggesting that grading should not be used except where it is essential.* Assessment should provide incentives for students, both in respect of the depth and mix of learning. To this end, grading, offering grades above pass, is argued to encourage the search for excellence. However, in practice it has proved difficult to establish grading criteria that are clear and meaningful, and grading practice that is consistent. Inconsistent grading practice will inevitably be demotivating. Grading may sometimes be needed for other reasons – for example, in T-levels to yield UCAS points for entry into higher education, but here too consistency will be essential. Detailed disaggregated data on grading should be published as a means of scrutinising the consistency of grading practice, given anecdotal evidence of its variability.

5 HOW CAN WE BE SURE OF OCCUPATIONAL COMPETENCE IN CERTIFICATIONS?

This section argues that the overall objective of quality assurance in assessment should be assurance that certification guarantees occupational competence. In delivering that guarantee, quality assurance of assessment should be seen as complementary to the quality assurance of learning. This section also examines the widely encouraged principle that assessment should be independent of the individuals and bodies responsible for technical training.

QUALITY ASSURANCE OF ASSESSMENTS AND CERTIFICATION

Quality assurance in technical assessment in England (and elsewhere) often focuses on internal coherence

As discussed in section 2, there is little or no evidence-based guidance on assessment practice. As a result quality assurance is often limited to ensuring the internal coherence of the assessment process, such as failure to follow one's own rules. In England, IfATE identify assessment weaknesses in terms of failure to follow assessment plans, weak internal management, and lack of recognition of the needs of the apprentice (IfATE, 2020¹⁶). Similarly, Ofqual's technical evaluation of assessment materials used by assessment bodies identified eight categories of weakness in the materials, including weak guidance for assessors, inconsistencies, and errors (Ofqual, 2020⁶⁰). While such measures help to eliminate incoherencies in assessment, they leave open the possibility of other more fundamental challenges to reliability and validity. A good performance on the selected assessment tasks might turn out to bear little relationship to subsequent occupational performance – a problem of lack of validity. This is no criticism of these quality assurance bodies because it reflects the best that can be done with little data or evidence. But it underlines the fallibility of assessment despite the best endeavours of quality assurance.

Quality assurance of certification is best conceived broadly

Happily, the quality of certification, and its capacity to confirm occupational competence usually depends more than on just assessment. As discussed in section 2, certification very often requires candidates not only to pass an assessment, but also to have undertaken a training programme or have specific types of employment experience, or have met other conditions. Consequently, quality assurance of certification is best seen in the round, recognising the complementary contributions of regulations covering assessment, and those covering the learning process. Moreover the fallibility of assessment, as noted above, powerfully reinforces the argument that certification should be underpinned by quality assurance of the learning pathway, as well as in relation to the final assessment. In the case of apprenticeship, the requirements of the apprenticeship programme on the apprentice, on the training employer, on the off-the-job training provider, and on the assessment itself offer collective assurance that a qualifying apprentice is occupationally competent. In England, this broader view of quality assurance of certification can sometimes be obscured by the separate responsibilities of bodies such as Ofsted, ESFA, Ofqual and IfATE.

This broader view allows the balance between the regulation of training and that of assessment to be addressed

This more strategic perspective makes it easier to see where best to invest limited quality assurance and improvement resources – such as when deciding whether occupational competence among graduates is best ensured by investing energy in the final assessment, or by strengthening training arrangements. Quality assurance arrangements for assessment and learning may also be organised so as to be complementary. For example, some assessments only cover a subset of all the competences necessary for occupational competence. This can make sense when there is some other form of reassurance that the full range of competences required for the occupation has been acquired through training or employment experience. In illustration of this point, the apprenticeship assessment arrangements in the Netherlands rely heavily on a set of assessment projects, which inevitably will only assess certain of the competences required in an occupation (see Box 9). The assumption is that other elements of regulation will supplement the end-point assessment to ensure occupational competence. These other elements in the Netherlands include accreditation of the employers offering placements, linked to requirements that the workplace mentors/supervisors of apprentices be themselves trained - see (Hoftijzer, Stronkowski and Rozenbaum, 2018⁸⁴).

England has a relatively complex system of quality assurance of assessment

England has quite an elaborate system of quality assurance covering EPAs for apprenticeship, involving four separate agencies – IfATE, ESFA, Ofqual, and OfS. IfATE is primarily responsible for standards, and agrees the assessment plan when the standard is established. ESFA accredits assessment bodies, and Ofqual is now gradually taking on responsibility for regulating the quality of the EPAs for these standards (other than EPAs for degree standards, which are regulated by the OfS) (Ofqual, 2018⁸²) (Ofqual, 2020⁸⁵). This contrasts with the arrangements in many countries, where lighter touch regulation is in place: a CEDEFOP survey reports that in several countries assessors often validate their own assessments and that this is no more than a formality (CEDEFOP, 2015³¹).

This raises the question of whether, in apprenticeship in England, the emphasis on assessment, as opposed to the quality of learning is right

At the same time England has, by international standards, a very relaxed system of quality assurance of apprenticeship learning and some of the basic regulations on the system are not enforced. Recent surveys suggest that either around 70% (DfE, 2020⁸⁶) or 38% (IfATE, 2020⁸⁷) of apprentices receive less than the 20% off-the-job training required by law, so basic training requirements are being widely ignored. The implication is that devoting more attention to this large gap in enforcement would yield much greater returns, in terms of the credibility of certification, than any plausible enhancement of the end-point assessment, where the outcome would presumably be marginal changes in the current 97.5% pass rate, for those who undertake an EPA.

INDEPENDENCE

Teachers and trainers have close knowledge of their students

The requirement of independence in assessment is a form of quality assurance. As in the domain of general education, the responsible teachers and trainers will have most direct knowledge of the performance of trainees, and sometimes play a big role in assessment. In Spain, teachers responsible for the modules of technical

programmes also take part in assessment board discussions of grading individual students, and deciding whether students can continue to the second year of the programme (CEDEFOP, 2015³¹).

But they may also have biases in an assessment

Teachers and trainers also have biases – perhaps positive because the performance of the trainee can reflect on their own performance as trainers, or perhaps negative because of conflicts with the trainee. In the context of academic assessments there is evidence of different types of bias (Lee and Newton, 2021⁸⁸). This means that while it can be valuable to take trainer views into account in assessments, assessments also need a degree of independence from the training process. Such independence can help to avoid reliability, such as when the same external assessor helps to ensure consistent grading of students in different local contexts.

Independence may be realised by various means

Sometimes independence is realised through nationally or regionally organised examinations, including practical assessments, which are independent of local teachers and trainers. But, as argued earlier, it is often desirable to assess occupational competence in real-world contexts, where local actors are necessarily involved. Some degree of assessment independence can then be realised by mixing an external independent assessment with an internal element. Some examples follow:

- In Australia, the training organisations who train apprentices usually also provide assessments, creating a risk of bias. In the Australian state of Victoria, an independent assessment body has been established to provide end-point assessments which are independent of providers (Australian Skills Quality Authority, 2021⁸⁹).
- In Austria, the final assessment for apprentices includes both a practical and a theoretical component. The practical part of the examination is organised by the regional apprenticeship offices and managed by a board of examiners including a chairperson appointed by the regional advisory board on apprenticeship, one representative of employers and one of employees. The chairperson must be an authorised apprenticeship trainer and at least one other member of the board must be a professional expert (European Alliance for Apprenticeships, 2021⁹⁰).
- In Hungary, responsibility for assessment is shared between the technical school which organises the assessment and an independent examination committee (CEDEFOP, 2015³¹).
- In Korea, technical qualifications are typically awarded after an internal assessment undertaken by a training institute, and an external assessment undertaken by the awarding body (Human Resource Development Korea), with the award depending on an adequate score in both assessments (Coles and Bateman, 2017⁹¹).

Often it is not independence that is sought, but rather a balance of perspectives

The terminology of independence can sometimes be misleading. All assessors and assessment bodies will have different perspectives and incentives. As discussed earlier, in some countries employers are systematically represented, alongside other stakeholders, in the assessment. The function of the mix of individuals found in German apprenticeship assessment is not to provide independence, but rather

to ensure that there is a balanced representation of different perspectives in the assessment (see Box 14). Given the limitations of quality assurance, it is worth noting the existence of an international standard on assessments as part of technical certification (see Box 18).

BOX 18. IS THERE SCOPE TO MAKE MORE USE OF AN INTERNATIONAL STANDARD FOR ASSESSMENTS AND CERTIFICATION? ISO/IEC 17024:2012

This ISO standard has been established by a working body through the International Standards Organisation. It aims to set out clear standards governing the integrity, impartiality and credibility of examination systems used for professional certification. While it is primarily concerned with certification rather than assessment, it naturally covers a lot of the ground relative to assessment. It addresses the definition of competences, the consistency and transparency of examination criteria, the knowledge and skills of examiners, the impartiality of the examiners and avoidance of conflicts of interest. It also looks at how to ensure that the examination is a valid test of competence. (UNIDO, 2018⁹²) (National Accreditation Board for Education and Training, 2021⁹³).

In the United States, ANSI (The American National Standards Institute) provides accreditation for many organisations delivering professional certifications following an assessment, and, as part of this process, confirms that these organisations meet ISO/IEC 17024 (Kuczera and Field, 2013⁹⁴).

How can we be sure of occupational competence? Implications for England

This section has looked at the quality assurance of assessment, and the need to place it in the context of certification for occupational competence. It has also considered how different countries are managing the requirement for independence in assessments. This discussion has three implications for England:

- *5.1 Quality assurance of technical assessment is best considered, alongside other conditions of certification such as approved learning programmes, in ensuring that a certification provides a robust demonstration of occupational competence.* This approach rests on the fact that certification depends not only on assessments, but also on other criteria, such as the candidate having followed a quality-assured learning route. In England, this implies a need for close co-ordination, at least, of the different agencies responsible for different elements of technical programmes and assessments, so that the quality assurance of certification – the key objective – can be secured in the round, recognising the complementary contributions of different elements of regulation and quality assurance to this end.
- *5.2 In apprenticeship, the balance between a resource-intensive EPA and weaker programme requirements might be reviewed.* Currently, many apprentices do not receive the training to which they are entitled, which is not only unfair to these apprentices and a potential misuse of levy funding by training provider; it also raises questions about their occupational competence on graduation. Resolving this problem would achieve far more in terms of assurance of occupational competence among qualified apprentices than any conceivable improvements in EPAs, which currently pass nearly all candidates.

- 5.3 *The desirability of independence in assessment must be qualified by the arguments for involving key stakeholders.* Independence, in the sense of an element in the assessment process which is independent of local training providers, is an important means of removing local biases and increasing consistency in standards. However independence is not absolute. There are good reasons for allowing employers collectively to influence the shape of assessments, as they are best placed to judge the competences necessary for the job, and the level of competence that might reasonably be expected in a newly qualified person. In England, continued efforts are needed to involve employers directly in assessments in the way envisaged in the Richard review.

ANNEX A. ASSESSMENT IN SWISS APPRENTICESHIP

Introduction

In Switzerland around two-thirds of young people pursue apprenticeship at upper secondary level (one of the highest rates of participation in the world), following early school tracking at around the age of 11 into either a general or pre-vocational school tracks. (https://eacea.ec.europa.eu/national-policies/eurydice/content/organisation-vocational-upper-secondary-education-79_en. <https://www.fr.ch/sites/default/files/2020-12/2020-11-16%20Statistiques%202020-2021-corr.pdf>).

The apprenticeship programme involves around 240 different occupations, and is undertaken typically in companies with around one or one and a half days each week in a vocational school, over 3-4 years. The programme leads to the Federal VET Diploma *Berufslehre* (apprenticeship), in one of the occupational fields. The federal government and its agencies work closely with industry associations in the development of “ordinances”, “*Bildungspläne*” (training plans), which define the required competences in the different occupations. They also work closely in the training of trainers and examiners in the different occupational fields.

Form of assessment

During the apprenticeship, assessment is carried out during the theoretical (classroom-based) and practical (work-based learning) training and is individualised according to objectives (set out in a national regulation relative to the target occupation) and agreed between the learner and vocational trainer:

During the apprenticeship, a company vocational trainer records the candidate’s technical, methodological, social, and personal skills in a monthly learning report (“*Bildungsbericht*”) (<https://swisseducation.educa.ch/en/vet-learning-objectives-curriculum-and-assessment>) throughout the course of the programmes. These notes on practical experience are then taken into account in the final assessment.

Assessment involves both the general education element which has been pursued in the vocational school, and an apprenticeship examination specific to the profession.

For apprentices, the qualification procedure for general education comprises a performance grade, an in-depth assignment and (in the 3- and 4-year apprenticeships) a final examination.

The responsible cantonal authority decides on the form of the final apprenticeship examination, which depends on the profession. (<https://www.sbf.admin.ch/sbf/en/home/services/publications/data-base-publications/vocational-and-professional-education-and-training-in-switzerland.html>, page 10). It may include:

1. Practical work, where the individual must demonstrate the ability to carry out defined professional activities successfully. This is done through an individual practical assignment linked to an obligatory professional interview, or as a predefined practical assignment (standardised) with an optional interview.

2. Professional knowledge assessed via an exam. This can either take the form of a written exam, or both a written and oral examination (dependent on the type of practical assessment and whether or not the candidate has had to do a professional interview). The cantons ensure that this examination (according to profession) takes place at the same time everywhere. The development of Swiss-wide examination papers also requires harmonisation of examination dates, so the results of the examinations are not distorted by exchange of information between trainees.
3. Optional: An interim exam, undertaken part-way through the programme.
4. 'General culture' assessed through a school grade, personal in-depth work, and a final examination <https://www.fedlex.admin.ch/eli/cc/2006/510/fr>: Its weighting in the overall mark for the VET Certificate and Diploma is at least 20%.
5. Language, communication and society, which provides young people with basic skills needed to orient themselves both in their personal lives and within society as well as to overcome personal and professional challenges.

(<https://www.sbf.admin.ch/sbf/fr/home/formation/formation-professionnelle-initiale/projet-procedures-de-qualification.html>, page 10).

In order to pass, candidates must achieve a score of 4 or higher both overall, (reflecting the average of the weighted scores for the different qualification areas) and specifically for the qualification area "practical work".

Those responsible for organising and delivering the assessment

The Swiss federal government regulates the assessments, either directly or through federal agencies – <https://ibw.at/bibliothek/id/413/>.

The awarding bodies and those responsible for carrying out the examination procedures for the VET Certificate and Diploma are the cantons. The trainers in companies (employers/companies) undertake the assessments that take place during training, but not the final examinations. The examiners are trained by SFIVET (a federal agency) and are not employees of particular companies. (<https://www.oecd.org/switzerland/42578681.pdf> page 17) but examination experts with a cantonal mandate.

The final exams are supervised by the cantons and the professional organisations (<https://www.apprenticeship-toolbox.eu/standards-matching/examination-certification/92-examination-certification-in-switzerland>)

The organisation of the final exams involves the cantonal authority (responsible for monitoring the legality of the processes), a head of examination (implementation of exams), a chief expert (responsible for planning and guaranteeing the quality of the final exam) and an examination expert (those responsible for the qualification procedure, undertaking the exam and assessing the exams).

Grading and certification

Assessments are graded. Performance across the 3–4-year VET Diploma is evaluated by marks ranging from 1 to 6.

If the VET Diploma was obtained through the qualification procedure with a final examination, the report card reports: the overall grade, the grades for each qualification area of the final examination and the grade for “experience”.

The final examination may be retaken twice, and usually the sections passed do not need to be retaken https://www.berufsbildung.ch/DYN/bin/21423-23922-1-sdbb_wegweiser_en.pdf

Quality assurance and oversight

The apprenticeship system is managed and quality-assured by three partners: the federal government (responsible for strategic management and development), the cantons (responsible for implementation and supervision) and the professional organisations (responsible for training content and apprenticeships).

Alternative routes to assessment and qualification

For adults, there are several ways to obtain a Federal VET Certificate or a Federal VET Diploma, the first of which is enrolment on a standard dual-track VET programme. However, for adults who have already gained a certain amount of employment experience, they can skip some elements of the dual-track VET programme (effectively, a shortened version). Adults with 5+ years of employment experience are eligible for direct admission to the qualification procedure (final exam) without having to complete training. They can also obtain validation of non-formal and informal learning (VNIL), although this is only available for a limited number of professions.

An example – optician

Optician: <https://www.becc.admin.ch/becc/public/bvz/beruf/show/85504>

For each qualification ‘module’ (practical work, professional knowledge etc...), at least two examination experts assess performance.

The final assessment involves a compilation of grades from five elements.

- A “partial examination” (4 hours) that takes place towards the end of the second year of training. This covers basic professional work and is taken into account in the final qualification procedure (at the end of the 3rd or 4th year). (20% of the final overall score).
- “Practical work” is assessed through prescribed practical work (2 hours) that takes place towards the end of the programme. The trainee must show that they are able to perform the tasks required in a professional manner and according to the needs and the situation. (20% of the final overall score).
- “Professional knowledge” (4 hours) is assessed through a written (or written and oral) examination taken towards the end of the programme. If an oral examination is organised, it can be a maximum duration of 1 hour. (25% of the final overall score).

- “General culture” relates to the school-based part of the apprenticeship, and is assessed through a school grade for personal in-depth work and a final examination (20% of the final overall score).
- There is also a grade for “experience”. The experience score is the average of the grades for: “teaching of professional knowledge” (average grades from biannual bulletins) and the “inter-company courses” (average of the grades from the proficiency tests). (15% of the final overall score).

The candidate is successful if:

- a. The grade awarded for “practical work” is equal to or higher than 4
- b. The grade awarded for “professional knowledge” is greater than or equal to 4
- c. The overall grade (the average of the weighted grades as indicated above) is greater than or equal to 4.

The weighting for each individual area of competence depends on the profession and is quite variable. In the case of a potter, for example, “practical work” is weighted at 50 % of the overall grade; “professional knowledge” is 20 %, “general culture” is 20 % and the grade for “experience” is 10 % <https://www.becc.admin.ch/becc/public/bvz/beruf/show/39506>

ANNEX B. A BTEC CERTIFICATE FOR ART AND DESIGN AND A DANISH CERTIFICATE FOR A HEALTH CARE COOK

6> BTEC

Pearson
3> BTEC Level 1/Level 2 <4
First Award


2> in ART AND DESIGN
 WITH GRADE DISTINCTION AT LEVEL 2
 is awarded to
1> MINNIE MOUSE
 who has completed an approved programme at
ERNESFORD GRANGE COMMUNITY ACADEMY

5> AWARDED: JULY 2016

THIS DOCUMENT CONSISTS OF MORE THAN ONE PAGE
 20442 : BHX40 :GA43021:000498539:02:02:01:ISSUED 15-OCT-2016 : QN 600/4781/1
7


This BTEC Certificate meets the minimum requirements as it shows:

1. Full name of Apprentice
2. Qualification name
3. Qualification type
4. Qualification level
5. Date of achievement
6. Awarding Organisation identifier: BTEC/Edexcel/EDI/Pearson acceptable
7. Date of Birth

Regulated by

 Sharon Hague
 Responsible Officer
 Pearson Education Ltd


 Recognised awarding body


 Accreditation



ERHVERVSUDDANNELSE
Erhvervsfaglig fællesindgang MAD TIL MENNESKER
VOCATIONAL EDUCATION AND TRAINING Vocational common base FOOD FOR HUMANS

BEVIS

DIPLOMA

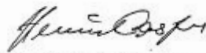
for gennemført uddannelse til
for completion of the

ERNÆRINGSASSISTENT

HEALTH CARE COOK education programme

Navn Name	[REDACTED]
CPR-nr. Civil reg. No.	[REDACTED]
Skoleophold på Theoretical programme carried out at	Uddannelsescenter Holstebro
Praktikophold på Practical programme carried out at	Mercantec
Afsluttet den and completed	31. januar 2013
Svendeprøven er aflagt på The final examination is taken at	Uddannelsescenter Holstebro
Dansk kvalifikationsramme for livslang læring: The European Qualification Framework (EQF)	Niveau 4 Level 4
Der er opnået karakteren Mark achieved	7/C i den selvvalgte praktiske opgave for the practical examination of own choice 12/A i den bundne praktiske opgave for the set practical examination 10/B i den mundtlige prøve for the oral examination
Gennemsnittet af den selvvalgte praktiske og den bundne praktiske opgave Average mark for the practical examination of own choice and the set practical examination	9,5

FAGLIGT UDVALG FOR ERNÆRINGSASSISTENTUDDANNELSEN
THE PROFESSIONAL COMMITTEE FOR THE HEALTH CARE COOK EDUCATION PROGRAMME



udvalgsformand
Chair of the Committee

REFERENCES

- 79 Adecco (2017), *AUTONATION. analysing the potential risks and opportunities automation could bring to Britain's labour market*. https://www.economicmodelling.co.uk/wp-content/uploads/2019/11/Autonation-Report-Final_WEB.pdf
- 39 Ahn, S. and S. Nyström (2020), "Simulation-based training in VET through the lens of a sociomaterial perspective", *Nordic Journal of Vocational Education and Training*, Vol. 10/1, pp. 1-17, <https://www.diva-portal.org/smash/get/diva2:1463505/FULLTEXT01.pdf>.
- 54 Allais, S., P. Molebatsi and C. Marock (2014), *The Development of Occupational Standards in English-speaking Countries*, https://www.researchgate.net/publication/274316003_The_Development_of_Occupational_Standards_in_English-speaking_countries.
- 34 Alphaplus Consultancy Ltd. (2014), *Validation of Vocational Qualifications: Final Report*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/606033/2014-02-11-validation-of-vocational-qualifications-final-report.pdf.
- 2 Assessment in Education: Principles, Policy & Practice (2021), *Assessment in Education: Principles, Policy & Practice*, <https://www.tandfonline.com/toc/caie20/current>.
- 89 Australian Skills Quality Authority (2021), *Clauses 1.8 to 1.12—Conduct effective assessment*, <https://www.asqa.gov.au/standards/training-assessment/clauses-1.8-to-1.12#:~:text=The%20RTO%20informs%20the%20learner,and%20be%20reassessed%20if%20necessary.&text=Assessment%20is%20flexible%20to%20the,reflecting%20the%20learner's%20needs>.
- 9 Baethge, M. (2006), *PISA-VET A Feasibility Study*, Franz Steiner Verlag, https://books.google.cm/books?id=nUzafEq8tf4C&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.
- 52 BIBB (2021), *Informationen zu Aus- und Fortbildungsberufen: Klempner*, https://www.bibb.de/dienst/berufesuche/de/index_berufesuche.php/profile/apprenticeship/110512.
- 44 Bjaelde, O., K. Lauridsen and A. Lindberg (2018), *Current Trends in Assessment in Europe: the Way Forward*, <https://www.coimbra-group.eu/wp-content/uploads/WP-Trends-in-assessment-FINAL.pdf>.
- 46 Black, P. and D. Wiliam (1998), "Assessment and Classroom Learning", *Assessment in Education: Principles, Policy and Practice*, Vol. 5/1, pp. 7-7, <https://www.gla.ac.uk/t4/learningandteaching/files/PGCTHE/BlackandWiliam1998.pdf>.
- 71 BMBF (2019), *Report on Vocational Education and Training 2019*, https://www.bmbf.de/upload_filestore/pub/Berufsbildungsbericht_2019_englisch.pdf.
- 72 BMBF (2003), *Germany's Vocational Education at a Glance*, https://blogs.sch.gr/atrisvas/files/2016/06/germanys_vocational_education_at_a_glance.pdf.
- 45 Bradley, S. (2016), *Managing plagiarism in programming assignments with blended assessment and randomisation*, <https://dl.acm.org/doi/10.1145/2999541.2999560>.

- 36 Canadian Council of Directors of Apprenticeship (2016), *2016 Annual Review*, http://www.red-seal.ca/docms/2016ar_eng.pdf.
- 4 CEDEFOP (2021), *VET in Europe country reports*, <https://www.cedefop.europa.eu/en/events-and-projects/networks/refernet/vet-in-europe-country-reports>.
- 24 CEDEFOP (2018), *Vocational education and training in Europe: France 2018*, <https://www.cedefop.europa.eu/de/publications-and-resources/country-reports/vocational-education-and-training-europe-france-2018>.
- 23 CEDEFOP (2016), *Vocational education and training in the Netherlands*, <https://www.cedefop.europa.eu/en/publications-and-resources/publications/4142>.
- 31 CEDEFOP (2015), *Ensuring the quality of certification in vocational education and training*, <https://op.europa.eu/en/publication-detail/-/publication/b0543d49-c977-11e5-a4b5-01aa75ed71a1/language-en>.
- 80 Centre on Educational Policy (2013), *Career Readiness Assessments across States: A Summary of Survey Findings*, <https://files.eric.ed.gov/fulltext/ED554578.pdf>.
- 26 Chaparro, T. (2012), *Recognition of prior learning and the role of quality assurance agencies. Accreditation of prior learning in France (VAE) as a case study.*, https://www.cti-commission.fr/wp-content/uploads/2012/03/CTI_ENQA_Mars2012.pdf.
- 38 Chłoń-Domińczak, A. (2019), *Vocational Education and Training in Europe: Poland*, https://cumulus.cedefop.europa.eu/files/vetelib/2019/Vocational_Education_Training_Europe_Poland_2018_Cedefop_ReferNet.pdf.
- 65 Clayton, B. (1995), *Focusing on Assessment: Strategies for Off-the-Job Teachers and Trainers*, https://www.ncver.edu.au/__data/assets/file/0015/16530/td_tnc_42_24.pdf.
- 25 Colardyn, D. and J. Bjornavold (2004), *Validation of Formal, Non-Formal and Informal Learning: policy and practices in EU Member States*, https://www.competences.info/ibak/root/img/pool/docs/open/bjornavold_colardyn_example_en.pdf.
- 91 Coles, M. and A. Bateman (2017), *Towards Quality Assurance of Technical Vocational Education and Training*, https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000259282&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_3ba90c4c-1cb8-407d-a917-81cb8106ce4b%3F_%3D259282eng.pdf&locale=en&multi=true&ark=/ark:/48223/p.
- 61 Constantinou, F. (2020), "What is synoptic assessment? Defining and operationalising an as yet non-mainstream assessment concept", *Assessment in Education: Principles, Policy & Practice*, pp. 27:6, 670-686., <http://dx.doi.org/10.1080/0969594X.2020.1841734>.
- 14 Coussin, S. (2017), *Documentation 2e cycle Médecine Epreuves Classantes Nationales (ECN) FIN DES ECN : ENTRÉE EN VIGUEUR EN 2019*, <http://www.remede.org/documents/fin-des-ecn-entree-en-vigueur-en-2019.html>.
- 43 Danish Ministry of Education (2008), *The Danish Vocational Educational System*, https://www.apprenticeship-toolbox.eu/files/144/Competent-Bodies/133/The_Danish_VET_System.pdf.

- 12 Darr, C. (2005), "A hitchhiker's guide to reliability", <https://assessment.tki.org.nz/Using-evidence-for-learning/Working-with-data/Concepts/Reliability-and-validity#:~:text=The%20reliability%20of%20an%20assessment,it%20was%20designed%20to%20measure.>
- 13 Darr, C. (2005), "A hitchhiker's guide to validity", <https://assessment.tki.org.nz/Using-evidence-for-learning/Working-with-data/Concepts/Reliability-and-validity#:~:text=The%20reliability%20of%20an%20assessment,it%20was%20designed%20to%20measure.>
- 70 DfE (2021), *Apprenticeships and Traineeships: Academic Year 2020/21*, <https://explore-education-statistics.service.gov.uk/find-statistics/apprenticeships-and-traineeships/2020-21>.
- 49 DfE (2021), *T Level industry placements: delivery guidance (updated 5 July 2021)*, <https://www.gov.uk/government/publications/t-level-industry-placements-delivery-guidance/t-level-industry-placements-delivery-guidance>.
- 86 DfE (2020), *Apprenticeships Evaluation 2018-19 - Learners*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/875553/Apprenticeships_Evaluation_-_Learner_Report.pdf.
- 5 DfE (2020), *Guidance. Introduction of T levels. Updated 4 September 2020*, <https://www.gov.uk/government/publications/introduction-of-t-levels/introduction-of-t-levels#grading-and-certification>.
- 51 DfE (2020), *Introduction of T Levels: Updated September 2020*, <https://www.gov.uk/government/publications/introduction-of-t-levels/introduction-of-t-levels#industry-placements>.
- 58 DfE and BIS (2013), *The Future of Apprenticeships: Next Steps from the Richard Review*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/190632/bis-13-577-the-future-of-apprenticeships-in-england-next-steps-from-the-richard-review.pdf.
- 37 Electrical Workers Registration Board (2021), *Tuition Courses and Practical Assessments*, <https://www.ewrb.govt.nz/becoming-an-electrical-worker/tuition-courses-and-practical-assessments/>.
- 90 European Alliance for Apprenticeships (2021), *Examination & Certification in Austria*, <https://www.apprenticeship-toolbox.eu/standards-matching/examination-certification/88-examination-certification-in-austria>.
- 95 European Alliance for Apprenticeships (2021), *Examination & Certification in Denmark*, <https://www.apprenticeship-toolbox.eu/standards-matching/examination-certification/89-examination-certification-in-denmark>.
- 6 European Alliance for Apprenticeships (2021), *Examination & Certification in Germany*, <https://www.apprenticeship-toolbox.eu/standards-matching/examination-certification/90-examination-certification-in-germany>.
- 53 European Alliance for Apprenticeships (2021), *Examination and Certification in Luxembourg*, <https://www.apprenticeship-toolbox.eu/standards-matching/examination-certification/91-examination-certification-in-luxembourg>.
- 7 Eurydice (2020), *Assessment in Vocational Upper Secondary Education*, https://eacea.ec.europa.eu/national-policies/eurydice/content/assessment-vocational-upper-secondary-education-24_en.

- 66 Federal Institute for Vocational Education and Training (2017), *Quality assurance of company based training in the dual system: An overview for practitioners and VET experts*, http://www.nuv.cz/uploads/EQAVET/soubory/BIBB_Quality_Assurance_2017.pdf.
- 27 Field, S. (2020), *Beyond the Missing Middle: Developing Higher Technical Education. A report to the Gatsby Foundation*, <https://www.gatsby.org.uk/uploads/education/beyond-the-missing-middle-pvw.pdf>.
- 21 Field, S. (2020), *Strengthening Skills in Scotland. OECD Review of the Apprenticeship System in Scotland.*, http://www.oecd.org/skills/centre-for-skills/Strengthening_Skills_in_Scotland.pdf.
- 22 Field, S. (2018), *Taking Training Seriously: Lessons from an international comparison of off-the-job training for apprenticeships in England*, Gatsby Charitable Foundation, <https://www.gatsby.org.uk/uploads/education/final-apprenticeships-and-off-the-job-training-may-2018.pdf>.
- 28 Field, S. et al. (2012), *A Skills Beyond School Review of Denmark*, <http://dx.doi.org/10.1787/9789264173668-en>.
- 75 Field, S. and A. Guez (2018), *Pathways of Progression: Between Technical and Vocational Education and Training and Post-Secondary Education.*, UNESCO, Paris, <http://unesdoc.unesco.org/images/0026/002659/265943e.pdf>.
- 18 Finnish Ministry of Education and Culture (2017), *Reform of vocational upper secondary education*, <https://minedu.fi/en/reform-of-vocational-upper-secondary-education>.
- 56 Fuller, A. and L. Unwin (2013), *Apprenticeship and the Concept of Occupation*, Gatsby Charitable Foundation, <https://www.gatsby.org.uk/uploads/education/reports/pdf/apprenticesfullpvs.pdf>.
- 16 Government of Western Australia (2016), *Assessment in the VET sector*, <https://www.dtwd.wa.gov.au/sites/default/files/uploads/Assessment%20in%20the%20VET%20Sector%20-%202016%20-%20Final.pdf>.
- 81 Grant, D. and W. Green (undated), "Grades as Incentives", <https://www.shsu.edu/~dpg006/gincentives2.pdf>.
- 84 Hoftijzer, M., P. Stronkowski and J. Rozenbaum (2018), *Getting Out of School and into the Workplace: Strengthening Work-Based Learning in Upper Secondary Technical Education in Poland's Świętokrzyskie Region.*, <http://dx.doi.org/10.1596/978-1-4648-1322-1>.
- 59 IfATE (2021), *Apprenticeship Standards*, <https://www.instituteforapprenticeships.org/apprenticeship-standards/?includeApprovedForDelivery=true>.
- 30 IfATE (2021), *End-point assessment plan for Pipe Welder apprenticeship standard*, https://www.instituteforapprenticeships.org/media/3325/st0851_pipe_welder_13_ap_final_for-publication_17072019.pdf.
- 87 IfATE (2020), *Apprenticeship Panel Survey - August 2020*, <https://www.instituteforapprenticeships.org/reviews-and-consultations/reports/apprentice-panel-survey-august-2020/>.

- 35 IfATE (2020), *Developing an End-Point Assessment Plan*, <https://www.instituteforapprenticeships.org/developing-new-apprenticeships/developing-an-end-point-assessment-plan/>.
- 69 IfATE (2020), *End-point assessment plan for Advanced Upholsterer apprenticeship standard*, https://www.instituteforapprenticeships.org/media/3662/st0656_advanced_upholsterer_l3_ap_for-publication_10102019.pdf.
- 63 IfATE (2020), *End-point assessment plan for Knitted product manufacturing technician apprenticeship standard*, https://www.instituteforapprenticeships.org/media/4464/st0910_knitting_technician_l3_ap_for-publication-2700720.pdf.
- 17 IfATE (2020), *External Quality Assurance Annual Report 2020*, <https://www.instituteforapprenticeships.org/media/4724/eqa-annual-report.pdf>.
- 50 Independent Panel on Technical Education (2016), *Report of the Independent Panel on Technical Education*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/536046/Report_of_the_Independent_Panel_on_Technical_Education.pdf.
- 42 International Labour Organisation (2020), *ILO Toolkit for Quality Apprenticeships: Volume 2: Guide for Practitioners*, https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/publication/wcms_751116.pdf.
- 68 Kantar Public (2019), *DfE Learners and Apprentices Study: Reasons for non-Completion*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/912618/DfE_Learners_and_apprentices-Reasons_for_non-completion.pdf.
- 19 Karttunen, A. (2018), *The big VET reform in Finland*, <https://nvl.org/content/the-big-vet-reform-in-finland>.
- 94 Kuczera, M. and S. Field (2013), *Skills beyond School Review of the United States; OECD Reviews of Vocational Education and Training*, <http://dx.doi.org/10.1787/9789264202153-en>.
- 77 Kuczera, M. and S. Jeon (2019), *Vocational Education and Training in Sweden, OECD Reviews of Vocational Education and Training*, <http://dx.doi.org/10.1787/g2g9fac5-en>.
- 88 Lee, M. and P. Newton (2021), *Systematic divergence between teacher and test-based assessment: literature review*, <https://www.gov.uk/government/publications/systematic-divergence-between-teacher-and-test-based-assessment/systematic-divergence-between-teacher-and-test-based-assessment-literature-review#contents>.
- 78 Lerman, R. (2013), "Are employability skills learned in U.S. youth education and training programs?", *IZA J Labor Policy*, Vol. 2/6, <https://doi.org/10.1186/2193-9004-2-6>.
- 55 Lester, S. and J. Religa (2017), "'Competence' and occupational standards: observations from six European countries", *Education + Training*, 59 (2), pp. 201-214.

- 29 Ministère de l'Éducation Nationale et de la Jeunesse (2018), *Note d'Information 18, 30*, https://cache.media.education.gouv.fr/file/2018/94/5/depp-ni-2018-18-30-Dispositif-academique-validation-acquis-12_700-diplomes-delivres-en-2017_1038945.pdf.
- 15 Ministère de Solidarité et de Santé (2018), *Pour Les Étudiants Et Leurs Futurs Patients, Des Études Médicales Rénovées*, https://solidarites-sante.gouv.fr/IMG/pdf/180705_-_dp_-_etudes_medicales_renovees.pdf.
- 32 Misko, J. (2014), *Quality assessments: practice and perspectives*, <https://www.ncver.edu.au/research-and-statistics/publications/all-publications/quality-assessments-practice-and-perspectives>.
- 93 National Accreditation Board for Education and Training (2021), *Accreditation Services*, <https://nabet.qci.org.in/accreditation-services/iso-17024>.
- 33 National Electrotechnical Training (2021), *AM2*, <https://www.netservices.org.uk/am2/>.
- 11 New Zealand Ministry of Education (2021), *Reliability and Validity*, <https://assessment.tki.org.nz/Using-evidence-for-learning/Working-with-data/Concepts/Reliability-and-validity#:~:text=The%20reliability%20of%20an%20assessment,it%20was%20designed%20to%20measure>.
- 1 Newton, P. (2018), *Grading Vocational & Technical Qualifications: Recent Policies and Current Practices*, Ofqual, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755681/Grading_vocational_and_technical_qualifications_-_FINAL.pdf.
- 73 Norwegian Centre for International Cooperation in Education (2016), *Vocational education and training in Europe: Norway*, https://cumulus.cedefop.europa.eu/files/vetelib/2016/2016_CR_NO.pdf.
- 48 Norwegian Ministry of Education and Research (2006), *Education Act*, https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_5-2#KAPITTEL_5-2.
- 67 OECD (2021), *Minimum relative to average wages of full time workers*, <https://stats.oecd.org/Index.aspx?DataSetCode=MIN2AVE>.
- 74 OECD (2021), *Strictness of Employment Protection: individual dismissals (regular contracts)*, https://stats.oecd.org/Index.aspx?DataSetCode=EPL_R.
- 3 OECD (2021), *Vocational Education and Training and Adult Learning*, <https://www.oecd.org/education/innovation-education/vet.htm>.
- 85 Ofqual (2020), *Apprenticeship end-point assessments*, <https://www.gov.uk/government/collections/apprenticeship-end-point-assessments>.
- 60 Ofqual (2020), *Findings from Ofqual's technical evaluation of apprenticeship end-point assessment materials.*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895234/Technical_Evaluation_findings_-_EPA_Materials_25.06.2020.pdf.

- 83 Pearson (2020), *BTEC Level 2 Technicals*, <https://qualifications.pearson.com/en/qualifications/btec-technical.html>.
- 64 Penelope Peterson, E. (ed.) (2010), *The Multiple Purposes of Assessment*, Elsevier.
- 57 Richard, D. (2012), *Richard Review of Apprenticeships*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/34708/richard-review-full.pdf.
- 62 Skills Development Scotland and Centre for Workbased Learning (2018), *Skills 4.0 A Skills Model to Drive Scotland's Future*, https://www.skillsdevelopmentscotland.co.uk/media/44684/skills-40_a-skills-model.pdf.
- 40 Stanley, G. (2017), *Accreditation and Assessment in VET*, OUP.
- 8 Strategic Development Network (2018), *Under the bonnet of "high-stakes" end-point assessment*, <https://www.strategicdevelopmentnetwork.co.uk/reality-of-end-point-assessment/>.
- 76 Swiss State Secretariat for Education, R. (2020), *Vocational and Professional Education and Training in Switzerland: Facts and Figures*, <https://www.sbf.admin.ch/sbf/en/home/services/publications/data-base-publications/vocational-and-professional-education-and-training-in-switzerland.html>.
- 92 UNIDO (2018), *Guidelines on Conformity Assessment ISO/IEC 17024:2012*, https://www.unido.org/sites/default/files/files/2018-07/Guidelines_on_Conformity_Assessment.pdf.
- 47 University of Exeter (2021), *Improving Formative Assessment in Vocational Education and Literacy, Language and Numeracy*, <http://education.exeter.ac.uk/ifa/>.
- 20 Virolainen, M. (2018), *Changing patterns of transition to VET and from VET to higher education : the on-going Finnish VET reform*, <https://jyx.jyu.fi/bitstream/handle/123456789/67047/1/virolainenchanging%20patterns%20of%20transition.pdf>.
- 82 Wolf, A. (2011), *Review of vocational education: the Wolf report*, <https://www.gov.uk/government/publications/review-of-vocational-education-the-wolf-report>.
- 10 WorldskillsUK (2021), *National Competitions*, <https://www.worldskillsuk.org/skills/national-competitions/>.
- 41 Yu, K. and G. Frempong (2012), "Standardise and individualise – an unsolvable tension in assessment?", *Education as Change*, pp. 143-157, <http://dx.doi.org/10.1080/16823206.2012.692210>.

The Gatsby Charitable Foundation
The Peak, 5 Wilton Road, London SW1V 1AP
T +44 (0)20 7410 0330 www.gatsby.org.uk
Registered Charity number 251988

Copyright © Gatsby Charitable Foundation
September 2021