How 'CASLO' qualifications work

ofqual

Authors

- Paul E. Newton and Charlotte Lockyer, from Ofqual's Standards, Research, and Analysis Directorate.

- Colleagues from Ofqual's Vocational and Technical Qualifications Directorate were responsible for much of the data collection: Sue Bentley, Andrew Gregg, Michael Larkin, Sarah Purdell, Caroline Richards, Zeek Sweiry, Liz Walters, Nic Wroe.

- Project administration by Jackie Domingue.

With thanks to

- We are very grateful to officers from EAL Awards, University of the Arts London, Qualsafe Awards, Chartered Management Institute, Highfield Qualifications, and City & Guilds, all of whom kindly agreed to be interviewed for this study and provided us with very helpful insights into the effective operation of CASLO qualifications.

- Thanks also to members of Ofqual's Standards Advisory Group and to Safety Training Awards for helpful feedback on the penultimate draft of this report.

Contents

Foreword

This report focuses on a family of (mainly) vocational and technical qualifications that is defined very precisely in terms of how its unit standards are specified and assessed. We recognise that acronyms abound within the education sector, and that caution should be exercised before introducing a new one, but we were keen to give these qualifications a clear identity in order to be able to understand and regulate them effectively. With some advice therefore we have decided to label them 'CASLO' qualifications because of the way in which assessors are required to Confirm the Acquisition of Specified Learning Outcomes.

Many of the qualifications that Ofqual regulates are part of this family. Many BTECs are CASLO qualifications – constructed entirely or mainly from CASLO units – making BTEC their largest and most familiar brand. But all sorts of qualifications adopt this approach, from the ASDAN Level 1 Award in Personal and Social Effectiveness, to the SFJ Awards Level 3 Diploma in Policing, to the DAO Level 5 Diploma in Leadership and Management, and beyond.

The CASLO family has links to various movements, including Competence-Based Assessment (CBA), Outcomes-Based Assessment (OBA), Mastery-Based Assessment (MBA), and so on. However, the historical roots of this family – including, most recently, the impact of the Qualifications and Credit Framework – have led to CASLO qualifications in England, today, sharing a very distinctive set of characteristics, despite being very divergent in many ways. As such, they are very hard to categorise.[1] This report attempts to characterise them by introducing the idea of an 'archetypal' CASLO qualification.

Partly because this is an unusual family of qualifications – given its parochial historical roots – we know far less about how CASLO qualifications work than about how more classical qualifications work, including GCSEs and A levels. This is therefore an unusual Ofqual report because it attempts to document, in a purely functional sense, how CASLO qualifications work. This is analogous to setting out how an electric engine works in contrast to an internal combustion engine or a clockwork engine. For this reason, the report adopts a contrastive style, comparing the more widely understood classical approach – exemplified via General

---

[1] In England, these qualifications tend to be vocational, although there is nothing inherently vocational about the approach. Some could be described as CBAs, although many do not actually assess competence in a narrow, occupational sense. Similarly, although they are both outcomes-based and mastery-based, their particular approach to specifying outcomes in tandem with their particular approach to defining mastery means that neither the OBA nor MBA designation quite hits the spot.

Qualifications like GCSEs and A levels – with the less widely understood CASLO approach.

Adopting a functional approach to explaining how CASLO qualifications work means not considering thornier evaluative questions – such as whether certain CASLO qualifications work better or worse than others, or whether CASLO qualifications work better or worse than classical qualifications. It assumes simply that CASLO qualifications have the potential to work and seeks to explain how. This report helps to construct a foundation upon which evaluative work can be based.

Finally, this report does not attempt to explain everything about how CASLO qualifications work. Indeed, our focus narrowed somewhat as the project progressed, as did our core research question, which is best expressed like this: how is it possible for an awarding organisation to remain fully accountable for each CASLO qualification that it awards despite devolving substantial responsibility for assessment processes to centres? In other words, how does quality assurance work for qualifications in this family? This is one of Ofqual's key questions when it comes to regulating these qualifications. It was the lens through which we examined how CASLO qualifications work, and it is the main focus of this report.

Executive Summary

Many of the Vocational and Technical Qualifications (VTQs) that Ofqual regulates belong to a broad family that shares a particular assessment model. A distinguishing feature of this model is that it requires assessors to Confirm the Acquisition of Specified Learning Outcomes. Since there is no generic name for qualifications that share this particular characteristic, we decided to call them **CASLO** qualifications.

Core characteristics of CASLO qualifications

What sets CASLO qualifications apart from other qualifications is the way in which their learning domains and assessment standards are specified and assessed. Imagine, for instance, a Level 1 Award in Business Administration. If it were to be configured as a CASLO qualification, the first thing that its designer would need to do would be to identify the knowledge and skill requirements of the role in question. Next, these would be decomposed into components (units) and elements (outcomes). Criteria for having met each outcome would then be specified.

Assuming that its first unit concerned Communicating Electronically, its **Learning Outcomes** (LOs) and **Assessment Criteria** (AC) might be specified like this:

> The learner will be able to communicate electronically (LO1)
>
> The learner can:
>
>> Send, receive and forward emails (AC1.1)
>>
>> Prepare and send a clear and accurate text message (AC1.2)
>
> The learner will be able to use the internet securely (LO2)
>
> The learner can:
>
>> Log on to the internet (AC2.1)
>>
>> Access an appropriate website showing awareness of security (AC2.2)

In order to pass this unit, a learner would need to have acquired all of the specified LOs, as demonstrated by having met their associated AC. More formally, standards for passing CASLO units, are:

1. **specified analytically** (element by element) and **comprehensively** (to capture all relevant elements)

2. **defined conjunctively** (requiring the acquisition of all elements, that is, complete mastery)

3. **assessed exhaustively** (to cover all elements) and in a **confirmatory** fashion ('acquired' as opposed to 'not-yet-acquired')

Technically, it would be possible for a qualification to be built partly from CASLO units and partly from more classical components. In practice, though, Awarding Organisations (AOs) have tended to build qualifications that only (or predominantly) comprise CASLO units, which is why it makes sense to talk of CASLO qualifications.

The present report describes an investigation into how CASLO qualifications work. It considers how it is possible for an AO to remain fully accountable for each CASLO qualification that it awards despite devolving substantial responsibility for assessment processes to centres.

Too little is known about CASLO qualifications

In comparison with many General Qualifications (GQs), including GCSEs and A levels, CASLO qualifications are documented less comprehensively, researched less thoroughly, and theorised less well. Although individual AOs may know how their qualifications work, knowledge of this sort is not easy to access. Where information has been made available – in qualification specifications, handbooks, and suchlike – this has tended to be at a very practical level. Furthermore, although such practices have been documented over the years, the principles underpinning their effective operation – especially in relation to quality assurance – appear never to have been articulated formally, as far as we can tell. The research that is described in the present report was intended to rectify this, by explaining how CASLO qualifications work in greater depth and detail than has previously been achieved.

CASLO qualifications are different

Our report sets the scene for this research by describing how this family of qualifications evolved and by comparing CASLO qualifications with traditional GQs, which adopt a classical approach to educational assessment.

CASLO qualifications work quite differently from classical GQs. Although the mechanics of delivering and working towards a CASLO qualification are well understood by the many teachers, trainers, and students who take them each year, their technical characteristics are not as widely understood, even among assessment experts. This is largely as a result of inadequate documentation, research, and theorisation. This lack of transparency presents a significant challenge to policy-making. It presents the risk that individuals view these qualifications unproductively, through a more classical (GQ) lens, with insufficient attention paid to their distinctive characteristics.

Ofqual recently faced this challenge when reviewing its General Conditions of Recognition in relation to external quality assurance. It took time to clarify exactly how CASLO qualifications departed from classical GQs in respect of quality

assurance, and even this clarification left certain questions unanswered. The present report is part of Ofqual's response to this lacuna.

Centre-assessed components within classical GQs are marked by the centre according to a national mark scheme. The accuracy and consistency of centre-based marking is quality assured externally by a moderator. Moderation often involves fine-tuning mark totals awarded by the centre – scaling them downwards to rectify undue lenience, or upwards to rectify undue harshness – although sometimes a full re-mark of all work at the centre is required.

External quality assurance for CASLO qualifications works differently. This is partly because these qualifications are often 100% centre-assessed. This increases the external quality assurance burden for both AOs and centres. It also alters the potential risk profile. But the main reason why quality assurance works differently is because the idea of fine-tuning does not translate into a context in which centres judge candidate work directly against assessment criteria and where complete mastery of the learning domain is required. So, how does external quality assurance work for CASLO qualifications? This question provided a useful point of focus for the present study, helping us to pursue our broader aim, which was to explain more generally how CASLO qualifications work.

We decided to investigate how they work

Our research was originally intended to involve 4 strands, with each strand interrogating the same small sample of CASLO qualifications:

1. desk-based review of documentary evidence on each qualification

2. semi-structured face-to-face interview with AO officers

3. observation of external quality assurance in 2 centres

4. semi-structured follow-up interview with 2 External Quality Assurers

Unfortunately, completion of the Strand 2 interviews coincided with the COVID-19 lockdown. Fortunately, though, insights arising from the first 2 strands proved to be sufficiently rich to construct an account of how CASLO qualifications work.

Our original intention was to explore both principles and practices associated with CASLO qualifications, through the window of quality assurance. In response to COVID-19, however, we decided to restrict our focus to matters of principle. Strands 3 and 4 were replaced by an appraisal stage, in which insights from Strands 1 and 2 were presented to officers from the AOs in our sample, who provided feedback on their plausibility. This feedback was used to shape the account that is provided in this final report. Although this appears to be a plausible account, it is still a tentative one, having been based upon a small sample of disparate qualifications and AOs.

Owing to the very wide variety of CASLO qualifications in England, insights from this project were framed in terms of how the **archetypal CASLO qualification** appears to operate. We constructed this hypothetical qualification to capture commonly occurring features and processes, given the circumstances under which CASLO qualifications typically work.

The archetypal CASLO qualification

All CASLO qualifications have in common the fact that standards for passing CASLO units are: specified analytically and comprehensively, defined conjunctively, and assessed exhaustively and in a confirmatory fashion. However, beyond these structural requirements – which are common by definition – we identified additional features, which helped to paint a picture of the archetypal qualification within this broad family.

In England, the archetypal CASLO qualification is a VTQ, designed with **flexibility** in mind to accommodate learners studying under a variety of circumstances, for instance:

- learners following different pathways within the same qualification

- learners developing alternative (customised) skillsets within the same qualification

- learners studying in different locations (college-based, work-based, and so on)

- learners demonstrating their competence in different workplace settings

- learners starting and finishing their learning on different timelines

Flexibility is delivered in a variety of ways, from learners achieving different learning outcomes (from different units) within the same qualification, to learners being assessed on the same learning outcome using different kinds of evidence.

CASLO standards relate to domains of learning that are typically broken down into units – with each unit comprising a set of LOs and with each outcome characterised in terms of a set of AC. We can think of this as a **graphical** approach to specifying a domain of learning, in the sense that each unit specification functions like the map of a learning journey, where each learning outcome corresponds to a significant landmark along the way. The main purpose of the learning journey is to traverse the entire map and to visit each landmark.

Because standards for CASLO qualifications are specified comprehensively and assessed exhaustively – and because they often emphasise practically oriented outcomes – they carry a heavy assessment burden. The need to accommodate both flexibility and burden is significant, from a qualification design perspective, disposing

the archetypal CASLO qualification towards a **continuous centre-assessment** (or continuous workplace-assessment) model. Other common design features include:

- only awarding the passing grade

- completion of the qualification not being heavily time constrained

Identifying distinctive features along these lines helps to distinguish the archetypal CASLO qualification from the archetypal classical qualification. In fact, some qualifications have features that resemble the CASLO archetype and features that resemble the classical archetype. The 2 models are not necessarily in tension with each other and it is possible to imagine a variety of hybrid approaches.

Unfortunately, certain features associated with the archetypal CASLO qualification present significant threats to accurate and consistent assessment. For instance:

- a heavy assessment burden increases the potential scale of assessment error

- high flexibility makes it hard to manage the risk of assessment error

- the mastery approach to assessment increases the stakes of assessment error

Given the substantial devolution of responsibility for assessment processes to centres, AOs need to be very confident in the ability of each and every centre to assess accurately and consistently. This helps to explain why quality assurance, in particular, needs to be conceived quite differently for CASLO qualifications, even when compared to how centre-based assessment is quality assured within GQs.

How CASLO qualifications work

We set out to construct, from first principles, an account of how CASLO qualifications work – that is, to explain how they could work, in theory, rather than to evaluate how well they actually work, in practice. We attempted to make sense of how certain critical processes operate for CASLO qualifications – quality assurance processes in particular – by drawing comparisons with how ostensibly similar processes work for classical GQs, which tend to be understood far better.

We concluded that the circumstances in which these qualifications operate have given rise to a quite different model of accountability, with respect to centre-assessment, when compared with classical GQs. We explained this metaphorically, by contrasting a **hands-off sub-contracting** model (GQ) with a **hands-on partnership** model (CASLO). According to this metaphor, AOs appoint centres as apprentices, with the aim of enculturating them into the community of practice for the qualification in question. Yet, however experienced a centre becomes, the AO will always maintain a watching brief over its practices, via ongoing monitoring processes. The centre may transition from 'high risk' to 'low risk' but it should never be considered 'no risk'.

We observed that the archetypal CASLO qualification operates a tripartite centre-assessment quality assurance system, which comprises:

1. centre approval

2. internal quality assurance

3. external quality assurance

This implies a cradle-to-grave approach, which begins once a centre seeks approval to offer the qualification. It does not end until it has stopped offering the qualification and the last student has been certificated. Indeed, whereas quality assurance is end-loaded under the GQ model (with a particular emphasis upon external quality assurance) it is front- and middle-loaded under the CASLO model (with a particular emphasis upon internal quality assurance).

Underpinning principles

Information from our interviews with AO officers revealed sufficient commonalities across qualifications to propose several principles that appear to underpin the effective operation of CASLO qualifications:

1. **structural integrity** (centres need to demonstrate structural integrity from the outset – from a teaching, assessment, quality assurance, and management perspective – to be judged worthy to deliver the qualification)

2. **self-regulation** (centres need to operate as self-regulating systems, having internalised the AO's quality standards and quality assurance practices)

3. **comprehensive monitoring** (quality assurance needs to focus upon the effective delivery of a qualification in the round – teaching, assessment, quality assurance, and management)

4. **risk-based sampling** (given the potential scope of quality assurance activities, sampling needs to be driven by a risk-based model)

5. **incremental improvement** (centres need to be supported to improve incrementally – from centre approval onwards and for as long as a centre continues to offer a qualification)

6. **supportive surveillance** (the relationship between an AO and its centres – and between an Internal Quality Assurer and their assessors – needs to remain fundamentally supportive)

7. **conditional (evidence-based) trust** (with increasing competence comes increasing trust – but, if a centre were to abuse that trust, then this would need to be dealt with via an appropriate sanction)

These principles are implemented within CASLO qualifications through a variety of **scaffolding mechanisms** – including centre approval, training, standardisation, internal and external scrutiny of both processes and outcomes – which together are intended to ensure the effective operation of these qualifications.

One of the most important scaffolding mechanisms concerns teaching and learning. This mechanism, which is often described as **formative assessment**, is specifically designed to ensure that learners remain continually on target to pass the qualification. With appropriate course entry requirements, an effective selection process, and adequate formative assessment arrangements, the expectation is that learners who exert enough effort ought to be able to reach competence sooner or later.

Under the principle of self-regulation, the more effectively an Internal Quality Assurer (IQA) coordinates quality assurance, the less frequently an External Quality Assurer (EQA) will need to intervene. Effective IQAs keep qualification delivery rolling smoothly. Ideally, if an IQA has done their job effectively, then there should be no need for EQA intervention. The IQA is therefore a lynchpin, at the heart of this holistic system.

Whereas external standards scrutiny is at the heart of the centre-assessment quality assurance model for classical GQs, it is perhaps better understood as the final safety net for the archetypal CASLO qualification. It is premised on the assumption that an IQA will already have coordinated a range of quality assurance activities, including a substantial amount of internal standards scrutiny. This helps to explain why AOs do not describe the way an EQA scrutinises qualification standards – that which Ofqual now refers to as Centre Assessment Standards Scrutiny – as fundamentally interventionist (which might be a reasonable description of Moderation for GQs).

Having said that, external standards scrutiny still serves a number of important functions for CASLO qualifications. This includes a **corrective** function, based upon what might be described as a **sporadically interventionist** approach. This involves identifying and correcting glaring errors of professional judgement or serious anomalies or inaccuracies due to unintentional maladministration or intentional malpractice. Identifying and correcting only glaring errors is quite different from fine-tuning marking, which typifies external standards scrutiny for classical GQs. It is consistent with external standards scrutiny for CASLO qualifications functioning more like a safety net. It is also consistent with an account of how judgemental **tolerance** might be exercised within CASLO qualifications.

Finally, even adopting a sporadically interventionist approach, external standards scrutiny will still serve an important **deterrent** function – a deterrent to maladministration attributable to inattentiveness, as well as a deterrent to malpractice arising from bad intent.

Introduction

GCSEs and A levels are widely known and well understood. They attract a great deal of attention, are tightly regulated, and their features and processes are fairly typical of classical educational assessments internationally. Their underpinning principles are well established and their technical characteristics have been thoroughly researched. Unfortunately, the same cannot be said for the many Vocational and Technical Qualifications (VTQs) that operate alongside GCSEs and A levels. Their practices might be well known to the AOs, centres, and students who work with them, and well known to the employers who benefit from them. Yet their underpinning principles are far less well established and their technical characteristics are far less thoroughly researched.

Ofqual is keen to help to reverse this trend and has committed to regulating VTQs with the same seriousness and focus as it affords to General Qualifications (GQs).[2] An important consequence of this new commitment is a need to understand how VTQs work in far greater depth and detail than has been customary in the past. The research that is reported below is intended to support this goal, by investigating how a broad family of VTQ works. For reasons that will soon be elaborated, we decided to refer to them as 'CASLO' qualifications.

VTQs in England tend to differ from GQs

Because VTQs are documented less comprehensively, researched less thoroughly, and theorised less well than GQs, this raises the risk that policy-making may sometimes pay insufficient attention to important differences. Just prior to the onset of COVID-19, Ofqual revisited its General Conditions of Recognition, in relation to external quality assurance, in order to address a concern of this sort.

An investigation into licence-linked qualifications used in the private security industry had identified a number of risks associated with the award of these high stakes qualifications (Ofqual, 2017). In particular, this included the risk of not identifying and correcting inappropriate centre-assessment decisions before a qualification is awarded. It also included the associated risk of negative consequences from withdrawing a qualification once it has been awarded. This, in turn, implicated a practice that is widespread across the industry, whereby a centre may be granted Direct Claims Status (DCS).[3]

---

[2] Ofqual formally adopted this strategy in 2018. The present research project began in 2019, although its progress was delayed by the onset of COVID-19.

[3] Although this term is widely used in the industry, it is operated in different ways by different AOs. The term does not appear in the Ofqual regulatory framework.

When an Awarding Organisation (AO) is satisfied that a centre can effectively deliver a qualification and assess it with consistent accuracy, it may decide to confer DCS upon that centre for the qualification in question. Once granted, the centre can request certification for an individual or group of candidates – once they are satisfied that those candidates have met the required qualification standards – without the need for their assessment decisions to be externally quality assured before the award of each certificate. DCS is particularly useful for short 'roll-on-roll-off' qualifications, where candidates require rapid certification in order to be able to use their certificates straight away.

In response to the security industry investigation, Ofqual concluded that DCS was likely to conflict with its requirement that AOs remain fully accountable for the awards that they make. Ofqual embarked upon a series of consultations to investigate the adequacy and appropriateness of its existing General Conditions of Recognition. This resulted in a tighter, but also broader, set of conditions related to external quality assurance (Ofqual, 2020a; 2020b). The new regulations permitted DCS, as long as some form of Centre Assessment Standards Scrutiny (CASS) was undertaken and certain baseline requirements were met.

The new regulations also indicated that, for certain qualifications, Ofqual would require a specific form of external quality assurance, Moderation, which was defined as:

> A particular form of Centre Assessment Standards Scrutiny through which the marking of assessments by Centres is monitored to make sure it meets required standards and through which adjustments to a Centre's marking are made, where required, to ensure that results are based on the required standard. Moderation takes place before final results are issued under Condition H6.1.

(Ofqual, 2020a, p.6)

Moderation therefore specifies a form of standards scrutiny, where the aim is: to identify centres that have failed to apply standards with sufficient accuracy and consistency and to put this right before certificates are awarded. The traditional GQ approach to external quality assurance would be classified as Moderation.

Whereas Moderation requires that problems are identified and rectified prior to any certificate being awarded, other varieties of CASS are permitted to take place either before or after certificates have been awarded. Ofqual's definition of CASS is therefore more encompassing than its definition of Moderation. CASS is defined as:

> The process through which an Awarding Organisation –
>
> (a)   periodically scrutinises the marking of assessments by a Centre to ensure that it has not deviated from required standards,

(b) considers whether it is appropriate to correct any mark and, if appropriate, corrects that mark (including where changes are required under Condition H2.5(b)),

(c) in line with Condition H6.3(b), considers whether it is appropriate to correct any incorrect result and, if appropriate, corrects that result, and

(d) takes action to prevent such deviation from recurring.

(Ofqual, 2020a, p.6)

CASS is now the minimum requirement for all regulated qualifications that include centre-assessments. It means that an AO must periodically scrutinise samples of assessed work from all centres. This is to assure itself that all centres are assessing to the national standard and to exercise accountability for the award of all qualifications.

In short, consultations that followed the publication of the security industry report highlighted that many VTQs operate differently from the classical GQ model of external quality assurance. This led to Ofqual's regulations being changed in order to require standards scrutiny across the board. The new regulations permitted different 'levels' of CASS, depending on the nature and circumstances of the qualification in question.

This recent demonstration of the fact that VTQs may differ significantly from GQs provides an important part of the rationale for the present study. It also helps to explain why we decided to use external quality assurance as a window through which to explore such differences in greater depth and detail.

Recent history of VTQs in England

Understanding how VTQs work is no small feat. The landscape of occupational, vocational, technical, and professional qualifications in England – even just the subset that is regulated by Ofqual – is extremely diverse.[4] Around 140 AOs offer Ofqual-regulated VTQs (figure correct September 2021). Some of these are first and foremost professional institutes, offering specialist qualifications in a particular sector for development and accreditation purposes (for example, the Chartered Institute of Professional Development). Others are general awarding bodies, offering a wide range of qualifications across a wide range of sectors (for example, Pearson, or City & Guilds).

---

[4] We will use the abbreviation VTQ to include occupational, vocational, technical, and professional qualifications, without drawing hard-and-fast distinctions between these 4 categories.

During the twentieth century, as AOs responded to a multiplicity of market demands, qualifications of all shapes and sizes emerged. The evolutionary forces at work resulted in what was to become known (disparagingly) as England's qualification 'jungle'.

Although there are many similarities between qualifications in this complex VTQ landscape (as they have been developed in response to similar evolutionary forces), there are also many differences between them (as they have not been subjected to as much regulatory control as GQs have). Indeed, although it makes sense to talk of a generic GQ model, this is not really true of VTQs. Instead, it would be more correct to say that there are a number of families of VTQs, which operate quite differently from classical GQs. Within each family, certain features and processes may be common, while others may overlap in a less consistent and less formal manner. The largest family of VTQs in England – which, for reasons that will soon be explained, we will refer to as 'CASLO' qualifications – is the focus of the present research.

Since the 1980s, there have been numerous attempts to bring order to the landscape of qualifications in England, to make it easier to navigate. A key initiative was the introduction of National Vocational Qualifications (NVQs), followed subsequently by the National Qualifications Framework (NQF), and then the Qualifications and Credit Framework (QCF). The following subsections introduce CASLO qualifications – which have tended to be VTQs of one sort or another – showing how they relate to and emerged from NVQs and subsequent initiatives.

National Vocational Qualifications (NVQs)

National Vocational Qualifications were introduced during the late 1980s and early 1990s. They were originally conceived as work-based qualifications, with assessment undertaken in the normal course of work, by managers or supervisors (Jessup, 1991). They focused upon what an individual needed to have learned in order to be considered competent in an occupational role. This 'outcomes-based' approach to qualification development contrasted with previous 'input-based' approaches, which had been criticised for focusing unduly upon curriculum delivery (inputs) rather than competence acquisition (outcomes). No time limits were specified for achieving NVQ learning outcomes – neither minimum nor maximum – and this provided an 'antidote' to existing apprenticeships, which were all too frequently awarded almost as a reward for 'time served' by the apprentice. They also provided an 'antidote' to traditional qualifications, by focusing upon the practical demonstration of competence, as opposed to how candidates performed in written exams.

Central to the design of NVQs was the idea that, if qualification standards could be articulated with sufficient transparency, then this would render the assessment process unproblematic. In other words, if we can specify what it means to be

competent in a certain role – in terms of the set of elements that comprise what it means to practise competently – then assessment ought to reduce to the simple and straightforward matter of asking the learner to demonstrate their competence, element-by-element. In other words, if we ask a learner to demonstrate an element of competence and they respond with a successful performance, then we can conclude that they are competent on that element. Once they have demonstrated all of the specified elements of competence, then we can conclude that they are competent in the occupational role for which they have been trained.

NVQ qualification standards were intended to encapsulate what it means to be competent in a particular occupational role and were to be derived from National Occupational Standards (NOS). They were articulated, on a unit-by-unit basis, via statements that identified both elements of competence (now known as Learning Outcomes, LOs) and performance criteria (now known as Assessment Criteria, AC). The NVQ statement of competence was intended to provide a target for learners to aim at. The NVQ certificate was intended to provide a statement of what that learner had achieved (Jessup, 1991).[5]

NVQs were developed at a range of levels, corresponding to the loose hierarchy of roles that could be identified within most occupational areas. Lower-level qualifications were associated with routine, predictable, foundational competencies. Higher-level qualifications were associated with broader, more complex competencies, requiring a greater degree of autonomy.

NVQs were originally related to each other via the NVQ Framework, which was a 2-dimensional structure, mapping area of competence against qualification level. Some years later, Ron Dearing's 'Review of Qualifications for 16-19 Year Olds' added impetus to existing proposals for a more encompassing framework (see, for example, FEU, 1992; 1995), which would establish coherence across the full range of qualifications available to learners (Wilson, 2010). This came into being as the

---

[5] The idea of specifying learning outcomes in detail – a core characteristic of NVQs and of the Competence-Based Assessment (CBA) movement more generally – was widely embraced within England's technical and vocational education and training communities, as early as the 1970s. For instance, City & Guilds pioneered research and development into what was termed the 'process-competence-based' approach to technical education, as an alternative to the traditional subject-based approach. By September 1980, most of its Engineering Craft Studies schemes had been re-stated in terms of learning outcomes (Stevens, 1993). Qualifications developed by the Business Education Council, from 1979 onwards, were also modelled on the learning outcomes approach (Fisher, 2003). The BTEC approach evolved directly from these qualifications, following the 1983 merger of the Business Education Council and the Technician Education Council, to become the Business and Technician Education Council (BTEC). Having said that, the manner in which these predecessor qualifications operationalised CBA was not quite as strong as the NVQ approach. BTEC courses, in particular, were characterised by a heavy emphasis upon curriculum design considerations, which the NVQ approach tended to eschew.

National Qualifications Framework (NQF). Its goal was to achieve greater coherence by locating a wide variety of existing qualifications at different levels on the new framework. This was intended to improve transparency by revealing how individual qualifications related to each other within a broader system.

The NVQ approach to unit (and qualification) design required standards to be:

1. specified **analytically** and **comprehensively** (in terms of the set of elements – LOs – that collectively comprised the domain of competence)

2. defined **conjunctively** (in terms of having achieved the full set of LOs)

3. assessed **exhaustively** and in a **confirmatory** fashion (in terms of confirming the achievement of each and every LO in the set)

The present research is focused upon the broad family of qualifications (which are typically VTQs) that shares all of these features in common.

Competence-Based Assessment (CBA)

The approach that underpinned the development of NVQs was intended to be revolutionary, that is, to overthrow the traditional GQ approach that was based upon the teaching of syllabuses and (in the main) assessment by exam. In the VTQ context, this new approach became known as Competence-Based Assessment (CBA), emphasising that VTQs ought to be linked to professional or occupational standards of competence.

In fact, the underpinning model was not restricted to professional or occupational contexts.[6] Indeed, Gilbert Jessup – Director of Research at the National Council for Vocational Qualifications, which was responsible for NVQ rollout – saw these new qualifications as merely the vanguard of an emerging model of education and training, premised upon clarity in specifying LOs and AC. The following selection of quotations from the concluding section of his 1991 book helps to explain how he understood this wider revolution in education and training:

> The common feature […] would be that all forms of learning provision would be stated in terms of outcomes. Thus there would be predetermined statements in the form of competences or attainments which serve as targets and guide the course of learning. (p.134)

---

[6] For instance, the outcomes-based approach was simultaneously being rolled out through National Curriculum Assessment arrangements (see Shorrocks-Taylor, 1999).

Assessment is being brought into the real world and de-mystified within the new model of education and training. We shall, I hope, see the demise of the last minute swotting of information soon to be forgotten for examinations. (p.135)

What is assessed and the standards of performance required are open to both the assessor and the candidate alike. […] Self-assessment will become an important component in learning. (p.135)

Assessment will be continuous, normally integrated with the process of learning. (p.135)

Assessment will be far more thorough, it will focus on what needs to be assessed and will cover more systematically a far wider range of outcomes than can be achieved by traditional methods. As a consequence it will be more valid. (pp.136)

The overall model is designed to promote learning. It incorporates many features which make learning more attractive and easier to access. (p.138)

(Jessup, 1991)

Drawing its inspiration from North American educational movements related to mastery learning and criterion-referenced assessment, this new approach to developing qualifications was intended to improve teaching and learning, as well as to improve assessment.[7]

Qualifications and Credit Framework (QCF)

The desire to bring even greater coherence to the VTQ landscape led to the development of the Qualifications and Credit Framework (QCF) during the mid-2000s. Whereas the NQF was a 'weaker' descriptive framework – developed primarily to help make sense of a complex qualification landscape – the QCF was a 'stronger' prescriptive framework. In other words, it exerted greater regulatory force, requiring that the units and qualifications that were to populate the framework must share common design characteristics.

This level of prescription was intended to be strong enough to enable learners to 'mix and match' units across the framework, even permitting them to combine units from different AOs. To make this work, all units and qualifications needed to comply with a plethora of design rules relating to level, titling, credit, aggregation, and so on.

Design rules were even specified at the level of how to write LOs and AC for qualification units. High level requirements were detailed in the 'Regulatory

---

[7] This is true, even though the NVQ approach was focused specifically upon assessment (that is, outcomes) rather than upon curriculum (that is, inputs).

Arrangements for the Qualifications and Credit Framework' (Ofqual, 2008). For instance:1.4 All units must contain learning outcomes that:

a   set out what a learner is expected to know, understand or be able to do as the result of a process of learning

b   are clear and coherent, and expressed in language that is understandable by the learners for whom the unit is intended or by a helper or adviser where the learners themselves are not able to understand the learning outcomes

c   are expressed in a manner that addresses individual learners in the third person and will make sense to a learner both before a unit is offered and after the learning outcomes have been achieved

d   are capable of assessment and, in conjunction with the assessment criteria related to that outcome, set a clear assessment standard for the unit.

1.5 All units must contain assessment criteria that:

a   specify the standard a learner is expected to meet to demonstrate that the learning outcomes of that unit have been achieved

b   relate to an individual learning outcome in language consistent with it

c   are sufficiently detailed to support reliable, valid and consistent judgements that a learning outcome has been achieved, without creating an undue assessment burden for learners or assessors

d   do not include any explicit references to the methods or instruments of assessment to be used.

<div align="right">(Ofqual, 2008, pp.11-12)</div>

Further guidance was provided within a variety of supporting documents, for example,

• 'The Quality and Standardisation of Units in the Qualifications and Credit Framework' (QCDA, 2009)

• 'Guidelines for Writing Credit-Based Units of Assessment for the Qualifications and Credit Framework. Version 4' (QCDA, 2010a)

• 'Guidance on Combining Knowledge and Skills in Units of Competence-Based Vocational Qualifications in the Qualifications and Credit Framework. (QCF)' (QCDA, 2010b)

The CASLO designation

When regulatory arrangements for the QCF were published, the 3 key features of NVQ units identified above were specified as design rules (see Ofqual, 2008). Consequently, by the mid-2000s, the group of qualifications that we are presently interested in had become far broader than the label 'NVQ' might imply. Furthermore, many of these QCF qualifications did not assess 'competence' against professional or occupational standards, suggesting that even the 'CBA' label might be unduly narrow. So, how ought we to label the family of qualifications that we are interested in?

One option was to describe them as 'QCF' qualifications. However, this seemed to imply broader connotations, associated with the plethora of design rules mentioned earlier (for example, titling, or credit). Furthermore, the Framework itself was withdrawn in 2015, whereas qualifications designed according to this template are still common.

It was therefore decided to introduce a new label – **CASLO** – to characterise the underpinning assessment model, which requires assessors to Confirm the Acquisition of Specified Learning Outcomes (CASLO). This captures the critical requirement that standards for passing CASLO units are:

- specified **analytically** and **comprehensively**

- defined **conjunctively**

- assessed **exhaustively** and in a **confirmatory** fashion

Most regulated CASLO qualifications in England not only specify standards analytically and comprehensively, but also highly **atomistically**. They specify a set of AC for each and every LO and require that all AC are satisfied to confirm the acquisition of each outcome. This exhaustive, confirmatory approach to assessment tends to drive an exhaustive, confirmatory approach to managing the teaching and learning process.[8]

A key idea underpinning the CASLO model is its expectation that, by participating in the programme of learning that leads up to the qualification, a learner will acquire **all** of its specified learning outcomes, at least to a satisfactory standard. In the GQ context, there is no expectation – neither explicit nor implicit – that most learners should acquire all of the outcomes associated with a domain of learning. Indeed, GQs normally adopt a compensatory approach to aggregating assessment

---

[8] The term 'exhaustive' is not meant to be pejorative. It is simply intended to indicate that the assessment – and, in preparation for assessment, the teaching – covers each and every one of the LOs (and typically each and every one of the AC too).

judgements. This means that failure to demonstrate a satisfactory level of performance in relation to one outcome can be compensated for by an exceptional level of performance in relation to another.[9]

The CASLO model therefore defines 'certification-worthiness' quite differently from the classical GQ model. In particular, the attainment profile of a learner who just passes a CASLO qualification is likely to differ considerably from that of a learner who just passes a GQ. It is likely to be more balanced, with no missing competencies, consistent with having achieved threshold competence across all LOs. Conversely, because GQ candidates can compensate for worse performance in one area via better performance in another, their attainment profiles often end up being quite 'spiky'. This happens at all levels of attainment, including at the pass-fail boundary.

It is important to emphasise that this is a very broad family of qualifications. The characteristics that we have identified as quintessential to the CASLO model are structural. They were specified as basic structural requirements for all units (and qualifications) that were to be recognised within the QCF. Consequently, all sorts of qualifications, serving all sorts of learners, within all sorts of contexts, came to be specified in this manner and became part of the CASLO family. This family of qualifications is therefore hard to describe definitively, other than via the CASLO designation itself and the implications that follow directly from this. Consequently, in order to explain how CASLO qualifications work, we decided to use our research and analysis to construct a picture of an **archetypal CASLO qualification** (see below).

Classical GQs versus CASLO qualifications

The major differences between classical GQs and CASLO qualifications relate to how their learning domains and assessment standards are specified and, consequently, assessed. These differences are explored in detail in Annex 1, which compares an A level geography qualification with a BTEC Diploma in pet services. We chose a BTEC for this introductory illustration because many BTECs are CASLO qualifications – constructed entirely or mainly from CASLO units – making BTEC their largest and most familiar brand.

The following 2 sections explore structural differences between classical GQs and CASLO qualifications. This is to set the scene for a deeper analysis of how CASLO qualifications work based on document analysis and interviews with AOs. The first

---

[9] There are isolated examples of assessments based upon a CASLO model within certain GQs. The A level science endorsement of practical skills is one such example. GQs are mainly based upon a compensatory approach, however, and their LOs and AC tend to be specified far less analytically and comprehensively.

section considers high level structural features and the second focuses more specifically upon external quality assurance. Quality assurance operates quite differently for CASLO qualifications. Exploring these differences seemed likely to reveal important insights into the principles that underpin how CASLO qualifications work.

Comparing classical GQs and CASLO qualifications

What particularly sets CASLO qualifications apart from classical GQs is the explicitness with which their learning domains and assessment standards are specified, in terms of LOs and AC, respectively. To illustrate this, Table 1 presents the LOs and AC that comprise Unit 3 of the pet services BTEC (see Annex 1). All AC need to be met, across all of the LOs, in order to pass this unit.

Learning outcomes	Assessment criteria
LO 1. Explore how effective customer service contributes to business success	1.1 Describe the different approaches to customer service delivery 1.2 Examine ways that customer service in a selected business can meet the expectations and satisfaction of customers and adhere to relevant current legislation and regulations
LO 2. Demonstrate customer service in different situations, using appropriate behaviours to meet expectations	2.1 Demonstrate the appropriate communication and interpersonal skills to meet customer needs in different situations 2.2 Review own customer service skills, identifying gaps where improvements could be made 2.3 Present a clear, effective development plan for own customer service skills
LO 3. Investigate the methods used to improve customer service in a business	3.1 Research the methods that a business can use to monitor and make improvements to their customer service provision

Table 1. Unit 3: Customer Care in a Pet Service Business

Feature	CASLO Qualifications	Classical GQs
Content	Content organised by unit, but not necessarily specified in detail.	Content organised by subject area and specified in detail.
Learning Outcomes	LOs defined explicitly, for each unit, via a small number of high-level statements. Content may be elaborated via notes for each LO.	LOs defined implicitly, for each area, via a detailed content specification.
	Learners required to achieve all LOs.	Learners not required to achieve all LOs.
	Prescriptive LOs – range indicates required (minimum) attainment profile.	Aspirational LOs – range indicates ideal (maximum) attainment profile.
Assessment Objectives	AObs defined explicitly at LO and AC level, via command verbs – for example, explore, describe, examine.	AObs defined explicitly at qualification level, via command verbs – for example, demonstrate, apply, use.
		AObs operationalised at task level, to reflect a specified balance between them at the qualification level.
Assessment Criteria	AC defined explicitly, for each LO, via a small number of high-level statements.	AC written specifically for each task, as presented in a mark scheme.
	Prescriptive AC – statements define standards, which are applied as competence thresholds.	AC typically descriptive – exemplifying rather than defining standards (for example, via 'best fit' attainment level descriptions).
Aggregation method	Conjunctive – passing the qualification indicates having achieved all AC, for all LOs, across all units (mastery).	Compensatory – passing the qualification indicates having exceeded a threshold (overall) level of attainment.

Table 2. Structural comparison of CASLO versus classical GQ model

Critical structural differences between CASLO qualifications and classical GQs are summarised in Table 2. As noted above, classical GQs differ from CASLO qualifications primarily in terms of how they specify domains of learning, including subject content. Classical GQs adopt a more 'narrative' approach, while CASLO qualifications adopt a more 'graphical' one.

CASLO qualifications specify their domain of learning **graphically** in the sense that each of their unit specifications functions like the map of a learning journey. Each learning outcome corresponds to a significant landmark along the way. The map includes snapshots of these landmarks – learning outcomes with associated assessment criteria – and these snapshots enable learners to tell that they have arrived at a particular landmark. The main purpose of the learning journey is to traverse the entire map and to visit each landmark.

Classical GQs, on the other hand, specify their domain of learning **narratively** in the sense that their qualification specification functions more like a city guide. The guide provides a rich description of each of the districts in the city – subject content areas with assessment objectives – although without necessarily identifying precise boundaries between districts. The specification typically supports a less prescriptive learning journey, where the main purpose is to absorb the culture of the city and not (necessarily) to visit each and every landmark.

Under the graphical approach, subject content is implicit in both learning outcomes and assessment criteria. It may also be articulated explicitly, in greater or lesser detail, although this additional information is not always available. (Content implicit, outcomes explicit.) Under the narrative approach, learning outcomes are implicit in both subject content and assessment objectives, but they tend not to be articulated explicitly. (Content explicit, outcomes implicit.)

Centre Assessment Standards Scrutiny

There is a long tradition of incorporating centre-assessed components within GQs. This was a prominent feature of Certificate of Secondary Education (CSE) qualifications, which influenced the design of GCSEs. Indeed, some early GCSEs, notably in English, were assessed exclusively by teacher assessment. However, concerns related to workload and cheating, during the 2000s, led to controlled assessment replacing GCSE coursework. During the 2010s, subsequent concerns over controlled assessment led to centre assessment being permitted by exception only.[10] Geography fieldwork represents one such exception.

Quality assuring centre-assessed components requires some form of standards scrutiny, which involves scrutinising samples of centre-assessed work. Approaches

---

[10] This is partly why we now refer to coursework in the negative, that is, Non-Exam Assessment (NEA).

to quality assuring GQ centre-assessed components are well documented and well researched (for example, Taylor, 2005; Wilmut, 2005; Wilmut & Tuson, 2005; Williamson, 2016; Cuff; 2017; Cuff, 2018; Chambers, Williamson, & Child, 2019). They are consistent with the definition of Moderation, presented earlier.

Moderation of GQs

Once candidates have completed all of their work for an NEA and this work has been marked internally – which includes an internal standardisation process if there is more than one teacher marking the subject – the marking is subsequently quality assured by the AO. This involves appointing a Moderator, who receives an AO-specified sample of NEA work from the centre and who second marks that sample according to the mark scheme. Although AOs operate slightly different practices – compare AQA (2016) and Gill (2015) for instance – the following account captures the essence of GQ Moderation.

During GQ Moderation, quality is assured by second marking a sample of centre-assessed work for the NEA. It involves comparing the total mark awarded to each piece of work by the centre with the total mark subsequently awarded by the Moderator.[11] The Moderator begins by marking a sub-sample (of the sample) and compares differences between the marks they award and the marks awarded by the centre. That is, on a candidate-by-candidate basis, they compare their mark total with the centre's mark total. As long as none of the mark total differences exceeds a prespecified threshold, all marks awarded by the centre will be accepted. This threshold – which is known as the **tolerance** threshold – is typically 6% of the maximum mark. The approach means that small differences between the centre and Moderator are tolerated, while larger differences require further investigation.

Further investigation involves second marking additional students from the sample and comparing differences again. Possible outcomes include: accepting all centre marks, adjusting centre marks up or down, or re-marking the work of all candidates from the centre. Centre marks are adjusted when further investigation concludes that the centre has marked too leniently or harshly, but consistently so. This often leads to a tapered adjustment to candidates' marks, for example: candidates at the top of the mark range have their total reduced by 4 marks, candidates lower down the mark range have their total reduced by 3 marks, then 2 marks, then 1 mark, with centre marks standing for candidates right at the bottom of the mark range. Note that this process always preserves the centre's rank ordering and treats candidates in the

---

[11] Even when the centre and the Moderator award the same total mark to a piece of work, the profile of marks awarded across the tasks that comprise that piece of work may differ. Any such differences are ignored, for the purposes of quality assurance.

sample in exactly the same way as candidates who are not in the sample.[12] Consequently, the Moderator is not actually **re-marking** the work of individual candidates in the sample, they are **second marking** a sample of work from the centre. This is to evaluate the overall accuracy and consistency of assessment at the centre and to apply an adjustment if necessary.

CASS for CASLO VTQs

Given the structural differences identified above, it would not be possible for CASLO qualifications to be externally quality assured in exactly the same way as for GQs. Table 3 presents some high-level information comparing the pet services BTEC and the geography A level, which helps to explain why. As the following points make clear, looking at Centre Assessment Standards Scrutiny for CASLO qualifications through a classical GQ lens simply does not work.

First, even for GQs that include NEA – and many do not – NEA tends to comprise only a relatively small component of the qualification. Conversely, CASLO qualifications are often 100% internally assessed. This means that external quality assurance is potentially a far larger undertaking for the AOs that offer these qualifications and for the centres that deliver them. In addition, there is potentially far more variation over the nature of the assessment evidence that needs to be scrutinised, making judgemental processes more complex as well.

Second, the nature of the judgements that are quality assured is quite different across GQs and CASLO qualifications. A key feature of CASLO judgements is that they are supposed to confirm having satisfied an assessment criterion – and, by extension, having mastered a learning outcome – modelling competence as a straightforward binary (either-or) outcome. Conversely, a typical feature of GQ judgements – particularly when they relate to multi-mark tasks – is that they are supposed to represent the degree to which learning outcomes have been mastered, modelling competence as a more continuous outcome (with gradations).

Third, because the acquisition of **all** specified learning outcomes needs to be confirmed under the CASLO approach, each and every AC judgement is critical to the overall judgement of competence, which is far from true of GQs. Compare, for example, the 2 hypothetical external quality assurance outcomes for A level geography versus BTEC pet services, which are presented within Table 4.

For the A level component, 4 judgements need to be made and the 4 resultant marks are summed to a single mark total. The example in Table 4 shows how discrepant

---

[12] For example, if a Moderator marked a piece of work from the sample 4 marks lower than the centre mark, but the overall adjustment to centre marks was -3, then the candidate would have their mark reduced by 3 marks, in exactly the same way as for all other candidates in the centre.

judgements can average out (to no difference at all) in the overall mark total. Conversely, for the BTEC unit, 6 judgements need to be made and each of them impacts independently on the final decision. The example in Table 4 shows how just one discrepant judgement leads to failing the unit (and the qualification).[13]

---

[13] The idea of 'failing' a unit is used informally, here, although candidates are rarely deemed to have 'failed' a unit, as they are generally permitted to submit new evidence of competence in order to pass.

Feature	AQA geography A level	Pearson pet services BTEC
Assessment structure	1 NEA component. 2 Exam components. [All mandatory, with some optional tasks.]	6 centre-assessed components. [5 mandatory. 1 optional, from choice of 7.]
Centre-assessment format *	1 report.	3 separate assignments.
Structure of centre-assessed component *	Separate mark for each of 4 areas (10, 15, 20 15), that is, 60 marks in total.	Separate AC judgements for each of 3 LOs (2, 3, 1), that is, 6 AC judgements in total.
Criteria *	4 separate levels-based mark schemes.	List of 6 AC.
Judgemental process	Holistic best-fit judgements.	Met vs. not met judgements.
Purpose of marks or judgements	To estimate the learner's overall level of attainment in the component.	To confirm the acquisition of all specified LOs for the unit.
Qualification grading scheme	Unclassified, E-A*.	Unclassified, Pass.
Aggregation method	Compensatory (sum total).	Conjunctive (all required).

Table 3. Key features of the BTEC and A level assessment models (BTEC features in rows marked with an * refer to an illustrative unit from a range)

Qualification	Original centre-assessment of piece of work	External quality assurance assessment of piece of work
A level NEA	5+10+15+10 = 40	7+8+18+7 = 40
BTEC Unit 3	met+met+met+met+met+met = pass	met+met+met+NOT+met+met = fail (technically 'not yet met')

Table 4. Hypothetical 'second marking' across CASLO VTQ versus GQ

exactly the same reason, it is unclear how the GQ concept of tolerance – which ignores small differences in assessment judgements between the centre and the Moderator – could be operationalised within CASLO qualifications. In Table 4, despite small differences across the 4 separate judgements, there is actually no difference at all in the overall mark total. But, even if there were a small difference at the overall mark total level, this would still be tolerated under the GQ Moderation process. Tolerance of this sort appears not to be possible under a model whereby every AC judgement counts.

Fifth, it must be appreciated that candidates taking CASLO units are usually only submitted for certification once the centre is satisfied that they have met all AC across all LOs (and have, in effect, passed the unit). As such, any external quality assurance for a pass-fail unit could only ever lead to candidates being downgraded to a fail (failing the unit and therefore the qualification). The pet services BTEC described above requires the successful completion of 6 units, comprising (around) 18 LOs and (around) 39 AC (depending on which optional unit is chosen). If the entire work of a candidate across all 6 units were to be sampled just prior to certification, that would amount to 39 potential opportunities for a single difference in judgement (between centre and quality assurer) to imply that the candidate ought to have failed rather than passed. Any intervention (at all) at this stage would therefore have extremely high stakes. After all, we are imagining a situation in which students have finished their learning and are waiting to progress.

Sixth, even the underpinning rationale for CASS is a little unclear when viewed through a classical GQ lens. GQ Moderation aims to identify centres that have failed to apply standards with sufficient accuracy and consistency, and to put this right before certificates are awarded. The compensatory nature of assessment means that GQ assessment standards are defined at a high level – the level of the total mark that is awarded to a piece of work that is produced for an NEA. In this situation, a candidate-level sampling approach seems sensible, enabling an AO to detect centre-level problems with assessment judgements and then to correct them. If

significant bias[14] were to be detected, then centre mark totals could be scaled accordingly. If substantial unreliability[15] were to be detected, which tends to be less frequent in practice, then the full set of student work could be re-marked. Following intervention of this sort, students would typically end up with marks that were higher or lower, and some would end up with different grades. However, these impacts would not be huge and the process would be entirely manageable.

The situation is different for CASLO qualifications. For pass-fail units, where candidates are not submitted for certification until they are deemed to have passed, the possibility of a harsh centre-assessment judgement does not arise (because they will already have been judged by the centre to have achieved all learning outcomes). The only issue is potential lenience.[16] But, what if lenience were to be detected via the sample? Looking at this question through a classical GQ lens, the only option would be to fail all candidates and require the entire cohort to supply new assessment evidence for the criterion (or criteria) in question to ensure that the new evidence clearly satisfied that criterion (or criteria). Even if this were judged to be the right thing to do, if it were to happen more than just very occasionally, it would render external quality assurance for CASLO qualifications unmanageable.

In fact, it is not at all clear that this would be the right thing to do. To start with, even if evidence of consistent lenience were to be detected via the sample – affecting multiple candidates across a single AC – this would still not provide a reasonable warrant for concluding that **all** candidates at the centre ought to be downgraded to not met (analogous to a downward mark adjustment for GQ NEA). Some candidates might well have met the standard in question by a country mile. Failing all CASLO candidates on the basis of evidence of consistent bias in a sample would surely not be considered acceptable. It cannot be considered analogous to subtracting a mark or 2 from all GQ candidates based on evidence of consistent bias in a Moderation sample. More fundamentally, it would seem morally questionable to fail an entire cohort based on their performance on a single AC for a single LO within a single unit. Especially if this were a terminal assessment, quality assured right at the end of a course. Again, looking at external quality assurance for CASLO qualifications through a classical GQ lens simply does not work.

---

[14] Consistent harshness or leniency.

[15] Inconsistent harshness or leniency.

[16] For CASLO qualifications, consistent lenience might be observed in relation to a specific AC – for example, if multiple candidates were to exhibit work that failed to satisfy a criterion, despite all candidates having been judged to have met it. Alternatively, inconsistent lenience might be observed in relation to different AC – for example, if multiple candidates were to exhibit work that failed to satisfy a specific AC (a different one for each candidate) despite all candidates having been judged to have met those criteria.

The point of this introductory analysis is to set the scene for a deeper investigation into how CASLO qualifications work. It suggests that CASLO qualifications may not differ from GQs simply in terms of the practices they employ, but potentially also in terms of their underpinning principles. Superficially similar processes – such as external standards scrutiny – may need to be conceived quite differently for CASLO qualifications for them to work at all.

VTQ Quality Assurance

We began by noting that the ways in which VTQs differ from GQs are poorly documented, poorly researched, and poorly theorised. Although it is true that the features and processes that characterise VTQs are poorly documented, it is not true that there is no documentation at all. For instance, insights into the operation of internal and external verification processes can be gleaned from a variety of published books and reports, within discussions that typically concern NVQ or GNVQ practices (for example, Raggatt, 1991; Wolf, et al, 1994; FEFC, 1995; Goff & Leimanis, 1995; Beaumont, 1996; Ecclestone, 1996; FEFC, 1997; NCVQ, 1997; Wolf, 1998; Greatorex, 2000; Konrad, 2000; JAB, 2001; Warmington & Wilmut, 2001; Greatorex, 2002; QCA, 2006; Harth & van Rijn, 2010; Harth & Hemker, 2011; Ollin & Tucker, 2010; Curcin, et al, 2014).

One important source of information comes from materials that are produced to support training for Assessor and Quality Assurer qualifications. Admittedly, even these resources are few and far between (Read, 2012). However, they do give us something to focus on. A book by Ann Gravells, entitled 'Principles & Practices of Quality Assurance', is particularly useful (Gravells, 2016). Written for the UK Further Education and Skills sector just as the QCF was in the process of being withdrawn, it would certainly have been written with CASLO VTQs in mind.

What comes across clearly from this book is that quality assurance is about far more than standards scrutiny. Furthermore, although internal standardisation is certainly a feature of GQ NEA, internal quality assurance appears to play a far more fundamental role for VTQs, where the process is formally overseen by an Internal Quality Assurer (IQA). Gravells (p.44) explains how the IQA function is pivotal, with responsibilities that include:

1. sampling assessed learners' work

2. observing trainer and assessor practice

3. talking with learners

4. arranging team meetings

5. arranging standardisation activities

6. preparing for external quality assurance visits

Gravells (p.56) explains the work of an External Quality Assurer (EQA) in a similar fashion, with responsibilities that include:

1. looking at assessment (and related) documents

2. meeting the team at the centre

3. observing assessor and IQA practices

4. sampling learners' work

5. talking to learners and (assessment) witnesses

In short, we do have some insight into how VTQs are quality assured – including CASLO VTQs. This is based upon research reports, published training resources, regulatory, guidance documents, and other materials. Unfortunately, most of these are out of date now.

Following in the wake of 3 related research projects, conducted by research colleagues at Ofqual (see Annex 2), the present study aimed to provide more detailed insights into how CASLO qualifications work. It attempted to unpick the principles that underpin quality assurance practices and the operation of CASLO qualifications more generally.

Research questions

Our primary research question was very simple: How do CASLO qualifications work? However, there are all sorts of regulated CASLO qualifications and we know from previous research that they operate in a variety of different ways. This challenge affected both how we framed our research question and how we attempted to answer it.

To narrow our focus, we decided to use external quality assurance as a window through which to view CASLO qualifications. We originally anticipated paying special attention to the role of the External Quality Assurer and to the nature of the judgements that they make. This would enable us to focus upon how AOs ensure that they are fully accountable for the awards they make, in respect of their centre-assessed units.

We identified 3 sub-questions, to guide our evidence gathering:

1. how do AOs develop and specify standards for CASLO qualifications?

2. what controls do AOs build into CASLO qualifications to ensure the consistent and accurate application of their standards?

3. how do AOs quality assure the application of standards for CASLO qualifications (centres' assessment judgements)?

Our intention was not so much to answer these 3 questions specifically, but to use them to elicit broader, deeper, and more general insights into the operation, and thereby the conceptual underpinnings, of CASLO qualifications.[17] By trying to understand, for instance, the nature of the judgements that EQAs make when they scrutinise samples of assessed work, this would lead us to further questions concerning why judgements of that sort might need to be made. This, in turn, would help us to make sense of the features and processes built into the qualification prior to CASS being undertaken, and the conditions and circumstances that need to be obtained for those features and processes to work effectively.

Earlier, we referred to CASLO qualifications as a broad family with standards that are specified analytically and comprehensively, defined conjunctively, and assessed exhaustively and in a confirmatory fashion. Beyond these core features, we also noted that their features and processes are likely to overlap in a less consistent and less formal manner. To characterise how CASLO qualifications work, we decided that this could best be achieved by using insights from our research and analysis to paint a picture of an archetypal qualification within this family. This hypothetical qualification would be constructed so as to capture commonly occurring features and processes, given the circumstances under which CASLO qualifications typically work. Where the actual features, processes, and circumstances of any particular CASLO qualification might diverge from our construction of the archetypal qualification, this might raise questions for subsequent consideration.

This approach rendered the research project less of a descriptive exercise and more of an interpretive one. It was based upon evidence collected during the research, informed by insights from prior research and pre-existing documentary evidence.

To determine whether our archetypal CASLO qualification represented a plausible characterisation of the CASLO family – as opposed to something more like a Frankenstein's monster of a qualification – we needed to conduct a certain amount of 'member checking'. This is technical way of saying that we needed to check the plausibility of our account with key stakeholders, in particular, with members of the qualifications industry. In the first instance, this meant colleagues from the AOs that contributed to the study.

---

[17] In fact, our interviews provided more information in response to question 3 than to either question 1 or question 2, but responses to all 3 questions helped us to understand how CASLO qualifications work more generally.

Method

We decided to investigate how CASLO qualifications work by studying a small and varied sample in depth. This was originally intended to include 4 main strands of work:

1. desk-based reviews of documentary evidence on each qualification

2. semi-structured face-to-face interviews with AO officers

3. observations of external quality assurance in 2 centres

4. semi-structured follow-up interviews with 2 EQAs

Overview

In our letter to the AOs that were invited to participate in this study, we explained that these strands would be implemented in the following manner:

> We will assign two Ofqual officers to you, and they will share responsibility for gathering all of the evidence. They will engage with you as follows.
>
> First, they will contact you to introduce themselves, and to make sure that they have all of the necessary documentary evidence. This will include documents specific to the qualification (for example, Qualification Specification, Qualification Handbook, Unit Specifications, etc.). It may also include generic documents related to the qualification (for example, Centre-Based Assessment Strategy; Quality Assurance Handbook, etc.). We will aim to complete this stage by end December 2019.
>
> Second, they will visit your offices, to find out more about how the selected qualification works, and about how you design, develop, and deliver QCF-style qualifications more generally. The two Ofqual officers would come at the same time, armed with an interview schedule, aiming to spend no more than a couple of hours in conversation with you. The interview would be audio recorded, for subsequent analysis by another member of the project team. We will aim to complete this stage by end February 2020.
>
> Third, we would like you to nominate two External Quality Assurers who work on the selected qualification. We envisage that each Ofqual officer would shadow one of these EQAs, on a quality assurance visit to a centre of your choice. Again, we would like to be as clear as possible (both for you and for the EQA or centre) that the purpose of this observation would be purely to understand how quality assurance works, not to evaluate the work of the centre nor the EQA. We would not make any audio/visual recording of the visit; but it is likely that we will devise

some kind of written observation schedule, to help us to identify salient events, for subsequent discussion. We will aim to complete this stage by end June 2020.

Fourth, we would like to have a follow-up telephone/Skype interview with each of the EQAs, in order to help us to understand what we observed during the visit, as thoroughly and as accurately as possible. We would use this interview to help us to understand the quality assurance process in detail. We would audio record each interview, for subsequent analysis by another member of the project team. We will aim to complete this stage by end July 2020.

Finally, once we have collated and analysed all of this information, we will invite all of the awarding organisations who participate in the study to a workshop, where we can share what we have found, and explore with you any possible implications, or areas for further research.

Since the abbreviation 'CASLO' was invented specifically for the purpose of this study and participants would therefore have been unaware of it in advance, we explained that we were interested in understanding how 'QCF-style' qualifications work. We highlighted the core features that we were interested in, as described earlier, distinguishing them from less relevant features of QCF qualifications (for example, credit, titling) that we were not specifically interested in.

We anticipated that:

1. documentary evidence would provide us with a general understanding of how each of the sampled qualifications worked

2. interviews with key officers from each AO would provide insights into how AOs understand the operation of each of these qualifications (and would also give us insight into their views on why we can have confidence in their validity)

3. observing EQA visits would provide us with further insights into how QA processes associated with each qualification worked

4. interviews with EQAs would provide us with deeper insights into their own understandings of QA processes – especially standards scrutiny – and their views on why we can have confidence in them

Sample

We recognised that it would only be possible to study a very small number of qualifications in the required amount of depth and detail. Furthermore, our (original) decision to conduct 2 separate EQA visits, for each qualification, effectively halved our sample size. We decided to select 5 qualifications, plus another 'pilot' qualification that we would approach first, in order to test and refine our instruments. In the end, we did not radically refine any of our instruments, so we decided to

combine evidence from the pilot qualification alongside evidence from the broader sample.

Until 2015, when the QCF was withdrawn, most regulated VTQs were part of it and were therefore required to adopt the CASLO model. Some VTQs that were regulated beyond the QCF also followed it too. Nowadays, no qualifications are formally required to follow the CASLO model, but many still do. Some of these qualifications follow the CASLO model for some units, but not all. Most seem to adopt the same model across all units, as the QCF would have required. Although Ofqual does not record qualification model on its database, the high prevalence of CASLO qualifications (mid-2019) meant that it was fairly straightforward to identify a sample.

Qual. Name	Type	Sector	AO Rank	Qual. Certs.	GLH TQT
City & Guilds Level 2 Diploma in Health and Social Care (Adults) for England [501/1306/9]	OQ	Health, Public Services and Care	1	3250	317 (GLH) 460 (TQT)
CMI Level 5 Diploma in Management and Leadership [603/2392/9]	VRQ	Business, Administration and Law	48	2295 (prev. spec.)	175 (GLH) 370 (TQT)
Highfield Level 2 Certificate In Customer Service (RQF) [600/5472/4]	VRQ	Business, Administration and Law	4	8220	115 (GLH) 130 (TQT)
UAL Level 3 Diploma In Creative Media Production & Technology [601/3986/9]	OGQ	Arts, Media and Publishing	23	4670	600 (GLH) 1040 (TQT)
QA Level 3 Award in Emergency Paediatric First Aid (RQF) [603/0786/9]	VRQ	Health, Public Services and Care	7	5090	6 (GLH) 7 (TQT)
EAL Level 2 Certificate In Applying Business-Improvement Techniques [601/1686/9]	VRQ	Engineering and Manufacturing Technologies	14	3555	70 (GLH) 130 (TQT)

Table 5. Details of the sampled qualifications.

We decided that the sample should include as wide a variety of qualifications as possible, from multiple sectors. We also felt, however, that it was important to restrict the sample to prominent, well-established qualifications, which might be relied upon to operate like well-oiled machines. Consequently, we focused upon qualifications that certificated a large number of candidates – more than 2,000 during the 4 quarters up to June 2019 – from AOs ranked in the 'top 50' in terms of annual certifications.[18]

Details of the sampled qualifications are provided in Table 5. These included a minimally vocational Other General Qualification (OGQ) – the only graded qualification in the sample – and a maximally vocational Occupational Qualification (OQ). The remaining 4 were classified as Vocationally Related Qualifications (VRQ). The Diploma In Creative Media Production & Technology was the largest qualification in the sample: 8 units and 600 GLH, designed to be achieved over a year. The Award in Emergency Paediatric First Aid was the smallest: 1 unit and 6 GLH, designed to be achieved during a day.

For brevity, we will refer to each of the qualifications in the sample as follows:

City & Guilds Level 2 Diploma in Health and Social Care (Adults) for England	**C&G-L2-Health**
CMI Level 5 Diploma in Management and Leadership	**CMI-L5-Manage**
Highfield Level 2 Certificate In Customer Service (RQF)	**Highfield-L2-Customer**
UAL Level 3 Diploma In Creative Media Production & Technology	**UAL-L3-Creative**
QA Level 3 Award in Emergency Paediatric First Aid (RQF)	**QA-L3-FirstAid**
EAL Level 2 Certificate In Applying Business-Improvement Techniques	**EAL-L2-Business**

Approach

The project was designed and led by 2 researchers from Ofqual's Strategy, Risk and Research directorate, supported by 8 colleagues from our VTQ directorate. The researchers developed instruments for gathering information, piloted them, and then trained their colleagues in how to use them. The team then split into pairs, with each

---

[18] This information was taken from the dataset underlying the statistical release: 'Vocational and Other Qualifications Quarterly: April to June (quarter 2) 2019' [Ofqual/19/6542/1, 19 September 2019]

pair taking responsibility for a particular qualification, to gather evidence from across all 4 strands of the study. The 2 researchers were responsible for analysing the evidence.

The first 2 strands involved collating documentary evidence and conducting an interview with AO officers. The documents were scrutinised in order to determine as much as possible about each qualification, from publicly available sources, before meeting with AO colleagues. Information from the documents was entered into an Excel workbook, with tabs for background, course, assessment, and quality assurance. This provided a resource for those working on the project, to support the evidence gathering and subsequent analysis.

An interview schedule was developed to explore the research questions identified earlier. This is attached as Annex 3. It was split into 4 sections:

1. how unit standards are established

2. how assessors apply unit standards

3. how quality assurers quality assure

4. the QCF-style approach

A final 'background information' section enabled the interviewers to fill in any remaining gaps in the Excel workbook.

The interviews were audio recorded and these recordings were transcribed for analysis. Interviews with AO officers typically lasted around a couple of hours, involving 2 Ofqual interviewers and 2 or more AO interviewees. It was explained in advance that the write-up would use extracts from the interviews, although the interviewees would not be identified in any reports.

The project was interrupted by the outbreak of COVID-19 during early 2020. Shortly after the final Strand 2 AO interview had been conducted, the country went into lockdown. Subsequent stages of the project were put on hold. Only the first author remained on the project, while his colleagues were deployed to other roles.

During lockdown, the AO interviews were transcribed. They were uploaded to NVIVO and coded for common themes. In the first instance, these largely reflected responses to the various questions that were asked. Further interrogation of the data led to the identification of various principles that appeared to underpin the operation of CASLO qualifications.

Findings from the first and second strands of the study were written up. This provided a wide-ranging, apparently coherent account of the principles that underpin how CASLO qualifications work. Given the continuing disruption owing to COVID-19, we decided to use this account as the basis for member-checking conversations with AOs, rather than undertaking further empirical research with EQAs.

A Microsoft Teams meeting was held, in September 2021, with officers from 5 of the 6 AOs that participated in the study. They fed back on the plausibility and utility of the first draft of the present report – dated 27 November 2020 – and, in particular, on the plausibility and utility of its account of CASLO qualifications.

On the basis of this member-checking exercise, the general plausibility and utility of the report and its account of CASLO qualifications was established. This included support for the idea of constructing an archetypal CASLO qualification and for the way in which it was characterised. A variety of more controversial issues were raised, and discussed, leading to recommendations for minor revisions to the report. These recommendations were taken on board and have been incorporated in the following sections.

Insights

The following sections attempt to paint a picture of:

1. what CASLO qualifications look like
2. how CASLO qualifications work

What do CASLO qualifications look like?

Information from documents relating to the qualifications in our sample provided insights into certain common characteristics of CASLO qualifications, while also providing some indication of how CASLO qualifications may diverge across these dimensions.

Learners

Even from our small sample, it was clear that CASLO qualifications accommodate a wide variety of learners, studying under a wide variety of circumstances. Consistent with its classification as an Other General Qualification[19], UAL-L3-Creative catered primarily for learners studying full-time in a college-based environment. Conversely, consistent with its classification as an Occupational Qualification, C&G-L2-Health was designed for learners who would be expected to hone their skills (and to have them assessed) in the workplace and who were college-based only part-time. Highfield-L2-Customer and EAL-L2-Business catered for learners in a similar situation, often studying the qualification as apprentices. CMI-L5-Manage was designed for practising or aspiring managers across a variety of sectors. It accommodated centres that adopted quite different approaches to delivery, including classroom-based, distance, and weekend learning. Finally, whereas QA-L3-FirstAid was designed to be studied by a variety of different learners, its day-long course duration restricted it to a very specific centre-based delivery approach.

Unit structure and programme flexibility

Four out of the 6 qualifications in the study comprised mandatory units only. The other 2 illustrated how programme flexibility can play an important role in CASLO qualifications. Flexibility enabled these qualifications to be tailored to alternative

---

[19] These are qualifications that assess a particular subject area, such as music or art, other than GCSEs, AS and A levels. These qualifications are not directly work related but may support career development.

occupational pathways, divergent local circumstances, and to the personal requirements of individual learners.

The qualification with the most complex structure was C&G-L2-Health, which learners achieve via 1 of 3 pathways: Generic, Dementia, Adults with Learning Disabilities. All learners must pass the 9 mandatory Group A units, but they are permitted to choose units from both Group B and Group C, bearing in mind certain pathway-specific rules. As explained in Annex 4, the qualification can be achieved via 9 Group A units (from 9), 1 Group B unit (from 36), and 5 Group C units (from 67). This clearly provides a lot of flexibility within Group B and Group C.

The other qualification that promoted programme flexibility was CMI-L5-Manage. It is the third largest of 4 qualifications grouped under the same syllabus – Award, Certificate, Diploma, and Extended Diploma – each of which corresponds to a different qualification size. The structure of this qualification: "has been designed to support Learners to select combinations of units to address own development needs and interests. The units developed for each theme will also enable organisations to tailor the [qualification] to address specific organisational development needs."[20] Its units are arranged within 2 themes: Foundations for Excellence (2 units), and Developing Capabilities, Delivering Results, Driving Best Practice (25 units).

Although there are a small number of proscribed unit combinations for each of the 4 CMI qualifications – given a certain amount of overlap in unit content – learners have more or less free rein over their choice of units. This is as long as they achieve the minimum Total Unit Time specified for the qualification.[21] As illustrated in Annex 4, the Diploma can be achieved by passing any 8 of the 27 units within this syllabus (excluding a small number of proscribed combinations).

Outcomes and criteria

As these were all CASLO qualifications, their LOs and AC were inevitably specified in the same way. This is illustrated below, for the first unit of C&G-L2-Health ('Introduction to communication in health, social care or children's and young people's settings'). Note that the following LOs were prefaced by "The learner will" and the AC were prefaced by "The learner can" (which is typical of CASLO qualifications):

---

[20] CMI Level 5 Management and Leadership (RQF) Syllabus | August 2019 | Version 6.

[21] Each unit has a specified TUT, for example, unit 501, Principles of Management and Leadership in an Organisational Context, which has a TUT of 62.

LO1. Understand why communication is important in the work setting

AC1. identify different reasons why people communicate

AC2. explain how effective communication affects all aspects of own work

AC3. explain why it is important to observe an individual's reactions when communicating with them.

LO2. Be able to meet the communication and language needs, wishes and preferences of individuals

AC1. find out an individual's communication and language needs, wishes and preferences

AC2. demonstrate communication methods that meet an individual's communication needs, wishes and preferences

AC3. show how and when to seek advice about communication

LO3. Be able to reduce barriers to communication

AC1. identify barriers to communication

AC2. demonstrate how to reduce barriers to communication in different ways

AC3. demonstrate ways to check that communication has been understood

AC4. identify sources of information and support or services to enable more effective communication.

LO4. Be able to apply principles and practices relating to confidentiality at work

AC1. explain the term 'confidentiality'

AC2. demonstrate confidentiality in day to day communication, in line with agreed ways of working

AC3. describe situations where information normally considered to be confidential might need to be passed on

AC4. explain how and when to seek advice about confidentiality

There was a great deal of similarity across the qualifications in terms of how LOs were formulated. They were framed almost exclusively in terms of 'understanding…', 'knowing…', or 'being able to…'. The only exception was Highfield-L2-Customer, which incorporated verbs that are more frequently associated with AC, such as 'describe', 'identify', 'apply', 'demonstrate', and 'communicate'.

Verbs like these were common across the AC of the qualifications in the sample, alongside many others, including 'suggest', 'compare', 'use', 'produce', 'show how', 'explain', 'carry out', 'assess', 'evaluate', 'organise', 'discuss', 'develop', and so on.

Qualification	No. units	No. 'Und.'	No. 'Know.'	No. 'Able'	All LOs
C&G - A	9	14	8	18	40
C&G - B	1	0	2	0	2
C&G - C	5	4	0	14	18
CMI	8	16	9	0	25
Highfield	2	1	0	0	9
UAL	8	18	0	5	23
QA	1	1	2	4	7
EAL	7	7	0	0	7
Total No.	**41**	**61**	**21**	**41**	**131**
Percent	-	46.6	16.0	31.3	-

Table 6a. Summary of data in Table A2

Qualification	Mean GLHs per unit	Mean LOs per unit	Mean AC per unit	Mean AC per LO	Mean GLHs per LO
C&G - A	21.2	4.4	13.0	2.8	4.8
C&G - B	14.0	2.0	8.0	4.0	7.0
C&G - C	21.0	3.6	15.6	4.5	6.5
CMI	21.1	3.1	9.1	3.0	7.0
Highfield	57.5	4.5	30.0	6.7	13.1
UAL	13.0	2.9	4.6	1.7	4.4
QA	6.0	7.0	19.0	2.7	0.9
EAL	10.0	1.0	6.4	6.4	10.0

Table 6b. Summary of data in Table A2.

Table A2 (Annex 4) provides an overview of how LO statements were formulated, for each unit within each of the qualifications. Tables 6a and 6b summarise this information. Data are presented separately for C&G-L2-Health, owing to its complex option structure. Table 6a indicates that almost half of the LO statements were formulated as 'understanding' requirements and just under one-third were formulated as 'be able to' requirements. Note that the CMI-L5-Manage LOs were formulated exclusively in terms of 'understanding' and 'knowing', and the EAL-L2-Business LOs were formulated exclusively in terms of 'understanding'. This contrasts with the QA-L3-FirstAid LOs which were formulated largely in terms of 'be able to', which was also true for the C units of C&G-L2-Health.

It is clear from Table 6a that there is no standard 'formula' for determining how many LOs to include within a unit. Thus, QA-L3-FirstAid comprised a single unit with 7 LOs, whereas EAL-L2-Business comprised 7 units, each with a single LO. Having said that, CMI-L5-Manage (8 units, 25 LOs) and UAL-L3-Creative (8 units, 23 LOs) had a similar structure.

Table 6b extends this analysis, revealing that the qualifications in the sample differed substantially in terms of their average number of LOs per unit. They also differed in terms of their average number of AC per unit. Similarly, the typical number of AC per LO ranged from 1.7 for UAL-L3-Creative to 6.7 for Highfield-L2-Customer, while the typical number of GLHs per LO ranged from 0.9 for QA-L3-FirstAid to 13.1 for Highfield-L2-Customer. These are crude statistics, of course, with little intrinsic significance. But they do help to illustrate sizeable differences in approaches to specifying LOs and AC, whatever might be made of those differences.

Assessment

Again, as these were all CASLO qualifications, they were inevitably assessed directly against their LOs and AC. All units within all of the qualifications in the sample were centre-assessed. The single-unit QA-L3-FirstAid qualification also included an externally set (internally marked) multiple choice paper, in addition to observations of performance against criteria.

The qualification documents tended not to be very specific concerning assessment activities, with the implication that centres were expected to take primary responsibility for developing them. For instance, the CMI-L5-Manage syllabus document specified a range of assessment methods that are permitted for each unit, with almost all units permitting 4 or more possible methods. By way of example, a 'method' might involve asking the learner to write an account or assignment on the principles of management and leadership in an organisational context. Alternatively, it might ask them to respond to a scenario looking at a range of organisations and the approaches and styles of leaders and managers within them.

Documents relating to C&G-L2-Health also allowed centres some flexibility in their approach to assessment. Direct observations of performance and indirect judgements based on expert witness testimony were specified as the main sources of evidence for competence-based units. For knowledge-based units, learners had the option either to compile a portfolio of evidence, using a diverse range of assessment methods, or to complete externally set (centre-assessed) assignments.

The Highfield-L2-Customer specification document identified its single assessment method as 'portfolio of evidence'. However, it did not go into detail concerning the nature of the activities through which evidence might be provided, implying that flexibility was permitted. Guidance Notes were provided for unit 2, along the following lines:

> The following assessment criteria are based on candidate performance; however, simulation can be used if workplace evidence is not available.
>
> A scenario can be created to allow learners to achieve all criteria.
>
> Suitable evidence may include:
>
> 1.4 Completed customer record(s)
>
> 3.1 Various methods of communication such as letters, emails, producing posters, etc.
>
> 3.4 Use of a telephone, using the correct greeting, putting the customer on hold, transferring the call and wrapping up the call correctly

Highfield provided centres with an assessment pack for the unit, which included an evidence tracking sheet. An extract from this sheet (taken from the qualification specification document) is reproduced as Figure 1, below, to help explain the broad structure of the assessment process for this qualification. This helpfully illustrates an approach to managing evidence portfolios, via assessment records, which would seem to be fairly common among CASLO qualifications.

The critical elements of this evidence tracking sheet are the columns headed Evidence Type and Evidence Reference. These indicate the nature of the assessment activity – for example, Observation (Obs) – cross-referenced to the location within the portfolio that the relevant evidence can be found. This signposting enables participants (including learners, assessors, and quality assurers) to track the achievement of LOs, AC by AC, and to revisit assessment judgements where necessary.
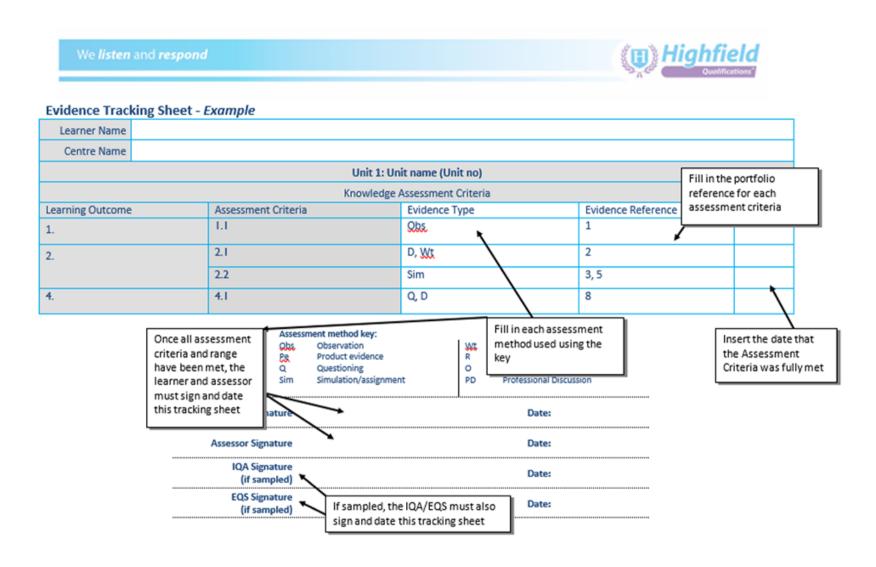
Figure 1. Highfield evidence tracking sheet (reproduced with permission from Highfield Qualifications).

Quality assurance

All 6 AOs had established procedures for both internal and external quality assurance. These were sometimes described in terms of 'Quality Assurance' and other times in terms of 'Verification' or 'Support' or 'Moderation'. The designation of key actors differed accordingly, for example:

- External Quality Assurer – QA, C&G

- Moderator – CMI, UAL

- External Quality Support Officer (or Visiting Officer) – Highfield

- External Verifier – EAL

Some AOs referred to both Verification and Quality Assurance, reflecting a gradual drift in terminology since the introduction of the QCF.

Most of the AOs offered Direct Claims Status (DCS) for the qualification that we studied. CMI did not offer DCS, having piloted it, but judged it not to be suitable for their style of delivery. Nor did UAL.

The archetypal CASLO qualification

All CASLO qualifications have in common the fact that standards for passing CASLO units are specified analytically and comprehensively, defined conjunctively, and assessed exhaustively and in a confirmatory fashion. However, beyond these structural requirements – which are common by definition – we were able to identify additional features, which helped to paint a picture of the archetypal qualification within this broad family.[22]

In England, the archetypal CASLO qualification is a VTQ, designed with flexibility in mind to accommodate learners studying under a variety of circumstances, for instance:

- learners following different pathways within the same qualification

- learners developing alternative (customised) skillsets within the same qualification

- learners studying in different locations (college-based, work-based, and so on)

- learners demonstrating their competence in different workplace settings

---

[22] We painted this picture on the basis of our analysis of the six focal qualifications and broader conversations with AO officers, in the context of the wider literature on VTQs in England and information from Ofqual's Register of Regulated Qualifications.

- learners starting and finishing their learning on different timelines

Flexibility is delivered in a variety of ways, from learners achieving different learning outcomes (from different units) within the same qualification, to learners being assessed on the same learning outcome using different kinds of evidence.

Because standards for CASLO qualifications are specified comprehensively and assessed exhaustively – and because they often emphasise practically oriented outcomes – they carry a heavy assessment burden. The need to accommodate both flexibility and burden is significant, from a qualification design perspective, disposing the archetypal CASLO qualification towards a continuous centre-assessment (or continuous workplace assessment) model. Other common design features include:

- only awarding the passing grade

- completion of the qualification not being heavily time constrained

Identifying distinctive features along these lines helps to distinguish the archetypal CASLO qualification from the archetypal classical qualification. It suggests that any particular qualification might be described in terms of the extent to which it aligns with the archetypal CASLO qualification and the extent to which it aligns with the archetypal classical qualification, rather than assuming that the 2 models are conceptually in tension. This paves the way for discussion of hybrid models too.

Unfortunately, certain features associated with the archetypal CASLO qualification present significant threats to accurate and consistent assessment. For instance:

- a heavy assessment burden increases the potential scale of assessment error

- high flexibility makes it hard to manage the risk of assessment error

- the mastery approach to assessment increases the stakes of assessment error

Given the substantial devolution of responsibility for assessment processes to centres, AOs need to be very confident in the ability of each and every centre to assess accurately and consistently. This helps to explain why quality assurance needs to be conceived quite differently for CASLO qualifications compared to quality assurance of centre-based assessment for GQs.

How do CASLO qualifications work?

Information from our interviews with AO officers revealed sufficient commonalities across qualifications to paint a picture of how an archetypal CASLO qualification might work. We describe this, below, in terms of:

1. the CASLO model

2. principles underpinning the model

3. what these principles mean for CASS

4. tensions that can arise when operating CASLO qualifications

The commonalities that we identified suggested that CASLO qualifications function quite differently from classical GQs, in respect of their centre-assessed components. There are obvious differences, of course, which have already been discussed (including the fact that CASLO qualifications are predominantly centre-assessed and the fact that assessment involves criterion judgements not marks). Beneath these obvious differences, however – and largely in order to accommodate them – CASLO qualifications appear to be based upon a somewhat different model of accountability, involving a somewhat different kind of relationship between the AO and its centres (in respect of their centre-assessed components). Whereas the GQ model of accountability for awards operates more like a subcontracting relationship, the CASLO model of accountability for awards operates more like a partnership between an apprentice and their master.

Model

Under the GQ (hands-off) **subcontracting model of accountability**, an AO devolves responsibility to a centre for what is typically a relatively small proportion of the assessment that underpins the qualification. It will typically specify centre-assessment processes in considerable detail, requiring all centres to follow the same procedures. It may also require all learners to undertake exactly the same task(s).[23] With minimum responsibility and heavy prescription, there seems to be an implicit assumption that all centres ought to be capable of delivering the assessment goods, to a reasonably satisfactory standard, with minimal guidance or supervision. In other words, there seems to an implicit assumption that, right from the outset of their relationship, the subcontractor (centre) will operate fairly independently of the employer (AO).

The GQ AO exercises accountability with respect to centre-assessment in different ways. For instance, when a centre signs up to delivering a qualification, it enters an 'assessment contract' with the AO, via centre approval and heavy procedural prescription. Subsequently, after each delivery round, the AO will 'review the quality of the assessment goods' produced by the centre – the accuracy and consistency of centre-assessment marks – via a separate subcontracting relationship (Moderation). If the AO is not satisfied that the centre has marked with sufficient accuracy and consistency, then this will need to be rectified. This means that sufficient time and resources will need to be set aside for this purpose.

---

[23] In the early days of GCSE coursework, there was much less prescription. This proved to be problematic, and prescription became much heavier.

Under the CASLO (hands-on) **partnership model of accountability**, an AO devolves responsibility to a centre for all (or a large proportion) of the assessment that underpins the qualification. It is far less likely to require all learners to undertake exactly the same task(s) and it may not specify centre-assessment processes in detail either. In contrast to the situation for GQ centre-assessment, maximum responsibility and light prescription necessitates a different kind of relationship with centres. This is to manage increased threats of assessment error.

Under these circumstances, it is likely the CASLO AO will need to provide more support for each centre and build a stronger relationship with them. This involves closer supervision, particularly at the outset. Consequently, for the archetypal CASLO qualification, the AO exercises accountability on an ongoing basis. This is as though the AO had appointed the centre as an apprentice, with the aim of enculturating it into the community of practice for the qualification in question.

Ongoing supervision has the potential to address 2 major challenges. First, it increases the likelihood that errors made by centres when developing and delivering assessment processes will be identified early enough for corrective action to be taken. With increased threats of assessment error, concentrating quality assurance at the end would not be acceptable. This is because it would risk identifying problems too substantial to be addressed straightforwardly within a limited timeframe. Second, it spreads the high burden of quality assurance across the entire duration of a course of learning, which makes the model manageable.

Ongoing supervision does not mean constant supervision, however. New apprentices may need to be monitored closely, but as they demonstrate their integrity and developing expertise, supervisory processes will change accordingly.

These 2 models are simply metaphors for how AOs exercise accountability for their awards, in respect of centre-assessed components. In addition, because they characterise the situation for the archetypal qualification, situations for actual CASLO qualifications may depart in a variety of different ways.[24]

Principles

The archetypal CASLO qualification operates a tripartite centre-assessment quality assurance system, based upon 3 supporting pillars:

1. centre approval

2. internal quality assurance

3. external quality assurance

---

[24] Equally, there may be circumstances in which certain GQs operate somewhat closer to the partnership model than to the subcontracting one.

These 3 functions comprise a holistic system, though which the quality of centre-assessment is ensured.

> […] in terms of the overall quality assurance, it's a package. It goes from start to finish. So it's not really, as far as any qualification is concerned, about one specific facet – it's about the whole thing.[25]

The archetypal CASLO qualification implements a cradle-to-grave approach, which begins once a centre seeks approval to offer the qualification. It does not end until it has stopped offering the qualification and the last student has been certificated.

Centre approval ensures that a centre has the wherewithal to deliver the qualification. This is not simply whether it is equipped to apply the qualification standard accurately and consistently, but whether it has appropriate and adequate systems, processes, and personnel in place to manage, teach, assess, and quality assure the qualification.

Internal quality assurance – led by a centre-appointed Internal Quality Assurer (IQA) – ensures that the centre delivers the qualification effectively, on an ongoing basis. Again, this is not simply whether it applies the qualification standard accurately and consistently, but whether it manages, teaches, and assesses the qualification effectively.

Finally, external quality assurance – led by an AO-appointed External Quality Assurer (EQA) – ensures that the IQA has effectively quality assured qualification delivery within the centre.

The expectation that centres should assume primary and substantial responsibility for quality assurance appears to be an important feature of how CASLO qualifications work. This is different from how GQs work. The more effectively an IQA coordinates quality assurance, the less frequently an EQA will need to intervene. Effective IQAs keep qualification delivery rolling smoothly. Ideally, if an IQA has done their job effectively, then there should be no need for EQA intervention. The IQA is therefore a lynchpin, at the heart of this holistic system.

External standards scrutiny (Moderation) is at the heart of the centre-assessment quality assurance model for traditional GQs. For the archetypal CASLO qualification, it is perhaps better understood as the final safety net. It is premised on the assumption that an IQA will already have coordinated a range of quality assurance activities, including a substantial amount of internal standards scrutiny.

---

[25] From here onwards, the report will reproduce direct quotations from our interviews with AO officers to illustrate key points. As these will be used to paint a picture of the archetypal CASLO qualification, they will be presented without attributing the quotation to a particular AO.

Information from our interviews with AO officers revealed sufficient commonalities across qualifications to propose several principles that appear to underpin the effective operation of CASLO qualifications:

1. structural integrity

2. self-regulation

3. comprehensive monitoring

4. risk-based sampling

5. incremental improvement

6. supportive surveillance

7. conditional (evidence-based) trust

These principles recognise that, compared with classical GQs, the archetypal CASLO qualification is associated with:

*   a far larger proportion of centre-assessment

*   greater flexibility in curriculum, teaching, learning and assessment arrangements

*   a more exacting assessment standard (that is, mastery of all LOs)

These considerations lead to the conclusion that quality assurance, within CASLO qualifications, needs to be **front- and middle-loaded** and **hands-on**. This must be throughout the entire period that a centre delivers a qualification. Conversely, a more hands-off, more end-loaded, approach (along the lines of the traditional GQ model) risks problems occurring too significant to be effectively addressed towards the end of a course of instruction.

1. Structural integrity

Even before a centre registers its first group of learners, an AO needs to ensure the centre has the wherewithal to deliver the qualification. In other words, it needs to establish the structural integrity of the centre, from a teaching, assessment, quality assurance, and management perspective.

> I think for me it actually starts at that centre approval, qualification approval, stage. I think making sure that centres have the resources, have the right people in place, have the kind of internal quality assurance strategies and policies and processes.

> You've got 3 elements. You've got the occupational competence […], you've got teaching qualification […], and then you've got an assessing qualification […]

Formal qualifications play a more significant role in establishing the credentials of a centre with respect to CASLO qualifications (and VTQs more generally) than with respect to traditional GQs. Not only is there typically an expectation the teacher or trainer has a teaching or training qualification, there is generally an expectation that any assessor will have (or will be working towards) an assessor qualification and an IQA will have (or will be working towards) an IQA qualification. Centre staff will also be expected to have competence and up-to-date experience in the domain they are teaching and assessing – particularly as a warrant for assuming that their basic understanding of the qualification standard is satisfactory.

AOs will also want to be confident that centre staff are familiar with the details of the qualification itself, including LOs, AC, and its broader learning domain – and that they remain so over time. This is why training, standardisation, and an expectation of Continuing Professional Development (CPD) are also important features of quality assurance arrangements for the archetypal CASLO qualification.

> So, [as part of an 'Engage Day'], we do a standardisation activity with them, and we tend to pick ones that may have been challenging for centres and we'll go through that together. So, it sets the scene again to say 'OK what is a good 'evaluate'?', or 'what is a good 'decide'?' […] we've done quizzes around command verbs, to say 'OK, these are the words, how do you match them to what they'll be expected to do in practice?'

> And the other thing I was going to add in is the requirement for CPD. So, assessors and EQAs and IQAs, there is an expectation that there is ongoing professional development happening.

Finally, the AO will want to know that the centre has put in place appropriate systems, processes, and tools for managing the delivery of the qualification – items such as reporting lines, meeting schedules, and activity checklists. To facilitate this, the AO will need to establish clear expectations and communicate them clearly via centre handbooks and related documents.

Although AOs will go to considerable lengths to ensure structural integrity at the centre from the outset, it is just as important that they continue to monitor its structural integrity throughout. It is therefore critical for the AO to have an open channel of communication with the centre. This is so the AO is fully appraised of any changes that might impact upon the effective delivery of the qualification – staff turnover, in particular.

2. Self-regulation

It is a principle of the partnership model of accountability that centres that wish to offer an archetypal CASLO qualification should operate as self-regulating systems. Just as apprentices of old would come to internalise their master's quality standards and quality assurance practices, so too should a centre internalise those of the AO.

> The key thing here, I think, is that internal quality assurance process before [the external quality assurance process]. You know, at that centre level, the internal people looking at those assessments and making that judgement, making sure that they're balanced and fair judgements and are reliable. And, then, we're just going in and checking that. The key thing is about that internal part.

> An IQA typically would […] assess all of their assessors and they would develop some kind of strategy of – we call it RAG rating – so a red-amber-green structure, risk rate your assessors. Who's competent, who's new, who's been shown to be not so competent? And then put in place a sampling strategy frequency based on that. For this qualification with it being 2 units, […] they might just decide to do both units, or they might do one unit. And, from that, they may then identify that there are commonalities across one assessor or a group of assessors [ie, issues to address] and they would use that intel to inform their standardisation activity.

Although accountability ultimately resides with the AO, the partnership approach devolves a considerable amount of responsibility to the centre to deliver the qualification effectively and to the national standard. In addition to high expectations of the rigour and quality of centre-based assessment, there are also high expectations of the rigour and quality of centre-based quality assurance.

3. Comprehensive monitoring

Quality assurance, for the archetypal CASLO qualification, concerns the effective delivery of the qualification from all perspectives – teaching or training, assessment, internal quality assurance, and management. In other words, quality assurance is not focused purely upon the accuracy and consistency of final results – it is far more comprehensive than this. This is true of both internal and external quality assurance functions.[26]

> So, we'll ask them key questions around the assessments, the assessment methodologies, the qualification as a whole, to make sure that it is, it remains fit

---

[26] In the following quotations, recall that some VTQ AOs use the term 'moderation' in a generic sense to describe external quality assurance, with different connotations from GQ Moderation.

for purpose… things haven't changed within industry to make sure that it can still be applied in that way.

And that's why checking the assignment brief is really essential. I always say to my moderators that that's your starting point, because if the students are being told to do the wrong thing […]

And, worth noting at moderation, if we find an issue with a particular area – particular learning outcome – we will then go back to the formative unit where that area should have been taught and assessed to see if there's an issue with how it's been taught in the year.

Comprehensive monitoring enables the IQA and centre, and in turn the EQA and AO, to identify and address any of a variety of issues. This includes those that threaten a learner's ability to reach competence, or to demonstrate their competence in the assessment, or threaten the centre's ability to recognise that competence. Early intervention helps to prevent nasty surprises at the end of the course, which might be impossible to rectify at that stage. The key point is that CASLO qualifications need to be understood as integrated teaching-learning-assessment systems. Quality assurance (both internal and external) needs to bring that entire integrated system into view.

4. Risk-based sampling

Because the archetypal CASLO qualification comprises multiple centre-assessed units, it would be hard, or impossible, to scrutinise assessed work for all units during any particular standards scrutiny cycle (either internally or externally). This contrasts with the situation for traditional GQs, most of which have just a single centre-assessed component. In the GQ case, it is reasonable to expect assessed work for its centre-assessed component to be scrutinised each cycle. Where a CASLO qualification has many centre-assessed components, these may need to be sampled for the purpose of standards scrutiny during any particular cycle.[27]

Because both internal and external quality assurance activities are based upon a principle of comprehensive monitoring, these functions stretch far beyond the scrutiny of centre-assessed work. This means that the burden of quality assurance is

---

[27] Consider, for example, the City & Guilds Level 2 Diploma in Health and Social Care, which has 9 mandatory Group A units, plus a choice of 1 Group B unit (from 36), and 5 Group C units (from 67). Particularly, where candidates within a centre have chosen different optional units, this represents a significant standards scrutiny challenge.

far greater than for traditional GQs. As a result, the entire quality assurance model will need to be premised upon sampling, for example, across:

- units (for teaching or training, learning, and assessment)

- teachers or trainers (assessors)

- centre sites (and potentially also candidates' work-based assessment sites)

- programmes (for nested qualification programmes including Award, Certificate, and Diploma)

- learner cohorts (where adjacent cohorts are taught in cycles, or where they are co-taught)

- activities (assessments, meetings, training, standardisations, and so on)

Where only a limited amount of sampling can take place, during any particular quality assurance cycle, it makes sense to adopt a targeted sampling approach, rather than a random (or stratified) one. This leads to the principle of risk-based sampling.

> For me then it is that ongoing monitoring of activity, based on the risk of the centres. We will look at the sorts of things that we know, you know, identify a risk. It can be volumes of learners, can be complexity of the delivery model. You've got a centre that has lots of assessment sites and locations and satellite sites.
>
> High turnover of staff. […]
>
> Yeah people, you know, new centres that are completely new to delivering this style of qualification in this sector. You know, potentially stuff from other complaints, malpractice investigations.

The risk-based sampling principle helps the IQA and centre, and in turn the EQA and AO, to determine how much monitoring needs to occur and how frequently, as well as where that monitoring ought to be focused.

When undertaking standards scrutiny for sampled units, there are further sampling decisions to take.

> So they would be looking at the decisions that an assessor has made for a sample of candidates. And that sample, well it depends on lots of things but they'd be doing CAMERA, a range of candidates, a range of assessors and a range of different methods, locations, looking at records […]

The 'CAMERA' framework is used by City & Guilds and other AOs, to ensure that assessed work for a unit is sampled robustly (see City & Guilds, 2016). It stands for:

- Candidates (Learners)

- Assessors

- Methods of assessment

- Evidence

- Records

- Assessment sites

5. Incremental improvement

Linked to the principles of comprehensive monitoring and risk-based sampling is the principle of continuous, incremental improvement. It is possible to think of the GQ subcontracting model of accountability as a one-off event, albeit one that is repeated each year. The CASLO partnership model, however, is premised upon there being an ongoing relationship between the AO and its centres. At the heart of this model is the idea of incremental improvement, right from the moment a centre begins delivering a qualification, and then year-in-year-out, for as long as the centre continues to offer it.

> You know, some of the advantages of verification over moderation to me are you're in there earlier. I think with moderation everything could go wrong and you're at the end of the 2 year programme and the centre's not got the hang of this and it's all a bit of a nightmare.

> Yeah, and it's not the cheapest way to run a qualification, but we feel that it produces the right results. Because we get good relationships with the centre, and we get them to understand the standard, because we spend quite a lot of time, we invest in them quite a lot in terms of time. So, it's good to, all those visits, all those events we talked about, they're all face to face, and that's where we're able to talk through in some detail the standard and therefore maintain it.

The incremental improvement principle highlights the idea of an AO investing in a centre and developing that centre over an extended period of time. It suggests that external quality assurance is as much about improving assessment practices for future learners as it is about checking assessment outcomes for current ones.

6. Supportive surveillance

If follows from preceding principles that the relationship between an AO and its centres, for the archetypal CASLO qualification, is fundamentally supportive. Only in very exceptional circumstances will an approved centre be deemed no longer worthy of support and be prevented from offering the qualification. In all other circumstances, the modus operandi will be one of support, to empower the centre to

deliver the qualification effectively. This applies to the role of both the EQA and the IQA.

> I think an IQA's support of the assessor is crucial and there's no 2 ways around that. […] I don't want to minimise the importance of the other [quality assurance] stuff when I say that support of the assessor is essential – but it is.

> This is something that we employ across the board because we are about supporting people to being the best that they can be. And then, in turn, it just, it spirals down to the learner. The learner then becomes the best that they can be.

> For me, I think, it's that interactivity in the support mechanism – having somebody there who is a point of contact. […] I think having that named EQA – you've got a face and a name of somebody, if you've got a problem, who's approachable. That's the key within the model that we run.

However, it is important to stress that support is provided with a strict expectation it must be acted upon and those actions remedy identified shortcomings.

> I think, as well, a good role model of an IQA is, if they're going to give some action to an assessor, it's that they supportively follow that through with them. Because they could turn round and say 'don't do that'. But if they don't say 'don't do that because of X, Y and Z' and 'the consequences could be A, B and C' and 'here is a better way to go about it'. So, we would be looking for that supportive action plan. We want to see that they're action planning people, but following it up and that the actions have been achieved.

It is also important to emphasise that both IQAs and EQAs carry out functions that are simultaneously formative and summative, that is, both supportive and judgemental. Ultimately, their role is to ensure that the qualification is delivered effectively and that candidate work is assessed with consistent accuracy. This is the bottom line.

> [The role] is one of support and guidance, but equally is it also one of pull your socks up otherwise there might be some repercussions.

7. Conditional (evidence-based) trust

This brings us to the final principle. Working in partnership with a centre requires a certain amount of trust. This is because a large amount of responsibility for assessment and quality assurance processes is devolved to the centre. For many CASLO qualifications, this is formally recognised by granting Direct Claims Status to the centre.

So, we don't grant DCS initially. We need assurances that centres are completing and following the assessment procedure correctly. To do that, they need to have 3 perfect audits – and we're talking absolutely everything is spot on. And that's not just in terms of the assessment paperwork, but also that internal quality assurance is taking place as well, and is to the standard required. Then, once we're satisfied, then DCS is likely to be granted. And then it falls into verification. We've got various different levels of verification. So, we 'centre risk-rate' centres – low, medium, and high. So, the high-risk centres get a physical visit, the medium ones get a remote Skype-type visit-call, and then the low ones get some more moderation activity associated with them as well.

It is important that this trust is evidence-based, earned, rather than granted automatically. Equally important, centres must be aware that penalties will follow from any abuse of trust, whether that be unintentional maladministration or intentional malpractice. The most extreme penalty is that an approved centre will be prevented from offering the qualification. However, to apply the conditional trust principle most effectively requires a sliding scale of penalties. This is to reflect the severity of the offence in question. This might mean raising the centre's risk status, the removal of DCS, or more extreme measures.

If the centre had DCS and [the EQA] uncovered something in their practice that they weren't happy with, they would remove DCS, so they could continue to support and sample the work before [certificates were] claimed.

And if they're not [following the standard] then it depends to what degree. So, if it's something that's affecting compliance with our guidance, then they would get an action plan that would be monitored. And that would involve some remedial work – maybe some standardisation training – and then that will be followed up to make sure that they have understood that information, in order for them then to go back out and assess to our standard. It's about supporting them, the best that we can, to make the right judgments.

Also, within our quality assurance models – based on risk – so if we come across the situation where our EQAs are not agreeing with assessment decisions, or IQA activity maybe not happening, or whatever, then that centre will be put on high-risk status and given additional support […] or additional quality assurance activities.

Interviewer: So what would happen if they had serious concerns about the marking of a particular assessor in a centre? […]

The main thing would be the risk would be identified. That gets reported back to us. We would raise the status of that particular centre and then we would look at which additional quality assurance measures needed to be put in place. And I'm talking about high-risk stuff, not small-risk stuff. And that centre wouldn't be able to register any additional learners, they wouldn't be able to certificate any learners on programme until such time as we'd worked through a kind of programme of additional quality assurance activities. That would move their status back down.

It is worth noting that Appendix 3 of the 'NVQ Code of Practice' (QCA, 2006) formally specified sanctions for non-compliance with centre approval criteria. Although this document is no longer in force, it will have influenced practices that continue to the present day for many regulated CASLO qualifications. For information, these sanctions and their rationales are presented below, in Table 7.

Tariff-Level of transgression	Sanction	Rationale
1	Entry in action plan	Non-compliance with centre approval criteria but no threat to the integrity of assessment decisions
2	Removal of direct claims status, that is, claims for certification must be authorised by the external verifier	Close scrutiny of the integrity of assessment decisions required
3	(a) Suspension of registration (b) Suspension of certification	(a) Threat to candidates (b) Loss of the integrity of assessment decisions – risk of invalid claims for certification
4	Withdrawal of centre approval of specific NVQs	Irretrievable breakdown in management and quality assurance of specific NVQs
5	Withdrawal of centre approval for all NVQs	Irretrievable breakdown in management and quality assurance of all NVQs run by the centre

Table 7. Levels of transgression (from QCA, 2006, Annex 3, Table 1)

CASS

In the Introduction section we showed it is very hard to make sense of Centre Assessment Standards Scrutiny (CASS) for CASLO qualifications through a classical GQ lens. We concluded that external standards scrutiny for CASLO qualifications would need to be conceived quite differently. The following sections provide an alternative account.

Active expectation of passing

A critical feature that sets the archetypal CASLO qualification apart from the classical GQ is the active expectation that a learner ought to pass it. This assumes they have the necessary credentials to be enrolled for the qualification, that they exert an appropriate amount of effort throughout the course, and that they supply the necessary assessment evidence along the way. In England, it is an empirical fact that most learners pass most GQs, too. Yet, the context is different for the archetypal CASLO qualification. Here the expectation of passing is more active and Pass is the highest available grade (and therefore the pinnacle of achievement) rather than the lowest grade available.[28]

In certain situations, particularly for secondary school students, there may be drawbacks associated with only awarding the passing grade. Most obviously, there is no incentive built into the qualification for learners to significantly exceed the threshold standard. However, there are potential advantages too. Targeting all teaching or training, learning, and assessment upon the same competence threshold provides a clear focus for managing all learners through to completion. For the archetypal (pass-fail) CASLO qualification, all learners need to jump all hurdles. As the course proceeds there is no doubt as to whether or not any particular learner is continuing to remain on track for success.

In addition, the archetypal CASLO learner is likely to be fairly motivated to acquire the requisite knowledge and skills assessed by the qualification. This is especially true of post-secondary learners who are studying a qualification to secure a particular job, or who may already be doing the job in question.

> You're being assessed on the job that you do and therefore there's a very good chance – if you're doing your job – that you can get through.

Learners who are unlikely to reach the competence threshold are likely to be counselled off the course long before any formal certification is sought. This helps to

---

[28] In graded systems, the passing grade is generally the lowest available grade. Thus, E is the passing grade at A level, and 1 is the passing grade at GCSE.

explain why it can be tricky for CASLO qualification AOs to judge what proportion of learners 'fail' their qualifications.

> So, I don't know that we necessarily can say that everybody that starts passes, or 80%, or 90%, or whatever, because we actually don't really have that information. But I think there were some good reasons why. You know, because it's about doing your job, because it's about continuous reflection and feedback. Because, if you've not met the requirements of a particular unit learning outcome, you can go away, you can re-learn, re-take without lots of re-sits and charges and burden and all of that, isn't there? So, I think there's some features in there that naturally lead you to that point, of most people succeed.

This quotation also indicates that the archetypal CASLO qualification includes a delivery mechanism that is specifically designed to ensure that learners will pass it. This is often described as formative assessment.

Formative assessment

Although formative assessment is not unique to qualifications that operate the mastery principle, it is fundamentally important to any such qualification. This is to ensure learners remain continually on target to pass. In other words, mastery teaching is critical to mastery assessment and formative feedback is key to mastery teaching.

Some AOs encourage or require centres to administer 'a formative assessment' before learners attempt their summative assessment (a bit like a mock exam). This enables them to judge their readiness, as well as to identify any outstanding learning needs. Many courses are not structured around formative assessment events, however, despite still having formative assessment at their heart. In such cases, the emphasis will be upon providing tailored feedback to students on how to improve. It is this feedback (rather than the assessment event itself) that is described as formative.

Again, the central issue is that a mastery approach to assessment requires a mastery approach to teaching or training. If all learners are to master all learning outcomes, then the teacher or trainer must manage the instructional process accordingly.

> They should be using formative assessment to gauge where the learner is within their journey and the formative assessment should then inform the summative assessment. That's not always 100% followed through. You can have learners that flunk even though they've done really well in formative. But it gives the assessors that indication that that learner is ready for a summative assessment.

It's quite an organic process, assessment. So, an assessor will sit with a learner and they will talk to them around their experience, or they will observe them doing something around certain assessment criteria quite organically, quite naturally in their working environment. And then, if there are additional pieces of evidence and criteria that need to be met, or their knowledge isn't quite there in certain areas, then the assessor is there to develop and support that learner and possibly send them tasks to go away and do, so that they can come back at a later date and be assessed.

I don't think of it as summative and formative assessment, I think of it as feedback, I think of it as support. So, it's not 'you're doing a poor job, I can't put you through'. It's 'I will work with you for as long as it takes to get you there'.

The CASLO model is therefore best suited to a continuous assessment approach, where outcomes are assessed regularly. Opportunities can be provided for learners to commit more time and effort to any particular learning outcome that they might initially struggle with.

Continuous quality assurance

The CASLO model is similarly best suited to a continuous quality assurance approach. Teachers or trainers (assessors) and learners are monitored regularly, throughout the period of delivering the qualification. This contrasts with the traditional GQ approach. In this case, quality assurance is loaded towards the end of the course, after all of the learning is completed.

[The IQA role is] about making sure that the assessment judgements are correct. That's the biggest thing – making sure that those assessments are valid, authentic, efficient and all that stuff. And, you know, the paperwork's been followed correctly. And where there are gaps, identifying those gaps, you know, and supporting the assessors to go back out and fill those gaps in.

Assuming that the IQA has done their job properly, the principle of self-regulation suggests that intervention by the EQA should be minimal. However, intervention may sometimes be necessary, for example, if the IQA had failed to spot that multiple learners had insufficient evidence of competence in relation to a particular criterion on a particular unit. Where intervention is required, it is better that this is identified early on, during the course of learning, enabling the problem to be rectified straightforwardly straightaway. If it were only identified at the end of the course, it would be much harder, if not impossible, to rectify.

By way of example, under a continuous quality assurance approach, an AO might arrange an EQA visit at the outset of term 3, of a 3-term (3-unit) course. For students who were to graduate that year, centre-assessed work would only be available, for

the purpose of external standards scrutiny, for 2 of the 3 units. If the centre had been granted DCS, then it would subsequently be able to claim certification for the 3rd unit and for the qualification overall, prior to the next EQA visit This is without the 3rd unit having been part of any external standards scrutiny sampling exercise for those students.

On the one hand, the risk associated with this approach is that problems on that 3rd unit might only be identified on a subsequent EQA visit. This is long after those students had been certificated. On the other hand, if all the external standards scrutiny were to have been end-loaded (if all 3 units had been scrutinised right at the end of the course, just days or weeks prior to certificates being awarded) then this might have led to a quite different set of problems. These relate to the unmanageability of rectifying problems related to such a lot of assessment evidence at such a late stage. This would be on top of needing to undertake external standards scrutiny for all centres at the same time each year, in a context where competent EQAs are a very limited resource.

Note that, in this example, external standards scrutiny would still be undertaken for all 3 units, during each EQA visit. It is simply that work scrutinised for Unit 3 would come from already-certificated candidates, from the previous cohort. As such, each year, the centre would receive feedback on their assessment of Unit 3 for the previous cohort at the same time as beginning to deliver Unit 3 to the present cohort. Clearly, this would help to prevent problems arising for the present cohort.[29]

The idea of continuous quality assurance is broader than simply the EQA visit, however, which might (as in the above example) be restricted to just once a year. It implies that an AO ought to be in regular communication with its centres and therefore on-hand to deal with issues whenever they might arise.

Effective sporadic intervention

The general expectation of passing, facilitated by a formative assessment mechanism and monitored via an internal quality assurance function, provides a reasonable warrant for assuming that most candidates who are put forward for certification are likely to have met the qualification standard. If so, then it becomes easier to see why the GQ approach to Moderation – which mainly involves fine-tuning centre-assessment judgements for general lenience or general harshness – is neither entirely necessary nor entirely appropriate for CASLO qualifications.

If the CASLO quality assurance system is functioning effectively, then IQA standards scrutiny should already have dealt with issues of fine-tuning of centre-assessment

---

[29] By varying exactly when, during the year, the EQA visit is scheduled, an AO can avoid the situation in which a particular unit (or set of units) is only ever scrutinised for already-certificated students.

judgements – at least, to within a degree of tolerance comparable to that exercised during GQ Moderation. This leaves EQA standards scrutiny (CASS) to focus more upon the possibility of glaring errors, anomalies, and inaccuracies. This might be due to unintentional maladministration, intentional malpractice, or the occasional serious error of judgement that does happen to slip through the internal quality assurance net.

Again, if the CASLO quality assurance system is functioning effectively, then these glaring errors, anomalies, and inaccuracies ought to be few and far between. This makes it more likely that they can be dealt with adequately when they do occur.

> On the other hand, they are undertaking a sample. So, they are selecting specific elements. Sometimes it could be a full portfolio, other times it will just be units. So it's, to me, it's a wider process. It's not a re-mark as such. It's a sample of pieces of work that they tend to… rather than sort of sitting there and sort of working… they will work through with an assessor. So, they will ask the assessor or the learner to explain how that evidence has been, how the piece of evidence has been generated. And how, particularly with the assessor, they feel it maps back to that assessment criteria.

> **Interviewer: […] there's maybe something you disagree with on this assessment decision, is there an element of trying to correct that, as an EQA?**

> I mean, there can be, but that would be in the minority cases. That's when you'd have expected that to have already been addressed through the IQA process. […] So, yeah, you wouldn't be expecting the EQA to have to identify an issue within individual learners' portfolios.

In relation to unintentional maladministration, an EQA engaged in CASS would check evidence had been provided for all the AC that the candidate needed to meet. If evidence were to be missing, then the centre would be expected to provide it. This might require an additional assessment activity. In relation to intentional malpractice, an EQA would be checking for evidence of plagiarism. They would also check that centre-assessment judgements were not consistently lenient, in a manner that hinted at intentional malpractice rather than unintentional maladministration.

> So, for example, if we've got 10 learners' worth of portfolios to sample, we may choose – depending on what intel we have on the centre, and what samples we've done previously, and what our findings are – to sample, say: unit one for one learner, unit 2 for the second learner, unit 3 for the other. And then go back and say 'right, so, unit one for the fourth', 'unit 2…', and so on. So, we're not second marking, but we will scrutinise the units that we see, because we're checking for authenticity, so whether there's plagiarism, […] If we find anything

we will advise on it, but ultimately we're gaining an overarching idea, if you will, of the competence of the assessor and the effectiveness of their assessment decisions across the sample that we've chosen.

[…] we've got examples of where centres are giving learners a helping hand, if you like, through assessment processes. Specific to this qualification, I don't think that's the case, but… we do get that on the odd occasion.

Based on insights like these, we propose that the concept of **tolerance** does have an important role to play when externally quality assuring CASLO qualifications. However, it is not explicit, as it is for marking tolerances within GQ Moderation. Instead, it will be exercised in a succession of implicit judgements – drawing upon the EQA's professional experience and expertise. An EQA would only query a 'met' judgement within a sample of assessed work if they disagreed it was a reasonable exercise of professional judgement. Hence, in addition to being on the look-out for critical anomalies or inaccuracies due to unintentional maladministration or intentional malpractice, EQAs would also be on the look-out for glaring errors of professional judgement.

To illustrate this point, if an EQA instinctively felt the evidence base for a candidate on a particular AC did not merit a 'met' judgement, but could still appreciate how others might see this differently, then the EQA would not take issue with it. Intervention, where it occurs, is likely to be quite **sporadic**, being restricted to cases that clearly represent an unreasonable exercise of professional judgement by an assessor.[30] This seems quite different from the fundamentally interventionist approach that characterises external standards scrutiny for traditional GQs, where intervention – that is, scaling marks up or down, or re-marking – is routine.

Tensions

Qualifications based upon different models will have different strengths and will face different challenges. The foregoing account, based upon the idea of an archetypal CASLO qualification, identified several principles that appear to underpin their effective operation. These principles lead to the establishment of mechanisms that carefully scaffold qualification delivery, empowering centres to be able to overcome challenges such as having to:

- assess most, if not all, components of the qualification

---

[30] This, incidentally, is consistent with how Ofqual has clarified the role of review markers, for GQs (see Ofqual, 2016),

- exercise considerable flexibility in curriculum, teaching, learning and assessment arrangements

- manage an exacting assessment standard (that is, mastery)

In other words, these scaffolding mechanisms enable centres to assume greater responsibility than might otherwise be reasonable to devolve. This is the core strength of the CASLO model. In turn, it is the effective operation of these scaffolding mechanisms that enables AOs to remain fully accountable for the awards that they make.

Of course, this system needs to be operating effectively, for an AO to be exercising accountability appropriately. This can be challenging, owing to several tensions that become salient under the CASLO model. The following 4 examples illustrate, without necessarily exhausting, these tensions.

Sampling

The sampling challenge is considerably larger for CASLO qualifications than for traditional GQs. This is particularly because CASLO qualifications tend to be 100% centre-assessed, meaning there may be many more centre-assessed components to scrutinise. Under these circumstances, especially given the mastery nature of these qualifications, it would be too risky to leave all of the quality assurance until the end of a course of instruction.

These requirements give rise to a quality assurance system that not only scrutinises centre-assessment outcomes, but that also scrutinises teaching, learning, assessment, and internal quality assurance processes too. This helps to ensure that centre-assessment outcomes are as accurate and consistent as possible before external standards scrutiny comes into play. This further reduces the risks that arise when this does happen to occur towards the end of a course of instruction.

However, this requirement to scrutinise both processes and outcomes renders the sampling challenge even greater. This is why the archetypal CASLO qualification will adopt a risk-based approach to sampling, rather than (say) a stratified sampling approach. The risk-based approach is critically dependent upon the AO having sufficient intelligence upon which to establish risk, across a variety of dimensions. Up-front centre approval, combined with ongoing scrutiny of teaching, learning, and assessment processes, will help to provide this kind of information. But there is always a risk that the AO may not have sufficient information to target the risk-based sampling effectively and may therefore miss critical process failures or outcome inaccuracies. Of course, even when the risk-based sampling approach is working effectively, there is still the possibility of missing critical failures or inaccuracies, were they to occur in low-risk areas. So, there is an inherent tension within this approach, between the need to focus limited resources upon high-risk areas and the possibility

that problems might still occur within low-risk areas. Challenges of this sort are not unique to CASLO qualifications, of course, and are a feature of risk-based approaches more generally.

As already noted, effective intelligence gathering and management is critical to the effective operation of a risk-based sampling approach, in order to distinguish between higher and lower areas of risk. However, it is equally important not to assume that low-risk areas can be ignored entirely and excluded from the sampling framework. This may be particularly challenging for qualifications with large numbers of optional units, each attracting small cohorts of learners.

Supporting

The idea of supportive surveillance captures a major tension within the CASLO model between the need to support and the need to judge. In order for the supportive role to work, those who are being supported need to feel sufficiently comfortable in their relationship to expose their weaknesses, as though revealing them to a critical friend.

This can be a challenging relationship for an IQA to establish with an assessor within their own centre, but it can be even more challenging for an independent EQA to establish with a centre. This is because the EQA has a more formal and final role to play in judging the centre. Sometimes this involves identifying process failures and outcome inaccuracies and adjusting the centre's risk rating accordingly. On the one hand, an EQA may become too friendly with a centre and fail to be sufficiently critical. On the other hand, an EQA may become too removed from a centre and fail to be sufficiently supportive. Achieving the right balance is no mean feat. An AO might try to facilitate this by ensuring that EQAs are not appointed to centres from their own region. Or it might re-appoint EQAs to different centres every few years.

Trusting

The idea of centres operating as self-regulating systems indicates how a supporting relationship ideally leads to a trusting one, under the CASLO model. For many CASLO qualifications, Direct Claims Status represents a formal recognition of this trust, as centres are empowered to claim certificates for candidates whose work (or some part thereof) has only been internally quality assured. This creates a tension between the need for an AO to trust the integrity of its centres and the risk certain centres abuse that trust, with more or less serious consequences. Sometimes an abuse of trust may arise from an unintentional lapse, which grows in significance over time. For example, when a centre loses an assessor or, worse still, an IQA and fails to inform their AO. Other times an abuse of trust may reflect straightforward

malpractice – such as claiming certificates for candidates who fail to provide any assessment evidence at all.

This tension means that the relationship between AO and centre can never be entirely based upon trust. As indicated by the idea of conditional trust, it must always be clear to a centre that any abuse of trust – however small it might be – will result in a sanction of one sort or another. Hence, the idea that levels of transgression should be associated with levels of sanction, as set out in the NVQ code of practice.

This tension is another reason why lower-frequency sampling, contingent upon a lower-risk rating, should not be allowed to degenerate into no sampling at all. Thus, even under DCS, there would need to be a comprehensive monitoring system in place. Even under DCS, the assessed work of candidates certificated prior to external standards scrutiny should still be available for scrutiny subsequently, albeit as part of a post-certification sample. More generally, even under DCS, there should be a channel of communication between an AO and its centres which is constantly open.

Investing

Finally, the CASLO model leads to a practical tension between the need to adopt a very hands-on approach and the need to offer a competitively priced service. The scaffolding mechanisms that ensure the effective operation of CASLO qualifications – centre approval, training, standardisation, internal and external scrutiny of both processes and outcomes – comprise a resource-intensive system. This can be challenging because of the financial costs associated with running CASLO qualifications effectively and being able to appoint suitably qualified and experienced personnel. The obvious risk is that shortcuts are taken, which compromise the validity of the qualification. This kind of tension is particularly important for a regulator to be aware of, where there is a risk that cost-cutting in one part of the industry might gradually become endemic.

Discussion

As far as we are aware, this is the first time anyone has attempted to articulate principles underpinning the effective operation of CASLO qualifications, from a quality assurance perspective. This is still a tentative account, based upon a small sample of disparate qualifications and AOs. The fact that it is possible to paint a picture of an 'archetypal CASLO qualification' based upon a small sample helps to support the plausibility of this account. This was also confirmed by officers from the AOs in our sample.

The following sections summarise our insights and contextualise them in relation to several outstanding questions concerning the operation of CASLO qualifications.

Summary of insights

Common to the family of qualifications at the heart of this report is the fact that they are designed to Confirm the Acquisition of Specified Learning Outcomes. We have therefore dubbed them 'CASLO' qualifications. Standards for passing CASLO units are:

1. specified analytically (element by element) and comprehensively (to capture all relevant elements)

2. defined conjunctively (requiring the acquisition of all elements, that is, complete mastery)

3. assessed exhaustively (to cover all elements) and in a confirmatory fashion ('acquired' as opposed to 'not-yet-acquired')

CASLO standards relate to domains of learning that, for VTQs in England, are typically broken down into units. Each unit comprises a set of learning outcomes and each learning outcome is characterised by a set of assessment criteria. We described this as a graphical approach to specifying a domain of learning. Each unit specification functions like the map of a learning journey, where each learning outcome corresponds to a significant landmark along the way. The main purpose of the learning journey is to traverse the entire map and to visit each landmark. Under the graphical approach, subject content is implicit in both learning outcomes and assessment criteria, although it may also be articulated explicitly via ancillary documents.

In England, the archetypal CASLO qualification is a VTQ. It is designed with flexibility in mind to accommodate learners studying under a variety of circumstances, for instance:

- learners following different pathways within the same qualification

- learners developing alternative (customised) skillsets within the same qualification

- learners studying in different locations (college-based, work-based, and so on)

- learners demonstrating their competence in different workplace settings

- learners starting and finishing their learning on different timelines

CASLO qualifications are naturally disposed to continuous centre-assessment (or continuous workplace-assessment). This is because they carry a heavy assessment burden, and they often need to accommodate flexibility in curriculum, teaching, learning, and assessment arrangements. It is therefore common for CASLO qualifications to rely on centre-assessment for all units. This substantial devolution of responsibility for assessment processes means that AOs need to be very confident in the ability of every centre to assess accurately and consistently. AOs therefore establish close relationships with centres to remain fully accountable for the awards that they make.

Requirements such as these have given rise to a different model of accountability with respect to centre-assessment compared with traditional GQs. We explained this metaphorically, by contrasting a hands-off sub-contracting model (GQ) with a hands-on partnership model (CASLO). According to this metaphor, centres are appointed as apprentices, to be enculturated into a community of practice for the qualification in question. Yet, however experienced a centre becomes, the AO will always maintain a watching brief over its practices via ongoing monitoring processes. The centre may transition from 'high risk' to 'low risk' but it should never be considered 'no risk'.

We observed that the archetypal CASLO qualification operates a tripartite centre-assessment quality assurance system, which comprises:

1. centre approval

2. internal quality assurance

3. external quality assurance

This implies a cradle-to-grave approach. which begins once a centre seeks approval to offer the qualification. It does not end until it has stopped offering the qualification and the last student has been certificated.

Information from our interviews with AO officers revealed sufficient commonalities across qualifications to propose several principles that appear to underpin the effective operation of CASLO qualifications:

1. structural integrity (centres need to demonstrate structural integrity from the outset – from a teaching, assessment, quality assurance, and management perspective – to be judged worthy to deliver the qualification)

2. self-regulation (centres need to operate as self-regulating systems, having internalised the AO's quality standards and quality assurance practices)

3. comprehensive monitoring (quality assurance needs to focus upon the effective delivery of a qualification in the round – teaching, assessment, quality assurance, and management)

4. risk-based sampling (given the potential scope of quality assurance activities, sampling needs to be driven by a risk-based model)

5. incremental improvement (centres need to be supported to improve incrementally – from centre approval onwards and for as long as a centre continues to offer a qualification)

6. supportive surveillance (the relationship between an AO and its centres – and between an IQA and their assessors – needs to remain fundamentally supportive)

7. conditional (evidence-based) trust (with increasing competence comes increasing trust – but, if a centre were to abuse that trust, then this would need to be dealt with via an appropriate sanction)

These principles are implemented within CASLO qualifications through a variety of scaffolding mechanisms. These include centre approval, training, standardisation, internal and external scrutiny of both processes and outcomes. Together they ensure the effective operation of these qualifications.

One of the most important scaffolding mechanisms relates to teaching and learning. This mechanism, which is often described as formative assessment, is specifically designed to ensure that learners remain continually on target to pass the qualification. With appropriate course entry requirements, effective selection processes, and adequate formative assessment arrangements, the expectation is that learners who exert enough effort ought to be able to reach competence.

The principle of self-regulation means that, if an IQA has done their job properly, then there should be no need for intervention by an EQA (following standards scrutiny). Equally, though, with this heavy emphasis upon formative assessment, if the teacher-assessor has done their job properly, then there should be no need for intervention by an IQA (following standards scrutiny). This helps to explain why AOs do not describe standards scrutiny for CASLO qualifications as fundamentally interventionist (which might be a reasonable description of Moderation for GQs).

Having said that, EQA standards scrutiny still serves a number of important functions. This includes a corrective function, based upon what might be described as a sporadically interventionist approach. The need to identify and rectify glaring errors, anomalies, and inaccuracies may still arise. This might include judgements of competence for which there is no available assessment evidence at all. Alternatively, it might include judgements of competence for which evidence is available but where

the decision is incompatible with any reasonable exercise of professional judgement. Such judgements ought, in theory, to be rare, given the number of scaffolding mechanisms in place.

Finally, even adopting a sporadically interventionist approach, EQA standards scrutiny will still serve an important deterrent function. It is a deterrent to maladministration attributable to inattentiveness, as well as a deterrent to malpractice arising from bad intent.

Outstanding questions

The intention of the present report was to construct, from first principles, an account of how CASLO qualifications work. We wanted to explain how they could work in theory rather than to evaluate how well they actually work in practice. We looked at how an AO could remain fully accountable for the awards it makes despite devolving substantial responsibility for assessment processes to centres. We attempted to make sense of how certain critical processes operate for CASLO qualifications – quality assurance processes in particular – by drawing comparisons with how ostensibly similar processes work for classical GQs, which tend to be understood far better.

Although we identified a number of tensions that could potentially arise when operating CASLO qualifications – under the headings of sampling, supporting, trusting, and investing – our intention was not to evaluate how effectively CASLO qualifications work.[31] As we noted at the outset, there is very little in the way of recent research into the operation of CASLO qualifications in England. This is especially from a quality assurance perspective. The research that does exist – based upon NVQs and GNVQs, with some based upon BTECs and related qualifications, tends to be dated.[32] It is open for debate whether conclusions based upon research into earlier qualifications should be expected to generalise to current CASLO qualifications, either the archetypal CASLO qualification or any particular suite of qualifications. However, there are certainly issues that are worth revisiting in relation to current qualifications and how effectively they may operate in practice.

Finally, the present report synthesised insights from conversations with AO officers to paint a picture – a hypothetical construction – of the archetypal CASLO

---

[31] Covid-19 prevented us from observing actual EQA practices. But, even before this set back, our intention was not specifically to evaluate those practices, merely to understand them.

[32] The following references are illustrative: Lester (1995); Eraut (1998); Garland (1998); Wolf (2001); Torrance, et al. (2005); Young (2007); Brockmann, et al. (2008); Johnson (2008); Ecclestone (2012); CAVTL (2013); Lum (2013); Hobley (2016); Wheelahan (2016); Carter & Bathmaker (2017); Winch (2021).

qualification. We noted, earlier, that where the actual features, processes, and circumstances of CASLO qualifications diverge from this hypothetical construction, this might raise questions for subsequent consideration. This research was unable to indicate much about the potential consequences of diverging from the archetypal model, given its very small sample size. However, we can still make suggestions concerning the kinds of divergences that might be worthy of comment and questions that might be asked.

Perhaps the most obvious example of divergence is the adaptation of the CASLO model to incorporate higher grades – that is, grades above the passing grade, typically including Merit and Distinction. A substantial number of CASLO qualifications do incorporate higher grades and an earlier Ofqual research project observed considerable differences in grading practices across CASLO qualifications (Newton, 2018). It remains to be determined whether there are significantly stronger or weaker approaches to incorporating higher grades within CASLO qualifications.

Another example of divergence concerns the use of CASLO qualifications within tightly time-bound contexts. The expectation that learners who exert enough effort ought eventually to reach competence is especially well suited to teaching and learning contexts that are no more than loosely time-bound. As such, if some learners need more time to reach competence than others, then this can be accommodated easily. The more tightly time-bound the context (for curriculum reasons, funding reasons, or other reasons) the more challenging it will become to accommodate learners departing significantly from nominal Total Qualification Time specifications. The issue, here, might concern the nature of the tensions that arise, the more tightly time-bound a CASLO qualification becomes, and how risks arising from these tensions can be mitigated.

Beyond divergences from the archetypal CASLO model, there will be all sorts of differences related to how particular AOs operate particular CASLO qualifications. Different circumstances will no doubt necessitate different trade-offs between the various scaffolding mechanisms that support their effective operation. For instance, some AOs will operate external standards scrutiny in a manner that is very close to (or that technically counts as) Moderation.[33] Other AOs will operate external standards scrutiny differently, within the broader definition of CASS. Some AOs will conduct most of their standards scrutiny remotely, via postal submission, or by accessing electronic portfolios. Other AOs will conduct all of their standards scrutiny in person, via face-to-face visits. The interesting and important issue here is how design decisions for the various features and processes within any particular qualification may complement (or fail to complement) each other. In other words,

---

[33] In the way that this is defined in Ofqual's General Conditions of Recognition – see earlier discussion.

how risks arising from the need to make one particular qualification design decision might be mitigated by compensatory decisions related to other aspects of its design. Once again, issues such as these are poorly documented, poorly researched, and poorly theorised for CASLO qualifications. The more that we can do to rectify this situation the better.

References

AQA (2016). *Moderation of Internal Assessments: An explanation for centres of how adjustments to marks are made (v1.10)*. Manchester: Assessment and Qualifications Alliance.

Beaumont, G. (1996). *Review of 100 NVQs and SVQs*. London: National Council on Vocational Qualifications.

Brockmann, M., Clarke, L. & Winch, C. (2008). Can performance-related learning outcomes have standards? *Journal of European Industrial Training*, 32(2/3), 99-113.

Carter, B., & Bathmaker, A. (2017). Prioritising progression over proficiency: limitations of teacher-based assessment within technician-level vocational education. *Journal of Further and Higher Education*, 41(4), 460-474.

Chambers, L., Williamson, J. & Child, S. (2019). Moderating artwork: Investigating judgements and cognitive processes. *Research Matters: A Cambridge Assessment publication*, 27, 19-25.

City & Guilds (2016). *Guidance on Internal Quality Assurance of Qualifications*. London: City & Guilds of the London Institute.

Commission on Adult Vocational Teaching and Learning (2013). *It's about work… Excellent adult vocational teaching and learning*. London: Learning and Skills Improvement Service.

Cuff, B.M.P. (2017). *An Exploratory Investigation Into How Moderators of Non-Examined Assessments Make Their Judgements*. Ofqual/17/6252. Coventry: Office of Qualifications and Examinations Regulation.

Cuff, B.M.P. (2018). *International Approaches to the Moderation of Non-Examination Assessments in Secondary Education*. Ofqual/18/6364. Coventry: Office of Qualifications and Examinations Regulation.

Cuff, B.M.P., Zanini, N. & Black, B. (2018). *An Exploration of Grade Inflation in 'Older Style' Level 3 BTEC Nationals: 2006 to 2016*. Coventry: Office of Qualifications and Examinations Regulation.

Curcin, M., Boyle, A., May, T. & Rahman, Z. (2014). *A Validation Framework for Work-Based Observational Assessment in Vocational Qualifications*. Ofqual/14/5374. Coventry: Office of Qualifications and Examinations Regulation.

Ecclestone, K. (1996). *How to Assess the Vocational Curriculum*. London: Kogan Page.

Ecclestone, K. (2012). Instrumentalism and achievement: A socio-cultural understanding of tensions in vocational education. In J. Gardner (Ed.). *Assessment and Learning* (pp.140-156). London: SAGE.

Eraut, M. (1998) Concepts of competence. *Journal of Interprofessional Care*, 12(2), 127-139.

FEFC (1995). *Assessment in Vocational Programmes in the Further Education Sector in England. National Survey Report*. Coventry: Further Education Funding Council.

FEFC (1997). *Standards and their Assurance in Vocational Qualifications. National Survey Report*. Coventry: Further Education Funding Council.

FEU (1992). *A Basis for Credit*? London: Further Education Unit.

FEU (1995). *A Framework for Credit: A common framework for post-14 education and training for the twenty-first century*. London: Further Education Unit.

Fisher, R. (2003). The golden age of BTEC: the business education curriculum in 1980s further education. *The Curriculum Journal*, 14(2), 253-277.

Garland, P. (1998) Assessment in GNVQs: learning the hard way. *Research in Post-Compulsory Education*, 3(3), 329-344.

Gill, T. (2015). The moderation of coursework and controlled assessment: A summary. *Research Matters: A Cambridge Assessment publication*, 19, 26-31.

Goff, P. & Leimanis, A. (1995). *Awarding Body Verification Procedures 1994-1995*. London: National Council on Vocational Qualifications.

Gravells, A. (2016). *Principles & Practices of Quality Assurance*. London: SAGE Learning Matters.

Greatorex, J. (2000). *What Research Could an Awarding Body Carry Out About NVQs*? A paper presented at the British Educational Research Association Conference, University of Cardiff, UK, 7-9 September.

Greatorex, J. (2002). *Two Heads are Better Than One: Standardising the judgements of National Vocational Qualification assessor*. A paper presented at the British Educational Research Association Conference, University of Exeter, UK, 12-14 September.

Harth, H. & Hemker, B. (2011). *On the Reliability of Results in Vocational Assessment: The case of work-based certifications*. Ofqual/11/4824. Coventry: Office of Qualifications and Examinations Regulation.

Harth, H. & van Rijn, P. (2010). *Reliability Issues in Competence-Based Assessment: Concepts and Estimates*. Coventry: Office of Qualifications and Examinations Regulation.

Hobley, J. (2016) 'Here's the iPad'. The BTEC philosophy: how not to teach science to vocational students. *Research in Post-Compulsory Education*, 21(4), 434-446.

Jessup, G. (1991). *Outcomes: NVQs and the Emerging Model of Education and Training*. London: The Falmer Press.

Johnson, M. (2008). Assessing at the borderline: Judging a vocationally related portfolio holistically. *Issues in Educational Research*, 18(1), 26-43.

Joint Awarding Bodies (2001). *Joint Awarding Body Guidance on Internal Verification of NVQs*. PP80D2/43926/0801/34. London: Department for Education and Skills.

Konrad, J. (2000). Assessment and verification of national vocational qualifications: policy and practice. *Journal of Vocational Education and Training*, 52(2), 225-243.

Lester, S. (1995). Professional pathways: A case for measurements in more than one dimension. *Assessment and Evaluation in Higher Education*, 20(3), 237-249.

Lockyer, C. & Cadwallader, S. (2020). *Internal Assessment in Existing National Technical and Vocational Qualifications*. Coventry: Office of Qualifications and Examinations Regulation.

Lum, G. (2013). Competence: A tale of two constructs. *Educational Philosophy and Theory: Incorporating ACCESS*, 45(12), 1193-1204.

NCVQ (1997). *The External Verification of NVQs: A guide for external verifiers*. London: National Council on Vocational Qualifications.

Newton, P.E. (2018). *Grading Vocational & Technical Qualifications: Recent policies and current practices*. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2008). *Regulatory arrangements for the Qualifications and Credit Framework*. Coventry: Office of the Qualifications and Examinations Regulator.

Ofqual (2016). *Changes to Regulations for Enquiries and Appeals*. Ofqual/16/6014. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2017). *Licence-linked Qualifications Used in the Private Security Industry: Regulatory Report.* Ofqual/17/6147. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2020a). *Awarding Organisation Controls for Centre Assessments – Regulations Consultation on Conditions, Requirements and Guidance.*

Consultation Decisions. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2020b). *Awarding Organisation Controls for Centre Assessments – Regulations Consultation on Conditions, Requirements and Guidance. Analysis Of Consultation Responses*. Coventry: Office of Qualifications and Examinations Regulation.

Ollin, R. & Tucker, J. (2012).*The Vocational Assessor Handbook* (5th edition). London: Kogan Page.

QCA (2006). *NVQ Code of Practice*. London: Qualifications and Curriculum Authority.

QCDA (2009). *The Quality and Standardisation of Units in the Qualifications and Credit Framework*. QCA/09/4281. Coventry: Qualifications and Curriculum Development Agency.

QCDA (2010a). *Guidelines for Writing Credit-Based Units of Assessment for the Qualifications and Credit Framework. Version 4*. QCDA/10/4725. Coventry: Qualifications and Curriculum Development Agency.

QCDA (2010b). *Guidance on Combining Knowledge and Skills in Units of Competence-Based Vocational Qualifications in the Qualifications and Credit Framework (QCF)*. QCDA/10/4856. Coventry: Qualifications and Curriculum Development Agency.

Raggatt, P. (1991). Quality assurance and NVQs. In P. Raggatt & L. Unwin (Eds.). *Change and Intervention: Vocational Education and Training* (pp.61-80). London: The Falmer Press.

Read, H. (2012). *The Best Quality Assurer's Guide*. Bideford: Read On Publications.

Shorrocks-Taylor, D. (1999). *National Testing: past, present and future*. Leicester: British Psychological Society Books.

Stevens, P. (1993). *City and Guilds of London Institute: A short history, 1878-1992*. London: City and Guilds of London Institute.

Taylor, M. (2005). *Teacher Moderation Systems*. London: National Assessments Agency.

Torrance, H., Colley, H., Garratt, D., Jarvis, J., Piper, H., Ecclestone, K. & James, D. (2005). *The Impact of Different Modes of Assessment on Achievement and Progress in the Learning and Skills Sector.* London: Learning and Skills Development Agency.

Warmington, P. & Wilmut, J. (2001). *The Roles of NVQ Internal Verifiers*. London: Department for Education and Employment

Wheelahan, L. (2016). Patching bits won't fix vocational education in Australia – a new model is needed. *International Journal of Training Research*, 14(3), 180-196.

Williamson, J. (2016). Statistical moderation of school-based assessment in GCSEs. *Research Matters: A Cambridge Assessment publication*, 22, 30-36.

Wilmut, J. & Tuson, J. (2005). *Statistical Moderation of Teacher Assessments*. London: Qualifications and Curriculum Authority.

Wilmut, J. (2005). *Experiences of Summative Teacher Assessment in the UK*. London: Qualifications and Curriculum Authority.

Wilson, P. (2010). *Big Idea, Small Steps: The making of credit-based qualifications*. Leicester: National Institute of Adult Continuing Education.

Winch, C. (2021). Learning outcomes: The long goodbye: Vocational qualifications in the 21st century. European Educational Research Journal.

Wolf, A. (1998). Portfolio assessment as national policy: the National Council for Vocational Qualifications and its quest for a pedagogical revolution. *Assessment in Education: Principles, Policy & Practice*, 5(3), 413-445.

Wolf, A. (2001). Competence-based assessment. In J. Raven & J. Stephenson (Eds.). *Competence in the Learning Society* (pp.453-466). New York: Peter Lang.

Wolf, A., Burgess, R., Stott, H., & Veasey, J. (1994). *GNVQ Assessment Review Project. Final Report.* London: Employment Department.

Young, M. (2007). Qualifications Frameworks: some conceptual issues. *European Journal of Education*, 42(4), 445-457.

Annex 1

The key differences between classical GQs and CASLO qualifications relate to how their learning domains and assessment standards are specified (and, consequently, assessed). Annex 1 illustrates these differences, using 2 Level 3 qualifications of similar size:

> AQA Level 3 Advanced GCE in Geography (601/8940/X) [34]

> Pearson BTEC Level 3 Diploma in Pet Services (603/2549/5) [35]

Both of these qualifications were available to learners, circa mid-2020, and both were classified as requiring 360 Guided Learning Hours. The A level in geography is assessed primarily by external written exam and is recorded as 360 hours Total Qualification Time. It is graded A* to E, else Unclassified. The Diploma in pet services is a Specialist BTEC, which is assessed exclusively by internal centre-devised assessments and is recorded as 430 hours Total Qualification Time. It only awards a Pass grade, else Unclassified. It is a work-related qualification, for learners who work in, or who want to work in the Pet Services sector.

Pearson BTEC Level 3 Diploma in Pet Services

The pet services BTEC comprises 6 units, of equal size (60 GLH).

Five units are mandatory:

1. Professional Practice for Pet Service Businesses

2. Business Management for Pet Service Businesses

3. Customer Service in a Pet Service Business

4. Animal Welfare

5. Industry based Research Project

One unit must be chosen from the following options:

1. Pet Boarding

2. Rehoming Animals

3. Professional Pet Services in Clients' Homes

---

[34] Information on this qualification was taken from: 'A-Level Geography (7037) Specification. For teaching from September 2016 onwards. For A-Level exams in 2018 onwards. Version 1.2 2 September 2019.'

[35] Information on this qualification was taken from: 'Pearson BTEC Level 3 Diploma in Pet Services. Specification – Issue 1 – September 2017 © Pearson Education Limited 2017.'

4. Animal Encounters

5. Pet Retail and Merchandising

6. Care of Aquatics

7. Care of Herptiles

The qualification specification document (see footnote) identifies the learning domain for the mandatory and optional units of this qualification, on a unit-by-unit basis. This domain is specified in 2 ways. First, via a set of LOs (with AC) for each unit. Second, via additional notes on "what needs to be learned" for each unit. These are illustrated below (first, the LOs for Unit 3, then, the notes for LO3 of Unit 3):

Unit 3: Customer Care in a Pet Service Business

Learning outcomes	Assessment criteria
LO 1. Explore how effective customer service contributes to business success	1.1 Describe the different approaches to customer service delivery 1.2 Examine ways that customer service in a selected business can meet the expectations and satisfaction of customers and adhere to relevant current legislation and regulations
LO 2. Demonstrate customer service in different situations, using appropriate behaviours to meet expectations	2.1 Demonstrate the appropriate communication and interpersonal skills to meet customer needs in different situations 2.2 Review own customer service skills, identifying gaps where improvements could be made 2.3 Present a clear, effective development plan for own customer service skills
LO 3. Investigate the methods used to improve customer service in a business	3.1 Research the methods that a business can use to monitor and make improvements to their customer service provision

LO 3: Investigate the methods used to improve customer service in a business

3.1 Monitoring and evaluating customer service provision

- Using research from customers to identify improvements and monitor complaints.

- Monitoring, using:

  o customer profiles; data, e.g. types of customer, products or services provided, customer care and service

  o sources of information, e.g. customers, colleagues, management

  o methods, e.g. questionnaires, comment cards, quality circles, suggestion boxes, staff surveys, mystery shoppers, recording and sharing information.

- Evaluating customer service, including:

  o analyse responses, e.g. level of customer satisfaction, quality of product or service, meeting regulatory requirements, balancing costs and benefits

  o planning for change, resolving problems/complaints.

3.2 Indicators in improved performance

- Reduction in numbers of complaints.

- Increase in profits.

- Reduction in turnover of staff.

- Repeat business from loyal customers.

After acquiring the learning outcomes for each unit, candidates are given an assignment, which may take a variety of forms. This includes practical and written formats. Assignments are issued to learners with a defined start date, completion date, and clear requirements for the evidence that they need to provide. By way of example, Pearson recommends the following formats for assignments targeting each of the 3 LOs of Unit 3:

LO1: A leaflet for employees outlining what customer care is, how to maintain it and the benefits of carrying it out effectively.

LO2: A reflective essay explaining how a range of customer care situations were handled and how this could be improved in future situations.

LO3: A survey of customers and an action plan to implement customer care improvements.

The following quotations from the specification document illustrate the non-compensatory approach to aggregating assessment judgements:

> The assessor needs to make a judgement against each criterion that evidence is present and sufficiently comprehensive.

> When a learner has completed the assessment for a unit then the assessor will give an assessment outcome for the unit. To achieve a Pass, a learner must have satisfied all the assessment criteria for the learning outcomes, showing appropriate coverage of the unit content and therefore attainment at the stated level of the qualification. The award of a Pass is a defined level of performance and cannot be given solely on the basis of a learner completing assignments. Learners who do not satisfy the assessment criteria for the units should be reported as Unclassified.

AQA Level 3 Advanced GCE in Geography

Part of the learning domain for this qualification – its core content, which is common across all geography A levels – is specified centrally, by the Department for Education, in a document entitled: 'Geography: GCE AS and A Level subject content. December 2014'. This includes 4 core knowledge and understanding themes (plus geographical skills and fieldwork requirements):

1. water and carbon cycles

2. landscape systems

3. global systems and global governance

4. changing place, changing places

The notes for each content theme run to a couple of sides of A4. This core content is required to represent 60% of the A level. AOs develop their own specifications but must incorporate the core content.

The AQA qualification specification document (see footnote) identifies the full learning domain for this qualification. It is laid out by area, rather than on a unit-by-unit basis. The 11 content areas are listed below, alongside additional fieldwork investigation and skills requirements. Some of this content is mandatory. Thus, all learners are required to study '1. Water and carbon cycles'. Other content is optional. Thus, learners are required to study either '5. Hazards' or '6. Ecosystems under stress'.

The AQA specification lists the full A level learning domain as follows:

Physical geography

1. Water and carbon cycles

2. [Option set 1] Hot desert systems and landscapes

3. [Option set 1] Coastal systems and landscapes

4. [Option set 1] Glacial systems and landscapes

5. [Option set 2] Hazards

6. [Option set 2] Ecosystems under stress

Human geography

7. Global systems and global governance

8. Changing places

9. [Option set 3] Contemporary urban environments

10. [Option set 3] Population and the environment

11. [Option set 3] Resource security

Geography fieldwork investigation

12. Fieldwork requirements

13. Investigation requirements

Geographical skills

14. Geographical skills checklist

The AQA specification exemplifies each content area like this (for area '5. Hazards'):

3.1.5 Hazards This optional section of our specification focuses on the lithosphere and the atmosphere, which intermittently but regularly present natural hazards to human populations, often in dramatic and sometimes catastrophic fashion. By exploring the origin and nature of these hazards and the various ways in which people respond to them, students are able to engage with many dimensions of the relationships between people and the environments they occupy. Study of this section offers the opportunity to exercise and develop observation skills, measurement and geospatial mapping skills, together with data manipulation and statistical skills, including those associated with and arising from fieldwork.

3.1.5.1 The concept of hazard in a geographical context Nature, forms and potential impacts of natural hazards (geophysical, atmospheric and hydrological). Hazard perception and its economic and cultural determinants. Characteristic human responses – fatalism, prediction, adjustment/adaptation, mitigation, management, risk sharing – and their relationship to hazard incidence, intensity, magnitude, distribution and level of development. The Park model of human response to hazards. The Hazard Management Cycle.

3.1.5.2 Plate tectonics Earth structure and internal energy sources. Plate tectonic theory of crustal evolution: tectonic plates; plate movement; gravitational sliding; ridge push, slab pull; convection currents and seafloor spreading. Destructive, constructive and conservative plate margins. Characteristic processes: seismicity and vulcanicity. Associated landforms: young fold mountains, rift valleys, ocean ridges, deep sea trenches and island arcs, volcanoes. Magma plumes and their relationship to plate movement.

3.1.5.3 Volcanic hazards The nature of vulcanicity and its relation to plate tectonics: forms of volcanic hazard: nuées ardentes, lava flows, mudflows, pyroclastic and ash fallout, gases/acid rain, tephra. Spatial distribution, magnitude, frequency, regularity and predictability of hazard events. Impacts: primary/secondary, environmental, social, economic, political. Short and long-term responses: risk management designed to reduce the impacts of the hazard through preparedness, mitigation, prevention and adaptation. Impacts and human responses as evidenced by a recent volcanic event.

3.1.5.4 Seismic hazards The nature of seismicity and its relation to plate tectonics: forms of seismic hazard: earthquakes, shockwaves, tsunamis, liquefaction, landslides. Spatial distribution, randomness, magnitude, frequency, regularity, predictability of hazard events. Impacts: primary/secondary; environmental, social, economic, political. Short and long-term responses; risk management designed to reduce the impacts of the hazard through preparedness, mitigation, prevention and adaptation. Impacts and human responses as evidenced by a recent seismic event.

3.1.5.5 Storm hazards The nature of tropical storms and their underlying causes. Forms of storm hazard: high winds, storm surges, coastal flooding, river flooding and landslides. Spatial distribution, magnitude, frequency, regularity, predictability of hazard events. Impacts: primary/secondary, environmental, social, economic, political. Short and long-term responses: risk management designed to reduce the impacts of the hazard through preparedness, mitigation, prevention and adaptation. Impacts and human responses as evidenced by two recent tropical storms in contrasting areas of the world.

3.1.5.6 Fires in nature Nature of wildfires. Conditions favouring intense wild fires: vegetation type, fuel characteristics, climate and recent weather and fire behaviour. Causes of fires: natural and human agency. Impacts: primary/secondary, environmental, social, economic, political. Short and long-term responses; risk management designed to reduce the impacts of the hazard through preparedness, mitigation, prevention and adaptation. Impact and human responses as evidenced by a recent wild fire event.

3.1.5.7 Case studies Case study of a multi-hazardous environment beyond the UK to illustrate and analyse the nature of the hazards and the social, economic

and environmental risks presented, and how human qualities and responses such as resilience, adaptation, mitigation and management contribute to its continuing human occupation. Case study at a local scale of a specified place in a hazardous setting to illustrate the physical nature of the hazard and analyse how the economic, social and political character of its community reflects the presence and impacts of the hazard and the community's response to the risk.

In addition to this content specification, AQA also provides additional teaching resources, including approved textbooks, schemes of work, and lesson plans.

Assessment requirements for A level geography are specified by Ofqual, in a document entitled: 'GCE Subject Level Conditions and Requirements for Geography. March 2015.' These include requirements concerning the amount of Non-Exam Assessment that is permitted (20%, for the fieldwork investigation) and requirements concerning the nature and balance of Assessment Objectives (AObs): demonstrating knowledge and understanding (30-40%), applying knowledge and understanding (30-40%), and using skills (20-30%).

Assessment criteria	Level 4 10-9 marks	Level 3 8-6 marks	Level 2 5-3 marks	Level 1 2-1 marks	Level 1 0 marks
To define the research questions which underpin field investigation. (AOb3)	A research question(s) is effectively identified and is completely referenced to the specification	A research question(s) is securely identified that is explicitly linked to the specification	A research question(s) which is partial. Links to the specification are imprecise	A research question(s) is generalised. Links to the specification are tentative	Does not meet criteria
To research relevant literature sources and understand and write up the theoretical or comparative context for a research question. (AOb3)	Well-supported by thorough use of relevant literature sources. Theoretical and comparative contexts are well-understood and well-stated	Supported by focused use of relevant literature sources. Theoretical and comparative contexts are consistently understood and stated	Supported by some use of relevant literature sources. Theoretical and comparative contexts are inconsistently stated	Limited or basic use of relevant literature sources. Theoretical and comparative contexts are isolated	Does not meet criteria

Table A1. Mark scheme for Area 1 of AQA geography NEA.

The AQA A level geography qualification is assessed (at the end of the course) via 3 components:

1. physical geography – written exam (120 marks)

2. human geography – written exam (120 marks)

3. fieldwork – investigation report of 3,000 to 4,000 words (60 marks)

The written components assess a similar balance of AObs (focused mainly on the first and second objective). This is different from the balance assessed by the fieldwork investigation (focused mainly on the third objective).

The fieldwork investigation is a centre-assessed, Non-Exam Assessment (NEA), which is externally moderated. Teachers are required to provide a separate mark for

each fieldwork area – introduction, investigation methods, analysis methods, conclusions – using a levels-based mark scheme for each area. This is illustrated in Table A1, for Area 1 (introduction), which is marked out of 10.

Annex 2

Three related research projects undertaken by researchers at Ofqual provide additional context for the present investigation:

1. grade inflation in (CASLO) BTEC qualifications

2. grading practices within VTQs

3. internal assessment within VTQs

BTEC grade inflation

Ofqual's grade inflation research (Cuff, Zanini, & Black, 2018) was undertaken in the context of a steep rise in uptake of Level 3 BTECs over the preceding decade. This was at the same time as a steep rise in the percentage of the BTEC cohort achieving the highest grades. Especially in the wake of action taken to halt grade inflation for A levels, from 2010 onwards, this prompted the question of whether grade inflation might also have been occurring for BTECs. The report concluded fairly definitively that there had been an element of grade inflation for BTECs.

Although it had not investigated this directly, the report speculated that grade inflation might be explicable in terms of the heavy use of internal assessment within (CASLO) BTECs. For instance:

> These findings might be explained by differences in the marking/awarding processes that exist for 'older style' L3 BTECs and A levels. For example, while a compensatory approach is taken for A levels, 'older style' L3 BTECs are graded according to firm criterion referencing (firm in the sense that candidates must be deemed to have achieved all pass criteria to achieve a pass, and all merit criteria to achieve a merit, etc.). As this approach does not allow for any adjustment of grade boundaries (there are no 'marks'), these criteria set the standard, and so become the method for standards maintenance. Arguably, because of accountability measures, teachers involved in grading have a vested interest in increasing outcomes over time, which this method cannot control for. Ultimately, this method is vulnerable to pressures of grade inflation.

> (Cuff, et al, 2018, p.14)

> Indeed, while this research has focused upon 'older style' L3 BTECs, it is possible that these findings may generalise to other similar qualifications showing grading increases (ie those operating similar models of standards maintenance in similar contexts).

> (Cuff, et al, 2018, p.7)

Thus, the report raised general questions concerning how CASLO qualifications work, including their robustness to standards drift.

Grading practices

Ofqual's research into grading VTQs (Newton, 2018) was conducted largely in parallel with the BTEC grade inflation research. It was published at around the same time. As the first survey of its kind, it adopted a 'deep-dive' approach. It explored in detail how each of 18 sampled qualifications operated – many of which were CASLO qualifications – based upon documentary analysis, supplemented by conversations with AO officers.

The research identified a wide variety of practices, which were classified and discussed in terms of their underlying measurement models and how they represented their measurement standards. Questions were identified relating to a variety of fundamental technical issues. These included: standardisation, grading and levelling, comparability, weighting, burden and backwash, and transparency. The main conclusion was that VTQ grading, in England, is not underpinned by a straightforward, generally accepted, set of principles governing good practice. This raised the question of what such principles might look like.

Standardisation of assessor judgement was foregrounded as a major potential threat to validity for CASLO qualifications – both at the passing grade and at higher ones too. This raised questions concerning how AOs might address such threats, when running CASLO qualifications.

Internal assessment

Ofqual's internal assessment project (Lockyer & Cadwallader, 2020) aimed to understand how internal assessment works for regulated VTQs, based upon interviews with teacher-assessors. This was based upon a sample of 6 qualifications, from Level 2 and Level 3, in Construction, Hair & Beauty, and Information Technology. It included both 'strongly' and 'softly' vocational qualifications, that is, those leading directly to employment and those that tended to lead to further study.

Very much in line with the grading practices study, one of the most striking observations from this study was just how diverse, complex and dynamic the VTQ sector is. The sampled qualifications reflected different learners, studying in different settings, with very different purposes. There was also variation in how centres delivered the same VTQs.

The report made several observations related specifically to quality assurance:

Various assessors delivering this particular course explained the purpose of EQA is not to quality assure summative assessments as such, but instead to ensure that a centre is capable of making and assuring accurate and reliable judgements.

(Lockyer & Cadwallader, 2020, p.45)

When we asked assessors to describe EQA, we anticipated responses would focus on its monitoring function. In fact, assessors perceived EQA to have another vital role. In 14 of the 20 centres, assessors discussed the support they receive as part of EQA. Seven assessors believed centre support was a central aim of the EQA process. Some felt the EQA was less of a check of centre performance and more of a form of centre reassurance and assistance.

(Lockyer & Cadwallader, 2020, p.47)

Assessors in 15 centres discussed whether they had ever had their assessment decisions adjusted or overturned. Nine told us they had no experience of changes to assessment decisions in these specific VTQs.

"She could actually overturn something if one of us didn't make the right decision. But I don't think that's ever happened. I mean all the times I've been here and at my last college we had [AO name] and they never really overturned anything we've decided on."

(Lockyer & Cadwallader, 2020, p.49)

Annex 3

The following questions formed the basis for the interviews with AO officers.

1.   How unit standards are established

You've adopted the QCF-style approach to outlining units for this qualification, by laying them out in terms of learning outcomes and assessment criteria, and by indicating that all learning outcomes need to be achieved to pass each unit…

1.1.1   Is that simply because this used to be a QCF qualification, so you had to adopt this approach, or because you think that this is the best approach for a qualification like this (or both, or neither)?

The learning outcomes and assessment criteria specified for each unit outline their content and standards…

1.2.1   How was it decided what the learning outcomes and assessment criteria for each unit needed to be?

2.   How assessors apply unit standards

Assessors need to have a deep (and consensual) understanding of the standards for each unit, in order to make accurate and consistent assessment judgements…

2.1.1   Is it essential/desirable for assessors to have occupational or sector expertise, in order to be able to apply the unit standards accurately and consistently?

2.1.2   Is there a professional 'community of practice' for this qualification (an organisation or network with a sense of ownership of its standards) and would assessors consider themselves to be members of this community?

2.1.3   In addition to the outcomes and criteria for each unit, do you provide assessors with any other 'tools' for helping them to apply the unit standards (eg guidance, exemplification materials, marking schemes)?

Where assessors judge students' performances using the assessment criteria that are associated with each learning outcome…

2.2.1   Are assessors expected to exercise judgement primarily at the level of the learning outcome, or primarily at the level of the assessment criteria?

2.2.2   Are assessors permitted or encouraged to give students the 'benefit of the doubt' under any circumstances (if so, then under what circumstances, and why might that be a useful thing to do)?

2.2.3   Is it an expectation that most if not all students will have achieved competence in a particular learning outcome (or set of learning outcomes) before they are formally assessed on it (them)?

2.2.4 What is it that provides you with confidence that all assessors are able to understand and apply unit/qualification standards in the same way?

3.   How quality assurers quality assure

Quality assurance often involves a variety of inter-related processes (centre-approval, internal standardisation, external scrutiny, etc.)…

3.1.1 If you were to list the quality assurance processes that you considered to be most critical to the credibility of the qualification, what would they be?

3.1.2 What are the 'secrets' to how quality assurance works for this qualification (ie, features/reasons that might not be obvious to the casual observer)?

One of the key components of quality assurance is (what Ofqual now refers to as) Standards Scrutiny, which involves an external quality assurer scrutinising samples of centre-assessed work…

3.2.1 When an external quality assurer scrutinises samples of centre-assessed work, for this qualification, how would you describe what they are intending to achieve through this process?

3.2.2 Would you say that the process is quite similar to, or quite different from, what GCSE moderators do when they scrutinise samples of assessed work (based upon 'second marking' samples of centre-assessed work)?

3.2.3 How often do external quality assurers tend to identify significant, or serious, problems with assessors' judgements, for this qualification?

3.2.4 How does the Standards Scrutiny process work for centres that have DCS (ie, if certificates can be claimed before Standards Scrutiny has happened)?

3.2.5 What is it that provides you with confidence that all external quality assurers are able to understand and apply unit/qualification standards in the same way?

4.   The QCF-style approach

Reflecting on the QCF-style approach more generally (its emphasis on internal assessment, its specification in terms of LOs or AC, and its emphasis on mastery)…

4.1.1 What are its main strengths (for teaching, or learning, or assessing)?

4.1.2 Does it present any significant challenges (for teaching, or learning, or assessing), and how are these best overcome?

4.1.3 Is the QCF-style approach especially well-suited to certain circumstances (eg, certain purposes, certain learners, certain contexts)?

5.   Background information

We'll use the final section of our interview to make sure that we've captured all of the background information that we're hoping to gather, related to your qualification and

how it operates. Most of this has already been captured – using an Excel template – based upon the documents that you've already provided. We'll use the final section of this interview to fill-in any remaining gaps.

Annex 4

Table A2 describes the structure of each of the sampled qualifications in terms of units and LOs, identifying the precise cognitive or behavioural requirements associated with each LO statement. These summaries were based on the full set of units for each qualification, although, where an option choice was required, example units were chosen.

Qual. Name	Type
City & Guilds Level 2 Diploma in Health and Social Care (Adults) for England	Nine mandatory units: • 23 GLH, 4 LOs: 1 x Understand, 3 x Be able to [14 AC] • 23 GLH, 4 LOs: 1 x Understand, 3 x Be able to [13 AC] • 20 GLH, 3 LOs: 1 x Understand, 1 x Know how to, 1 x Be able to [8 AC] • 9 GLH, 3 LOs: 2 x Understand, 1 x Know how to [6 AC] • 26 GLH, 5 LOs: 2 x Understand, 3 x Know how to [15 AC] • 14 GLH, 3 LOs: 1 x Understand, 2 x Be able to [9 AC] • 33 GLH, 6 LOs: 1 x Understand, 5 x Be able to [20 AC] • 33 GLH, 9 LOs: 4 x Understand, 2 x Know how to, 3 x Be able to [26 AC] • 10 GLH, 3 LOs: 1 x Understand, 1 x Know how to, 1 x Be able to [6 AC] One from Optional B (example illustrated below): • 14 GLH, 2 LOs: 2 x Know [8AC] Five from Optional C (example illustrated below): • 24 GLH, 3 LOs: 1 x Understand, 2 x Be able to [12 AC] • 19 GLH, 2 LOs: 2 x Be able to [8 AC] • 23 GLH, 5 LOs: 1 x Understand, 4 x Be able to [20 AC] • 19 GLH, 5 LOs: 1 x Understand, 4 x Be able to [18 AC] • 20 GLH, 3 LOs: 1 x Understand, 2 x Be able to [20 AC]
CMI Level 5 Diploma in Management and Leadership	8 optional units (examples illustrated below): • 25 GLH, 3 LOs: 3 x Understand [12 AC] • 25 GLH, 3 LOs: 2 x Understand, 1 x Know how to [8 AC] • 18 GLH, 4 LOs: 2 x Understand, 1 x Know how to, 1 x Know [8 AC] • 16 GLH, 2 LOs: 2 x Understand [6 AC] • 24 GLH, 4 LOs: 1 x Understand, 3 x Know how to [13 AC] • 22 GLH, 4 LOs: 3 x Understand, 1 x Know how to [9 AC] • 23 GLH, 3 LOs: 2 x Understand, 1 x Know how to [9 AC] • 16 GLH, 2 LOs: 1 x Understand, 1 x Know how to [8 AC]

Qual. Name	Type
Highfield Level 2 Certificate In Customer Service (RQF)	Two mandatory units: • 50 GLH, 5 LOs: 1 x Describe, 2 x Understand, 2 x Identify [32 AC] • 65 GLH, 4 LOs: 2 x Apply, Demonstrate, Communicate [28 AC]
UAL Level 3 Diploma In Creative Media Production & Technology	Eight mandatory units: • 12 GLH, 2 LOs: 2 x Understand [3 AC] • 10 GLH, 3 LOs: 2 x Understand, 1 x Be able to [4 AC] • 10 GLH, 3 LOs: 2 x Understand, 1 x Be able to [5 AC] • 9 GLH, 2 LOs: 2 x Understand [5 AC] • 9 GLH, 3 LOs: 3 x Be able to [5 AC] • 9 GLH, 3 LOs: 3 x Be able to [5 AC] • 9 GLH, 3 LOs: 3 x Be able to [5 AC] • 36 GLH, 4 LOs: 1 x Understand, 3 x Be able to [5 AC]
QA Level 3 Award in Emergency Paediatric First Aid (RQF)	One mandatory unit: • 6 GLH, 7 LOs: 1 x Understand, 2 x Know how to, 4 x Be able to [19 AC]
EAL Level 2 Certificate In Applying Business-Improvement Techniques	Seven mandatory units: • 6 GLH, 1 LO: 1 x Understand how to contribute to [3 AC] • 12 GLH, 1 LO: 1 x Understand how to contribute to [7 AC] • 14 GLH, 1 LO: 1 x Understand how to contribute to [7 AC] • 12 GLH, 1 LO: 1 x Understand how to contribute to [7 AC] • 8 GLH, 1 LO: 1 x Understand how to contribute to [7 AC] • 8 GLH, 1 LO: 1 x Understand how to contribute to [7 AC] • 10 GLH, 1 LO: 1 x Understand how to contribute to [7 AC]

Table A2. Structure of LOs in CASLO units.

February 2022   Ofqual/22/6895